Plant Genetics and Genomics: Crops and Models 10

Ray Ming
Paul H. Moore *Editors*

Genetics and Genomics of Papaya

Springer

Plant Genetics and Genomics: Crops and Models

Volume 10

Series Editor
Richard A. Jorgensen

For further volumes:
http://www.springer.com/series/7397

Ray Ming • Paul H. Moore
Editors

Genetics and Genomics of Papaya

Springer

Editors
Ray Ming, PhD
Department of Plant Biology
University of Illinois at Urbana-Champaign
Urbana, IL, USA

Paul H. Moore, BA, MS, PhD
Hawaii Agriculture Research Center
Kaneohe, HI, USA

ISBN 978-1-4614-8086-0 ISBN 978-1-4614-8087-7 (eBook)
DOI 10.1007/978-1-4614-8087-7
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013942494

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

We dedicate this book to all the teachers, mentors, and colleagues who instilled in each of us the thrill of scientific discovery. This thrill has helped us strive to understand the natural world around us while constantly reminding us how clever nature has been in solving complex biological problems.

Preface

At first glance, papaya might seem an unlikely crop for inclusion in the Springer Series on Plant Genomics. Although papaya has long been widely grown and consumed in the tropical world and its popularity as a mainstream fruit crop has grown in recent years, much of the temperate world still regards papaya as a relatively exotic and obscure food. On the other hand, papaya has the distinction of being the fifth angiosperm genome to be sequenced and the first transgenic crop to be characterized at the whole genome level. Why and how has papaya achieved these scientific milestones? This book reveals that story, which begins with a virus disease for which there was no natural resistance.

In the early 1990s, the Hawaiian papaya industry was threatened with collapse when the crop was found to be heavily infected by papaya ringspot virus (PRSV), for which the crop has no resistance. In one of the most widely publicized success stories of genetically modified crops (see Chap. 7 of this book), papaya was transformed and the Hawaiian industry was saved. Transgenic papaya became a kind of poster child for the safety and efficacy of genetic transformation of a consumed fruit. That popularity, in addition to several other qualities, led to papaya becoming a focus for genomic research. Included among those other qualities are its nutritional value (one fruit provides 122 % of the U.S. Recommended Daily Allowance for vitamin A and 314 % for vitamin C); its medicinal applications (including recent findings of its potential in cancer treatment); and its small genome of 372 Mb. In the course of genetic research, it was also discovered that papaya has nascent sex chromosomes that make it a model for studying sex chromosome evolution in flowering plants. In short, this relatively obscure plant is far more interesting and scientifically valuable than anyone might have guessed.

This book is intended to provide the most up-to-date knowledge of papaya genetics and genomics. We hope it will stimulate current and future researchers to explore papaya's fundamental biology and its nutritional and medicinal properties for further improvement of this undervalued crop.

We thank Richard Jorgenson for envisioning this book and Liz Corra and Hannah Smith for their assistance.

Urbana, IL, USA Ray Ming
Kaneohe, HI, USA Paul H. Moore

Contents

Contributors

Rishi Aryal Department of Plant Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA

Anjanabha Bhattacharya Bench Bio Pvt Ltd, Vapi, Gujarat, India

Fernanda Antunes Carvalho Department of Systematic Botany and Mycology, University of Munich, Munich, Germany

Manash Chatterjee Bench Bio Pvt Ltd, Vapi, Gujarat, India

Nancy J. Chen Department of Tropical Plant and Soil Science, University of Hawaii at Manoa, Honolulu, HI, USA

David A. Christopher Department of Molecular Bioscience and Bioengineering, University of Hawaii at Manoa, Honolulu, HI, USA

Geo Coppens d'Eeckenbrugge Centre de Coopération Internationale en Recherche Agronomique pour le Développement, UMR 5175 CEFE, Montpellier, France

Rod Drew School of Biomolecular and Physical Sciences, Griffith University, Nathan, QLD, Australia

O.P. Dutta Department of Research and Development, Namdhari Seeds Pvt Ltd, Bangalore, Karnataka, India

Rebecca Ford Department of Agriculture and Food Systems, Melbourne School of Land and Environment, Melbourne, VIC, Australia

Gabriela Fuentes Unidad de Biotecnología, Centro de Investigación Científica de Yucatán AC (CICY), Colonia Chuburná de Hidalgo, Mérida, México

Dennis Gonsalves US Department of Agriculture, Agricultural Research Service, Pacific Basin Agricultural Research Center, Hilo, HI, USA

Marco V. Gutiérrez-Soto Estación Experimental Agrícola Fabio Baudrit Moreno, Universidad de Costa Rica, Alajuela, Costa Rica

Jennifer Han Department of Plant Biology, University of Illinois at Urbana Champaign, Urbana, IL, USA

Jiming Jiang Department of Horticulture, University of Wisconsin-Madison, Madison, WI, USA

Victor M. Jiménez CIGRAS, Universidad de Costa Rica, San Pedro, Costa Rica
Food Security Center, University of Hohenheim, Stuttgart, Germany

Chutchamas Kanchana-udomkan School of Biomolecular and Physical Sciences, Griffith University, Nathan, QLD, Australia

P.K. Anish Kumar Bench Bio Pvt Ltd, Vapi, Gujarat, India

Tina Kyndt Department of Molecular Biotechnology, Ghent University, Ghent, Belgium

Eric Lyons School of Plant Sciences, University of Arizona, Bio5 Institute, Tucson, AZ, USA

Richard Manshardt Department of Tropical Plant and Soil Sciences, University of Hawaii at Manoa, Honolulu, HI, USA

Ray Ming Department of Plant Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA

Paul H. Moore Hawaii Agriculture Research Center, Texas AgriLife Research, Kaneohe, HI, USA

Jong-Kuk Na Department of Agricultural Bio-Resources, National Academy of Agricultural Science, RDA, Suwon, Republic of Korea

Niranjan Nagarajan Department of Computational and Systems Biology, Genome Institute of Singapore, Singapore

Rafael Navajas-Pérez Facultad de Ciencias, Departamento de Genética, Universidad de Granada, Granada, Spain

Eric Mora-Newcomer Estación Experimental Agrícola Fabio Baudrit Moreno, Universidad de Costa Rica, Alajuela, Costa Rica

Timothy J. O'Hare QAAFI, Centre for Nutrition and Food Sciences, University of Queensland, Gatton, QLD, Australia

Robert E. Paull Department of Tropical Plant and Soil Sciences, University of Hawaii at Manoa, Honolulu, HI, USA

Brad W. Porter Department of Molecular Biosciences and Bioengineering, University of Hawaii at Manoa, Honolulu, HI, USA

Susanne S. Renner Department of Systematic Botany and Mycology, University of Munich, Munich, Germany

Jorge M. Santamaría Unidad de Biotecnología, Centro de Investigación Científica de Yucatán AC (CICY), Colonia Chuburná de Hidalgo, Mérida, Mexico

Xavier Scheldeman Regional Office for the Americas, Bioversity International, Cali, Colombia

Haibao Tang Center for Plant Genomics and Biotechnology, Fujian Agriculture and Forestry University, Fuzhou, China

Savarni Tripathi US Pacific Basin Agricultural Research Center, US Department of Agriculture, Hilo, HI, USA

Robert VanBuren Department of Plant Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA

Marisa M. Wall US Pacific Basin Agricultural Research Center, US Department of Agriculture, Hilo, HI, USA

Jianping Wang Department of Agronomy, Genetics Institute, Plant Molecular and Cellular Biology Program, University of Florida, Gainesville, FL, USA

David J. Williams Agri-Science Queensland, Department of Agriculture, Fisheries and Forestry, Brisbane, Australia

Pingfang Wu Department of Tropical Plant and Soil Sciences, University of Hawaii at Manoa, Honolulu, HI, USA

Qingyi Yu Texas A&M AgriLife Research Center at Dallas, Texas A&M University System, Dallas, TX, USA

Wenli Zhang Department of Horticulture, University of Wisconsin-Madison, Madison, WI, USA

Yun J. Zhu Hawaii Agriculture Research Center, Kunia, HI, USA

Part I
Natural History and Genetic Diversity

Chapter 1
Papaya (*Carica papaya* L.): Origin, Domestication, and Production

Gabriela Fuentes and Jorge M. Santamaría

Papaya Origin

Given that no direct archeological evidence is available, it is rather difficult to define a precise origin for *C. papaya* L. We will revise in this chapter a series of indirect pieces of evidences that might support the idea that *C. papaya* L. has its origins in South Mexico and/or in Central America. We will discuss a geobotanical approach and certain evidence based on the current existence of wild populations in southern Mexico and Central America. Studies based on molecular markers will be discussed in other chapters (Chap. 5 in this volume) of this book.

It is difficult to know the precise origin of *Carica papaya* L. because nowadays, given its wide distribution by the Spaniards and its great capacity of adaptation to the conditions of subtropical and tropical environments, it is widely distributed around most of the subtropical and tropical regions of the world. However, Vavilov (1987) described three centers of origin of the majority of the species: the Mesopotamic center, the Mesoamerican Center, and the North Chinese center. The Mesoamerican Center is the center of origin of important tropical crops, and it has been suggested as being a good candidate to also be the center of origin of *C. papaya* L. (Harlan 1971).

C. papaya L., described by von Linnaeus (1753), belongs to the Caricaceae family that is formed by 6 genera and 35 species. The genera belonging to this family are Carica (1 specie), Jarilla (3 species), Horovitzia (1 specie), Jacaratia (7 species), Vasconcellea (21 species), and Cylicomorpha (2 species), according to Badillo (1971, 1993, 2000). Except for the latter genus, all the other five genera of this family have an American origin (Scheldeman et al. 2011). Some authors had suggested that *C. papaya* L. originated in the north of South America (Badillo 1971; Prance 1984).

G. Fuentes • J.M. Santamaría (✉)
Unidad de Biotecnología, Centro de Investigación Científica de Yucatán AC (CICY), Calle 43 No. 130, Colonia Chuburná de Hidalgo, Mérida 97200, México
e-mail: jorgesm@cicy.mx

R. Ming and P.H. Moore (eds.), *Genetics and Genomics of Papaya*, Plant Genetics and Genomics: Crops and Models 10, DOI 10.1007/978-1-4614-8087-7_1,

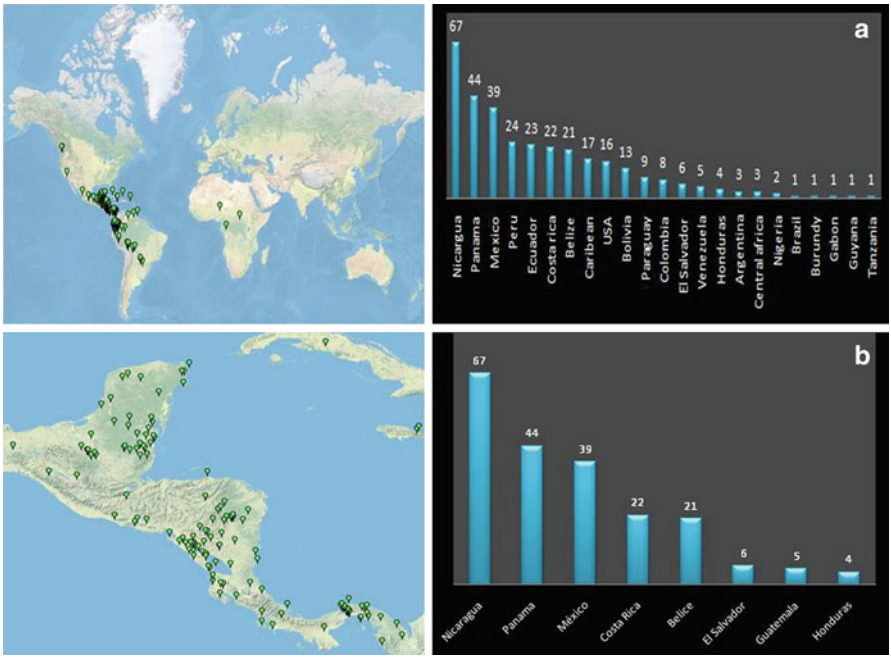

Fig. 1.1 Map distribution of records from Herbarium specimens of *Carica papaya* L. from the Missouri Botanic Gardens Herbarium database (2011). (**a**) Shows worldwide records (339 specimens) and (**b**) shows records from specimens collected at Central America (particularly Nicaragua) and the South of Mexico (208 specimens), representing 61 % of all the *C. papaya* L. records reported in the database (http//www.tropicos.org)

However, while this seems true for other members of the family Caricaceae (i.e., Vasconcellea), it is increasingly accepted that *C. papaya* L. had its origins in the south of Mexico and Central America. For instance, as early as 1833, De Candolle (1883) and Solms-Laubach (1889) suggested that *C. papaya* L. originated in Mexico. Besides, two other genera of the *Caricaceae* family are considered endemic to Mexico: *Horovitzia*, which is considered endemic to Mexico (Lorence and Colin 1988; Badillo 1993), and *Jarilla*, which is reported as endemic from Mexico and Guatemala (McVaugh 2001).

Taking records of *C. papaya* L. specimens from herbarium collections from around the world, it is interesting that most of the herbarium specimens had been collected in Central America and/or the south of Mexico. For instance, the collection of the Missouri Botanic Gardens Herbarium (2011) had 339 specimens of *C. papaya* L. from various parts of the world (Fig. 1.1a); however, more than 60 % of those specimens (208) were collected in Mexico and Central America (particularly in Nicaragua) (Fig. 1.1b). Moreover, in a more extensive collection at the Global Biodiversity Facility (GBIF) (2011), a total of 1297 specimens of *C. papaya* L. are reported (Fig. 1.2a), and again, more than 50 % of them (659) are occurrences from Central America and the south of Mexico (Fig. 1.2b).

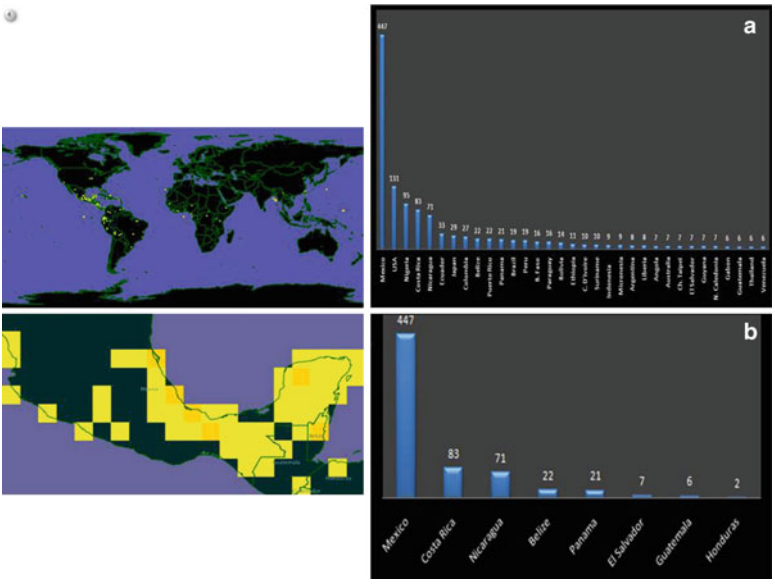

Fig. 1.2 Map distribution of occurrences (1297; *yellow dots*) recorded from around the world for *C. papaya* L., according to the Global Biodiversity Information Facility (GBIF) (**a**). The high concentration of records in the south of Mexico and Central America (659) represents 51 % of all the *Carica papaya* L. records reported in the database (**b**). GBIF data portal: data.gbif.org

Another piece of evidence that *C. papaya* L. may have had a Mesoamerican origin is the fact that it is still possible at present to find natural wild populations of this specie in isolated sites at the south of Mexico and in Central America. For instance, Manshardt and Zee (1994) found wild papayas in the Caribbean coastal lowlands of southern Mexico and northern Honduras. On the other hand, the Yucatan Scientific Research Center (CICY, México) has had an active program of collections from the flora of the Yucatan peninsula (México), including the States of Quintana Roo, Yucatán, and Campeche. From 1978 to 2003, the CICY Botanical Garden Herbarium had collected 58 geo-referenced specimens corresponding to *C. papaya* L. wild populations around the Yucatan Peninsula (Fig. 1.3a).

More recently, various *C. papaya* L. wild populations have been detected by our group in nonpopulated, isolated sites in the south of the Yucatan State (Fig. 1.3b–i). The plants from this living wild population found in the south of the Yucatán State differ greatly from the papaya plants from commercial plantations grown mostly in the northern part of the State. In Yucatán, as in most of the papaya growing areas in Mexico, the predominant commercial cultivar is Maradol (imported from Cuba), which has very distinctive morphological features when compared to those of the native wild population. Morphometric measurements have been taken of these plants from these wild populations, and they were compared with plants from commercial plantations of papaya, cultivar Maradol, grown in the north of Yucatan. Fruit length, width, and fresh weight from fruits from 5 plants from each of the 20

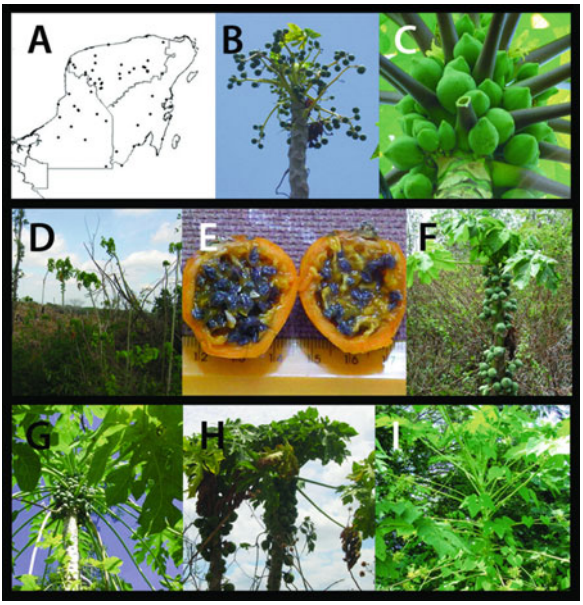

Fig. 1.3 Recent wild papaya populations localized in Yucatan Peninsula. (**a**) Map location of the collection sites along the Yucatan Peninsula, México, of wild populations of *C. papaya* L. from the Yucatan Scientific Research Center Herbarium "U Najil Tikin Xiw." (**b–i**) Photographs of *Carica papaya* L. (Ch'ich'put) found in different isolated sites from the south of the Yucatan State, taken in May 2006 and in September and November 2011: (**b**) male plant from wild populations found near Motilá. (**c**) Female plant from wild populations found near Xocchel. (**d**) Female plants from wild populations found near Libre Union. (**e**) Fruits from female plants from wild populations. (**f**) Female plant from wild populations found near Nenela. (**g**) Female plant (6 m tall) plant from wild populations found near Tixcacaltuyub. (**h**) Female plants from wild populations found near Mopilá. (**i**) Male plant from wild populations found near Mopila

different wild populations of *C. papaya* L. were compared with fruits from 5 plants from commercial papaya plants cultivar Maradol. The dendograms shown in Fig. 1.4 suggest that the 20 wild populations are related, although two distinct groups are formed between individuals from the population from Xocchel and those from the Tixcacaltuyub site, but individuals from the commercial cultivar Maradol clearly group out in a different clade. Currently, molecular studies using AFLPs are in progress in our lab, to analyze the genetic diversity between the different native wild populations and the commercial cultivars of papaya.

Papaya Domestication and Distribution

Again, it is also difficult to demonstrate the precise time and place where *C. papaya* L. was domesticated. Evidence exists for domestication of species such as maize by Aztecs and Mayas; however, early written evidence of domestication in papaya by this

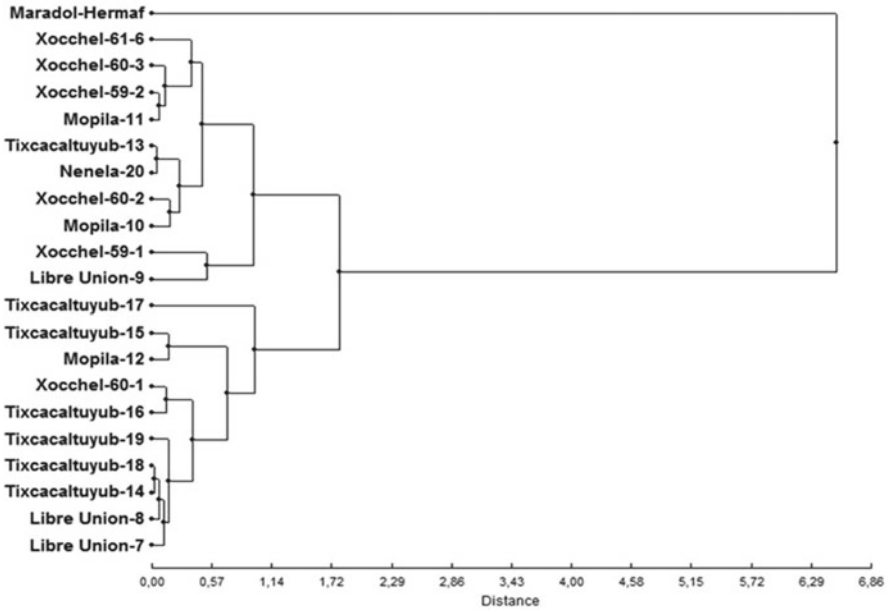

Fig. 1.4 Dendogram constructed from the analysis of morphometric characters (fruit length, width and weight) from 5 plants from each of 20 different wild populations of *Carica papaya* L. (Ch'ich'put) found in different isolated sites at the south of the Yucatan State, taken in May 2006 and September 2011. Equivalent data from five hermaphrodite papaya commercial (cultivar Maradol) plants were also included in the analysis

pre-Hispanic civilization is less documented. Therefore, various approaches have been put forward as indirect evidence of papaya domestication by early civilizations from Mexico and Central America.

No apparent direct mention of *C. papaya* L. appears in the best known pre-Colombian (Aztec and Mayan) codices, Badiano and Florentino (Emmart 1940; Galarza 1997). However, various studies include papaya in the list of plants that the Mayas would have used in their well-developed agricultural systems (de Oviedo 1959; Dunning et al. 1998; Terán and Rasmussen 1995; Colunga-GarcíaMarín and Zizumbo-Villarreal 2004), indicating that papaya ("put" or "puut," in Maya) was cultivated in Mexico and Belize before the arrival of the Spaniards. Certainly, lowland Mayas reached a very successful civilization, being an agricultural society apparently since at least 1300 BC (Pope et al. 2001). There is evidence that they had a well-developed agriculture, and the use of maize and *Agave fourcroydes* from food and fiber is well documented. Evidence of early agriculture has been found in Belize as the earliest archeological macrofossil of maize plant remains dated 1000 BC (Miksicek et al. 1991). However, Terán and Rasmussen (1995) and Colunga-GarcíaMarín and Zizumbo-Villarreal (2004) argue that this Mayan groups must have used some other plant species that grew naturally in the wild as well and include papaya among them.

Another interesting approach towards defining whether pre-Colombian native plants had been used by early civilization is one based on linguistics put forward by Brown (2010), and among the 41 species analyzed, they include papaya. From the linguistic point of view, the word papaya comes from a Caribbean word, again suggesting that the plant was known and used in this area before the Spaniards would have referred to it in later chronicle. However, among the Aztecs in Mexico papaya was known as chichihualtzapotl, a word that in Nahuatl means nurse fruit ("Zapote Nodriza") as it was related to fertility concepts, and among the Mayas it was known as "Ch'iich'puut" or "put" (Alvarez 1980). Nowadays, the papaya is also known as fruta bomba, lechosa (Venezuela, Puerto Rico, the Philippines and the Dominican Republic), mamão, papaw (Sri Lanka), Papol\Guslabu (tree melon) in Sinhalese, pawpaw, or tree melon; this reflects its ample distribution along various tropical areas of the world.

It is believed that papaya fruits (and seeds) were distributed from Central America to South America and other parts of the world by the Spaniards during the sixteenth century. The first written report of this specie appears to be that made by Oviedo in 1526, in the coastal areas of Panama. According to Morton (1987), papaya seeds were taken to Dominican Republic before 1525 and cultivation spread to warm elevations throughout South America, the West Indies and Bahamas, and then to Bermuda in 1616. Spaniards carried papaya seeds to the Philippines by approximately 1550, and by 1611, papaya was cultivated in India. From 1800 on, papaya was distributed among various islands in the south of the Pacific Ocean.

Strictly speaking, a domestication process implies that the domesticated plant species will not be able to survive without human intervention (Smith 2001). The commercial papaya cultivars, at least Maradol, certainly comply with this precept as they need irrigation to survive and to produce high yields, as they are highly susceptible to drought. On the other hand, their wild relatives are able to survive long seasons with low rain and high temperatures, for instance, from March to June at the Mexican State of Yucatán. Another feature in the domestication process of plant species is the change in the germination patterns; clearly the commercial papaya cultivars have been selected for high germination rates, while the wild relatives had low germination rates (Fuentes, unpublished). Another important feature in the domestication process is the presence of hermaphrodite plants in high proportion in commercial cultivars, whereas in plants from wild populations it is, from our experience, an extremely rare feature to be found in plants from wild native populations, at least in the ones present in Yucatán. Other contrasting features between the commercial cultivars and the papaya trees from wild populations are the number and size of fruits. While commercial papaya plants produce around 25 fruits, with an average weight of 1.5 kg, their wild counterparts are able to produce at least 70 fruits, with an average weight ranging from 20 to 35 g (Fuentes, unpublished).

Commercial Papaya Production, Harvested Area, Yields, and Trade

Papaya Production

As it can be observed in Fig. 1.5a, the worldwide production of papaya has increased from 1 to 10 million metric tons (t) per year, in the last 50 years. The latest data available show that the worldwide production of papaya reached 10.5 million t in

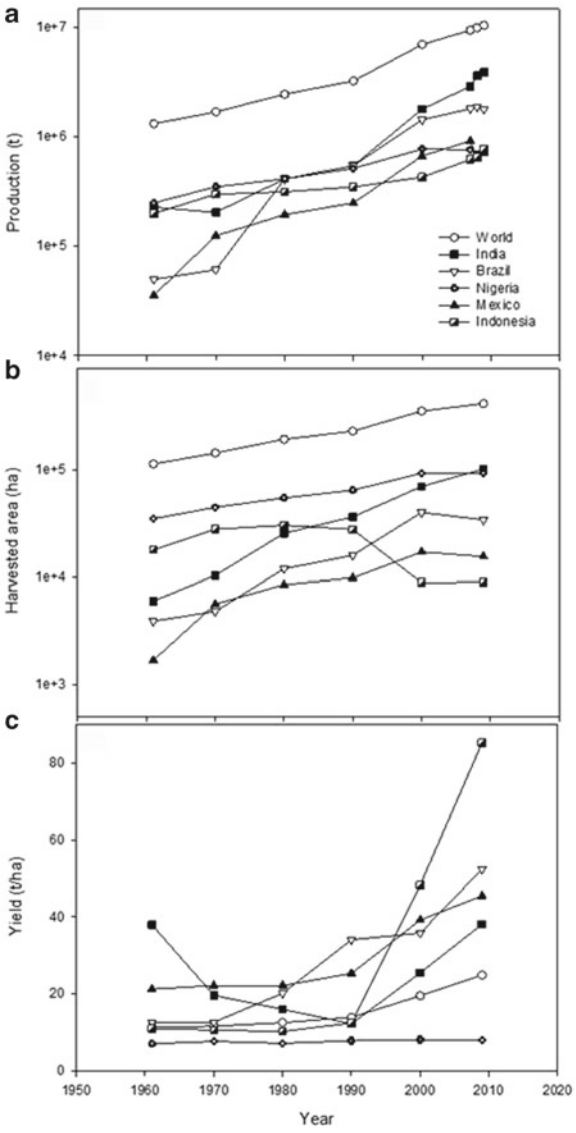

Fig. 1.5 Papaya worldwide production, harvested area and yields of papaya in the last 50 years. (**a**) Worldwide production (t) and production trends of the five top country-producers of papaya in the world. (**b**) Worldwide papaya harvested area (ha) and trends of cultivated area at the five top papaya country-producers in the world. (**c**) World average of yields (t/ha) obtained in papaya plantations, and those obtained in the five top country-producers in the world during the last 50 years (data from FAOSTAT 2011)

2009 (FAOSTAT 2011). The gross value of this worldwide production of papaya in 2009 is estimated by FAOSTAT to be 3.5 billion USD. In general, a boom in the papaya production worldwide occurred between 1990 and 2000 and later increased at slightly lower rates in the last decade.

The main five countries that have been papaya producers during the last 50 years are India, Nigeria, Brazil, Mexico, and Indonesia (Fig. 1.5a). In the 1960s, most of the total papaya production of the world was concentrated in India, Indonesia, and Nigeria; however, in the 1980s and 1990s, Brazil and Mexico increased their production to the extent that they produced equivalent quantities to those of India and Nigeria from 1980 to 1990.

In the last decade, India and Brazil have maintained their continuous increase in papaya production to rank first and second, after producing almost 4 and 1.8 million t, respectively, in 2009 (FAOSTAT 2011). In the case of Nigeria's and Mexico's production, it has reached a plateau with values ranging between 700 and 766 thousand t in the last year. In the case of Indonesia, their papaya production had remained fairly stable to rank fourth or fifth, with a production in the last 3 years between 650 and 765 thousand t (Fig. 1.5a).

The other five countries that complete the top ten papaya producers in the world are Ethiopia, Congo, Thailand, Guatemala, and Colombia (depending on the year, other countries such as Venezuela and China have been also part of the ten main papaya producers); these countries have lower productions ranging between 189 and 260 thousand t per year (FAOSTAT 2011). Taken as regions, Asia produced 5.4 million t (52 % of the world production) of papaya in 2009, America produced 3.6 million t (34 %), and Africa produced 1.4 million t (14 %) of papaya during 2009.

Papaya Harvested Area

In terms of harvested area, we will be using the SI unit hectare (ha) as used in FAOSTAT 2011. The area where papaya has been grown has increased continuously during the last 50 years from 114,192 ha in 1961 to 420,279 ha (438,239 ha in 2010) around the world in 2009 (Fig. 1.5b). Nigeria has ranked first in the number of ha cultivated with papaya in the last 50 years, although in 2009 it ranked second. Indonesia ranked second in the number of ha cultivated with papaya from 1960 to 1980, but it has declined from 1990 and presently ranks fifth. In the case of India, Brazil, and Mexico, the area cultivated with papaya was modest in the 1960s but increased continuously since then, and India now (2009) ranks first. In the case of Brazil and Mexico, the area of land cultivated with papaya has declined slightly in the last decade, but these countries still remain ranked third and fourth, respectively (Fig. 1.5b).

Papaya Yields

In terms of yields, we will be using the SI units (t/ha) as used in FAOSTAT, where t is metric ton (1 t = 1 mt = 1 tonne = 1,000 kg) and ha is hectare (ha = 10,000 m^2). The worldwide yield average obtained in papaya plantations has increased from 11.58 to 24.95 t/ha in the last 50 years (Fig. 1.5c). In Indonesia, the yields increased remarkably from 12.75 in the 1960s to 85.13 t/ha in 2009, after a sharp increase in the year 2000 where yields of 42.3 t/ha were obtained. In Brazil, yields increased from 12.75 t/ha in the 1960s to 52.39 t/ha in 2009. In the case of Mexico, yields also increase from 21 t/ha in the 1960s to yields of 45.42 t/ha obtained in 2009. In the case of India, yields have been rather erratic starting at high yields during the 1960s (38 t/ha) and then experiencing a marked drop in the 1980s and 1990s (12.36 t/ha), but it had an important increase in the last two decades, reaching again, in 2009, 38.12 t/ha. In contrast, Nigeria has maintained more modest yields during the 50-year period obtaining yields of 7.14 t/ha in the 1960s and 8.10 t/ha in 2009 (Fig. 1.5c).

Papaya Production Relative to That of Other Important Tropical Fruits

The high production volumes reached by papaya in 2009 position papaya among the five top tropical fruits around the world (Fig. 1.6a, b). However, the world production of papaya, 10 million t, is still far from the top-ranked tropical fruits (mangoes, mangosteen and guavas, MMG), grouped together by FAOSTAT 2011 and reaching values of 35 million t in 2009, and far from the banana, which reached more than 25 million t in the same year. In terms of production value, papayas had a worldwide production valued at 3.5 billion USD that is still low when compared to that of banana, reaching 28 billion USD (Fig. 1.6a, b). Therefore, there is still an opportunity to increase the production capacity of papaya. The social value of this important crop is something to be mentioned, as it is also a good source of jobs in the papaya producing countries during the harvesting season, which in some countries can reach one harvest a week during 1.5 years.

Papaya Trade

In 2009, the total of papaya exports among the 20 top exporter countries around the world was 215 thousand t, with a value of 119 million USD (FAOSTAT 2011). The top five exporter countries are Mexico, Brazil, Belize, Malaysia, and India

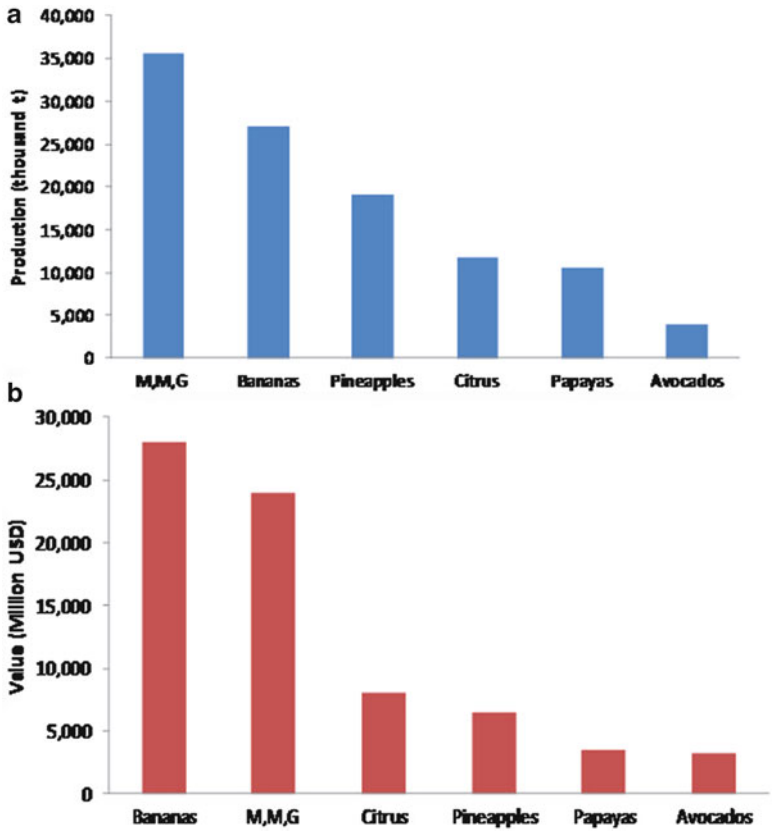

Fig. 1.6 Situation of the worldwide papaya production reached in 2009 when compared to that from other important tropical fruits. (**a**) Worldwide production (thousand t) of the six main tropical fruits. (**b**) Value of that production in millions of USD (data from FAOSTAT 2011)

(Fig. 1.7a, b). In terms of imports, the total volume of papaya imported in 2009 among the top 20 importer countries around the world was 232 thousand t, with a value of 199 million USD. The top five importer countries of papaya in 2009 were the USA, European Union, Singapore, Canada, and the Netherlands (Fig. 1.7a, b).

Conclusions

In terms of papaya's origin, we have reviewed additional facts based on the herbarium specimen records and the existence of local populations to support the hypothesis that the south of Mexico and Central America are the center of origin of *C. papaya*

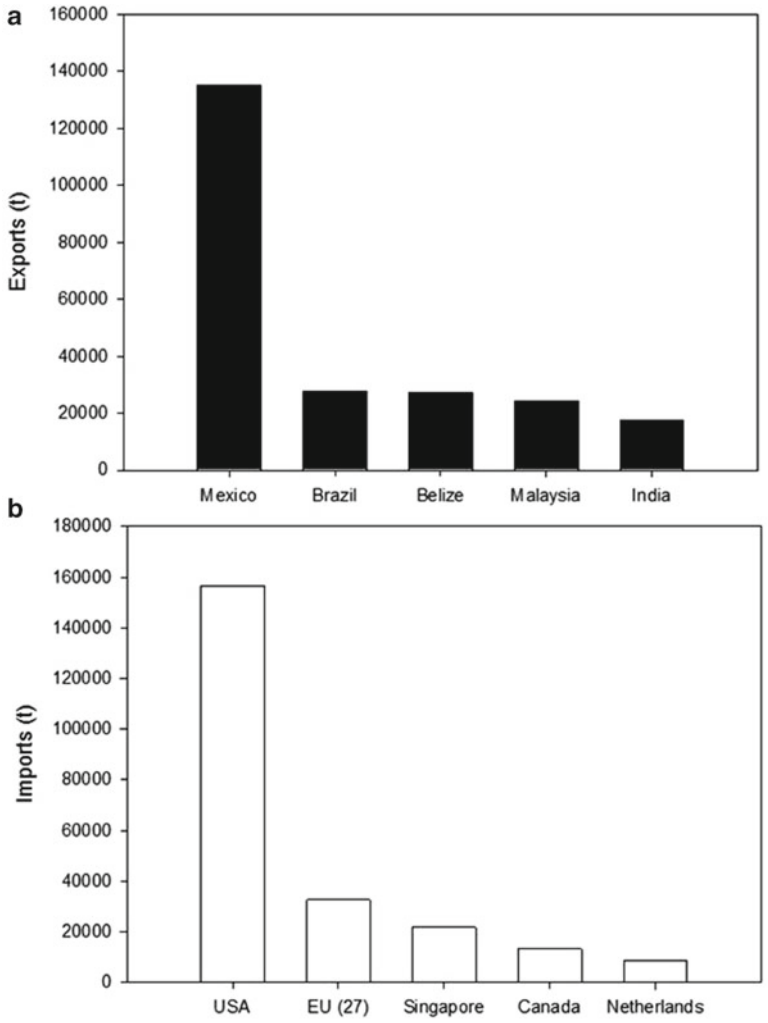

Fig. 1.7 (**a**) Papaya exports in 2009 from the top five exporter countries in the world. (**b**) Papaya imports by the top five importer countries in the world (data from FAOSTAT 2011)

L. However, molecular approaches are required to undoubtedly confirm this idea. The excellent, recent work by Carvalho and Renner is discussed later in Chap. 5 of this book.

In the present chapter, we have also reviewed some of the arguments for belief that domestication of *C. papaya* L. occurred within the lowland Mayas by at least 1500 AC. Preliminary results were shown of recent comparative morphological and molecular studies between native wild populations found in Yucatan, México, and

those of commercial papaya cultivars. This study might prove useful in further discussions on the domestication process for papaya.

Finally, we conclude that the worldwide commercial papaya production has experienced an important rise in the last 50 years, in terms of yields and production volumes. The current production of commercial papaya has reached important worldwide levels, reaching in 2009 a production of 10.5 million t with a value of 3.5 billion US dollars. Asian countries (mainly India and Indonesia) produce 52 % of the total papaya produced in the world; America (mainly Brazil and México) contributes with 34 %, while Africa produces 14 % of the world production with large land areas but low yields. This high worldwide production volume positions papaya among the top five tropical fruits produced around the world, although there is still opportunity to reach equivalent production levels to those met by mangoes, mangosteens, guavas, and bananas. It would be very important to define the reason (perhaps diseases, abiotic factors, production cost) for the plateau or even the decrease in papaya harvested areas, observed in some of the top producing countries in the last decade.

Acknowledgements The authors wish to acknowledge the expert advice from Dr. German Carnevalli, curator of the Scientific Research Center Herbarium. The assistance of Silvia Hernández (CICY's Herbarium database) is gratefully acknowledged. Expert advice in the field of domestication processes from Dr. Daniel Zizumbo from CICY is also acknowledged. Paulino Simá, Carlos Talavera, Arianna Chan, Humberto Estrella, Martha Chi, and Fatima Sosa (CICY) have participated in the collections of papaya in wild populations, while Francisco Espadas, Fernando Contreras, and Mariela Vázquez contributed some of the measurements of the Maradol counterparts (CICY).

References

Alvarez C (1980) Diccionario Etnolingüistico del Idioma Maya Yucateco Colonial. VI Mundo Fisico Inst Inv Filolog Cen Est Mayas. UNAM, Mexico City

Badillo VM (1971) Monografía de la familia *Caricaceae*. Publicado por la Asociación de profesores. Universidad Central de Venezuela, Maracay

Badillo VM (1993) *Caricaceae*. Segundo Esquema. Rev Fac Agron Univ Centr Venezuela 43:1–111

Badillo VM (2000) Carica vs Vasconcella St.-Hil. (*Caricaceae*) con la rehabilitación de este ultimo. Ernstia 10:74–79

Brown CH (2010) Development of agriculture in prehistoric Mesoamerica: the linguistic evidence. Pre-Columbian Foodways 1:71–107

Colunga-GarcíaMarín P, Zizumbo-Villarreal D (2004) Domestication of plants in Maya lowlands. Econ Bot 58:S101–S110

De Candolle A (1883) Origine des plantes cultivées. G Baillièreetcie, Paris

de Oviedo GF (1959) Historia General y Natural de las Indias. JP de Tudela. Ediciones Atlas, Madrid

Dunning N, Beach T, Farrell P, Luzzadder-Beach S (1998) Prehispanic agrosystems and adaptive regions in the Maya Lowlands. Cult Agric 20:87–101

Emmart EW (1940) Badianus manuscript (Codex Berberini, Latin 241). Vatican Library; an Aztec herbal of 1552. Introduction, translation and annotations by Emily Walcott Emmart. With a foreword by Henry E. Sigerist. The Johns Hopkins Press, Baltimore

FAOSTAT (2011) Food and Agriculture Organization of the United Nations Database. http://www.apps.fao.org. Accessed December 2011

Galarza J (1997) Los códices mexicanos. Arqueol Mexicana IV 23:6–15
Global Biodiversity Information Facility (2011) http://data.gbif.org/species/2874484/. Accessed December 2011
Harlan JR (1971) Agricultural origins: centers and non-centers. Science 174:468–474
Lorence DH, Colin RT (1988) *Carica cnidoscoloides* (sp. nov) and sect. Holostigma (sect. nov.) of *Caricaceae* from Southern Mexico. Syst Bot 13:107–110
Manshardt RM, Zee F (1994) Papaya germplasm and breeding in Hawaii. Fruit Variety J 48:146–152
McVaugh R (2001) Caricaceae. In: Flora Novo-Galiciana: a descriptive account of the vascular plants of Western Mexico, vol 3. Ochnaceaeto Losaceae. The University of Michigan Herbarium, Ann Arbor
Miksicek CH, Wing ES, Scudde SJ (1991) The ecology and economy of Cuello. In: Hammond N (ed) Cuelo: an early Maya community in Belize. Harvard University Press, Cambridge, pp 70–84
Missouri Botanical Garden (2011) Tropicos.org http://mobot.mobot.org/W3T/Search/vast.html. Accessed December 2011
Morton J (1987) Papaya. In: Morton JF (ed) Fruits of warm climates. Florida Flair Books, Miami, pp 336–346
Pope KO, Pohl ME, Jones JG, Lentz DL, vonNagy C, VegaFJ QIR (2001) Origin and environmental setting of ancient agriculture in the lowlands of Mesoamerica. Science 292:1370–1373
Prance GT (1984) The pejibaye, Guilielma gasipaes (H.B.K.) Bailey, and the papaya, Carica papaya L. In: Stone D (ed) Pre-Columbian plant migration. Peabody Museum, Cambridge, pp 85–104
Scheldeman V, Kyndt T, Coppens d'Eeckenbrugge G, Ming R, Drew R, Van Droogenbroeck B, Van Damme P, Moore PH (2011). Vasconcellea. In: Kole C (ed) Wild crop relatives: genomic and breeding resources. Springer Science+Business Media, New York, pp 213–249
Smith BC (2001) Documenting plant domestication the consilience of biological and archeological approaches. Proc Natl Acad Sci USA 98:1324–1326
Solms-Laubach (1889) Die Heimatundder Ursprungdes kultivierten Melonen baumes *Carica papaya* L. Botanische Zeitung 44:709–720
Terán S, Rasmussen C (1995) Genetic diversity and agricultural strategy in 16th century and present-day Yucatecan Milpa Agriculture. Biodiv Conserv 4:363–381
Vavilov NI (1987) Origin and geography of cultivated plants [English Translation by D Löve]. Cambridge University Press, Cambridge
von Linnaeus C (1753) Species plantarum, 2 vols [v 1: 1–560; v 2: 561–1200]. Imprensis Laurentii Salvii, Holmiae

Chapter 2
Biology of the Papaya Plant

Víctor M. Jiménez, Eric Mora-Newcomer, and Marco V. Gutiérrez-Soto

Introduction

The papaya plant (*Carica papaya* L.) has been described with a large variety of adjectives, which acknowledge the structural and functional complexity and the high phenotypic plasticity of this giant tropical herb (León 1987). *C. papaya*, with a somatic chromosome number of 18, is the sole species of this genus of the Caricaceae, a family well represented in the Neotropics, that includes six genera with at least 35 species (Fisher 1980; Ming et al. 2008; Carvalho and Renner 2013). Most likely, papaya originated along the Caribbean coast of Mesoamerica (Fitch 2005) and spread to many tropical and subtropical regions around the world (Kim et al. 2002), where its distribution is limited by chilling sensitivity (Allan 2002; Dhekney et al. 2007). Domestication eventually led to substantial changes in vegetative growth and sexual forms that distinguish wild populations from cultivated genotypes (Paz and Vázquez-Yanes 1998; Niklas and Marler 2007). Because of its high yield, nutritional value, functional properties, and year-round fruit production, the importance of this crop around the world is undeniable.

The papaya plant is a semi-woody, latex-producing, usually single-stemmed, short-lived perennial herb. The relatively small genome of this species shows peculiarities in major gene groups involved in cell size and lignification, carbohydrate economy, photoperiodic responses, and secondary metabolites, which place the papaya in an intermediate position between herbs and trees (Ming et al. 2008). Reproductive precocity, high photosynthetic rates of short-lived leaves, fast growth,

V.M. Jiménez (✉)
CIGRAS, Universidad de Costa Rica, 2060 San Pedro, Costa Rica

Food Security Center, University of Hohenheim, 70593 Stuttgart, Germany
e-mail: victor.jimenez@ucr.ac.cr

E. Mora-Newcomer • M.V. Gutiérrez-Soto
Estación Experimental Agrícola Fabio Baudrit Moreno,
Universidad de Costa Rica, Alajuela 183-4050, Costa Rica

R. Ming and P.H. Moore (eds.), *Genetics and Genomics of Papaya*, Plant Genetics and Genomics: Crops and Models 10, DOI 10.1007/978-1-4614-8087-7_2,

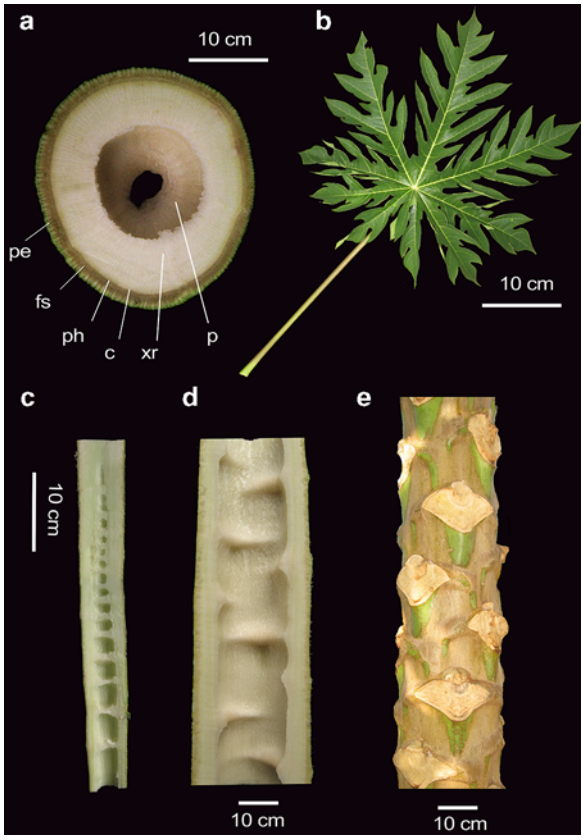

Fig. 2.1 Vegetative parts of the papaya plant. (**a**) Cross section of a 1-year-old papaya stem; periderm (pe), fiber sheath (fs), phloem (ph), cambium (c), xylem rays (xr), pith (p). (**b**) Leaf lamina and petiole. (**c**) Longitudinal section of a 3-month-old papaya stem showing hollow pith cavity. (**d**) Longitudinal section of a 1-year-old papaya stem showing complete pith cavity. (**e**) Stem of a 1-year-old papaya plant showing conspicuous petiole scars

high reproductive output, production of many seeds, and low construction cost of hollow stems (Fig. 2.1a–d), petioles, and fruits characterize this successful tropical pioneer. High phenotypic plasticity allows this plant to establish in recently disturbed sites, thriving during early stages of tropical succession and as members of diverse agroecosystems as well (Hart 1980; Ewel 1986), that constitute important genetic reservoirs (Brown et al. 2012). At any given time, adult papaya plants can sustain vegetative growth, flowering, and dozens of fruits at different stages of development, simultaneously.

Morphology, Architecture, and Anatomy of the Adult Plant

Papaya is usually a single-stemmed, semi-woody giant herb with fast, indeterminate growth (1–3 m during the first year). The plants may attain up to 10 m, although under modern cultivation height seldom surpasses 5–6 m. Occasionally, vigorous vegetative growth may induce axillary bud break and branching at the lower portions of the plant, which rarely exceeds a few centimeters in length. Some branching may

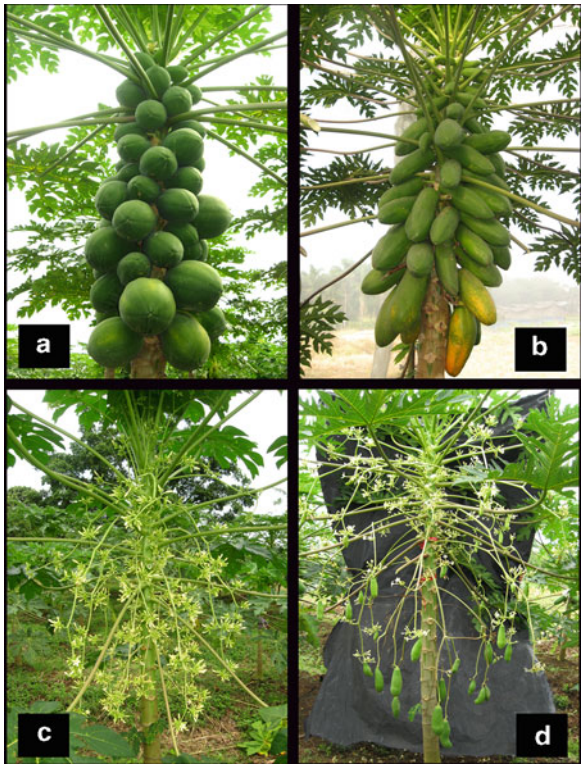

Fig. 2.2 Types of papaya plants according to sex forms. (**a**) Female. (**b**) Hermaphroditic. (**c**) Male. (**d**) Male fruit-bearing plant

also occur if apical dominance is lost due to tip damage, and, in tall plants, "distance" may release the lower buds from the dominant effect of the apex (Morton 1987).

The plant produces large palmate leaves (~0.6 m^2), with five to nine pinnate lobes (Fig. 2.1b) of various widths (40–60 cm), arranged in a spiral pattern (Fig. 2.1e) and clustered in the upper section of adult individuals (Morton 1987; Ming et al. 2008). Leaf blades are dorsiventral and subtended by 30–105 cm long, hollow petioles that grow nearly horizontal, endowed with a starch-rich endodermis, perhaps important for cavitation repair (Bucci et al. 2003; Posse et al. 2009; Leal-Costa et al. 2010). The leaf epidermis and the palisade parenchyma are composed of a single cell layer, while the spongy mesophyll consists of four to six layers of tissue. Reflective grains and druses are abundant throughout the leaf (Fisher 1980). Papaya leaves are hypostomatic, with anomocytic (no subsidiary cells) or anisocytic (asymmetric guard cells) stomata (Carneiro and Cruz 2009; Leal-Costa et al. 2010). Stomatal density of sunlit leaves is approximately 400/mm^2, which can adjust readily to environmental conditions of light, water, and heat. Important biologically active compounds have been identified in papaya leaves (Canini et al. 2007; Zunjar et al. 2011), where they function in metabolism, defense, signaling, and protection from excess light, among others (El Moussaoui et al. 2001; Konno et al. 2004). Adult plants may have three possible sexual forms: female, male, and hermaphroditic (Figs. 2.2a–d and 2.3a–f).

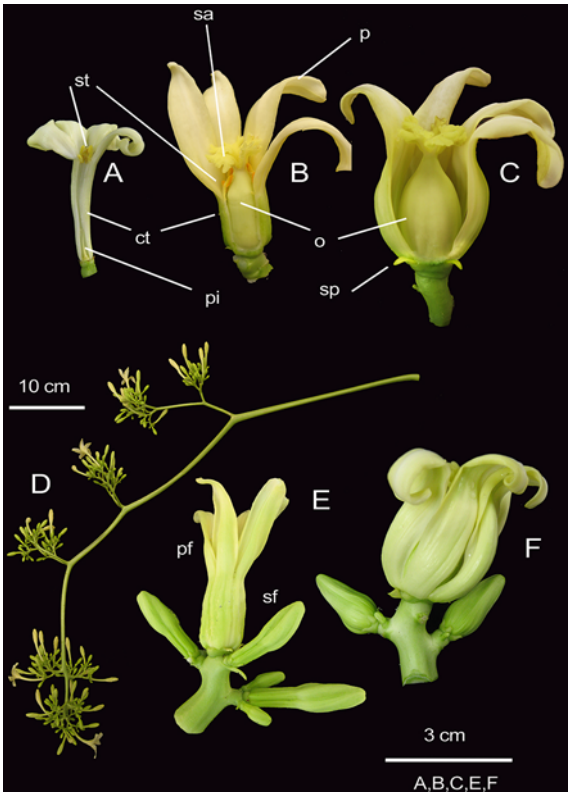

Fig. 2.3 Papaya flowers with one petal removed to show internal parts (**a**–**c**) and inflorescences (**d**–**f**). (**a**) Staminate flower showing stamens (st), pistillode (pi) and corolla tube (ct). (**b**) Perfect flower showing st, ct, stigmata (sa), petal (p) and an elongated ovary (o). (**c**) Pistillate flowers showing sepals (sp), petals and round ovary (o). (**d**) Long male inflorescence with dozens of staminate flowers. (**e**) Andromonoecious cyme showing one dominant perfect (pf) and five secondary staminate flowers (sf). (**f**) Female cyme with three pistillate flowers

Plant Growth and Development

Under appropriate conditions of water availability, light, oxygen, air temperature, and humidity, papaya seeds undergo epigeal germination (Fig. 2.4a); emergence is typically completed in 2–3 weeks (Fisher 1980). Primary leaves of young seedlings are not lobed (Fig. 2.4b) but become so after the appearance of the second leaf (Fig. 2.4c). Papaya leaves of adult plants are simple, large, and palmate (Fig. 2.1b). In tropical conditions, approximately two leaves emerge at the apex of the plant in a 3/8 spiral phyllotaxy every week (Fisher 1980). Leaf life commonly spans for 3–6 months under tropical conditions and persistent scars remain on the trunk as they abscise (Fig. 2.1e). The loss of leaves on the lower section of the plant and the continuous emergence of new ones at the apex give the canopy a sort of umbrella shape that casts a considerable amount of shade.

The papaya plant develops very fast, taking 3–8 months from seed germination to flowering (juvenile phase) and 9–15 months for harvest (Paterson et al. 2008). The plant can live up to 20 years; however, due to excessive plant height and pathological constraints, the commercial life of a papaya orchard is normally 2–3 years.

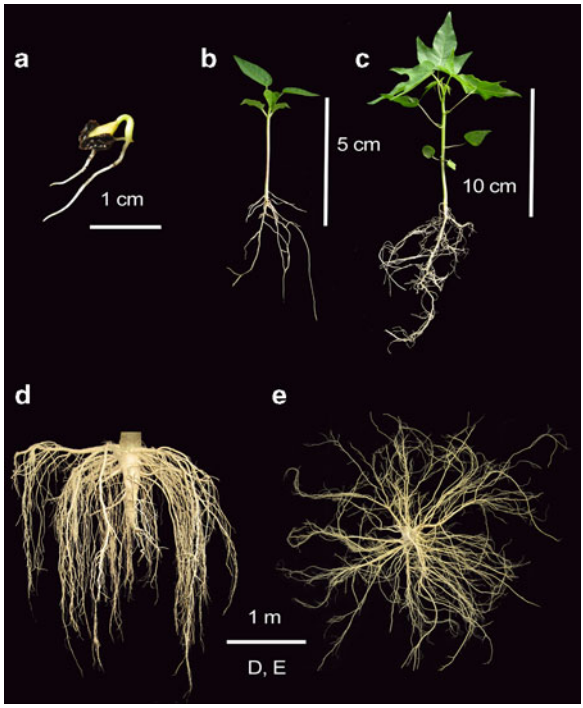

Fig. 2.4 Papaya seedlings and root system. (**a**) Germinating papaya seed. (**b**) Ten-day-old papaya seedling showing cotyledonary leaves and first true leaves. (**c**) Three-week-old papaya seedling with six true leaves. (**d**) Side view of an excavated 5-month-old papaya root system, showing the main and secondary roots. (**e**) Upper view of the same root system, showing horizontal distribution of secondary roots

Although papayas are considered sun-loving plants, morphological plasticity in the shade is high and involves changes in many characteristics such as leaf mass per area, chlorophyll *a*/*b* ratio, stomatal density, internode length, and degree of blade lobing (Buisson and Lee 1993). This plasticity is evidenced by the morphology adopted by papayas growing in multistoried agroecosystems and in high-density orchards as well (Marler and Discekici 1997; Iyer and Kurian 2006).

Papaya seedlings and adults are very responsive to mechanical stimuli and show strong thigmomorphic responses or touch-regulated phenotypes (Fisher and Mueller 1983; Porter et al. 2009). These responses could be essential to the success of papaya in harsh, early successional sites exposed to high winds, because it triggers hardening mechanisms that result in compact architecture, increased lignification, and the formation of petiole cork outgrowths (Clemente and Marler 2001; Porter et al. 2009).

The Stem: Support and Transport Systems

In papaya plants, the single stem provides structural support, body mass, storage capacity, defense substances, height, and competitive ability, and carries a bidirectional flow of water, nutrients, various organic compounds, and chemical and physical signals that regulate root and shoot relations (Reis et al. 2006). Stem diameters of

adult plants vary from 10 to 30 cm at the base to 5–10 cm at the crown. Stem density is only 0.13 g cm^{-3}. The lower internodes are compact and wider and seem to mechanically support the entire weight of the plant (Morton 1987).

In papaya stems, a thick, single layer of secondary phloem, rich in fibers, and two sclerenchyma layers located immediately inside the bark are responsible for most of the rigidity (Fig. 2.1a). The xylem is poorly lignified and aids in storage of water and starch (Fisher 1980). A well-developed pith is conspicuous from early stages of development. Young stems become progressively hollow by the dissolution of the pith at the internodes as they mature (Fig. 2.1c, d) and as fibers thicken and harden (Carneiro and Cruz 2009). Along with stem thickening, the fibers of the outer collenchyma layer yield, leaving wide spaces later occupied by parenchyma and periderm, so that rigidity shifts to the inner layers, which widen to allow for stem expansion. In addition to buffering daily water balance, stem water storage may also be a major structural determinant of the mechanical stability of these succulent, tall herbaceous plants (Fisher 1980).

Fisher (1980) described the features of papaya vascular tissues summarized here. The xylem is composed of wide vessels that can be seen with the naked eye, imbibed in non-lignified parenchyma tissues and rays. Pits are alternate, bordered, or unbordered, and perforation plates are simple and transverse. Phloem rays are wide, multiseriate, and tall. Sieve elements may be multistoried, and sieve plates are transverse and located on lateral walls. Recorded rates of xylem sap flow, source:sink ratios, and observations on the hydraulic architecture of papaya plants indicate that high water transport and phloem translocation capacity sustain the high rates of gas exchange and growth observed in the field. Phloem loading is probably symplastic, but because of the intermediate nature of this species, possible shifts in loading mechanisms in papaya plants growing along altitudinal gradients (ranging from 0 to 3,500 m above the sea level) should be examined.

Laticifer Conducts

Laticifers of papaya are complex tissue systems of the articulate-anastomosing type. In general, these conducts are multicellular columns with perforated transverse lateral walls, protoplast fusion, and intrusion of phloem cells, forming branched networks (Hagel et al. 2008). Damage to any aerial part of the papaya plant, where laticifers are widely distributed, elicits latex release, which is very typical for this species (Azarkan et al. 2003). This milky latex is a slightly acidic fluid composed of 80 % water (Rodrigues et al. 2009). It contains sugars, starch grains, minerals (S, Mg, Ca, K, P, Fe, Zn), alkaloids, isoprenoids, lipidic substances, and proteins, including enzymes like lipases, cellulases, and cysteine proteases (papain, chymopapain), important in defense against insect herbivores and in tissue and organ formation (pith differentiation) (Sheldrake 1969; El Moussaoui et al. 2001; Azarkan et al. 2003; Konno et al. 2004). In young fruits, laticifers develop near the vascular bundles and become septate through transverse walls that later dissolve, making these laticifers a series of

superposed fused cells. Young papaya fruits (>10 cm diameter) are tapped for papain by making long, straight cuts (Madrigal et al. 1980). Yields of crude papain are around 245 kg/ha the first year (Becker 1958).

Root System

Young roots show well-differentiated epidermis, cortex, and endodermis, enclosing an exarch vasculature in which six xylem and six phloem poles alternate. Cambium formation in a concentric ring triggers secondary growth and root thickening while maintaining succulence. The papaya root is predominately a non-axial, fibrous system, composed of one or two 0.5–1.0 m long tap roots. Secondary roots emerge from the upper sections and branch profusely (Fig. 2.4d, e). These second-order feeding roots remain shallow during the entire life of the plant and show considerable gravitropic plasticity. Many adventitious, lower-order categories of thick and fine roots are also observed in excavated specimens. Healthy roots are of a whitish cream color, and no laticifers have been observed in them (Marler and Discekici 1997; Carneiro and Cruz 2009).

Root phenotypic plasticity is also high. Root size, number, distribution, and orientation adjust readily across the soil profile, to various soil conditions, and throughout the life of the plant, making papayas preferred components of complex agroecological models and hillside vegetation (Fisher and Mueller 1983; Marler and Discekici 1997).

Papaya plants are dependent on mycorrhizas for their nutrition and benefit greatly from soil mulching and appropriate drainage that facilitate biotic interactions in the rhizosphere and water and nutrient uptake, especially phosphorus and nitrogen. Four to five genera and 11 species of arbuscular mycorrhizal fungi have been reported associated with papaya roots: *Glomus*, *Acaulospora*, and *Gigaspora*, among others (Walsh and Ragupathy 2007; Khade et al. 2010). Mycorrhizal interactions of male and female papaya plants may differ: females seem more responsive to changes in soil fertility and readily adjust mycorrhizal colonization accordingly (Vega-Frutis and Guevara 2009).

Sex Expression

Papaya has three sex forms (female, male, and hermaphrodite), regulated by an incipient X–Y chromosome system. Papayas can be either dioecious (with male and female plants) or gynodioecious (with hermaphrodite and female plants). Several studies suggest that the Y chromosome contains a small specific region that controls expression of male (Y) or hermaphrodite (Y^h) types. Female plants are of the XX form. All combinations among the Y and/or Y^h chromosomes are lethal; therefore, the male and hermaphrodite types are heterozygous (XY and XY^h, respectively) (Ming et al. 2007).

Flowers

Papaya flowers are produced profusely near the trunk apex and open between 7 and 9 a.m. Individual flower longevity may be of 3–4 days, but the actual period of pistil receptivity is unknown. Papaya flowers are actinomorphic cymes arranged in inflorescences on the leaf-stem junction. Cymes of hermaphroditic and female plants can bear a variable number of flowers (2 to 15) (Fig. 2.3e, f). Male plants produce very long inflorescences that contain dozens or even hundreds of flowers (Fig. 2.3d). Although papaya flowers are visited by hawkmoths, several beetles, skipper butterflies, bees, flies, and hummingbirds, among others, probably attracted by the sweet fragrance and the copious nectar they produce (the latter only in male and hermaphroditic flowers) (Ronse Decraene and Smets 1999), recent evidence suggests that hawkmoths are responsible for most of the pollen exchange (Martins and Johnson 2009; Brown et al. 2012). Wind pollination has also been cited (Sritakae et al. 2011). The time from pollination to first ovule penetration was 25 h at 28 °C (Cohen et al. 1989).

Hermaphroditic Flowers

The perfect flower of papaya, also referred to as the elongate type, consists of five petals, five pairs of anthers, and an ovary (Fig. 2.3b). The petals are fused on the lower part of the flower (connate), to the point where the stamens are inserted, forming the corolla tube. The upper parts of the petals are free and slightly twisted. The ovary is superior, elongated, and composed of five carpels. Each pistil has five broad and flattened stigmata joined at their base, which may bend slightly backwards when the flowers open. There are five pairs of anthers, inserted into two whorls (diplostemonous androecium), but each member of a pair belongs to a different whorl. Stamens belonging to the antesepalous have longer filaments than those in the antepetalous whorl (Ronse Decraene and Smets 1999).

Although the term hermaphrodite has been used to refer to papaya plants that bear perfect flowers, the correct term should be andromonoecy, which indicates the occurrence of staminate and hermaphroditic flowers on the same plant. Typically, the small inflorescences of hermaphroditic papaya plants bear one or two main perfect flowers and a few secondary female sterile (staminate) and intermediate type flowers (Fig. 2.3e). The ratio of perfect to staminate flowers within an inflorescence may vary greatly due to the effects of genetic and environmental factors and may range from totally perfect to totally sterile. Female sterility in andromonoecious papayas is often expressed progressively, leading to reductions in ovary size, carpel number, and associated tissues and ultimately may lead to completely staminate flowers which contain only a pistillode (Fig. 2.5a–d) (Nakasone and Lamoureux 1982).

When hermaphrodite papaya plants are subjected to stresses such as high temperatures and water and nitrogen shortages, female sterility is exacerbated (Awada and Ikeda 1957; Arkle and Nakasone 1984; Almeida et al. 2003). This may even affect the main flowers, leading in some cases to totally infertile (staminate) and

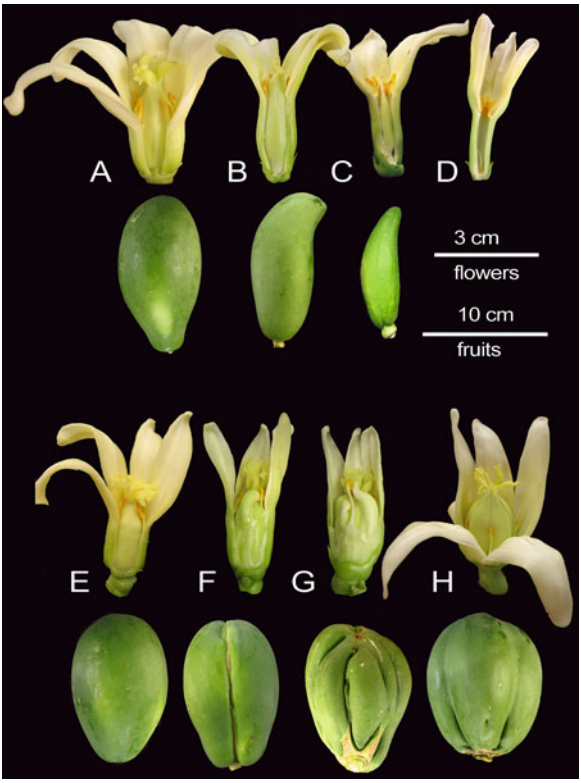

Fig. 2.5 Female sterility and carpellody of papaya. (**a**–**d**) Gradual reduction in pistil size due to female sterility of flowers and approximate phenotype of the corresponding fruit. (**a**) Normal elongata flower with five carpels. (**b**, **c**). Reduced ovary due to loss of carpels as a result of partial female sterility. (**d**) Completely female sterile flower. (**e**, **f**) Increasing levels of carpellody and approximate phenotype of corresponding fruit. (**e**) Normal elongata flower. (**f**–**g**) Fusion to the ovary and partial transformation to carpels and of one (**f**) and two (**g**) stamens, leading to misshapen fruits. (**h**) Complete transformation of the five antepetalous whorl of stamens into carpels, leading to the "pentadria" type of flower, with a rounded ovary (and fruit) and almost free petals

unproductive inflorescences. Perfect papaya flowers may also undergo variable degrees of fusion between their stamens and the ovary (carpellody; Fig. 2.5f, g) (Ronse Decraene and Smets 1999). In severe cases, the five antepetalous stamens are completely transformed into carpels, and the resulting flower resembles a female one, with a rounded ovary and free petals almost all along their length. This type of flower is also known as the "pentadria type" (Fig. 2.5h). Intermediate carpellodic states are also common, in which only some of the stamens are completely or partially fused with the ovary, resulting in the development of misshapen fruits (Figs. 2.5f, g and 2.6c). Although the tendency to produce carpellodic flowers has a strong genetic component (Storey 1953; Ramos et al. 2011), low temperatures, high soil moisture, and high nitrogen seem to favor this condition (Awada 1953, 1958; Awada and Ikeda 1957; da Silva et al. 2007).

Female Flowers

Female papaya flowers have five free petals and a rounded superior ovary (Fig. 2.3c) (Ronse Decraene and Smets 1999) that is five carpellate and hollow and exhibits

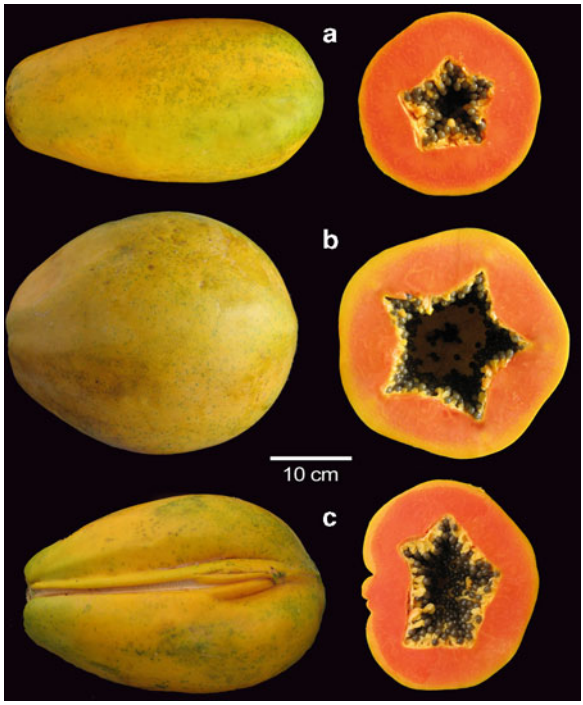

Fig. 2.6 Papaya fruits according to sex type. (**a**) Fruit of a hermaphroditic plant. (**b**) Fruit from a female plant. (**c**) Misshapen fruit from a hermaphroditic plant due to carpellody

parietal placentation (Fisher 1980). In contrast to the hermaphroditic plants, females are completely stable and their flowers do not appear to undergo sex reversal due to environmental fluctuations.

Male Flowers

Stamen arrangement in the male flowers is the same as in the hermaphrodite flower, surrounding a rudimentary pistil or pistillode (Fig. 2.3a). In some cases, due to genetic or environmental causes, some of the dominant flowers within the inflorescence may have fully developed pistils, resulting in a hermaphroditic flower and an overall male, fruit-bearing phenotype (Fig. 2.2d) (Storey 1953).

Fruits

The papaya fruit has been studied extensively (Roth and Clausnitzer 1972; Roth 1977). Papaya fruits are berries and show high diversity in size and shape. Fruits from hermaphroditic plants tend to be elongated and vary from cylindrical to pear shaped, while fruits of female plants tend to be round (Fig. 2.6a, b). Fruit size can

vary extensively, ranging from less than 100 g in some wild accessions to over 10 kg in certain landraces. The pericarp may become 2.5–3.0 cm thick in ripe fruits. A large cavity that hosts the seeds makes up most of the fruit volume. The ripe fruit is built mostly of parenchymatous tissue organized in three distinct layers: the outer layer is composed of smaller, plastid-rich cells; the median zone is composed of larger round cells, rich in intercellular spaces; and the inner region is a spongy parenchyma with stretched, branched cells and abundant air spaces. Two sets of five vascular bundles each enter the fruits, one dorsal (outer) and another ventral inner rings (Roth 1977).

The contribution of fruit photosynthesis to carbon gain is probably small, and significant only during early stages of growth, when the surface:volume ratio of young green fruits is high; in addition, the fruits develop in the shade under the dense papaya canopy, which also limits photosynthesis. Fruit photosynthesis could be more important in refixing respiratory CO_2 released in the fruit interior. Fruit development takes approximately 5 months to maturity.

Papaya fruit ripening is climacteric, and high ethylene production may start only hours after harvest at the recommended stage (appearance of one to two yellow stripes on the fruit). As they ripen, papaya fruits change color, firmness, carbohydrate composition, and production of secondary compounds, which are responsible for fruit color and fragrance. The color of ripe fruits may vary from yellow to salmon red. The most important carotenoids are lycopene and β-cryptoxanthin (Schweiggert et al. 2011a, b), and a rich list of more than 150 volatile esters and alcohols are responsible for the sweet aroma and flavor of the fruit (Pino et al. 2003).

Papaya fruits are sensitive to chilling injury, and storage temperatures below 10 °C can cause fruit shriveling, skin pitting, localized necrosis, softening, electrolyte leakage, and postharvest disease outbreaks (Chen and Paull 1986). Treating the fruits with hot water may reduce their sensibility to chilling damage (Paull and Jung Chen 2000), perhaps by inducing the production of plant stress and heat-shock proteins (McCollum et al. 1993). Other important physiological disorders of papaya fruits are bruising ("skin freckles") and translucent and lumpy pulp (Oliveira and Vitória 2011).

Seeds

Well-pollinated fruits can have 600 black seeds or more (Fig. 2.6a–c). The embryo is straight and ovoid, with flattened cotyledons (Fisher 1980). The seeds are coated by a mucilaginous mass derived from the pluristratified epidermis of the outer integument (Roth 1977). The embryo is enclosed in a gelatinous sarcotesta at physiological maturity. A compact mesotesta and outer and inner integuments can be observed underneath. The endosperm is composed of thin-walled cells with abundant oil bodies and aleurone grains, lacking starch at maturity (Fisher 1980; Teixeira da Silva et al. 2007). Photosensitive seeds of wild papayas are dormant at maturity, and their germination may be triggered by changes in light quality during forest gap formation (Paz and Vázquez-Yanes 1998).

Photosynthetic Carbon Gain, Water Use, and Source: Sink Relations

Papaya plants exhibit C_3 photosynthesis (Campostrini and Glenn 2007). Optimum temperature for growth is 21–33 °C, under which papayas can produce 2 leaves/week and 8–16 fruits/month. Temperatures below 10 °C are not well tolerated (Allan 2002, 2005). Light compensation point for leaf-level photosynthesis is ca. 35 μmol m^{-2} s^{-1}, and saturation is reached at ca. 2,000 μmol m^{-2} s^{-1} of photosynthetic photon flux density (PPFD) (Campostrini and Glenn 2007). High photosynthetic rates of 25–30 μmol CO_2 m^{-2} s^{-1} alternate with pronounced (midday) gas exchange depressions, apparently caused by direct stomatal and cuticular responses to air humidity (El-Sharkawy et al. 1985; Marler and Discekici 1997; Marler and Mickelbart 1998). This process may reduce productivity by 35–50 % (Campostrini and Glenn 2007).

A further reduction of 25–30 % in net efficiency of carbon assimilation can be caused by photorespiration, which could be ameliorated with improved water availability and PPFD conditions. The reduction observed in photosynthesis at high PPFD could also be a consequence of a decrease in stomatal conductance caused by the direct action of radiant energy on leaf heating. Chronic photo-inhibition through damage and replacement of the D1 protein in the reaction center of PSII also reduces photosynthesis at high PPFD levels (Reis et al. 2006; Campostrini and Glenn 2007). Another important factor in carbon balance is the cost of growth and maintenance respiration. Whole papaya plants growing in a greenhouse respired from 400 μL O_2 g^{-1} h^{-1}, about one third of daily carbon fixation, at 25 °C, to 1,600 μL O_2 g^{-1} h^{-1} when the plants were exposed to very high temperatures (50 °C) (Todaria 1986).

In contrast to other tropical fruits (like bananas), papayas do not store starch, and continuous flowering and fruiting requires a steady carbon flow from the leaves. Sugar accumulation is controlled by three key enzymes. Sugar content increases slowly during the first two thirds of fruit development, under sucrose synthetase regulation, and later increases substantially during ripening, controlled by apoplastic invertases. The third enzyme, sucrose-phosphate synthase, remains low but active throughout the whole course of fruit development (Zhou et al. 2000; Zhou and Paull 2001).

Maximum rates of sap flow in the xylem are close to 0.6–0.8 L H_2O m^{-2} h^{-1}, and transpiration rates are about 25 mmol H_2O m^{-2} s^{-1} (Reis et al. 2006). In general, 1 m^2 of leaves transpires 1 L H_2O daily but can rise substantially with increasing evaporative demands. A papaya plant with 35 leaves, equivalent to approximately 3.5–4 m^2 leaf area, can fix ca. 70 g of CO_2 and transpire ca. 10 L of water daily (Coelho Filho et al. 2007). For well-watered papaya plants, the crop irrigation coefficient (Kc) is close to unity but may reach values of 1.2, as a consequence of the strong dependence of canopy gas exchange, photosynthesis, and water use, on the solar radiation available (Campostrini and Glenn 2007; Coelho Filho et al. 2007).

Responses to water stress include dehydration postponement through strict stomatal regulation, cavitation repair, and intense osmotic adjustment (Marler and Mickelbart 1998; Marler 2000; Mahouachi et al. 2006). The succulent roots do not tolerate

excess water, and 2 days with hypoxia cause chlorosis, leaf shedding, and even death after 3–4 days with oxygen deprivation (Campostrini and Glenn 2007).

In modern cultivars, one papaya leaf can sustain the development of three to four fruits. However, there are indications of poor adjustment capacity of source:sink ratios in fruiting papaya plants, presumably because the fruits have low capacity to attract assimilates (Acosta et al. 1999; Zhou et al. 2000). This is important because in most crops, biomass allocation to the harvested organ is the yield component most susceptible to selection and breeding (Bugbee and Monje 1992).

Nutrient Economy

The nutritional requirements of papaya plants are high. Mineral nutrients are taken up by plants grown at full sunlight as follows: K>N>Ca>P>S>Mg (for macronutrients) and Cl>Fe>Mn>Zn>B>Cu>Mo (for micronutrients). Nitrogen, phosphorus, and potassium, very important in metabolism and frequently limiting in tropical soils, are extracted in high amounts: a ton of fresh harvested fruits contains 1,770, 200, and 2,120 g of each of these nutrients, respectively. High-density plantings may extract 110, 10, and 103 kg of N, P, and K per ha, respectively; however, this can be much higher, depending on yield. The fruits represent 20–30 % of the nutrients removed. Thus, the development of healthy rhizosphere and mycorrhizal associations should sustain locally tight biogeochemical cycles and guide the formulation of fertilizers and amendments to deal with multiple soil stresses in tropical environments, such as shallowness, compaction, poor aeration, mineral deficiencies, and nutrient imbalances (Villachica and Raven 1986; Arango-Wiesner 1999).

Conclusions

Papaya plants are superb vegetables. Their morphological and ecophysiological attributes are as impressive as the highly efficient mechanisms of resource capture, transport, and utilization that support them. High photosynthetic rates, carbon gain, reproductive output, and growth and plasticity occur at the expense of high rates of water use and mineral nutrient demand. Phenotypic plasticity is high at the shoot, root, and reproductive levels. All of these attributes have implications in the design of sustainable cropping systems for papaya. To achieve this, it is convenient to conduct interdisciplinary integration constructed around the whole plant, at various scales of observation (cells and tissues, organs, whole plants, and orchards), and along the soil–plant–atmosphere continuum. The papaya plant may also constitute a model for ecophysiological studies, linking the genomics of short-lived perennials, through studies of growth, metabolism, sex expression, and longevity, to population ecology and evolutionary questions.

References

Acosta C, González HV, Livera M, Matheis M (1999) Respuesta de las plantas de papayo al diferente número de frutos por planta. I. Distribución de biomasa. Rev Chapingo (México) Serie Hortic 5(2):131–136

Allan P (2002) *Carica papaya* responses under cool subtropical growth conditions. Acta Hortic 575:757–763

Allan P (2005) Phenology and production of *Carica papaya* "Honey Gold" under cool subtropical conditions. Acta Hortic 740:217–223

Almeida FT, Marinho CS, Souza EF, Grippa S (2003) Expressão sexual do mamoeiro sob diferentes lâminas de irrigação na Região Norte Fluminense. Rev Brasil Fruticult 25:383–385

Arango-Wiesner LV (1999) El cultivo de la papaya en los llanos orientales de Colombia. Manual de Asistencia Técnica No 4, vol 4. Villavicencio, Colombia

Arkle TD, Nakasone HY (1984) Floral differentiation in the hermaphroditic papaya. HortScience 19(6):832–834

Awada M (1953) Effects of moisture on yield and sex expression of the papaya plants (*Carica papaya* L.). Hawaii Agricultural Experiment Station Progress Notes 97, 4 pp

Awada M (1958) Relationships of minimum temperature and growth rate with sex expression of papaya plants (*Carica papaya* L.). Hawaii Agric Exp Station Tech Bull 38:1–16

Awada M, Ikeda WS (1957) Effects of water and nitrogen application on composition, growth, sugars in fruits, yield, and sex expression of the papaya plants (*Carica papaya* L.). Hawaii Agric Exp Station Tech Bull 33:3–16

Azarkan M, El Moussaoui A, van Wuytswinkel D, Dehon G, Looze Y (2003) Fractionation and purification of the enzymes stored in the latex of *Carica papaya*. J Chromatogr B 790(1–2): 229–238

Becker S (1958) The production of papain—an agricultural industry for tropical America. Econ Bot 12(1):62–79

Brown JE, Bauman JM, Lawrie JF, Rocha OJ, Moore RC (2012) The structure of morphological and genetic diversity in natural populations of *Carica papaya* (Caricaceae) in Costa Rica. Biotropica 44(2):179–188

Bucci SJ, Scholz FG, Goldstein G, Meinzer FC, Sternberg LDSL (2003) Dynamic changes in hydraulic conductivity in petioles of two savanna tree species: factors and mechanisms contributing to the refilling of embolized vessels. Plant Cell Environ 26(10):1633–1645

Bugbee B, Monje O (1992) The limits of crop productivity: theory and validation. Bioscience 42(7):494–502

Buisson D, Lee DW (1993) The developmental responses of papaya leaves to simulated canopy shade. Am J Bot 80(8):947–952

Campostrini E, Glenn DM (2007) Ecophysiology of papaya: a review. Braz J Plant Physiol 19:413–424

Canini A, Alesiani D, D'Arcangelo G, Tagliatesta P (2007) Gas chromatography–mass spectrometry analysis of phenolic compounds from *Carica papaya* L. leaf. J Food Compos Anal 20(7):584–590

Carneiro CE, Cruz JL (2009) Caracterização anatômica de órgãos vegetativos do mamoeiro. Ciênc Rural 39(3):918–921

Carvalho FA, Renner SA (2013) The phylogeny of Caricaceae. In: Ming R, Moore PH (eds) Genetics and genomics of papaya. Springer Science+Business Media, New York

Chen NM, Paull RE (1986) Development and prevention of chilling injury in papaya fruit. J Am Soc Hortic Sci 111(4):639–643

Clemente HS, Marler TE (2001) Trade winds reduce growth and influence gas exchange patterns in papaya seedlings. Ann Bot 88(3):379–385

Coelho Filho MA, Coelho EF, Cruz LL (2007) Uso da transpiração máxima de mamoeiro para o manejo de irrigação por gotejamento em regiões úmidas e sub-úmidas, vol 162, EMBRAPA. Cruz das Almas, Bahia

Cohen E, Lavi U, Spiegel-Roy P (1989) Papaya pollen viability and storage. Sci Hortic 40(4): 317–324

da Silva F, Pereira M, Junior P, Pereira T, Viana A, Daher R, Ramos H, Ferreguetti G (2007) Evaluation of the sexual expression in a segregating BC_1 papaya population. Crop Breed Appl Biotechnol 7(1):16–23

Dhekney SA, Litz RE, Moraga Amador DA, Yadav AK (2007) Potential for introducing cold tolerance into papaya by transformation with C-repeat binding factor (CBF) genes. Vitro Cell Dev Biol Plant 43(3):195–202

El Moussaoui A, Nijs M, Paul C, Wintjens R, Vincentelli J, Azarkan M, Looze Y (2001) Revisiting the enzymes stored in the laticifers of *Carica papaya* in the context of their possible participation in the plant defence mechanism. Cell Mol Life Sci 58(4):556–570

El-Sharkawy M, Cock J, Hernandez A (1985) Stomatal response to air humidity and its relation to stomatal density in a wide range of warm climate species. Photosynth Res 7(2):137–149

Ewel JJ (1986) Designing agricultural ecosystems for the humid tropics. Annu Rev Ecol Syst 17(1):245–271

Fisher JB (1980) The vegetative and reproductive structure of papaya (*Carica papaya*). Lyonia 1:191–208

Fisher JB, Mueller RJ (1983) Reaction anatomy and reorientation in leaning stems of balsa (*Ochroma*) and papaya (*Carica*). Can J Bot 61(3):880–887

Fitch MMM (2005) *Carica papaya* – Papaya. In: Litz RE (ed) Biotechnology of fruit and nut crops, vol 29. CABI, Cambridge, pp 174–207

Hagel JM, Yeung EC, Facchini PJ (2008) Got milk? The secret life of laticifers. Trends Plant Sci 13(12):631–639

Hart RD (1980) A natural ecosystem analog approach to the design of a successional crop system for tropical forest environments. Biotropica 12(2):73–82

Iyer CPA, Kurian RM (2006) High density planting in tropical fruits: principles and practice. International Book Distributing Co, Delhi

Khade SW, Rodrigues BF, Sharma PK (2010) Arbuscular mycorrhizal status and root phosphatase activities in vegetative *Carica papaya* L. varieties. Acta Physiol Plant 32(3):565–574

Kim M, Moore P, Zee F, Fitch MMM, Steiger D, Manshardt R, Paull R, Drew RA, Sekioka T, Ming R (2002) Genetic diversity of *Carica papaya* as revealed by AFLP markers. Genome 45(3):503–512

Konno K, Hirayama C, Nakamura M, Tateishi K, Tamura Y, Hattori M, Kohno K (2004) Papain protects papaya trees from herbivorous insects: role of cysteine proteases in latex. Plant J 37(3):370–378

Leal-Costa MV, Munhoz M, Meissner Filho PE, Reinert F, Tavares ES (2010) Anatomia foliar de plantas transgênicas e não transgênicas de *Carica papaya* L. (Caricaceae). Acta Bot Brasil 24:595–597

León J (1987) Botánica de los cultivos tropicales. IICA, San José

Madrigal SL, Ortiz NA, Cooke RD, Fernandez HR (1980) The dependence of crude papain yields on different collection ('tapping') procedures for papaya latex. J Sci Food Agric 31(3): 279–285

Mahouachi J, Socorro A, Talon M (2006) Responses of papaya seedlings (*Carica papaya* L.) to water stress and re-hydration: growth, photosynthesis and mineral nutrient imbalance. Plant Soil 281(1):137–146

Marler TE (2000) Water conductance and osmotic potential of papaya (*Carica papaya* L.) roots as influenced by drought. In: Stokes A (ed) The supporting roots of trees and woody plants: form, function and physiology, vol 87. Kluwer, Dordrecht, pp 239–244

Marler TE, Discekici HM (1997) Root development of 'Red Lady' papaya plants grown on a hillside. Plant Soil 195(1):37–42

Marler TE, Mickelbart MV (1998) Drought, leaf gas exchange, and chlorophyll fluorescence of field-grown papaya. J Am Soc Hortic Sci 123(4):714–718

Martins DJ, Johnson SD (2009) Distance and quality of natural habitat influence hawkmoth pollination of cultivated papaya. Int J Trop Insect Sci 29(3):114–123

McCollum TG, D'Aquino S, McDonald RE (1993) Heat treatment inhibits mango chilling injury. HortScience 28(3):197–198

Ming R, Yu Q, Moore PH (2007) Sex determination in papaya. Semin Cell Dev Biol 18(3):401–408

Ming R, Yu Q, Blas A, Chen C, Na JK, Moore PH (2008) Genomics of papaya, a common source of vitamins in the tropics. In: Moore PH, Ming R (eds) Genomics of tropical crop plants, vol 1. Springer Science+Business Media, New York, pp 405–420

Morton J (1987) Papaya. In: Fruits of warm climates. Julia F. Morton, Miami, pp 336–346

Nakasone HY, Lamoureux C (1982) Transitional forms of hermaphroditic papaya flowers leading to complete maleness. J Am Soc Hortic Sci 107(4):589–592

Niklas KJ, Marler TE (2007) *Carica papaya* (Caricaceae): a case study into the effects of domestication on plant vegetative growth and reproduction. Am J Bot 94(6):999–1002

Oliveira JG, Vitória AP (2011) Papaya: nutritional and pharmacological characterization, and quality loss due to physiological disorders. An overview. Food Res Int 44(5):1306–1313

Paterson A, Felker P, Hubbell S, Ming R (2008) The fruits of tropical plant genomics. Trop Plant Biol 1(1):3–19

Paull RE, Jung Chen N (2000) Heat treatment and fruit ripening. Postharvest Biol Technol 21(1):21–37

Paz L, Vázquez-Yanes C (1998) Comparative seed ecophysiology of wild and cultivated *Carica papaya* trees from a tropical rain forest region in Mexico. Tree Physiol 18(4):277–280

Pino JA, Almora K, Marbot R (2003) Volatile components of papaya (*Carica papaya* L., Maradol variety) fruit. Flavour Fragrance J 18(6):492–496

Porter BW, Zhu YJ, Webb DT, Christopher DA (2009) Novel thigmomorphogenetic responses in *Carica papaya*: touch decreases anthocyanin levels and stimulates petiole cork outgrowths. Ann Bot 103(6):847–858

Posse RP, Sousa EF, Bernardo S, Pereira MG, Gottardo RD (2009) Total leaf area of papaya trees estimated by a nondestructive method. Scientia Agricola 66:462–466

Ramos HCC, Pereira MG, Silva FF, Viana AP, Ferreguetti GA (2011) Seasonal and genetic influences on sex expression in a backcrossed segregating papaya population. Crop Breed Appl Biotechnol 11:97–105

Reis FO, Campostrini E, Sousa EF, Silva MG (2006) Sap flow in papaya plants: Laboratory calibrations and relationships with gas exchanges under field conditions. Sci Hortic 110(3):254–259

Rodrigues S, Da Cunha M, Ventura JA, Fernandes P (2009) Effects of the *Papaya meleira virus* on papaya latex structure and composition. Plant Cell Rep 28(5):861–871

Ronse Decraene LP, Smets EF (1999) The floral development and anatomy of *Carica papaya* (Caricaceae). Can J Bot 77(4):582–598

Roth I (1977) Fruits of angiosperms. Borntraeger, Berlin, xvi, 675 pp. (Handbuch der Pflanzenanatomie spezieller Teil, Band x, Teil 1). Anatomy and Morphology (KR, 197706808)

Roth I, Clausnitzer I (1972) Desarrollo y anatomia del fruto y de la semilla de Carica papaya L. (lechosa). Acta Bot Venezuelica 7:187–206

Schweiggert R, Steingass C, Heller A, Esquivel P, Carle R (2011a) Characterization of chromoplasts and carotenoids of red- and yellow-fleshed papaya (*Carica papaya* L.). Planta 234(5):1031–1044

Schweiggert RM, Steingass CB, Mora E, Esquivel P, Carle R (2011b) Carotenogenesis and physico-chemical characteristics during maturation of red fleshed papaya fruit (*Carica papaya* L.). Food Res Int 44(5):1373–1380

Sheldrake AR (1969) Cellulase in latex and its possible significance in cell differentiation. Planta 89(1):82–84

Sritakae A, Praseartkul P, Cheunban W, Miphokasap P, Eiumnoh A, Burns P, Phironrit N, Phuangrat B, Kitsubun P, Meechai A (2011) Mapping airborne pollen of papaya (*Carica papaya* L.) and its distribution related to land use using GIS and remote sensing. Aerobiologia 27(4):291–300

Storey WB (1953) Genetics of the papaya. J Hered 44(2):70–78

Teixeira da Silva JA, Rashid Z, Nhut DT, Sivakumar D, Gera A, Souza MT Jr, Tennant PF (2007) Papaya (*Carica papaya* L.) biology and biotechnology. Tree Forest Sci Biotechnol 1(1):47–73

Todaria N (1986) Respiration rates of some greenhouse cultivated tropical and subtropical species. Biol Plant 28(4):280–287

Vega-Frutis R, Guevara R (2009) Different arbuscular mycorrhizal interactions in male and female plants of wild *Carica papaya* L. Plant Soil 322(1):165–176

Villachica H, Raven K (1986) Nutritional deficiencies of pawpaws (*Carica papaya* L.) in the central tropical forest of Peru. Turrialba 36(4):523–531

Walsh KB, Ragupathy S (2007) Mycorrhizal colonisation of three hybrid papayas (*Carica papaya*) under mulched and bare ground conditions. Aust J Exp Agric 47:81–85

Zhou L, Paull RE (2001) Sucrose metabolism during papaya (*Carica papaya*) fruit growth and ripening. J Am Soc Hortic Sci 126(3):351–357

Zhou L, Christopher DA, Paull RE (2000) Defoliation and fruit removal effects on papaya fruit production, sugar accumulation, and sucrose metabolism. J Am Soc Hortic Sci 125(5): 644–652

Zunjar V, Mammen D, Trivedi BM, Daniel M (2011) Pharmacognostic, physicochemical and phytochemical studies on *Carica papaya* Linn. leaves. Pharmacognosy J 3(20):5–8

Chapter 3
Phenotypic and Genetic Diversity of Papaya

Paul H. Moore

Role of Diversity in Crop Improvement

Genetic polymorphism is the primary source for variation of the morphological and physiological appearance (phenotypes) of plants and animals. Natural selection for or against a particular phenotype leads to evolutionary changes within and among species and eventually to the level of genetic diversity in the species, which in turn provides the basis for adaptation to changes in environmental conditions (Hammer et al. 2003).

Knowledge of the diversity within a particular crop species is necessary for understanding evolutionary relationships, developing efficient strategies for preservation of genetic resources, and effectively directing the breeding and selection for crop improvement. When clear-cut traits such as color, organ morphology, pathogen interactions, or enzyme variants exist largely independent of environmental influences, they are reliable qualitative traits that can be easily evaluated to reflect the genetic diversity of the crop to a high degree. However, when the phenotypic traits are not so distinctive for a particular genotype or when they are influenced by the environment, they are quantitative in nature so that more costly direct molecular measures must be used to evaluate their genetic basis. Agronomically important traits, such as crop yield, plant height, leaf and fruit size, fruit quality, and tolerance to pests and diseases are generally quantitative traits whose diversity is evaluated at the molecular level. Discovering the relationship between a molecular genotype and a desired agronomic trait will facilitate the development and maintenance of germplasm repositories and enable breeders in the production of new and improved cultivars.

The magnitude and structure of the genetic diversity of a population determine the ability of that population to adapt to its environment through natural selection

P.H. Moore (✉)
Hawaii Agriculture Research Center, Texas AgriLife Research,
Kaneohe, HI 96744, USA
e-mail: pmoore@harc-hspa.com

R. Ming and P.H. Moore (eds.), *Genetics and Genomics of Papaya*, Plant Genetics and Genomics: Crops and Models 10, DOI 10.1007/978-1-4614-8087-7_3,

of the particular combination of genes conferring greatest fitness. Thus, natural populations need to have adequate genetic diversity to assure their continued existence in the face of changing biotic and abiotic factors of the environment. Similarly, crop improvement programs require identification and utilization of adequate genetic diversity in breeding populations to enable the recombination of different sets of genes in order to select progeny with desired traits. Just as in nature, the traits desired by breeders are those that enable the crop to be successful in a target environment, or against targeted pests and pathogens, or to have qualities such as good nutritional value or flavor important to consumers.

Published analyses on the degree of phenotypic and genetic diversity of papaya greatly reflect the breadth of the taxonomic group under consideration at the time. When the genus *Carica* included 22 species in two sections, *Carica* and *Vasconcellea*, there were reports of greater diversity than when the taxonomic group was constricted to a single species, *Carica papaya*. To be specific, common papaya (*C. papaya* L.) is a member of the small family Caricaceae, which is comprised of six genera some of which have changed taxonomic rank over time (Ming et al. 2005; Scheldeman et al. 2007). Notably, the genus *Carica* has always included the common papaya but has at times also included highland papayas of the genus *Vasconcellea* as a section within the genus *Carica*. However, *Vasconcellea* was recently restored as a genus (Badillo 2000) comprised of 21 species some of which are known as highland papayas because they generally occur at higher altitudes (Ming et al. 2005; Scheldeman et al. 2007). With the restoration of *Vasconcellea* as a genus, the genus *Carica* has been reduced to the single species *C. papaya* with less diversity than when it included *Vasconcellea*.

Phenotype Diversity

Common papaya germplasm shows moderate to high phenotypic variation for the morphological characters of leaf shape and size, types of inflorescences and flowers, fruit shape and size, and reactions to pests and diseases (descriptors provided by IBPGR 1988). The most distinctive and economically most significant phenotypic traits of papaya varieties are related to flower and fruit characteristics (Table 3.1).

The morphology of papaya inflorescences and flowers varies with sex of the tree. Papaya varieties are typically either dioecious (with unisexual flowers on exclusively male and female plants) or gynodioecious (with bisexual flowers on hermaphrodite plants and unisexual flowers on female plants); see chapters 16 and 17 and Ming et al. 2007 for a more complete discussion of papaya sex. The inflorescence stalks, or peduncle, of female plants are short (2.5–6 cm long) bearing only one or a few large bell-shaped flowers that have curvy separate petals. Fruit from the female flowers varies from spherical to ovoid. The inflorescence peduncles of male plants are 60–90 cm to over 150 cm long and pendulous and bear numerous smaller trumpet-shaped flowers with the petals and stamen filaments fused in a narrow tube which has flared lobes. The inflorescence peduncles of hermaphrodite plants are

Table 3.1 Common papaya varieties in commerce and breeding

Variety	Origin	Average fruit size, notable traits	Fruit characteristics (e.g., shape, color)
Bettina	Australia (Florida Betty × Queensland var.)	1.36–2.27 kg	Round-ovoid Well colored
Cariflora	Florida, USA	0.8 kg Tolerant to PRSV	Round, dark yellow to light orange flesh
Coorg Honey Dew[H]	India	2–3.5 kg	Long to ovoid, yellow
Eksotika[H]	Malaysia (Sunrise Solo × Sugang 6)	0.6–0.90 kg	Elongate (from hermaphrodite), orange–red flesh
Eksotika II[H]	Malaysia (Eksotika lines 19 × 20)	0.6–1.0 kg, higher yield than from Eksotika	Fewer freckles on skin and sweeter than Eksotika
Hortus Gold (selection: Honey Gold)	South Africa	1 kg, propagated from cuttings	Round-ovoid, golden yellow flesh
Kapoho Solo[H]	Hawaii, USA	0.45 kg	Pear shaped, but shorter neck than "Sunrise Solo"; orange–yellow flesh
Known You I[H]	Taiwan	1.6–3 kg, tolerant to PRSV	Very long and slender, yellow flesh
Maradol	Cuba	2.6 kg	Elongate, green or yellow skin
Rainbow[H]	Hawaii, USA (SunUp × Kapoho Solo)	0.65 kg, transgenic resistance to PRSV	Pear shaped to ellipsoid, yellow–orange flesh
Red Lady 786	Taiwan	1.5–2 kg, tolerant to PRSV	Elongate, red flesh
Red Maradol	Mexico	2.5–2.6 kg	Red flesh, yellow–orange skin
Sekaki[H]	Malaysia	1.0–2.5 kg	Long, cylindrical with smooth skin; red, firm flesh
Solo[H]	Developed in Hawaii, USA; originally from Barbados	0.5–1 kg, bisexual flowers, highly selfing	Pear shaped (from hermaphrodites), orange–yellow skin, golden orange flesh
Sunset Solo[H]	Hawaii, USA	0.4–0.55 kg,	Pear shaped, red firmer flesh
Sunrise Solo[H]	Hawaii, USA	0.57 kg	Pear shaped, reddish-pink flesh
SunUp[H]	Hawaii, USA	0.45–0.9 kg transgenic resistance to PRSV	Pear shaped, red flesh
Tainaung 1[H]	Taiwan	1.1 kg	Pointed blossom end (from hermaphrodite), red flesh
Waimanalo[H]	Hawaii, USA	0.6–0.9 kg	Pear shaped to round ovoid, yellow flesh

Revised with permission from Chan and Paull (2008); and from Anonymous (2005)
H hermaphrodite variety (i.e., gynodioecious)

intermediate to the unisexual types with peduncles less than 2.5–6 cm long and bearing bisexual flowers that can be sexually variable, but they are generally tubular with a midpoint or lower constriction below the larger petal lobes. Fruits of hermaphrodite plants are pear shaped (pyriform) with a variable neck constriction depending on the variety.

Although papaya sexuality is genetically controlled, sex expression can change under the influence of environmental conditions. Hermaphrodite plants can stop producing hermaphrodite flowers to produce male flowers through reductions in ovary size and function. Male plants may exhibit seasonal sex reversal by developing staminate, bisexual, and pistillate flowers (Storey 1958, 1976). Young hermaphrodite plants may have male flowers when under stress, but bisexual flowers with optimal conditions. Female plants have stable sex expression and are not known to develop flowers with masculine structures (Hofmeyr 1939; Nakasone and Lamoureux 1982).

Phenotype diversity in papaya is broad for many traits (IBPGR 1988), including the horticulturally important traits of the length of the juvenile phase, plant stature, stamen carpellody, carpel abortion, fruit characters, and reactions to pests and pathogens. In addition, commercial papaya cultivars may be inbred gynodioecious lines, typified by the Hawaiian Solo lines, or outcrossing dioecious populations, such as the Australian papaws from southern Queensland, F1 hybrids, the Tainung series (Taiwan), Eksotica II (Malaysia), Rainbow (Hawaii), and Hortus Gold (South Africa). As discussed previously, fruit shape reflects the flower type and flower type reflects plant sex. Preference of local consumers and markets also influence the fruit shape, size, flesh color, flavor, and sweetness (Nakasone and Paull 1998). The generally large fruits of certain varieties (Table 3.1) vary from spherical or ovoid to pear shaped or elongate from 10 to 50 cm in length. Fruit weight can vary substantially from 0.35 to 10 kg, or even 12 kg, again chiefly dependent on selection by growers for specific local markets. Storey (1969) reported regional preference for 2.5–6.0 kg fruits in South America and the South Pacific, 1.25–2.5 kg lobular fruits in South Africa, and small 0.4–0.5 kg for Solo-type fruits developed in Hawaii. Solo refers to a group of small-fruited, high sugar content, cultivars developed in Hawaii from introductions from Barbados in 1910 (Storey 1969).

Morphological and isozyme idiotypes have been reported on 32 feral, 7 home garden, and 7 commercially cultivated lines of common papaya in Costa Rica (d'Eeckenbrugge et al. 2007). As indicated previously, considerable morphological variation has been reported for fruit traits among the typical wild female-type plants (very small seedy fruits with thin yellow mesocarp) to commercial fruits (medium to large fruits with abundant orange or reddish pulp). The wild-type fruit lines were more common along the Pacific Coast, while more diverse fruit types were more common on the Caribbean side of the country indicating local introgression between the wild and cultivated accessions. All of the unimproved materials were morphologically differentiated from commercial cultivars by having larger leaves and petioles, with less marked leaf divisions, presence of bracts, and smaller flowers. Analysis of variance showed significant sex effect on length and lobulation of the leaves. Isozyme distinctions were less informative failing to indicate significant genetic diversity. Results of this initial study were largely confirmed by a

subsequent independent study in Costa Rica of morphological and microsatellite diversity among natural populations and cultivars (Rieger 2009). Morphological diversity was greatest among the natural populations for reproductive characteristics. The greater morphological diversity among natural papaya populations was seen also with microsatellite diversity (Pérez 2006; Eustice et al. 2008).

Genetic Diversity

The major advantage of molecular methods over phenotyping for characterizing plant diversity is the greater resolution at the DNA level and that sequences can provide a direct link to diagnostic genes (Table 3.2). In addition, since molecular methods exclude all environmental influences, they can be applied at any developmental phase. The earliest published molecular study of genetic diversity in papaya germplasm examined a small group of seven Hawaiian Solo cultivars and three unrelated papaya lines, using random amplified polymorphic DNA (RAPD) markers (Stiles et al. 1993). The objective of that study was to test the validity of RAPD data for revealing phylogenetic relationships among papaya lines of known pedigree, rather than to provide a broad overview of genetic diversity. Results indicated that RAPD analysis could be used to evaluate the relationships among the tested cultivars which showed only a moderate degree of genetic diversity with simple matching coefficients ranging from 0.7 to 0.95, indicative of a narrow genetic base for domesticated papayas.

Genetic relationships between common papaya and related *Vasconcellea* species within the family Caricaceae was also investigated using a combination of RAPD and isozyme markers (Jobin-Décor et al. 1997), restriction fragment length polymorphisms (RFLP) within a cpDNA intergenic spacer region (Aradhya et al. 1999), amplified fragment length polymorphism (AFLP) markers (Van Droogenbroeck et al. 2002), and microsatellite markers (Pérez 2006; Eustice et al. 2008). The most extensive study (Van Droogenbroeck et al. 2002) was based on AFLPs which analyzed 95 accessions belonging to three genera including *C. papaya*, at least eight *Vasconcellea* spp. and two *Jacaratia* spp., all collected in Ecuador. The cluster analysis, based on 496 polymorphic AFLP markers produced from five primer combinations, clearly separated the species of the three genera and showed a large genetic distance between the accessions of *C. papaya* and the *Vasconcellea* spp. Collectively, these diversity analysis studies, reviewed by Ming et al. (2005), suggested that *C. papaya* diverged from *Vasconcellea* spp. early in its evolution. Aradhya et al. (1999) and Van Droogenbroeck et al. (2002) noted that the wild South American *Carica* species are more closely allied with a member of the related genus *Jacaratia* than to *C. papaya.*

Genetic relationships among *C. papaya* cultivars, breeding lines, unimproved germplasm, and the *Vasconcellea* section of Carica (Jobin-Décor et al. 1997) were established using amplified fragment length polymorphism (AFLP) markers (Kim et al. 2002). Seventy-one *C. papaya* accessions and related *Carica* species were analyzed with nine *Eco*RI–*Mse*I primer combinations. A total of 186 informative AFLP markers was generated and analyzed. Cluster analysis suggested limited

Table 3.2 Comparison of methods for measuring genetic variation and diversity

Method	Variation detected	Throughput	Loci analyzed per assay	Reproducibility	Type of character	Inheritance of character	Technology level
Morphology	Low	High	Low no.	Medium	Phenotypic	Qualitative/quantitative	Low
Isozymes	Medium	Medium	Low no.	Medium	Proteins	Codominant	Medium
RFLP (low copy)	Medium	Low	Low no. (specific)	Good	DNA	Codominant	High
RFLP (high copy)	High	Low	High no. (specific)	Good	DNA	Dominant	High
RAPD	High to medium	High	High no. (random)	Poor	DNA	Dominant	Medium
AFLPs	Medium to high	High	High no. (random)	Medium	DNA	Dominant	High
Sequence-tagged SSRs	High	High	Medium no. (specific)	Good	DNA	Codominant	High
DNA sequencing (old)	High	Low	Low no. (specific)	Good	DNA	Codominant/dominant	High
DNA sequencing (new)	High	High	High no. (specific)	Good	DNA	Codominant/dominant	High

Modified with kind permission of Springer Science+Business Media from Hammer et al. (2003)

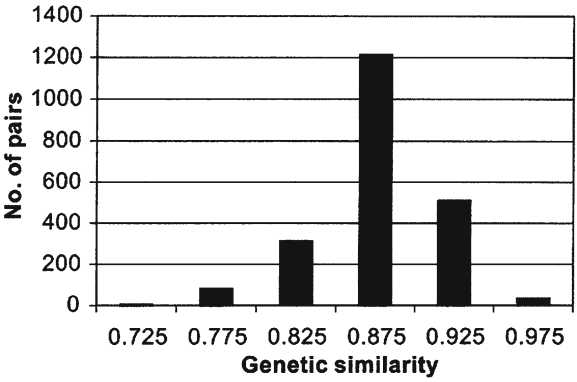

Fig. 3.1 Distribution of pair-wise comparisons of genetic similarity among papaya cultivars. (Used with permission from Kim et al. (2002). Copyright © 2008 Canadian Science Publishing.)

genetic variation within papaya, showing an average genetic similarity among 63 papaya accessions of 0.880 (Fig. 3.1). Genetic diversity among cultivars derived from the same or similar gene pools was smaller, as shown by the Hawaiian Solo hermaphrodite cultivars showing a genetic similarity at 0.921and the Australian dioecious cultivars showing a genetic similarity of 0.912 (Fig. 3.2). Surprisingly, the results indicated that the open-pollinated dioecious cultivars were no more variable than were the self-pollinated hermaphrodite cultivars, possibly due to the narrow genetic base from which both dioecious and hermaphrodite cultivars were derived (Ming et al. 2005).

Genetic diversity between *C. papaya* and six related Caricaceae species was also evaluated. *C. papaya* shared the least genetic similarity with these species, with an average genetic similarity of 0.432. The average genetic similarity among the six other species was 0.729. The results from AFLP markers provided detailed estimates of the genetic variation within and among papaya cultivars and supported the notion that *C. papaya* diverged from the rest of *Carica* species early in the evolution of this genus.

A more extensive evaluation of papaya genetic diversity (Pérez 2007) was conducted on 72 accessions from 13 geographical locations in Costa Rica (wild and cultivated lines), Colombia, Venezuela, and the Antillean islands. Analysis involved the use of 15 microsatellite markers to compute allelic richness and frequency, expected heterozygosity (*He*), genetic distances (Nei 1973, 1978), and a principle coordinate analysis (PCO). A total of 99 alleles were identified for the 15 loci. Although the papaya lines sampled from Costa Rica showed the highest number of rare and unique alleles, they showed the lowest genetic diversity with *He* values between 0.37 and 0.44. There was no differentiation between the wild and cultivated individuals in Costa Rica confirming introgression between them as previously reported (d'Eeckenbrugge et al. 2007) based on morphological and isozyme descriptors. Papaya accessions from Guadeloupe, Venezuela, Colombia, and Barbados showed highest diversity with *He* values between 0.69 and 0.50. The PCO analysis showed clear differentiation among sampled papaya accessions according to their geographical origin. Interestingly, the cultivar "Solo" plotted close to the

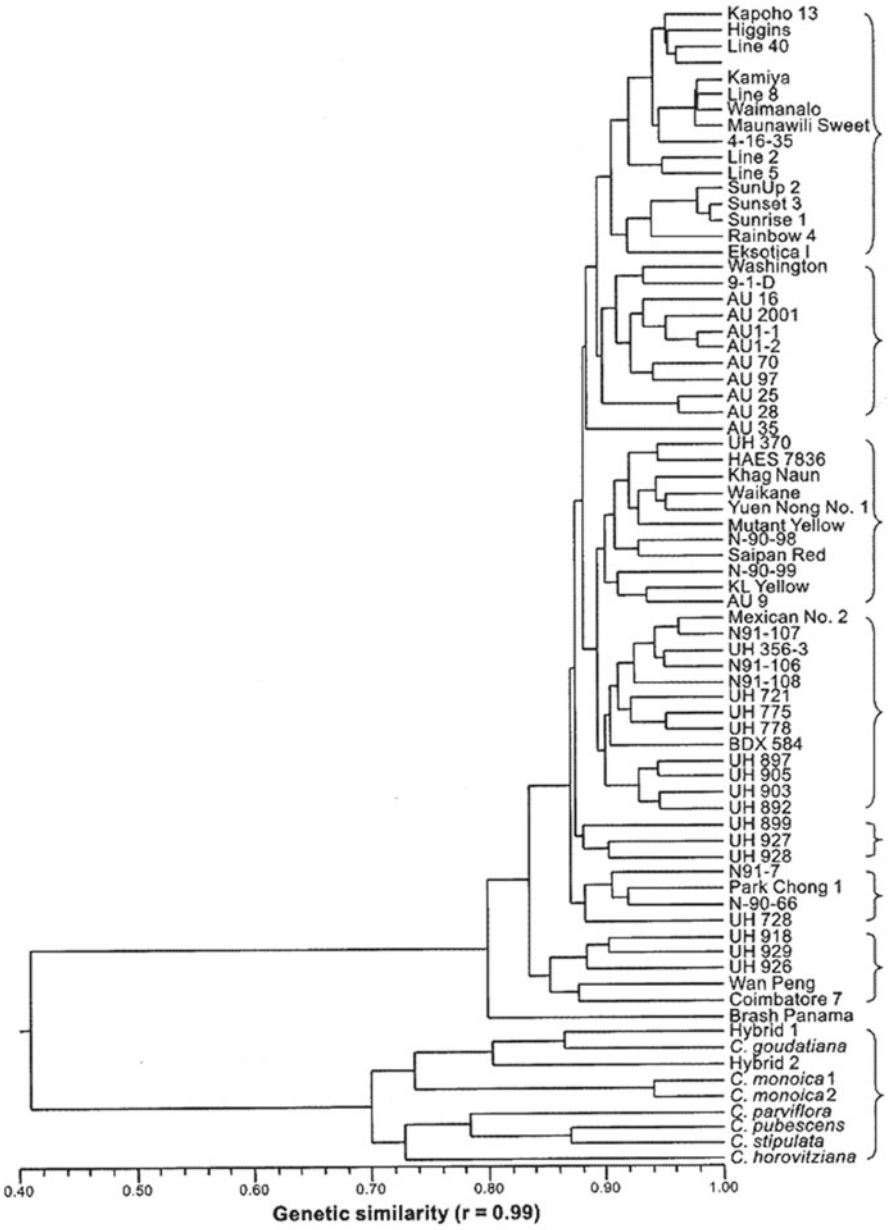

Fig. 3.2 Phenogram based on simple matching coefficient of similarity among 71 papaya accessions and related species. Cophenetic correlation coefficient = 0.99. (Used with permission from Kim et al. (2002). Copyright © 2008 Canadian Science Publishing.)

Barbados group, which seems logical since it was initially selected from Barbados germplasm. Reports from this study of rare alleles in the accessions of Costa Rica and Colombia and higher allelic diversity among the few cultivated accessions collected in Colombia, Venezuela, and Guadeloupe bring into question a conclusion from the AFLP survey of genetic diversity (Kim et al. 2002) that the present USDA germplasm collection in Hawaii includes the vast majority of papaya natural variation. Although genetic diversity of *C. papaya* is small, more may be found through future germplasm collecting and characterization from regions of Central America and northeastern South America as the presumed center of origin of the species.

In summary, genetic diversity studies within cultivated papaya indicate that cultural preference and geographic isolation have forced selection of cultivated papaya from a relatively narrow genetic base to result in low genetic diversity (Aradhya et al. 1999; Jobin-Décor et al. 1997; Kim et al. 2002; Van Droogenbroeck et al. 2002). Hawaiian "Solo" papayas, for instance, were developed from a single introduction from Barbados in 1910 (Storey 1969). Although a wide range of morphological characters is visible in the field, only ca. 12 % amplified fragment length (AFLP) polymorphisms among 63 accessions (Kim et al. 2002). Of the 63 accessions analyzed, 82 % of the pair-wise comparisons exhibited genetic similarity greater than 0.85 and fewer than 4 % showed less than 0.80. The genetic variation that exists in cultivated papayas is attributed to natural outcrossing events and the genetic similarity reported is attributed to each accession's origin according to a specific breeding or selection program (Kim et al. 2002). Van Droogenbroeck et al. (2002) included six accessions of cultivated papaya in a study of genetic relationships among 95 accessions representing three genera and 11 species of the Caricaceae from Ecuador. AFLP analysis of these accessions showed low genetic variation (0.99 average similarity) and also showed cultivated papaya to be very distinct from the other genera tested, with an average genetic similarity of only 0.23, supporting the idea that *C. papaya* diverged early from its wild relatives and proceeded to evolve in isolation (Aradhya et al. 1999; Van Droogenbroeck et al. 2002).

Prospects

The phenotypic diversity of *C. papaya* is difficult to estimate because diversity is not widely studied and most of it that has been recorded limited to plant growth habit, flower and fruit characteristics, and response to pathogens. For these agronomically important traits, the diversity in common papaya seems moderate. Likewise, the genetic diversity of common papaya is difficult to estimate, perhaps because the molecular methods used thus far have limited resolution or perhaps because common papaya lines are generally inbred and really do have only limited genetic diversity. For certain, the diversity between *C. papaya* and its wild relatives, especially the more than 20 species of the highland papayas of the genus *Vasconcellea,* formerly regarded as a section of *Carica* (Jobin-Décor et al. 1997), is much greater than among papaya cultivars.

The potential for obtaining greater diversity within papaya from its relatives indicates a need to introgress desirable traits from those wild relatives. However, such introgression has proved elusive. Although cultivated papaya and its wild relatives have the same somatic chromosome number ($2n = 18$) and similar genome sizes (~372 megabases), they differ considerably at the molecular level. Isozyme and RAPD analysis revealed 73 and 69 % dissimilarity, respectively, between cultivated papaya and highland papayas of the genus *Vasconcellea.* AFLP analysis of a chloroplast DNA (cpDNA) intergenic spacer region also grouped cultivated papaya in a separate clade away from *Vasconcellea* with a bootstrap analysis confidence level of 64 % (Aradhya et al. 1999). Whatever the cause, some form of highly expressed post-zygotic barriers preclude natural hybridization and makes forced hybridization between *Carica* and *Vasconcellea* difficult.

Sexual incompatibility is less common within *Vasconcellea.* Several *Vasconcellea* produce natural hybrids in areas of overlapping distributions (Badillo 1993; Scheldeman et al. 2013). Contrarily, cultivated papaya requires either embryo rescue (Manshardt and Wenslaff 1989; Manshardt and Drew 1998) or the use of a hybrid–backcross bridge (Drew et al. 2006) to introduce genetic diversity from *Vasconcellea.* Embryo rescue is laborious and inefficient. The hybrid–backcross bridge worked better consisting of the use of an intergeneric hybrid between *C. papaya* and *Vasconcellea quercifolia* backcrossed to *C. papaya*. This joint Australia/the Philippines project has thus far produced three backcrossed lines with resistance to Australian isolates of papaya ringspot virus (PRSV) and six lines with resistance to Philippine PRSV isolates. Infertility and incompatibility problems are no longer a concern by the second backcross generation (Drew et al. 2006). Refinement of these techniques for intergeneric hybridization and embryo rescue will further facilitate future introgression of desirable traits and genetic diversity into *C. papaya*.

References

Anonymous (2005) Consensus document on the biology of papaya (Carica papaya). Series on Harmonisation of Regulatory Oversight in Biotechnology. No. 33, ENV/JM/MONO(2005)16. Paris

Aradhya MK, Manshardt RM, Zee F, Morden CW (1999) A phylogenetic analysis of the genus *Carica* L. (Caricaceae) based on restriction fragment length variation in a cpDNA intergenic spacer region. Genet Res Crop Evol 46:579–586

Badillo VM (1993) Caricaceae. Segundo esquema. Rev Fac Agron Univ Cent Venezuela, Alcance 43, Maracay, Venezuela

Badillo VM (2000) *Carica* L. vs. *Vasconcellea* St. Hill. (Caricaceae) con la rehabilitación de este último. Ernstia 10(2):74–79

Chan YK, Paull RE (2008) Caricaceae: *Carica papaya* L.; papaya. In: Janick J, Paull RE (eds) The encyclopedia of fruit & nuts. CABI, London, pp 237–247

d'Eeckenbrugge GC, Restrepo MT, Jiméz D (2007) Morphological and isozyme characterization of common papaya in Costa Rica. In: Chan YK, Paull RE (eds) Proceedings of the 1st international symposium on papaya. Acta Hortic 740:109–120

Drew RA, Siar SV, O'Brien CM, Sajise AGC (2006) Progress in backcrossing between *Carica papaya* × *Vasconcellea quercifolia* intergeneric hybrids and *C. papaya*. Aust J Exp Agric 46:419–424

Eustice M, Yu Q, Lai CW, Hou S, Thimmapuram J, Liu L, Alam M, Moore PH, Presting GC, Ming R (2008) Development and application of microsatellite markers for genomic analysis of papaya. Tree Genet Genomes 4:333–341

Hammer K, Arrowsmith N, Gladis T (2003) Agrobiodiversity with emphasis on plant genetic resources. Naturwissenschaften 90:241–250

Hofmeyr JDJ (1939) Sex-linked inheritance in *Carica papaya* L. South Afr J Sci 36:283–285

International Board for Plant Genetic Resources (IBPGR) (1988) Descriptors for papaya. IBPGR, Rome, 34 pp

Jobin-Décor MP, Graham GC, Henry RJ, Drew RA (1997) RAPD and isozyme analysis of genetic relationships between *Carica papaya* and wild relatives. Gene Res Crop Evol 44:471–477

Kim MS, Moore PH, Zee F, Fitch MMM, Steiger DL, Manshardt RM, Paull RE, Drew RA, Sekioka T, Ming R (2002) Genetic diversity of *Carica papaya* as revealed by AFLP markers. Genome 45:503–512

Manshardt RM, Drew RA (1998) Biotechnology of papaya. Acta Hortic 461:65–73

Manshardt RM, Wenslaff TF (1989) Inter-specific hybridization of papaya with other species. J Am Soc Hortic Sci 114:689–694

Ming R, Van Droogenbroeck B, Moore PH, Zee FT, Kynd T, Scheldeman X, Sekioka T, Gheysen G (2005) Molecular diversity of *Carica papaya* and related species. In: Sharma AK, Sharma A (eds) Plant genome: biodiversity and evolution, vol 1B: Phanerograms. Science Publishers, Enfield, pp 229–254

Ming R, Yu Q, Moore PH (2007) Sex determination in papaya. Semin Cell Dev Biol 18:401–408

Nakasone HY, Lamoureux C (1982) Transitional forms of hermaphroditic papaya flowers leading to complete maleness. J Am Soc Hortic Sci 107:589–592

Nakasone HY, Paull RE (1998) Papaya. In: Tropical fruits. CABI, New York, pp 239–269 (Chapter 10)

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci USA 70:3321–3323

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583–590

Pérez JO (2006) Microsatellite markers in Carica papaya L.: isolation, characterization and transferability to Vasconcellea species. Mol Ecol Notes 6:212–217

Pérez JO (2007) Papaya genetic diversity assessed with microsatellite markers in germplasm from the Caribbean region. In: Chan YK, Paull RE (eds) Proceedings of the 1st international symposium on papaya. Acta Hortic 740:93–101

Rieger JE (2009) Genetic and morphological diversity of natural populations of *Carica papaya*. Masters of Science Thesis, Miami University of Ohio

Scheldeman X, Willemen L, Coppens d'Eeckenbrugge G, Romeijn-Peeters E, Restrepo MT, Romero Motoche J, Jiménez D, Lobo M, Medina CI, Reyes C, Rodríguez D, Ocampo JA, Van Damme P, Goetgebeur P (2007) Distribution, diversity and environmental adaptation of highland papayas (*Vasconcellea* spp.) in tropical and subtropical America. Biodiv Conserv 16:1867–1884

Scheldeman X, Kyndt T, d'Eeckenbrugge GC, Ming R, Drew R, Van Droogenbroeck B, Van Damme P, Moore PH (2011) *Vasconcellea*. In: Kole C (ed) Wild crop relatives: genomic and breeding resources, tropical and subtropical fruits. Springer-Verlag Berlin Heidelberg. pp 213–249

Stiles JI, Lemme C, Sondur S, Morshidi MB, Manshardt R (1993) Using randomly amplified polymorphic DNA for evaluating genetic relationships among papaya cultivars. Theor Appl Genet 85:697–701

Storey WB (1958) Modification of sex expression in papaya. Hortic Adv 2:49–60

Storey WB (1969) Papaya. In: Ferwerda FP, Wit F (eds) Outlines of perennial crop breeding in the tropics. H Veenman, Zonen NV, Wageningen, pp 21–24

Storey WB (1976) Papaya. In: Simmonds NW (ed) The evolution of crop plants. Longman, London, pp 21–24

Van Droogenbroeck B, Breyne P, Goetghebeur P, Romeijn-Peeters E, Kyndt T, Gheysen G (2002) AFLP analysis of genetic relationships among papaya and its wild relatives (Caricaceae) from Ecuador. Theor Appl Genet 105:289–297

Chapter 4
Vasconcellea for Papaya Improvement

Geo Coppens d'Eeckenbrugge, Rod Drew, Tina Kyndt, and Xavier Scheldeman

Introduction

Most early attempts at widening the genetic basis for papaya improvement were based on a poor understanding of the relationship between the common papaya and mountain papayas or highland papayas, i.e., between *Carica* and *Vasconcellea*. Breeders were faced with the taxonomical confusion between the two genera, which led them to serious misconceptions. Several experts located the center of papaya diversity in the Andes of South America (e.g., Prance 1984; Brücher 1977, 1989), which is in fact the center of diversity for *Vasconcellea*. Their knowledge of available genetic resources was then poor, and the very limited samples they had at hand did not help them in identifying the problem. Indeed, the first genetic studies were most often based on one representative per *Vasconcellea* species. In addition *Carica* sampling was not satisfactory, as wild common papayas have remained mostly unexplored, despite a description by Manshardt and Zee (1994).

It has been the extensive work of Prof. Badillo that has progressively clarified the relationship between *Carica* and *Vasconcellea*. A "classical taxonomist" with an acute sense of observation and extensive field experience, Badillo was very open to

G. Coppens d'Eeckenbrugge (✉)
Centre de Coopération Internationale en Recherche Agronomique pour le Développement, UMR 5175 CEFE, Campus CNRS, 1919 Route de Mende 34293 Montpellier Cedex 5, France
e-mail: geo.coppens@cirad.fr

R. Drew
Department of Biological and Physical Sciences, Griffith University, Nathan 4111, QLD, Australia

T. Kyndt
Department of Molecular Biotechnology, Ghent University, Ghent 9000, Belgium

X. Scheldeman
Regional Office for the Americas, Bioversity International, Cali, Colombia

R. Ming and P.H. Moore (eds.), *Genetics and Genomics of Papaya*, Plant Genetics and Genomics: Crops and Models 10, DOI 10.1007/978-1-4614-8087-7_4,

incorporating the observations and data from other disciplines, first from Professors Horovitz and Jiménez and their team at the Universidad Central de Venezuela and, more recently, from teams in Hawaii (Aradhya et al. 1999), Venezuela, Colombia, Ecuador, and Belgium (Badillo et al. 2000; Leal and Coppens d'Eeckenbrugge 2003; National Botanic Garden of Belgium 2008). Thus, Badillo's morphological observations have been supplemented with the cytological and genetic data obtained from the first field experiments in Venezuela, and later with morphological and molecular characterization, as well as modern phylogenetic studies, carried out on relatively diverse samples. Badillo's research culminated in 2000 and 2001 with the taxonomic restoration of the genus *Vasconcellea*. The separation of the two genera has been widely confirmed in subsequent molecular studies, and there is now clear recognition that, despite considerable morphological similarity, *Vasconcellea* is not the closest relative of *Carica* (see Chap. 5 in this text). Obviously, this has considerable impact when analyzing its potential for improvement of the common papaya. To emphasize the importance of this distinction, we first present highland papayas by themselves, demonstrating their specificity and their own potential for production, even as competitors of the common papaya (latex production and medicinal uses). Secondly, we examine the relationship between *Carica* and *Vasconcellea*. Thirdly, we present the history of intergeneric hybrid breeding in papaya, reviewing the potential of species explored to date. The fourth section is devoted to conservation of highland papaya germplasm, before conclusions on the potential of *Vasconcellea* for genetic improvement of the common papaya.

Highland Papayas: Agro-Economic Potential and Current Uses

Vasconcellea species show interesting potential and scope for domestication in different regions. The fruits have appealing organoleptic properties (e.g., aroma, taste, color), and commercial crops could be developed from either the diversity within existing species or varieties or based on new trait combinations obtained by hybridizing different *Vasconcellea* species. Secondly, highland papayas are potential sources of papain and other biologically active components of pharmaceutical interest. According to the National Research Council (1989), smallholders in Andean countries are currently using highland papayas for these purposes; however, given adequate quality control, it should be possible to develop both fresh fruit exports and an Andean papain extraction industry.

Highland Papayas as Fruit Crops

Vasconcellea species are commercially and socially important in their local setting; however, they could supply specific niches in the international market. In households

Fig. 4.1 Three highland papayas exploited as fruit crops in the Andes. From left to right: babaco (*V.*×*heilbornii*) and two papayuelas (*V. cundinamarcensis* and *V. goudotiana*)

in the Andes, their fruits are consumed fresh, they are roasted and processed into juices, they are made into marmalades and preserve or added to dairy products, and they are prepared in sauces, pie fillings, and pickles (National Research Council 1989; CAF 1992; Van den Eynden et al. 1999, 2003). The two most important species are babaco [*V.*×*heilbornii* "Babaco" (Badillo) Badillo] in Ecuador and southern Colombia and the papayuela (*V. cundinamarcensis* Badillo) in all Andean countries but more particularly in Colombia and Chile (Fig. 4.1).

Only babaco (*V.*×*heilbornii* "Babaco"), the cultigen with the largest fruits, has been the object of active commercial development, albeit on a small scale, except in Ecuador, where it is always available on the market. Endt (1981) describes the taste of babaco as a unique blend of strawberry, pineapple, and papaya, although due to its low content of soluble solids, sugar must be added to processed and even fresh fruits (Kempler and Kabaluk 1996). It was introduced as a crop in New Zealand, Australia, South Africa, Italy, Spain, Switzerland, the Netherlands, and Canada between 1973 and 1994; however, most introductory trials have failed, due to problems in commercialization (Scheldeman et al. 2011). Breeding for fruit traits is possible. Spontaneous hybridization of *V.*×*heilbornii* varieties with *V. stipulata* (Badillo) Badillo resulted in numerous forms with different flavors and varying quantities of seeds (Horovitz and Jiménez 1967; National Research Council 1989). More recently, higachos or baby babacos, fruits from other botanical varieties of *V.*×*heilbornii*, have gained importance. Because of their aroma and smaller size, they are better adapted to the urban market.

V. cundinamarcensis is mainly grown in the Colombian altiplano in the vicinity of Bogotá, either in home gardens or in small plots that are intensively cultivated for commercial purposes. Although a minor fruit crop, it ensures a significant market as the population of this region is similar to that of Ecuador. Taking into account the small garden production in Venezuela and Peru (National Research Council 1989)

and the small but intensive production and canning industry in Chile (225 ha; Carrasco et al. 2009), the economic importance of *V. cundinamarcensis* may compare to that of babaco in Ecuador. Other *Vasconcellea* species are consumed locally (Badillo 1993; Van den Eynden et al. 1999; De Feo et al. 1999; Scheldeman 2002). They include *V. candicans* (Gray) A.DC., *V. crassipetala* (Badillo) Badillo, *V. goudotiana* Tr. et Pl. (Fig. 4.1), *V. microcarpa* (Jacq.) A.DC., *V. monoica* (Desf.) A.DC., *V. palandensis* Badillo et al., *V. parviflora* A.DC., *V. quercifolia* St.-Hil., *V. sphaerocarpa* (García-Barr. et Hern.) Badillo, and *V. stipulata*. However, these species have not yet received the same level of commercial importance as babaco.

Pomological studies (Scheldeman et al. 2003) confirm that the aromatic *V. stipulata* has the highest levels of soluble solids of all consumed *Vasconcellea* species. Its smaller size makes it more attractive than the bigger and less aromatic babaco. In southern Ecuador, local consumers prefer this species over the more commercial ones (Scheldeman 2002). *V. stipulata* should therefore be a priority species for further domestication. A well-targeted promotion of *V. cundinamarcensis* and *V. goudotiana* could increase their market significantly, especially in Colombia. The development of new products more suited to urban culture, such as preserves and health products, should be considered.

Some of the noncommercial *Vasconcellea* species could be used to improve babaco cultivation, either directly as rootstocks or indirectly as breeding materials. In Ecuador, babaco production is seriously threatened, especially in greenhouses, by the root fungus *Fusarium oxysporum* and the nematode *Meloidogyne incognita*. Resistant and tolerant rootstocks, such as *V. monoica* and *V. weberbaueri* (Harms) Badillo (Scheldeman et al. 2003), may provide a solution in the absence of adequate chemical treatments (Ochoa et al. 2000). Evaluation of grafting techniques with different combinations of species showed good compatibility with success rates ranging from 50 to 100 %, confirming earlier results obtained by Jiménez (1958), who successfully tested several combinations involving *Carica* and *Vasconcellea* species. The use of *V. cauliflora* (Jacq.) A.DC. as a rootstock allowed maintenance of highland species under tropical lowland conditions.

Highland Papayas as Sources of Papain

Apart from the preliminary work of Scheldeman et al. (2003), little research has been applied to commercial production of latex from *Vasconcellea* species. FAO (1992) refers to the use of latex of *V. cundinamarcensis* as a meat tenderizer while the National Research Council (1989) speculates that *V. stipulata* and other highland papayas might be grown as a source of papain. The proteolytic activity of freeze-dried latex of *V. cundinamarcensis* was between five- and eightfold higher than that of *C. papaya* latex (Baeza et al. 1990). They also reported a water content in fresh latex of 80 % (weight percentage), which is 8 % lower than that of papaya, and demonstrated that stem extracts had even higher proteolytic activities. Horovitz and Jiménez (1967) advocated the use of whole plants and reported a particularly

high proteolytic activity in a *V. stipulata* × *V. monoica* hybrid. Preliminary papain analysis of some Ecuadorian *Vasconcellea* species revealed very high proteolytic activity when compared to *C. papaya*. Proteolytic activity of babaco latex is equivalent to or slightly higher than that of the common papaya (Dhuique Mayer et al. 2001). Some accessions of *V.* × *heilbornii* and *V. stipulata* showed particularly promising papain activity levels (Scheldeman et al. 2002; Kyndt et al. 2007).

Highland Papayas' Medicinal Uses

Little is known about medicinal values of *Vasconcellea*, although FAO (1992) reports the use of fruit to treat arterial sclerosis and the use of latex to cure skin mycosis and verruca plana. Clearly, *Vasconcellea* species have as many or more potential medicinal uses than *C. papaya*. Vernacular names of several species (*tapaculo*) refer to the antidiarrheal properties of their seeds. Recently, proteinases from *V. cundinamarcensis* were shown to heal chemically induced gastric ulcers in experimental rodent models (Mello et al. 2008). One of the cysteine proteases of *V. cundinamarcensis* latex, CMS2MS2, could potentially be used in wound healing treatments since the recombinant protein was shown to induce a mitogenic response in fibroblasts (Corrêa et al. 2011).

Wild and Feral Populations

Most highland papaya species still exist in the wild, and some have become naturalized out of their original range. In particular, *V. cundinamarcensis* has been reportedly grown and consequently become naturalized in tropical highlands and cool subtropical regions of Sri Lanka, India, New Zealand, Puerto Rico (Morton 1987), and Zambia (Kew herbarium data). *V.* × *heilbornii* is the only cultigen in the family, as it is propagated artificially by vegetative methods.

Geographic Distribution of Vasconcellea *Diversity and Climatic Preferences*

Scheldeman et al. (2007) analyzed the distribution of *Vasconcellea* species and their climatic preferences. The results clearly confirmed the association of the genus with the Andes and designated the center of its diversity as southern Ecuador which holds 16 of the 21 species. Other countries with high species diversity are Colombia and Peru. Analysis of climate data on collection sites shows that most *Vasconcellea* species are adapted to zones with mild temperatures (hence, their preference for higher altitudes) with average yearly precipitation ranging from 800 to 1,400 mm.

Exceptions are *V. chilensis* (Planch. ex A.DC.) A.DC., found in desert conditions, *V. pulchra* (Badillo) Badillo which shows ecological preference for extremely wet conditions, *V. parviflora* which occurs on the Ecuadorian and Peruvian coasts in warm and relatively dry climates, and *V. cauliflora*, a tropical lowland species. Only the distribution of the latter shows an appreciable overlap with that of the wild *C. papaya*, as it extends northward to Central America.

Organization of *Vasconcellea* Diversity and Its Relation with *Carica*

Taxonomy

With 21 species *Vasconcellea* is the largest genus in the Caricaceae (Table 4.1). Because of the strong similarity in their morphology and uses, common papaya and highland papayas have long been treated as close relatives. The taxonomical

Table 4.1 List of *Vasconcellea* species and their distribution in Latin American countries

Species	Country	Cultivation status
V. candicans	Ecuador, Peru	Wild, home gardens
V. cauliflora	Colombia, Costa Rica, Venezuela, Nicaragua, Panama, Guatemala Mexico, Honduras, El Salvador	Wild, home gardens
V. chilensis	Chile	Wild
V. crassipetala	Colombia, Ecuador	Wild
V. cundinamarcensis	Colombia, Ecuador, Venezuela, Peru, Bolivia, Panama, Chile, Costa Rica	Commercial, home garden, tolerated, wild?
V. glandulosa	Bolivia, Argentina, Peru, Brazil	Wild
V. goudotiana	Colombia, Ecuador (recent introduction), Venezuela	Home garden, wild
V. horovitziana	Ecuador	Wild
V. longiflora	Ecuador, Colombia	Wild
V. microcarpa	Ecuador, Peru, Colombia, Venezuela, Brazil, Bolivia, Costa Rica, Panama	Wild
V. monoica	Ecuador, Bolivia, Peru, Colombia	Wild, home gardens
V. omnilingua	Ecuador	Wild
V. palandensis	Ecuador	Wild
V. parviflora	Ecuador, Peru	Wild, occasionally used as ornamental
V. pulchra	Ecuador, Colombia	Wild
V. quercifolia	Argentina, Bolivia, Brazil, Paraguay, Peru	Wild, home gardens, fences
V. sphaerocarpa	Colombia	Wild
V. sprucei	Ecuador	Wild
V. stipulata	Ecuador	Wild, home gardens
V. weberbaueri	Ecuador, Peru	Wild
V.×heilbornii	Ecuador, Peru	Cultigen, commercial, home gardens

distinction between *Carica* and *Vasconcellea* has been supported essentially by the former presenting a one-celled ovary and the latter a five-celled one (Saint Hilaire, cited in de Mello and Spruce 1869). This distinction was questioned very early (de Mello and Spruce 1869). In 1867, Bentham and Hooker concluded that the differences between *Carica* L. and *Vasconcellea* Saint Hilaire were so slight that they incorporated the latter into the former as a section. The number of sections was later increased to three (Solms-Laubach 1889, 1894; Harms 1925). In a preliminary presentation of his work, Badillo (1967) focused on the reduction of the genus *Carica* from 57 to 21 species and proposed to reduce infrageneric divisions to two, without much conviction on the number of ovary divisions. However, this trait was the first one used in his key and he formalized this division in 1971, reestablishing two sections, i.e., *Carica*, with one species, and *Vasconcellea*, with 21 species. In 1993, he suggested the possible rehabilitation of the latter as a distinct genus, as well as the reevaluation of two complex species, *C. microcarpa* Jacquin and *C. glandulosa* A.DC. His comments on macro- and micromorphological traits and their variation at both inter- and intraspecific levels called for caution in the application of identification keys. He emphasized also that interspecific barriers were often more labile within section *Vasconcellea*, resulting in numerous observations of spontaneous hybridization. Finally, following the molecular phylogenetic study of Aradhya et al. (1999), he fully restored *Vasconcellea* as a genus (Badillo 2000, 2001).

Morphology and Floral Biology

In addition to the relative difference in ovary division, *Carica* is differentiated from *Vasconcellea* by a hollow stem, leaves with more than seven principal veins, and lamelliform protuberances on the seed sclerotesta. Unlike papaya, where different floral types coexist, ranging from strictly pistillate to strictly staminate flowers including different intermediate forms (Lassoudière 1969), most *Vasconcellea* species are strictly dioecious and do not have bisexual flowers (Horovitz 1954). Only *V. monoica* is, as its name indicates, monoecious.

V. cundinamarcensis is polygamous, as it presents rare cases of andromonoecy, a trait shared with *C. papaya*. Such plants are male, bearing numerous flowers on long and highly ramified inflorescences. Most of these flowers are male, but a varying proportion, at the extremity of the ramifications, bear ovules that are not completely sterile, so that they can develop into a few small fruits hanging far from the plant stem. Such polygamous plants also occur in *V. parviflora* (X. Scheldeman, pers. obs.). Horovitz and Jiménez (1967) showed that this trait is under the double control of a Ypm chromosome and a *pm* cytoplasmic factor. Other species may bear the latter factor but lack the former.

As for the common papaya, all mountain papayas show some degree of parthenocarpy (Badillo 1993). This trait is expressed particularly in babaco [*V.* × *heilbornii* (Badillo) Badillo], whose fruits are formed independently from pollination and/or seed formation (Horovitz and Jiménez 1967). Indeed, seeds are rarely formed in this cultigen.

Cytogenetics and Interspecific Hybridization

All *Vasconcellea* species investigated to date (*V. cundinamarcensis*, *V. cauliflora*, *V.* × *heilbornii*, *V. monoica*, *V. goudotiana*, *V. quercifolia*, and *V. microcarpa*) present the somatic chromosome number of $2n = 18$, characteristic of the family (Darlington and Wylie 1955; de Zerpa 1959; Magdalita et al. 1997a; Costa et al. 2008). Similarly, preliminary studies with DAPI staining (T. Kyndt, unpublished) showed that *V. candicans*, *V. stipulata*, and *V. weberbaueri* had 18 chromosomes. When compared to *C. papaya*, similar or slightly larger genome sizes have been observed for some *Vasconcellea* species. *V. longiflora* (Badillo) Badillo appears to have a genome size about twice that of *C. papaya* (T. Kyndt, unpublished). Fluorescent in situ hybridization of the rDNA showed that *V. goudotiana and V. cundinamarcensis* hold only one 5S rDNA locus, whereas in papaya three loci were found. Conversely, one locus of 18S rDNA was detected in papaya, compared to four and five in *V. goudotiana* and *V. cundinamarcensis*, respectively (Costa et al. 2008). These cytogenetic observations indicate substantial genetic differences between *Carica* and *Vasconcellea*.

A high degree of sexual incompatibility is ensured by post-zygotic barriers between *Carica* and *Vasconcellea*. Strong interspecific incompatibility is less common within *Vasconcellea*, and cases of spontaneous hybridization have been reported (Badillo 1993). Meiosis is often regular in hybrids (de Zerpa 1959). *V.* × *heilbornii* and a hybrid between *V. monoica* and *V. cundinamarcensis* were described in detail by Horovitz and Jiménez (1967) and Badillo (1971). More hybrids may be expected to occur in regions where species distributions overlap.

Organization of Genetic Diversity and Interspecific Relationships Within Vasconcellea

Despite the relatively limited number of species, taxonomic analysis of highland papayas is difficult. They are often identified by staminate flowers, and this can be a problem in a mostly dioecious group. New species may not have been described in poorly collected areas, like southern Colombia, while taxonomical rearrangements may be necessary in poorly understood species and species complexes (e.g., *V. microcarpa*; Romeijn-Peeters 2004). Systematic work is further complicated by frequent interspecific compatibility (Jiménez and Horovitz 1958; Horovitz and Jiménez 1967; Mekako and Nakasone 1975) that affects the distribution and organization of *Vasconcellea* diversity, allowing spontaneous hybridizations and introgressions in areas of sympatry (Badillo 1971) and resulting in a mosaic of complex hybrid populations with a wide range of different recombinants and segregating progenies (Barton and Hewitt 1985). Finally, it has been suggested that highland papayas show high levels of phenotypic plasticity (Kyndt et al. 2005a), since morphological data are not always directly correlated with genetic relationships.

This difficult situation has lead researchers to extend their studies with genotypic information. In the last decade, several phenetic and phylogenetic studies have been performed on DNA markers and sequences. Incongruences between nuclear and chloroplast datasets (Kyndt et al. 2005a, 2005b) and intraindividual ITS sequence heterogeneity (Kyndt et al. 2005b) have confirmed the frequent occurrence of interspecific gene flow and hybridization and subsequent chloroplast capture in *Vasconcellea*. Based on all molecular and morphological studies, *Vasconcellea* seems to be a young complex network of closely related species. Therefore, it might be useful to introduce the concept of hybridizing species complexes or "syngameons" into future taxonomical revisions of the genus. Such complexes consist of numerous genetically distinct species or lineages, which periodically split and/or fuse as they extend through time. During this process, morphological intermediates form and species overlap. Such a rapid diversification and speciation is characteristic of plant evolution on the South American continent (Burnham and Graham 1999), probably under the influence of the strong vertical gradients in the Andean mountains which create diverse ecological regions and countless microclimates. Many examples of recent hybridization and subsequent chloroplast capture have been reported in *Vasconcellea* (see further); intraspecific diversification of the chloroplast genomes within *Vasconcellea* taxa has been observed with PCR-RFLP (Restrepo et al. 2004b), the sequencing of *psb*A–*trn*H region (Kyndt et al. 2005b) and cpDNA SSRs (Kyndt et al. 2006).

Based on all morphological, molecular, and distribution data, the following species complexes may be recognized within *Vasconcellea*: (1) *V. weberbaueri*, *V. stipulata*, *V.* × *heilbornii*, and *V. parviflora;* (2) a basal lineage with *V. chilensis*, *V. candicans*, *V. quercifolia*, and *V. glandulosa*; and (3) a clade holding all other taxa of the genus. Strikingly, these three clades present particular geographical distributions, where the southern Ecuador diversity center appears to be pivotal.

Clade 1: *V.* × *heilbornii*, *V. weberbaueri*, *V. stipulata*, and *V. parviflora*

Both phylogenetic and phenetic studies at the nuclear and chloroplast level [Van Droogenbroeck et al. 2002; Restrepo et al. 2004b; Kyndt et al. 2005b, 2006; Carvalho and Renner 2013 (also Chap. 5 in this text)] have revealed a strong genetic linkage between *V. stipulata*, *V. parviflora*, *V. weberbaueri*, and *V.* × *heilbornii*. This small clade is endemic to southern Ecuador and northern Peru and seems to have a history of introgression leading to an exchange of nuclear and chloroplast genomes. Probably under human influence, *V. stipulata* has hybridized with *V. cundinamarcensis* (see clade 3), and the resulting hybrids are called *V.* × *heilbornii* (Horovitz and Jiménez 1967; Badillo 1971). *V.* × *heilbornii* includes the babaco and a number of other cultigens producing smaller fruits ("higachos" or "baby babacos"). These hybrids present a low level of sexual fertility and germination rates below 5 % (Jiménez et al. 1998) and are hence propagated vegetatively. The incongruence between nuclear ITS and chloroplast psbA-trnH datasets provided molecular

evidence of the hybrid nature of the *V.*×*heilbornii* varieties (Kyndt et al. 2005b). Moreover, a very likely consequence of these hybridizations is significant ITS sequence heterogeneity that was found within the genome of southern Ecuadorian specimens of *V. stipulata*, *V. cundinamarcensis*, and *V.*×*heilbornii*. Different nuclear molecular marker studies on sample sets from southern Ecuador revealed a high genetic diversity in *V.*×*heilbornii* and placed its accessions at an intermediate position between the presumed parental species, with one group of accessions closer to *V. stipulata* and one group closer to *V. cundinamarcensis*, suggesting bidirectional introgression with the putative parents of the hybrid species (Van Droogenbroeck et al. 2002, 2006; Kyndt et al. 2005a, 2006). Also, an impressive morphological variation has been found in *V.*×*heilbornii*, blurring the distinction with *V. stipulata* and suggesting that introgression processes are still very active in its evolution, despite the predominantly vegetative reproduction (Scheldeman et al. 2002; Restrepo et al. 2004a). Molecular studies at the chloroplast level have confirmed this view (Aradhya et al. 1999; Restrepo et al. 2004b; Van Droogenbroeck et al. 2006) and revealed that *V. cundinamarcensis* only acts as a pollen donor in this cross, as no *V.*×*heilbornii* accession bears its chlorotype or mitotype (Restrepo et al. 2004b; Van Droogenbroeck et al. 2006). Remarkably, the predominant chloroplast haplotype of *V.*×*heilbornii* has not been found in any other specimen and seems to be closely related to that of *V. weberbaueri* (Restrepo et al. 2004b; Van Droogenbroeck et al. 2004, 2006; Kyndt et al. 2006), while only a few specimens of this taxon show the typical haplotype of *V. stipulata* (Restrepo et al. 2004b; Kyndt et al. 2005a). In summary, *V.*×*heilbornii* does not result from one simple hybridization event but is rather the product of an evolving hybridizing species complex, in which natural processes like introgression, chloroplast capture, selection, differentiation, mutations, and human selection have all played a part.

Clade 2: Basal Lineage with *V. chilensis*, *V. candicans*, *V. quercifolia*, and *V. glandulosa*

Although only a few specimens of these taxa have been included in molecular studies, *V. chilensis*, *V. candicans*, *V. quercifolia*, and *V. glandulosa* seem to be related genetically, and members of this lineage are always found at the base of the *Vasconcellea* genus [Jobin-Décor et al. 1997; Kyndt et al 2005a, b, 2006; Carvalho and Renner 2013 (also Chap. 5 in this text)]. Nuclear and cpDNA studies (Kyndt et al. 2005a, b, 2006) reveal a close relationship between *V. quercifolia* and *V. candicans* and a more distant one between these sister species and *V. chilensis*. While the molecular information for *V. glandulosa* is limited, both the specimen studied by Aradhya et al. (1999) and the one of Kyndt et al. (2005b) were related to *V. quercifolia* on cpDNA level, but ITS sequences did not clarify their nuclear relationship (Kyndt et al. 2005b). This group of taxa has its northern geographical limit in southern Ecuador (Scheldeman et al. 2007). *V. quercifolia* is widespread, while the other three members of this lineage have a narrower distribution; for example, *V. candicans* is only found in the drier regions of southern Ecuador and coastal Peru.

The fact that *V. quercifolia* appeared in a basal position in early genetic studies (Jobin-Décor et al. 1997) and the first success obtained in introgressive breeding (see section on Traits of Interest for Introgression into the Common Papaya later in this chapter) have led to the proposal that this species was genetically less distant from *C. papaya* and hence the "best" option for intergeneric hybridizations (Siar et al. 2011). The basal clade of genus *Vasconcellea* may contain other species that could be crossed to *C. papaya,* for example, *V. chilensis, V. candicans*, and *V. glandulosa*. However, being basal to *Vasconcellea* means that these taxa from clade 2 have diverged early in evolution from the other species of the genus. There is no evidence that they have evolved to be particularly closely related to the genetically distinct *C. papaya*. Other *Vasconcellea* species, for example, *V. parviflora*, might be as good or even better for intergeneric crossing (O'Brien and Drew 2009), using the optimized methodologies of Drew et al. (2006a, b).

Clade 3: All Other Taxa of the Genus

Although the remaining *Vasconcellea* species seem to form a separate clade in many molecular analyses, their interspecific relationships are often contradictory at morphological, nuclear, and chloroplast levels (Kyndt et al. 2005a, b). The fact that for many of these species only one specimen or one population has been analyzed precludes us from drawing consistent conclusions about their interspecific relationships. Specimens from different taxa but belonging to sympatric populations often show a close genetic relationship but reveal high morphological diversity, related to hybrid segregation and presumed phenotypic plasticity (Kyndt et al. 2005a).

A sister relationship between the sympatric species *V. pulchra* and *V. longiflora* has been confirmed by morphological data and different molecular marker analyses (Van Droogenbroeck et al. 2004; Kyndt et al. 2005a, b; see also Chap. 5 in this text). Also, *V. sphaerocarpa* appears morphologically and genetically very closely related to *V. goudotiana*, as indicated by PCR-RFLP studies of cp DNA (Aradhya et al. 1999; Restrepo et al. 2004b) and chloroplast and nuclear sequences (Kyndt et al. 2005b; see also Chap. 5 in this text). The relationship between the former two taxa and *V. palandensis* seems slightly more distant, according to the SSR study of Kyndt et al. (2006) and the cp DNA studies of Restrepo et al. (2004b) and Kyndt et al. (2005a). In this clade, *V. cundinamarcensis* is the species with the widest distribution (Scheldeman et al. 2007). This taxon seems to be compatible in interspecific crosses with *V. stipulata*, *V. monoica*, *V. microcarpa*, and *V. horovitziana* (Horovitz and Jiménez 1967). Depending on the origin of the analyzed specimens and the used technique, specimens of this taxon either group with members of clade 1 (Van Droogenbroeck et al. 2002; Kyndt et al. 2006) or within clade 3 (Restrepo et al. 2004b; Kyndt et al. 2005b; see also Chap. 5 in this text). The relationship with clade 1 has only been found with nuclear markers in specimens from southern Ecuador, suggesting that this phenomenon is related to introgression of *V. stipulata* genes in populations of *V. cundinamarcensis* in southern Ecuador. Conversely, isozyme and cpDNA analysis on Colombian samples showed a link with other Colombian

species within this clade, namely, *V. goudotiana*, *V. sphaerocarpa*, and *V. crassipetala* (Jiménez 2002; Restrepo et al. 2004b). Based on phylogenetic analysis of ITS sequences, *V. cundinamarcensis* is close to *V. sprucei* and *V. horovitziana*, while chloroplast sequences show a close relationship with *V. monoica* and *V. cauliflora* (Kyndt et al. 2005b). On a morphological level, this taxon looks quite different depending on its geographic origin; for example, fruits that are green turning to yellow in Colombia and Venezuela may exhibit very different colors, from pink to dark red, and are smaller in Ecuadorian accessions. Carrasco et al. (2009) investigated the cultivated *V. cundinamarcensis* from Chile using ISSR markers and found remarkably low genetic diversity in comparison with studies performed on samples from other countries (e.g., Restrepo et al. 2004b). This is probably due to the fact that this species was introduced into Chile.

Genetic Divergence Between Carica *and* Vasconcellea

The molecular studies described previously have not only supported an early evolutionary separation of genera *Carica* and *Vasconcellea*, but they have shown that they are relatively distant genera in the family and that *Jacaratia* appears more closely related to *Vasconcellea* (Aradhya et al. 1999; Van Droogenbroeck et al. 2002; Kyndt et al. 2005b, 2006). Indeed, Carvalho and Renner (2013; also Chap. 5 in this text) have shown that the closest relatives of papaya belong to the genera *Horovitzia* and *Jarilla*.

Use of *Vasconcellea* Germplasm in Papaya Breeding

The first systematic hybridization program was carried out in Venezuela by Horovitz and Jiménez (1967), who proposed to explore the potential for both fruit and papain production. Unfortunately, it was discontinued after the retirement of Prof. Horovitz (de Zerpa 1980). Most subsequent hybridization programs have focused on the introgression of resistance to the papaya ringspot virus, PRSV-P, although other potentially useful traits have been observed. PRSV-P is by far the most widespread and severe disease for the commercial production of papaya, as it is present in all major regions where papaya is grown. Although the development of virus-resistant plants via genetic engineering has been a model for incorporation of viral genes in the plant genome (see Chap. 15 in this text), they have not met all the requirements of the papaya industry worldwide (Fermin et al. 2010). To silence viral genes, a high sequence homology (>98 %) is needed between the virus and the transgene (Gonsalves 1998; Tennant et al. 2001). Furthermore, certain variants of the HC-Pro viral gene may be involved as a silencing suppressor and/or virulence enhancer (Tripathi et al. 2008). Given the radiative evolution of PRSV-P strains and the

correlation of their genetic divergence with their geographic distribution (Wang et al. 1994; Wang and Yeh 1997; Olarte Castillo et al. 2011), the development of transgenic cultivars for different papaya growing regions is problematic. Another difficulty lies in the reluctance to accept transgenic fruits in several key markets, such as the EU and Japan (Kent 2004). Thus, despite the undeniable technical success of genetic transformation in several regions, a few teams have persevered with the more difficult task of introducing a broader form of PRSV-P resistance, or even immunity, from *Vasconcellea* to *Carica* through intergeneric introgression. This approach requires the identification of sources of resistance; the comprehension of resistance/susceptibility mechanisms, including their inheritance; overcoming the intergeneric incompatibility barriers; the development of markers to assist the selection process; and a series of backcross generations to eliminate unwanted *Vasconcellea* traits.

Traits of Interest for Introgression into the Common Papaya

Horovitz and Jiménez (1967) reported genetic resistance against PRSV-P in four *Vasconcellea* species of their Venezuelan collection (*V. cauliflora*, *V. cundinamarcensis*, *V. stipulata*, *V. candicans*) and susceptibility in six others (*V. monoica*, *V. quercifolia*, *V. microcarpa*, *V. horovitziana*, *V. goudotiana*, *V. parviflora*). They underlined the divergence between their results and those of Conover (1964a) for *V. cauliflora* and *V. quercifolia* (respectively susceptible and resistant in the Florida collection), suggesting intraspecific variation in resistance of the plant and/or in its specific response to local PRSV-P strains. Genetic resistance against PRSV-P in *V. cauliflora* and susceptibility of *V. parviflora* and *V. goudotiana* were confirmed in later studies (Mekako and Nakasone 1975; Alvizo and Rojkind 1987; Magdalita et al. 1988; Amaral et al. 2006), whereas González and Trujillo (2005) identified a susceptible population of *V. cauliflora* in Venezuela.

Highland papayas may be sources of genetic resistance against other diseases. Jiménez and Horovitz (1958) reported resistance to papaya mosaic virus in *V. cauliflora* and *V. cundinamarcensis*. However, Conover (1964b) reported susceptibility of these two species, as well as *V. goudotiana* and *V. monoica*. Sawant (1958) reported resistance for papaya bunchy top in *V. monoica*. Drew et al. (1998) also reported resistance against *Phytophthora* in *V. goudotiana*, against papaya die back (a phytoplasma disease) in *V. parviflora*, and against black spot (*Asperisporium caricae*) in *V. cundinamarcensis*.

Other traits of potential interest for introgressive breeding include proteolytic capacity in *V. stipulata* × *V. monoica* (Horovitz and Jiménez 1967), high sugar content in *V. quercifolia* (Drew et al. 1998), cold tolerance in *V. cundinamarcensis* (Horovitz and Jiménez 1967; Manshardt and Wenslaff 1989b; Drew et al. 1998), drought tolerance in *V. chilensis*, and monoecy in *V. monoica* (Warmke et al. 1954; Horovitz and Jiménez 1967).

Reproductive Barriers Between Carica *and* Vasconcellea

Most early attempts at interspecific hybridization among Caricaceae were empirical, essentially reporting failure in fruit and seed development and or seed germination, particularly when *C. papaya* was involved (Warmke et al. 1954; Sawant 1958). The first systematic study was that of Jiménez and Horovitz (1958) who crossed representatives of *C. papaya* and five *Vasconcellea* species, following an incomplete diallel design. Observing the development of the hybrid embryos and endosperms, they distinguished several levels of interspecific barriers: obtention of viable seeds, cultivable embryos, and early embryo abortion, which led them to interpret their results in terms of gene pools. In the largest one, *V. cauliflora*, *V. candicans*, *V. microcarpa*, and *V. monoica* readily produced viable hybrid seeds. A second *Vasconcellea* gene pool was represented by *V. goudotiana*, which mostly produced hybrid seeds lacking an endosperm, when crossed with members of the first gene pool. A third gene pool, comprising *C. papaya*, was genetically distant from the others, as hybrid seeds were never viable, requiring the systematic use of embryo rescue and culture for the recovery of hybrids. D.M. de Zerpa developed culture techniques and recovered embryos 2.5 months old or older, depending on the cross. In addition she detailed cytogenetic observations on several of the hybrids obtained by Horovitz and Jiménez (de Zerpa 1959, 1980).

Manshardt and Wenslaff (1989a) further investigated the incompatibility barriers in the cross between *C. papaya* and *V. cauliflora*. No barrier was found during the programic phase. In the post-zygotic phase, the first clear symptom was the absence of endosperm, while early embryogenesis appeared normal. Deviations from the normal pattern of embryo development occurred after 45 days, with many embryos showing necroses or disorganized cellular proliferation. Some embryos continued to develop, although abnormally, had a marked tendency toward proliferation, and often comprised polyembryonic masses. The majority of the zygotic embryos did not grow, and only a small fraction produced culturable polyembryos. The success of the cross strongly depended on the parents that were used and the direction of the cross, as none of the four tested combinations gave positive results in both directions.

In a second paper, Manshardt and Wenslaff (1989b) presented very similar results from crosses of *C. papaya* with five other species, *V. monoica*, *V. parviflora*, *V. cundinamarcensis*, *V. quercifolia*, *V. stipulata*, and babaco, *V.*×*heilbornii*. However, all crosses where *C. papaya* was the male parent failed, except for *V. cundinamarcensis*×*C. papaya*. Rescuable embryos were produced with the five non-hybrid *Vasconcellea* species. Variation in the success of the crosses depended on which maternal lines of *C. papaya* were used. Similar variation between the crossability of maternal lines of the common papaya was reported by Vegas et al. (2003) in Venezuelan and Costa Rican populations and by Magdalita et al. (2001) in crosses between 15 papaya cultivars and *V. cauliflora*. However, the results of Magdalita et al. (2001) differed from those of Manshardt and Wenslaff (1989a) as they reported a complete failure of crosses where *C. papaya* was the male parent.

Another reproductive barrier between *Carica* and *Vasconcellea* is the general sterility of their hybrids, which seems to be related to a lack of homology between

Fig. 4.2 *V. cauliflora* (*left*) and *V. quercifolia* (*right*) are the two wild highland papayas most often hybridized with the common papaya [photographs by John A. Ocampo Pérez (*V. cauliflora*); and Luciano R. Soares—SEMA/RS (*V. quercifolia*), with kind permission]

the parental genomes (Manshardt 1992; Magdalita et al. 1997a, 1998). In Hawaii, backcrosses that involved the hybrids as female parents produced rare sesquidiploids (two copies of the *Carica* genome and one from *Vasconcellea*) from unreduced megaspores (Manshardt et al. 1995).

Potential of the Different Vasconcellea *Species for Introgression Breeding*

Vasconcellea cauliflora

V. cauliflora (Fig. 4.2) was one of the first sources of genetic resistance to PRSV-P that was identified, and transferred to the susceptible *V. monoica* and *V. horovitziana*, forming resistant F_1 generations, while the F2 generation segregated for resistance (Horovitz and Jiménez 1967). Genetic resistance to PRSV-P has been confirmed widely, including 114 glasshouse-planted and 20 field-planted hybrids in Australia (Magdalita et al. 1997b). There appears no quantitative variation in the level of resistance of the F_1 hybrids, suggesting again that it is a simple and dominant trait. Amaral et al. (2006) have found no resistance in *V. cauliflora* to papaya lethal yellowing virus (PYLV) in Brazil.

As *V. cauliflora* is a lowland-resistant species, similar to the common papaya in its climatic requirements, it has long appeared a natural first choice as a genitor. Most teams have reported success in the rescue of hybrid embryos from the intergeneric cross, in both directions (Jiménez and Horovitz 1958; Horovitz and Jiménez 1967; Manshardt and Wenslaff 1989a), but only when *C. papaya* was used as the female genitor in the subtropical climate of southeast Queensland (Magdalita et al. 2001). Magdalita et al. (1996) developed a highly efficient protocol to produce *C. papaya* × *V. cauliflora* hybrids, resulting in much lower losses of embryos and stronger plantlets in vitro.

Characterization was reported on 120 intergeneric hybrids between *C. papaya* and *V. cauliflora* by Magdalita et al. (1997a). These F_1 hybrids had morphological characteristics that resembled either parent or were intermediate between those of the two parents. Of the 16 characters examined, only the number of main leaf veins was exactly intermediate between *C. papaya* and *V. cauliflora*. Two characters, hermaphrodite flowers and low vigor, were distinguishing characteristics of the intergeneric hybrids. Petiole length, stem diameter, leaf length, leaf width, and flower color were similar to those of *C. papaya*. Leaf shape, type, serration, venation, as well as hairs on petioles and flower shape resembled those of *V. cauliflora*. All the interspecific hybrids showed cupping of the leaves, and three plants had a branched stem, which was not observed in either parent. In this study, where female *C. papaya* plants were crossed with male *V. cauliflora*, neither single sex form was observed in the five interspecific hybrid progenies that flowered (Magdalita et al. 1997a); however, it should be noted that because of lack of vigor, only 5 out of the 120 hybrid plants that were the subject of this study survived to the flowering stage (R. Drew, pers. obs.).

Variation in vigor of these generic hybrids has been reported. Horovitz and Jiménez (1967) lost their hybrids after transplantation to the field. Manshardt and Wenslaff (1989a) report satisfactory initial field growth of 25 hybrids from 5 embryos; however, only one *V. cauliflora*×*C. papaya* hybrid reached maturity and bore a few staminate flowers, with less than 1 % stainable pollen. According to Drew et al. (1998), many plantlets die in vitro or later in the glasshouse. One thousand hybrids were acclimatized after tissue culture and established in pots in a glasshouse in southeast Queensland, Australia. They grew well for 6–8 weeks in pots but then started to die and only five survived till the flowering stage. In a more tropical and humid climate at Los Baños in the Philippines, a few of these hybrids grew to maturity and flowered but were infertile (P. Magdalita, pers. comm.).

In Venezuela, Vegas et al. (2003) observed a slower growth in vitro and a longer acclimatization phase for the intergeneric hybrids under shade (6–8 months), as compared to the common papaya vitroplants (3 months). However, the comparison differed after field planting at altitudes of 500 and 1,500 m. At both levels, the hybrids flowered after 3–4 months. This time from planting to flowering was comparable to that of the common papaya at 500 m, despite a markedly slower growth of the hybrids. At 1,500 m, the two types of plants developed similarly; in fact, the common papaya trees presented limited growth and flowering and did not fruit. Altitude also influenced sex expression in one of the hybrid clones, at 500 m males that had normal male flowers, and at 1,500 m andromonoecious plants that had both male and hermaphrodite flowers.

Although there have been many reports of hybridization between *C. papaya* and *V. cauliflora*, the resultant F_1 hybrids were weak and infertile (Manshardt and Wenslaff 1989a; Magdalita et al. 1997a; Vegas et al. 2003). According to Magdalita et al. (1997a), pollen stainability of *C. papaya* and *V. cauliflora* was from 88 to 98 % and from 90 to 97 %, respectively. However, the intergeneric hybrids had only 0.5–1.4 % stainable pollen, and attempts at backcrossing to *C. papaya* failed. One cause of infertility in these hybrids was the high level of aneuploidy (Magdalita et al. 1997a). *C. papaya*, *V cauliflora*, and their putative hybrids had a somatic

chromosome number of $2n = 18$. However, some of the intergeneric hybrids had up to 48 % aneuploid cells, and the frequency was negatively associated ($r = -0.88$) with the number of RAPD bands from the male parent. Moreover, the low pollen fertility of the five hybrids that flowered and the failure of most hybrids to flower indicated that this may be due to chromosome imbalance related to incompatibility barriers and lack of homology between parental chromosomes (Magdalita et al. 1997a; Siar et al. 1998). The use of aneuploids after restoring fertility via chromosome doubling has been suggested (Magdalita et al. 1998). de Zerpa (1958) reported a majority of normal tetrads during meiosis in tetraploid plants of *C. papaya* and *V. cauliflora*, but there have been no further reports of their use in introgression programs.

Vasconcellea quercifolia

V. quercifolia (Fig. 4.2) is a mountain papaya with a particularly wide adaptation, being distributed from highlands to lowlands in tropical and subtropical South America, from the Brazilian Nordeste to Peru, Bolivia, northern Argentina, Paraguay, and Uruguay. It has been reported to be resistant to PRSV-P in Hawaii (Manshardt and Wenslaff 1989b) and Australia (Drew et al. 2006a) but susceptible to PRSV-P in Venezuela (Horovitz and Jiménez 1967). Another trait of interest is the high sugar content of *V. quercifolia* fruits.

The first recorded attempt to cross *C. papaya* and *V. quercifolia* was in 1914 by Higgins and Holt (1914), but the first successful report was that of Manshardt and Wenslaff (1989b) who obtained 28 embryos from 136 pollinations of *C. papaya* by *V. quercifolia*. All embryos were from the same genotype of *C. papaya*, although crosses to three papaya genotypes were attempted. The reciprocal cross yielded no fruits. The cross was repeated by Drew et al. (1998) who obtained 300 hybrids which grew vigorously in the glasshouse and later in the field. A few plants were tall and upright like *C. papaya*; however, most were branched like *V. quercifolia*. The intergeneric hybrids varied in their reaction to strains of PRSV-P in Australia and the Philippines. Of the hybrids, 75 % were resistant to PRSV-P, while 25 % produced virus symptoms. The hybrid plants had the sex ratio of 2 males: 49 hermaphrodite: 49 females (Drew et al. 2006a). All males and a few hermaphrodites (7 males and 11 hermaphrodites) produced some viable pollen that could be germinated on artificial medium (1.1–6.1 %); however, the backcross to female *C. papaya* was successful with pollen from only four male plants (Drew et al. 2005). Cytogenetic studies on one of these intergeneric hybrids which was grown in the Philippines revealed variable chromosome counts, presence of univalents, laggards and meiotic aberrations, and pollen fertility of 0.2–2.2 % (Drew et al. 2005).

Backcross populations were produced in both Australia and the Philippines; however, BC_1 plants lacked vigor, and mortality was high. In Australia, 1,426 papaya flowers were pollinated with pollen from hybrid plants, and 63 embryos were rescued from 112 fruits. Of these, only 50 developed into plants that could be micropropagated in vitro. Second backcross generations (BC_2) were easier to produce than BC_1 generations as pollen fertility was >80 % in male BC_1 plants

(Drew et al. 2006b). Segregating populations of BC_1 and BC_2 plants were evaluated for PRSV-P resistance, fertility, and morphological characteristics including fruit shape, size, and quality. Variable reactions to the two viral strains in the two countries were observed. One plant (clone 54) demonstrated a high level of tolerance against multiple inoculations of PRSV-P in the glasshouse and stood up against aphid transmitted PRSV-P in the field for 9 months in Australia; however, it showed little resistance to PRSV-P in the Philippines (Drew et al. 2005).

In the Philippines 114,839 seeds were dissected from 940 fruits of a BC_1 generation and yielded 1,011 embryos (Siar et al. 2011). Of these, 733 germinated in vitro and 700 developed into plantlets that were inoculated three times in a glasshouse with PRSV-P, and then any without symptoms were planted in a field under high disease pressure and exposure to inoculation by viruliferous aphids. One BC_1 male plant showed good resistance to PRSV-P and pollen fertility. From this BC_1 male plant, 1,465 plants [137 BC_2, 546 SbC_2 (BC_2 sib-crosses), 147 BC_3, 379 SbC_3, and 256 BC_4] were produced and inoculated with PRSV-P. Plants that did not show symptoms after three inoculations in a glasshouse were transplanted to the field. Resistant BC_3 and BC_4 plants were selected from these generations. The percentage of seedlings that were free of viral symptoms, after three inoculations in a glasshouse, increased with selection and backcrossing and was highest for the BC_4 and SbC_3 generations: 96.1 and 91.3 %, respectively. When transferred to soil, more plants developed symptoms of PRSV-P. The rate of symptom development in the BC and sib-cross generations was slower than that of the Davao Solo control which produced severe symptoms 2 months after planting. There were plants of BC_2 and SbC_2 that remained symptom-free for about 5–6 months. BC plants in contrast to cv. Davao Solo had the ability to recover from PRSV-P infection. Some BC_4 and SbC_3 plants showed only mild infection after 8 months in the field and had the ability to recover, as new growth had no viral symptoms. A few BC_3 and BC_4 plants were free from viral infection after 18 months in the field. After 18 months they developed very mild symptoms on their leaves and a few ringspots on their fruits. They continued to grow vigorously and produce fruit for 3 years under high disease pressure. Fruit quality of BC_4, BC_3, SbC_3, BC_2, and SbC_2 plants that had low levels or no virus was good, and fruits were generally sweet. This could, in part, be attributed to *V. quercifolia* since it is known for its high sugar levels (Siar et al. 2011). Figure 4.3 shows F_1 hybrids and their backcross progenies.

Siar et al. (2011) attributed the final success to the efficiency in producing large numbers of genotypes at each generation, through (1) the use of efficient embryo culture protocols, (2) a *C. papaya* female clone producing highly prolific immature embryos, (3) the restoration of pollen fertility following the progressive elimination of *V. quercifolia* DNA in the F_1 and BC_1 generations, (4) constant efforts in producing large numbers of plants, and (5) the relative genetic proximity of the *Vasconcellea* parental species. While the latter argument is not corroborated by phylogenetic analyses, the results demonstrate that efforts in wide hybridization to transfer PRSV-P resistance to *C. papaya* are better directed toward crosses between *C. papaya* and *V. quercifolia* than with most other *Vasconcellea* species.

Fig. 4.3 *C. papaya*, *C. papaya* × *V. quercifolia*, hybrids, and backcross plants in UPLB, Los Baños, Philippines. The trees in the front row are PRSV-P susceptible papaya trees. The row behind has an F_1 intergeneric hybrid tree on the right and some obtained from backcrosses from the hybrid onto *C. papaya*. Both the F_1 hybrid tree and the backcross trees show resistance to PRSV-P

The genetic control of PRSV-P resistance in *V. quercifolia* is still unclear. In the F_1 generation of *C. papaya* × *V. quercifolia*, plants segregated three resistant to one susceptible (Drew et al. 1998), which suggests that more than one gene was expressing; however, the interpretation is confounded by incomplete meiotic pairing of chromosomes and the probable preferential elimination of *V. quercifolia* DNA. This conclusion is consistent with the identification of only one PRSV-P-resistant BC_1

Fig. 4.4 *C. papaya* (*right*) and *C. papaya* × *V. cundinamarcensis* growing in close proximity in southeast Queensland, Australia. The papaya tree is badly infected with both PRSV-P and black spot, while the intergeneric hybrid presents no symptoms of either

plant although many others showed susceptibility to varying degrees (Drew et al. 2005). Sib-crossing was accomplished in an effort to produce plants that were homozygous for the PRSV-P-resistant gene(s); however, the results were inconclusive. The SbC_3 generation showed an increased percentage of resistant plants (91.3 %) compared to those of the BC_3 generation (65.3 %); however, the SbC_2 generation contained 48.9 % resistant plants compared to 72.3 % in the BC_2 generation (Siar et al. 2011).

Currently BC_4 and BC_5 generations have been planted in Thailand and Hawaii, and there are further plans to test the resistance of these genotypes against other strains of the virus in other papaya producing countries.

Vasconcellea cundinamarcensis

V. cundinamarcensis, also known as *V. pubescens*, is the only *Vasconcellea* species that has consistently been reported to be resistant to PRSV-P. This contrasts to *V. cauliflora* and *V. quercifolia* where there are reports of some genotypes in some countries being susceptible to PRSV-P (Horovitz and Jiménez 1967). *V. cundinamarcensis* and F_1 hybrids with *C. papaya* also demonstrated resistance to black spot (*Asperisporium caricae*) (Drew et al. 1998; Fig. 4.4). *V. cundinamarcensis* has been repeatedly cited as a source of cold tolerance. The latter trait, related to its

adaptation to the cool climates of its natural Andean habitat, is potentially useful to expand the papaya climatic range of adaptation, taking into account that, while the plant can adapt to cool conditions, it cannot support lowland permanently hot conditions. Indeed, Horovitz and Jiménez report that *V. cundinamarcensis* flowers are completely sterile in Maracay, near sea level, and fertile at 1,800 m. Similarly, Manshardt and Wenslaff (1989b) mention that *V. cundinamarcensis* was fertile in Hawaii lowland regions only in the cooler months. Dhekney et al. (2007) reported the presence of cold-inducible sequences in the genome of *V. cundinamarcensis*, similar to those that induce the cold acclimation pathway in *Arabidopsis*. Cold-inducible sequences were observed only in *Vasconcella*.

Both cold tolerance and PRSV-P resistance were transmitted to the hybrid progenies that were established under field conditions by several teams (Horovitz and Jiménez 1967; Manshardt and Wenslaff 1989b; Drew et al. 1998). According to Manshardt et al. (1995), *C. papaya* × *V. cundinamarcensis* hybrids showed good vigor and resistance in Hawaii. Drew et al. (1998) also report that these hybrids are vigorous and resistant to PRSV-P and that they grow slowly in the hot summer months in a subtropical climate in southeast Queensland but more vigorously in the winter months when the night temperatures can drop to 4 °C. Horovitz and Jiménez (1967) reported that the F_1 hybrids were sterile, as their pollen parent, in Maracay, which could be due to tropical lowland conditions, as they underline that the hybrids were not tested in cooler highlands. This provides another clear example of the importance of breeding sites and attendant ecological conditions on the success of *Carica* × *Vasconcellea* introgression programs. However, large populations of F_1 hybrids that grew vigorously during winter months in Australia were infertile (Drew et al. 1998; O'Brien and Drew 2009). Thus, despite the potential for transfer of PRSV-P resistance from *V. cundinamarcensis* to *C. papaya*, introgression has been prevented to date by infertility of the F_1 hybrids.

Concerning other traits observed in the F_1 hybrids, Jiménez and Horovitz (1958) report pubescence on young stems and petioles, angulous stem at cotyledons height, and anthocyanins on foliar veins. All plants bear only female flowers, colored either yellow or cream (Drew et al. 1998).

Some attempts have been made to hybridize *C. papaya* and *V. cundinamarcensis* by protoplast fusion; however, no progress could be made beyond the first phases of embryogenic development, and no viable structures could be obtained (Jordan et al. 1986). Similar difficulties have been experienced when protoplasts of *C. papaya* were fused with those of *V. cundinamarcensis, V. quercifolia*, and *V. stipulata* (Kanchana-Udomkan et al. 2013; also Chap. 19 in this text). It has been possible to fuse protoplasts of these crosses; however, they only result in a few cell divisions and then inability to progress further.

Vasconcellea parviflora

V. parviflora has a climatic adaptation similar to *C. papaya* (Scheldeman et al. 2007); however, it grows much more slowly under similar climatic conditions

(O'Brien and Drew 2009). It is susceptible to PRSV-P (Horovitz and Jiménez 1967; Magdalita et al. 1998) but resistant against papaya die back, a disease caused by a phytoplasma and which is a more serious disease than PRSV-P in southeast Queensland (Drew et al. 1998).

Manshardt and Wenslaff (1989b) reported the first successful hybridization with *C. papaya*. However, shoots did not produce roots, and attempts to graft the shoots onto papaya seedlings failed. Drew et al. (1998) obtained strong hybrid plants that grew vigorously in the glasshouse. In the field they appeared to be very sensitive to cool climatic conditions and like the *V. parviflora* parents were deciduous in winter. The sex ratio was 1:1 of male/female, and some males produced viable pollen. None of these 100 plants were affected by a severe outbreak of papaya dieback that destroyed hundreds of papaya plants in surrounding plantings.

Because *V. parviflora* × *C. papaya* hybrids presented pollen fertility up to 45 % and *V. parviflora* readily crosses with *V. cundinamarcensis*, it has potential to be used as a bridging species to introgress genes from *V. cundinamarcensis* to *C. papaya* (O'Brien and Drew 2009). Hybrid populations of *C. papaya* × *V. cundinamarcensis*, *C. papaya* × *V. parviflora*, and *V. cundinamarcensis* × *V. parviflora* were produced and evaluated for morphological characteristics and PRSV-P resistance. F_2 and F_3 populations were produced from the *V. cundinamarcensis* × *V. parviflora* F_1 hybrids. PRSV-P-resistant individuals with homozygous genotypes (RR) were selected by use of a codominant CAPS marker and backcrossed to *V. parviflora* or outcrossed to *C. papaya* (Dillon et al. 2005, 2006; O'Brien and Drew 2009). When F_2 and F_3 plants were outcrossed to *C. papaya*, resultant hybrids were infertile. Thus, a marker-assisted backcross program is being conducted to produce *V. parviflora* plants containing genes for PRSV-P resistance from *V. cundinamarcensis*. Attempts will be made to transfer this virus resistance to *C. papaya*.

Vasconcellea monoica

V. monoica is susceptible to PRSV-P (Horovitz and Jiménez 1967) but resistant to papaya bunchy top (Sawant 1958). Other traits of potential interest are monoecy, vigor, and the ability of *V. monoica* and its hybrids to retain its fruits on the trees for long periods of time, remaining firm for up to 12 months (Mekako and Nakasone 1975). Manshardt and Wenslaff (1989b) crossed *V. monoica* to *C. papaya*; however, shoots did not produce roots, and attempts to graft the shoots onto papaya seedlings failed.

According to Jiménez and Horovitz (1958), crosses between *V. cauliflora* and *V. monoica* produced viable seeds in both directions, with variable results, and F_1, F_2, and BC_1 hybrids were fertile. Sawant (1958) reported failure of the cross when *V. cauliflora* was the seed parent. Mekako and Nakasone (1975) reported that *V. monoica*, when used as the seed parent, reduced fruit set by half, as compared to the reciprocal cross. *V. monoica* also crossed successfully in both directions to *V. microcarpa* and *V. cundinamarcensis* and produced vigorous and fertile hybrid progenies (Warmke et al. 1954; Horovitz and Jiménez 1967). Thus, *V. monoica* should be

explored further as a potential bridging species to get access to genes of *V. cauliflora* and *V. cundinamarcensis*. *V. monoica* may also produce viable hybrid seeds when crossed to *V. goudotiana*, if used as the pollen donor (Warmke et al. 1954; Jiménez and Horovitz 1958; Sawant 1958; Mekako and Nakasone 1975).

Monoecy is determined by a specific dominant allele of the sex gene. The F_1 between *V. monoica* and *V. cauliflora* or *V. goudotiana* is intermediate between the two parents, vigorous and monoecious, bearing abundant fertile male and female flowers, while the F_2 displays wide segregation for morphological traits and segregates in a 3:1 ratio for monoecy (Warmke et al. 1954; Horovitz and Jiménez 1967). Different results were reported by Horovitz and Jiménez (1967) when the female parent was from *V. cundinamarcensis*, *V. stipulata*, *V. microcarpa*, or *V. horovitziana*. Their F_1 hybrids segregated in a 1:1 ratio for monoecy, which led these authors to propose that these species have a different allele of the sex gene, or a dominant gene with a pleiotropic effect on monoecy. This gene could even exist in *V. monoica*, as suggested by the existence of two rare specimens presenting all traits of *V. monoica*, except that they were female (Badillo, cited in Horovitz and Jiménez 1967). It is worth noting that Mekako and Nakasone (1975) mentioned male and female plants in the *V. goudotiana* × *V. monoica* F_1 progeny that they obtained, which suggests intraspecific variation for genetic factors affecting the expression of monoecy.

Vasconcellea stipulata

V. stipulata has been cited as a potential source of interesting flavor and color in addition to PRSV-P resistance (Scheldeman 2002). Manshardt and Wenslaff (1989b) crossed it to *C. papaya* and successfully established hybrids in the field. Climatic restrictions on its growth were similar to those for *V. cundinamarcensis*, and *V. stipulata* was fertile in Hawaii lowlands only in the cooler months. The two species produce germinable hybrid seeds in both directions (Horovitz and Jiménez 1967), so *V. stipulata* could be tested as a potential bridging species to access the interesting genes of *V. cundinamarcensis*. More generally, as it also forms viable seeds when crossed to *V. microcarpa*, *V. stipulata* appears to belong to gene pool I of Jiménez and Horovitz (1958).

Vasconcellea goudotiana

V. goudotiana is considered susceptible to PRSV-P (Horovitz and Jiménez 1967; Magdalita et al. 1988) but resistant to *Phytophthora* (Drew et al. 1998). Hybrids with the common papaya, both male and female trees, grew vigorously in the field and produced high yields of fruit (Drew et al. 1998; Fig. 4.5); however, they were prone to root-rots other than *Phytophthora*. This is consistent with a report by Warmke et al. (1954) who observed a root disease that severely affected their *V. goudotiana* × *V. monoica* F_2 generation. Other traits commonly observed in hybrids

Fig. 4.5 A high-yielding *C. papaya* × *V. goudotiana* F_1 hybrid growing in a subtropical climate in southeast Queensland bearing multiple fruit on each peduncle, similar to the papaya parent

from *V. goudotiana* are deeply lobulated leaves with foliar veins that frequently exhibit an intense red color.

Mekako and Nakasone (1975) mentioned limited success in the cross *V. parviflora* × *V. goudotiana* when the pollen came from a Colombian accession and complete failure when it came from a Venezuelan one, again indicating the importance of understanding infraspecific genetic variation in introgression programs.

Other Species

No attempts at hybridization have been reported concerning *V. candicans*, reported as PRSV-P resistant by Horovitz and Jiménez (1967). Other species have been neglected in introgression breeding, probably because *V. microcarpa* and *V. horovitziana* were reported as susceptible to PRSV-P by these authors. Both species are sexually compatible with species of gene pool I of Jiménez and Horovitz (1958), so they may have potential as bridging species with *V. cauliflora* and *V. cundinamarcensis*, and they should be tested for crossability with the common papaya. Furthermore, intraspecific variation for PRSV-P resistance may exist in these species, requiring characterization of a wider range of samples of the different species.

This is particularly true for a species with a very ample distribution, such as the highly polymorphic *V. microcarpa*.

We are not aware of any reports on the study of other species such as *V. glandulosa* that has a wide distribution and considerable polymorphism, as well as endemics, such as *V. crassipetala*, *V. horovitziana*, *V. longifolia*, *V. omnilingua*, *V. pulchra*, *V. sprucei*, or *V. weberbaueri*. Thus, most of the necessary characterization remains to be done to enable a reasonable assessment of the potential of these *Vasconcellea* for the secondary gene pool of the common papaya.

Vasconcellea Conservation and Genetic Erosion

The risk of genetic erosion in wild papaya species was long considered low because of their fast growth, their easy adaptation to disturbed habitats, numerous seeds and their breeding system which favors outcrossing (IBPGR 1986). However, this conclusion now appears simplistic and overoptimistic. Indeed, *Vasconcellea* species show significant differences in conservation status. Some species such as *V. cundinamarcensis* or *V. microcarpa* present wide distributions, while others are endemics, thus being more vulnerable to genetic erosion or even extinction.

No formal studies on genetic erosion of *Vasconcellea* species have been conducted to date. In southern Ecuador, farmers often eliminate local materials from home gardens and replace them by babaco. Increased pressure on forests, mainly driven by expansion of pastures, is having a significant and negative effect on the presence of wild germplasm, which often prefer forest margins (Scheldeman et al. 2002). A Red Listing exercise from Bolivia also reports that wild *Vasconcellea* populations are subject to increased pressure from grazing, burning, and extension of the agricultural frontier (N. de la Barra, pers. comm.).

Five of the 21 *Vasconcellea* species have been included in the Red List, with *V. omnilingua* and *V. horovitzia* listed as endangered, *V. pulchra* and *V. sprucei* as near threatened, and the recently described species *V. palandensis* as vulnerable (IUCN 2011). The latter should even be considered as critically endangered, as no specimens were found in the area from which the species was first collected and described by Badillo et al. (2000) and as most of the forest in which it occurred has been cut down (X. Scheldeman and T. Kyndt, pers. obs.). Furthermore, according to Scheldeman et al. (2011), two additional species (*V. chilensis* and *V. weberbaueri*) should be included in the Red List, on the basis of the extent of their occurrence. Thus, at least one third of all known *Vasconcellea* species are facing some degree of risk.

In Situ Conservation

From the revision of Scheldeman et al. (2011), it appears that 9 of the 21 species have not been observed in protected areas so far. Of concern is that of these nine species, four are Red List species.

Ex Situ Conservation

Vasconcellea germplasm collections exist in Colombia (CORPOICA), Ecuador (INIAP, Universidad Técnica de Ambato, UNL, Loja Botanical Garden), Peru (INIA, UNALM, UNSAAC), Brazil (EMBRAPA), France (CIRAD-Guadeloupe), the USA (Hawaii-USDA), and India (IARI, IIHR). The revision and update of available information by Scheldeman et al. (2011) indicates adequate ex situ conservation for *V. cundinamarcensis* and to a lesser degree for *V. cauliflora*, *V.* × *heilbornii*, *V. goudotiana, V. monoica*, and *V. stipulata*. No collections of the threatened (Red Listed) species, *V. horovitziana*, *V. omnilingua, V. palandensis, and V. sprucei,* are found in gene banks, and this is a reason for concern and corrective measures. In total, 6 of the 21 *Vasconcellea* species are not reported to be conserved ex situ. There is limited information on the methods used for conservation. Most collections are maintained either as field or seed collections. Only part of the USDA collection is conserved in vitro. Field collections often suffer from lack of long-term funding. There have been several reports of loss of field collections, mainly in Ecuador, the country with the highest diversity. Some *Vasconcellea* germplasm is maintained in botanical gardens (Botanical Gardens Conservation International 2008); however, these activities focus on conservation of the species, giving less emphasis on intraspecific diversity. Effective cryopreservation of most *Vasconcellea* species have not been reported although a vitrification-based shoot tip cryopreservation protocol was developed and successfully applied to *V. cundinamarcensis* (Ashmore and Drew 2006; Ashmore et al. 2007).

As very little is known on long-term conservation of *Vasconcellea* seeds, it is not known whether some of the reported seed collections still contain viable seeds. A study on viability of *V. cundinamarcensis* seeds (Vanhove 2000) indicates there was no significant loss of viability in seeds that had been conserved at 4 °C for 2 years, indicating a probable orthodox seed behavior for this species. No information on conservation for longer periods could be found in literature; however, limited experiments on *Vasconcellea* seed cryopreservation are encouraging (Coppens d'Eeckenbrugge, unpublished).

Neither *Carica* nor *Vasconcellea* are included in Annex I of the International Treaty on Plant Genetic Resources for Food and Agriculture (Fowler et al. 2003). Although countries can decide to provide access according to the terms of the Treaty to non-Annex I crops (SGRP 2007), in most cases access takes place according to the provisions of the Convention of Biological Diversity (CBD), based on bilateral agreements.

Concluding Remarks

The history of breeding at the intergeneric level in Caricaceae reminds us that plant breeding is both a science and an art. Ideally, the former should have preceded the latter; however, the reverse often happened. Early attempts, made at a time when all

species were included in a single genus, produced the first clear indication that the common papaya and the highland papayas belong to markedly different gene pools, and this was confirmed half a century later by thorough genetic studies. Presently, even the significant recent progress in our understanding of the organization of Caricaceae diversity does not provide us with a complete explanation of the observations and first successes of hybrid breeding programs. Thus, while we have a consistent picture of the genetic organization of *Vasconcellea*, containing three clades identified by different molecular tools, we cannot relate this to the accumulated experience of difficulty in hybridization. Crossability with *C. papaya* is highly variable, even within species, and success clearly depends on particular germplasm and ecoclimatic conditions of the breeding site. Some of the most encouraging results were obtained with *V. quercifolia*, *V. stipulata*, *V. parviflora*, and *V. goudotiana*, i.e., in species from all three clades. Even interspecific crosses within *Vasconcellea* do not conform to expectations based on genetic affinities. *V. cundinamarcensis* and *V. stipulata* are not closely related, yet represent the best documented case of gene flow within *Vasconcellea*. Similarly, F_1 and F_2 hybrids of *V. cundinamarcensis* and *V. parviflora* are vigorous, and most markers show no distortion in segregation. A better understanding of the genome organization would probably help in understanding and predicting hybridization success, as suggested by the first studies of de Zerpa (1959, 1980) on meiosis in the Venezuelan hybrids, as well as correlations observed between hybrid fertility and aneuploidy rates. Current advanced microscopical techniques like Fluorescence in situ hybridization and genomic in situ hybridization will certainly be helpful in the further elucidation of the genomic constitution and chromosomal interaction of interspecific and intergeneric hybrids.

Similarly, the distribution of the main trait of interest does not follow the logics of genetic affinities. PRSV-P resistance was found in *V. cauliflora*, *V. cundinamarcensis*, *V. stipulata*, *V. candicans*, and *V. quercifolia*, i.e., in species from all three clades, similarly for the species that are susceptible. Moreover, contradictory results have indicated that PRSV-P resistance varies even within species. There remains the possible speculation that the type of resistance may be clade related, with the strong, monogenic resistance found in *V. cauliflora* and *V. cundinamarcensis*, both from clade 3.

Thus, beyond the success in the introgression of resistance from *V. quercifolia*, the future prospects for *Vasconcellea* germplasm in papaya improvement are a very open question, and the breeder must still rely on his/her sense of observation and his/her inventivity; for example, in the use of bridging species for the introgression of new traits of great economic potential, such as monoecy or cold tolerance, stronger resistance to PRSV-P, or resistance to other threatening diseases (e.g., bacterial canker). Other approaches may involve isolating, mapping, and cloning useful genes from segregating *Vasconcellea* hybrids for purposes of genetic transformation.

While practical breeding has preceded modern botany in the planning of hybridization programs, it is still far behind in terms of knowledge and exploitation of genetic resources. Many *Vasconcellea* species have not yet been tested and/or characterized for traits of potential interest. Meanwhile, much remains to be discovered and protected from rapid biodiversity erosion. Several species or species complexes

remain to be described and/or characterized for phenotypic and genetic diversity. A significant effort of collection and conservation is still ahead, which will strongly depend on our capacity to establish international collaboration. Germplasm documentation should be improved as well as a knowledge of original environments seems particularly important for the success of hybridizations.

Finally, recent developments in research on the whole Caricaceae family open a range of possibilities well beyond the genus *Vasconcellea*, as *Jarilla* and *Horovitzia* appear to be more closely relative to *Carica*. The information available on these herbaceous species is very scarce. The fruits of the three *Jarilla* species are edible. Those of *J. heterophylla* (which reach sizes of 18 cm × 9 cm) are found in Mexican local markets (F.A. Carvalho, pers. comm.). The other two species have much smaller fruits (2–7 cm × 2–4 cm). According to the description provided by Tookey and Gentry (1969), "fruits of *J. chocola* have a faint flavor suggestive of lemon." *J. heterophylla* occurs at high altitudes (1,500–2,000 m), probably being more cold tolerant than *J. chocola*, which occurs at lower levels. There has been no investigation on the resistance of these plants to virus, bacteria, or any other pathogens or on the existence of reproductive barriers between them and *Carica*.

Acknowledgments The authors thank Fernanda Antunes Carvalho and Susanne Renner for information on *Jarilla* and *Horovitzia* and fruitful discussion on botanical and genetic affinity within Caricaceae.

References

Alvizo HF, Rojkind C (1987) Resistencia al virus mancha anular del papayo en *Carica cauliflora*. Rev Mex Fitopatol 5:61–62

Amaral PP, Resende RO, Júnior MTS (2006) Papaya Lethal Yellowing Virus (PLYV) infects *Vasconcellea cauliflora*. Fitopatol Bras 31:4004

Aradhya MK, Manshardt RM, Zee F, Morden CW (1999) A phylogenetic analysis of the genus *Carica* L. (Caricaceae) based on restriction fragment length variation in a cpDNA intergenic spacer region. Genet Resour Crop Evol 46:579–586

Ashmore SE, Drew RA (2006) The application of biotechnology in an integrated project of conservation and utilization of papaya and its wild relatives. Acta Hortic 725:89–94

Ashmore SE, Drew RA, Azimi M (2007) Vitrification-based shoot tip cryopreservation of *Carica papaya* and a wild relative *Vasconcellea pubescens*. Aust J Bot 55:541–547

Badillo VM (1967) Esquema de las Caricaceae. Agron Trop 4:245–272

Badillo VM (1971) Monografía de la familia Caricaceae. Associacion de Profesores, Univ Central Venezuela, Maracay

Badillo VM (1993) Caricaceae. Segundo esquema. Rev Fac Agron Univ Central Venezuela, Alcance 43, Maracay

Badillo VM (2000) *Carica* L. vs *Vasconcella* St. Hil. (Caricaceae) con la rehabilitación de este último. Ernstia 10:74–79

Badillo VM (2001) Nota correctiva *Vasconcellea* St. Hill. y no *Vasconcella* (Caricaceae). Ernstia 11:75–76

Badillo VM, Van den Eynden V, Van Damme P (2000) *Carica palandesis* (Caricaceae), a new species from Ecuador. Novon 10:4–6

Baeza G, Correa D, Salas C (1990) Proteolytic enzymes in *Carica candamarcensis*. J Sci Food Agric 51:1–9

Barton NH, Hewitt GM (1985) Analysis of hybrid zones. Annu Rev Ecol Syst 16:113–148
Bentham G, Hooker JD (1867) Genera plantarum: ad exemplaria imprimis in Herbariis Kewensibus servata definite. Reeve, London
Botanical Gardens Conservation International (2008) PlantSearch. London. Accessible on http://www.bgci.org/plant_search.php
Brücher H (1977) Tropische Nutzpflanzen: Ursprung, Evolution und Domestikation. Springer, Berlin
Brücher H (1989) Useful plants of neotropical origin and their wild relatives. Springer, Berlin
Burnham RJ, Graham A (1999) The history of neotropical vegetation: new developments and status. Ann Miss Bot Gard 86:546–589
CAF (1992) Manual técnico del cultivo de chamburo. Centro Agrícola de Quito, Corporación Andina de Fomento, Quito
Carrasco B, Avila P, Pérez-Díaz J, Muñoz P, García R, Lavandero B, Zurita-Silva A, Retamales JB, Caligari PDS (2009) Genetic structure of highland papayas (*Vasconcellea pubescens* (Lenné et C. Koch) Badillo) cultivated along a geographic gradient in Chile as revealed by Inter Simple Sequence Repeats (ISSR). Genet Resour Crop Evol 56:331–337
Carvalho FA, Renner SS (2013) The phylogeny of the Caricaceae. In: Ming R, Moore P (eds) Genetics and genomics of papaya. Springer Science+Business Media, New York
Conover RA (1964a) Distortion ringspot, a severe virus disease of papaya in Florida. Proc Fla State Hortic Soc 77:440–444
Conover RA (1964b) Mild mosaic and faint mottle ringspot, two papaya virus diseases of minor importance in Florida. Proc Fla State Hortic Soc 77:444–448
Corrêa NCR, Mendes IC, Gomes MTR, Kalapothakis E, Chagas BCA, Lopes MTP, Salas CE (2011) Molecular cloning of a mitogenic proteinase from *Carica candamarcensis*: its potential use in wound healing. Phytochemistry 72:1947–1954
Costa FR, Pereira TNS, Hodnett GL, Pereira MG, Stelly DM (2008) Fluorescent in situ hybridization of 18S and 5S rDNA in papaya (*Carica papaya* L.) and wild relatives. Caryologia 61:411–416
Darlington CD, Wylie AP (1955) Chromosome atlas of flowering plants. George Allen & Unwin Ltd, London
De Feo V, De Simone F, Arias Arryo G, Senatore F (1999) *Carica candicans* Gray (Mito), an alimentary resource from Peruvian flora. J Agric Food Chem 47:3682–3684
de Mello JC, Spruce R (1869) Notes on Papayaceae. J Linn Soc Bot 10:1–15
de Zerpa DM (1958) Tetraploidía en *Carica papaya* y *Carica cauliflora*. Agron Trop 8:67–73
de Zerpa DM (1959) Citología de híbridos interespecíficos en *Carica*. Agron Trop 8:135–144
de Zerpa DM (1980) Comportamiento meiótico de la descendencia híbrida producida al transferir el carácter bisexual de *Carica pubescens* a *Carica stipulata*. Rev Fac Agron (Maracay) 11:5–47
Dhekney SA, Litz RE, Moraga D, Yadav A (2007) Is it possible to induce cold tolerance in papaya suing genetic transformation? Acta Hortic 738:159–163
Dhuique Mayer C, Caro Y, Pina M, Ruales J, Dornier M, Graille J (2001) Biocatalytic properties of lipase in crude latex from babaco fruit (*Carica pentagona*). Biotechnol Lett 23:1021–1024
Dillon S, Ramage C, Drew R, Ashmore S (2005) Genetic mapping of a PRSV-P resistance gene in "highland papaya" based on inheritance of RAF markers. Euphytica 145:11–23
Dillon S, Ramage C, Ashmore S, Drew RA (2006) Development of a codominant CAPS marker linked to PRSV-P resistance in highland papaya. Theor Appl Genet 113:1159–1169
Drew RA, O'Brien CM, Magdalita PM (1998) Development of interspecific *Carica* hybrids. Acta Hortic 461:285–292
Drew RA, Siar SV, Villegas VN, O'Brien CM, Sajise AGC (2005) Development of PRSV-P resistant *Carica papaya* genotypes by introgression of genes from wild *Vasconcella* species. Acta Hortic 694:73–77
Drew RA, Siar SV, O'Brien CM, Magdalita PM, Sajise AGC (2006a) Breeding for papaya ringspot virus resistance in *Carica papaya* via hybridisation with *Vasconcellea quercifolia*. Aust J Exp Agric 46:413–418
Drew RA, Siar SV, O'Brien CM, Sajise AGC (2006b) Progress in backcrossing between *Carica papaya* × *Vasconcellea quercifolia* intergeneric hybrids and *C. papaya*. Aust J Exp Agric 46:419–424

Endt DW (1981) The babaco – a new fruit in New Zealand to reach commercial production. Orchardist New Zealand 54:58–59
FAO (1992) Cultivos marginados. Otra perspectiva de 1492. Colección FAO, Producción y protección vegetal, No. 26. FAO, Rome
Fermin G, Castro LT, Tennant PF (2010) CP-transgenic and non-transgenic approaches for the control of papaya ringspot: current situation and challenges. Transgenic Plant J 4:1–15
Fowler C, Moore G, Hawtin GC (2003) The International Treaty on Plant Genetic Resources for Food and Agriculture: a primer for the Future Harvest Centres of the CGIAR. International Plant Genetic Resources Institute, Rome
Gonsalves D (1998) Control of papaya ringspot virus in papaya: a case study. Annu Rev Phytopathol 36:415–437
González A, Trujillo G (2005) Susceptibilidad de *Vasconcellea cauliflora* al virus de la mancha anillada de la lechosa. Rev Fac Agron (Maracay) 22:229–240
Harms H (1925) Caricaceae. In: Engler E (ed) Die natürlichen Planzenfamilien. Bd 21, 2 Aufl. Von wilhelm Englemann, Leipzig
Higgins JE, Holt VS (1914) The papaya in Hawaii. Hawaii Agric Exp Sta Bull 32
Horovitz S (1954) Determinación del sexo en *Carica papaya* L.: estructura hipotética de los cromosomas sexuales. Agron Trop (Maracay) 3:229–249
Horovitz S, Jiménez H (1967) Cruzamientos interespecíficos e intergenéricos en Caricaceas y sus implicaciones fitotécnicas. Agron Trop (Maracay) 17:323–343
IBPGR (1986) *Carica papaya* (Papaya). In: Genetic resources of tropical and sub-tropical fruits and nuts (excluding *Musa*). International Board for Plant Genetic Resources, Rome
IUCN (2011) Red list of threatened species. Version 2011.1. www.iucnredlist.org. Downloaded on 17 Oct 2011
Jiménez H (1957) Injertos entre especies de Carica. Agron Trop 7:33–37
Jiménez D (2002) Caracterización y estudio de la diversidad genética de los géneros *Vasconcellea* y *Carica* (Caricaceae) en Colombia y Ecuador con marcadores isoenzimáticos. Dissertation. Universidad de Caldas, Manizales, Colombia
Jiménez H, Horovitz S (1958) Cruzabilidad entre especies de *Carica*. Agron Trop 7:207–215
Jiménez Y, Romero J, Scheldeman X (1998) Colección, caracterización y descripción de *Carica*×*heilbornii* nm. *pentagona* B.; *Carica pubescens* (A.DC) Solms-Laub. y *Carica stipulata* B., en la provincia de Loja. Rev Difusión Técnica y Científica Fac Ciencias Agric (UN Loja) 29:43–54
Jobin-Décor MP, Graham GC, Henry RJ, Drew RA (1997) RAPD and isozyme analysis of genetic relationships between *Carica papaya* and wild relatives. Genet Resour Crop Evol 44:471–477
Jordan M, Ciudad G, Rojas ML, Valverde F (1986) Isolation, culture and fusion of *Carica candamarcensis* and *C. papaya* protoplasts. Gartenbauwissenschaft 51:175–178
Kanchana-Udomkan C, Ford R, Drew R (2013) Molecular markers in papayas. In: Ming R, Moore P (eds) Genetics and genomics of papaya. Springer Science+Business Media, New York
Kempler C, Kabaluk T (1996) Babaco (*Carica pentagona* Heilb.): a possible crop for the greenhouse. Hortscience 31(5):785–788
Kent L (2004) What's the holdup? Addressing constraints to the use of plant biotechnology in developing countries. AgBioForum 7:63–69
Kyndt T, Romeijn-Peeters E, Van Droogenbroeck B, Romero-Motochi JP, Gheysen G, Goetghebeur P (2005a) Species relationships in the genus *Vasconcellea* (Caricaceae) based on molecular and morphological evidence. Am J Bot 92:1033–1044
Kyndt T, Van Droogenbroeck B, Romeijn-Peeters E, Romero-Motochi JP, Scheldeman X, Goetghebeur P, Van Damme P, Gheysen G (2005b) Molecular phylogeny and evolution of Caricaceae based on rDNA internal transcribed spacer (ITS) and chloroplast sequence data. Mol Phylogenet Evol 37:442–459
Kyndt T, Van Droogenbroeck B, Haegeman A, Roldán-Ruiz I, Gheysen G (2006) Cross-species microsatellite amplification in *Vasconcellea* and related genera and their use in germplasm classification. Genome 49:786–798
Kyndt T, Van Damme EJM, Van Beeumen J, Gheysen G (2007) Purification and characterization of the cysteine proteinases in the latex of *Vasconcellea* spp. FEBS J 274:451–462

Lassoudière A (1969) Le papayer (quatrième partie). Description des inflorescences et fleurs des papayers "Solo". Fruits 24:143–151
Leal F, Coppens d'Eeckenbrugge G (2003) Memorias del Primer Taller Internacional sobre Caricaeae, Maracay October 19–21 1999. FONTAGRO-IICA, Impresora Feriva, Cali
Magdalita PM, Villegas VN, Pimentel RB, Bayot RG (1988) Reaction of papaya (*Carica papaya* L.) and related species to ringspot virus. Philipp J Crop Sci 13:129–132
Magdalita PM, Adkins SW, Godwin ID, Drew RA (1996) An improved embryo-rescue protocol for a *Carica* interspecific hybrid. Aust J Bot 44:343–353
Magdalita PM, Drew RA, Adkins SW, Godwin ID (1997a) Morphological, molecular and cytological analysis of *Carica papaya* × *C. cauliflora* interspecific hybrids. Theor Appl Genet 95:224–229
Magdalita PM, Persley DM, Godwin ID, Drew RA, Adkins SW (1997b) Screening *Carica papaya* × *C. cauliflora* hybrids for resistance to papaya ringspot virus-type P. Plant Pathol 46:837–841
Magdalita PM, Adkins SW, Godwin ID, Drew RA (1998) Randomly amplified polymorphic DNA markers for a Carica interspecific hybrid. Acta Hortic 461:133–140
Magdalita PM, Drew RA, Adkins SW (2001) Breeding for resistance to papaya ringspot virus. J Trop Plant Pathol 37:1–15
Manshardt RM (1992) Papaya. In: Hammerschlag FA, Litz RE (eds) Biotechnology of perennial fruit crops. CAB International, Oxon
Manshardt RM, Wenslaff TF (1989a) Zygotic polyembryony in interspecific hybrids of *Carica papaya* and *C. cauliflora*. J Am Soc Hortic Sci 114:684–689
Manshardt RM, Wenslaff TF (1989b) Interspecific hybridization of papaya with other *Carica* species. J Am Soc Hortic Sci 114:689–694
Manshardt RM, Zee FTP (1994) Papaya germplasm and breeding in Hawaii. Fruit Var J 48:146–152
Manshardt RM, Zee F, Ferreira S, Wenslaff T, Fitch M, Gonsalves D, Slightom JL, Sanford J, Stiles J (1995) Papaya breeding for PRV resistance. Acta Hortic 370:27–32
Mekako HU, Nakasone HY (1975) Interspecific hybridization among 6 *Carica* species. J Am Soc Hortic Sci 100:237–242
Mello VJ, Gomes MTR, Lemos FO, Delfino JL, Andrade SP, Lopes MTP, Salas CE (2008) The gastric ulcer protective and healing role of cysteine proteinases from *Carica candamarcensis*. Phytomedicine 15:237–244
Morton J (1987) Papaya. In: Fruits of warm climates. Creative Resource Systems, Winterville, pp 336–346
National Botanic Garden of Belgium (2008) Living plant collections database (LIVCOL). Meise. Accessible on http://www.br.fgov.be/RESEARCH/COLLECTIONS/LIVING/LIVCOL/index.html.
National Research Council (1989) Highland papayas. In: Ruskin FR (ed) Lost crops of the Incas: little-known plants of the Andes with promise for worldwide cultivation. National Academy Press, Washington, DC, pp 252–261
O'Brien CM, Drew RA (2009) Potential for using *Vasconcellea parviflora* as a bridging species in intergeneric hybridisation between *V. pubescens* and *Carica papaya*. Aust J Bot 57:592–601
Ochoa J, Fonseca G, Ellis M (2000) First report of *Fusarium* wilt of babaco (*Carica* × *heilbornii* var. *pentagona*) in Ecuador. Plant Disease 84:199
Olarte Castillo XA, Fermin G, Tabima J, Rojas Y, Tennant PF, Fuchs M, Sierra R, Bernal AJ, Restrepo S (2011) Phylogeography and molecular epidemiology of Papaya ringspot virus. Virus Res 159:132–140
Prance GT (1984) The pejibaye, *Guilielma gasipaes* (H.B.K.) Bailey, and the papaya, *Carica papaya* L. In: Stone D (ed) Precolumbian plant migration, vol 76. Papers Peabody Museum, Cambridge, pp 85–104
Restrepo MT, Duval M-F, Coppens d'Eeckenbrugge G, Jiménez D, Vega J, Van Droogenbroeck B (2004a) Study of cpDNA diversity in mountain papayas and the common papaya using PCR-RFLP markers. Proc Interamer Soc Trop Hortic 48:101–107

Restrepo MT, Jiménez D, Coppens d'Eeckenbrugge G, Vega J (2004b) Morphological diversity of cultivated mountain papayas (*Vasconcellea* spp.) in Ecuador. Proc Interamer Soc Trop Hortic 48:119–123

Romeijn-Peeters E (2004) Biodiversity of the genus *Vasconcellea* (Caricaceae) in Ecuador: a morphological approach. Dissertation. Ghent University

Sawant AC (1958) Crossing relationships in the genus Carica. Evolution 12:263

Scheldeman X (2002) Distribution and potential of cherimoya (*Annona cherimola* Mill.) and highland papayas (*Vasconcellea* spp.) in Ecuador. Dissertation. Ghent University

Scheldeman X, Van Damme P, Romero Motoche J (2002) Highland papayas in southern Ecuador: need for conservation actions. Acta Hortic 575:199–205

Scheldeman X, Romero Motoche JP, Van Damme V, Heyens V, Van Damme P (2003) Potential of highland papayas (*Vasconcella* spp.) in southern Ecuador. Lyonia 5:73–80

Scheldeman X, Willemen L, Coppens d'Eeckenbrugge G, Romeijn-Peeters E, Restrepo MT, Romero Motoche J, Jiménez D, Lobo M, Medina CI, Reyes C, Rodríguez D, Ocampo JA, Van Damme P, Goetghebeur P (2007) Distribution, diversity and environmental adaptation of highland papaya (*Vasconcellea* spp.) in tropical and subtropical America. Biodivers Conserv 16:1867–1884

Scheldeman X, Kyndt T, Coppens d'Eeckenbrugge G, Ming R, Drew R, Van Droogenbroeck B, Van Damme P, Moore PH (2011) Vasconcellea and Carica. In: Kole C (ed) Wild crop relatives: genomic and breeding resources. Springer, Berlin, pp 213–249

SGRP (2007) A de facto certificate of source: the Standard Material Transfer Agreement under the International Treaty. SGRP, Rome

Siar SV, Geronimo SB, Sierra Z, Villegas VN (1998) Cytology of *Carica papaya*, *Carica cauliflora* and their F_1 interspecific hybrids. Philipp J Crop Sci 23:91–96

Siar SV, Beligan GA, Sajise AJC, Villegas VN, Drew RA (2011) Papaya ringspot virus resistance in *Carica papaya* via introgression from *Vasconcellea quercifolia*. Euphytica 181:159–168

Solms-Laubach H (1889) Caricaceae. Flora Bras 13:174–196

Solms-Laubach H (1894) Caricaceae. Engl Prantl Nat Pflanzenfam 3:94–99

Tennant P, Fermin G, Fitch M, Manshardt R, Slightom J, Gonsalves D (2001) Papaya ringspot virus resistance of transgenic Rainbow and SunUp is affected by gene dosage, plant development, and coat protein homology. Eur J Plant Pathol 107:645–653

Tookey HL, Gentry HS (1969) Proteinase of *Jarilla chocola*, a relative of papaya. Phytochemistry 8(6):989–991

Tripathi S, Suzuki JY, Ferreira SA, Gonsalves D (2008) Papaya ringspot virus-P: characteristics, pathogenicity, sequence variability and control. Mol Plant Pathol 9:269–280

Van den Eynden V, Cueva E, Cabrera O (1999) Plantas silvestres comestibles del sur del Ecuador – Wild edible plants of southern Ecuador. Ediciones Abya-Yala, Quito

Van den Eynden V, Cueva E, Cabrera O (2003) Wild foods from Ecuador. Econ Bot 57:576–603

Van Droogenbroeck B, Breyne P, Goetgebheur P, Romeijn-Peeters E, Kyndt T, Gheysen G (2002) AFLP analysis of genetic relationships among papaya and its wild relatives (Caricaceae) from Ecuador. Theor Appl Genet 105:289–297

Van Droogenbroeck B, Kyndt T, Maertens I, Romeijn-Peeters E, Scheldeman X, Romero-Motochi JP, Van Damme P, Goetghebeur P, Gheysen G (2004) Phylogenetic analysis of the highland papayas (*Vasconcellea*) and allied genera (Caricaceae) using PCR-RFLP. Theor Appl Genet 108:1473–1486

Van Droogenbroeck B, Kyndt T, Romeijn-Peeters E, Van Thuyne W, Goetghebeur P, Romero-Motochi JP, Gheysen G (2006) Evidence of natural hybridization and introgression between *Vasconcellea* species (Caricaceae) from southern Ecuador revealed by morphological and chloroplast, mitochondrial and nuclear DNA markers. Ann Bot 97:793–805

Vanhove W (2000) Generatieve vermenigvuldiging en genetische variatie bij het genus *Carica*. MSc Dissertation. Ghent University

Vegas A, Trujillo G, Sandrea Y, Mata J (2003) Obtención, regeneración y evaluación de híbridos intergenéricos entre *Carica papaya* y *Vasconcellea cauliflora*. Interciencia 28:710–714

Wang CH, Yeh SD (1997) Divergence and conservation of the genomic RNAs of Taiwan and Hawaii strains of papaya ringspot potyvirus. Arch Virol 142:271–285

Wang CH, Bau HJ, Yeh SD (1994) Comparison of the nuclear inclusion b protein and coat protein genes of five papaya ringspot virus strains distinct in geographic origin and pathogenicity. Phytopathology 84:1205–1209

Warmke HE, Cabanillas E, Cruzado HJ (1954) A new interspecific hybrid in the genus *Carica*. Proc Am Soc Hortic Sci 64:284–288

Chapter 5
The Phylogeny of the Caricaceae

Fernanda Antunes Carvalho and Susanne S. Renner

The Position of the Caricaceae and Moringaceae in the Brassicales

The Caricaceae and their sister family Moringaceae are part of the mustard-oil plant clade or Brassicales, which also comprises 15 other families, including Brassicaceae, Capparaceae, and Tropaeolaceae. Sister group relationships between the Caricaceae and the Moringaceae and the position of both among the early diverging Brassicales are well supported by molecular data (for a recent large-scale phylogeny of the Brassicales, see Beilstein et al. 2010). Moringaceae comprise one genus, *Moringa*, with 13 species that mostly occur in Africa and Madagascar, although a few are endemic in India. Moringaceae are woody, often stout-stemmed shrubs or trees with spiral, odd-pinnate, 1- to 3-compound leaves that have conspicuous glands at the leaflet articulations (Stevens 2001). Their flowers resemble those of legumes and their fruits are three-angled capsules. Especially striking is the growth form of Moringaceae, namely, "bottle trees" or tuberous shrubs, often with fleshy root/stem transitions called pachypodia (Olson and Rosell 2006). All species occur in more or less arid habitats. Some of these traits probably already occurred in the shared common ancestor of Caricaceae and Moringaceae.

F.A. Carvalho • S.S. Renner (✉)
Department of Systematic Botany and Mycology, University of Munich, Munich 80638, Germany
e-mail: renner@lrz.uni-muenchen.de

R. Ming and P.H. Moore (eds.), *Genetics and Genomics of Papaya*, Plant Genetics and Genomics: Crops and Models 10, DOI 10.1007/978-1-4614-8087-7_5,

Relationships Within Caricaceae Revealed by Morphological and Molecular Studies

Caricaceae include around 35 species in 6 genera and have a disjunct distribution between Africa and the Neotropics (Fig. 5.1). *Cylicomorpha*, the only African genus of Caricaceae, is restricted to humid premontane forests in East Africa (*Cylicomorpha parviflora*; Fig. 5.2a, c) and West Africa (*C. solmsii*; Fig. 5.2b). Both its species are large trees, often growing in clusters (Cheek 2004). The other five genera occur in South and Central America. *Vasconcellea*, the largest genus of Caricaceae, comprises 20 species plus a naturally occurring hybrid, *Vasconcellea* × *heilbornii* (Badillo 1993, 2000; Van Droogenbroeck et al. 2006; Coppens d'Eeckenbrugge et al. 2013; also, Chap. 4 in this text). The genus has a center of species diversity in northwestern South America, especially Ecuador, Colombia, and Peru, with representatives in both wet and seasonally dry habitats. The second-largest genus is *Jacaratia*, with seven species of trees; it is widespread in the lowlands of the Neotropics (Fig. 5.2d). *Horovitzia* and *Jarilla*, finally, are mainly herbaceous plants of tropical seasonal forests in Mexico. *Horovitzia cnidoscoloides*, the only species in the genus *Horovitzia*, is a small, thin tree, attaining up to 6 m in height (Lorence and Colin 1988). It is restricted to the Sierra de Juaréz in northern Oaxaca, Mexico (Lorence and Colin 1988). *Jarilla* comprises three species of herbs with perennial tubers that resprout

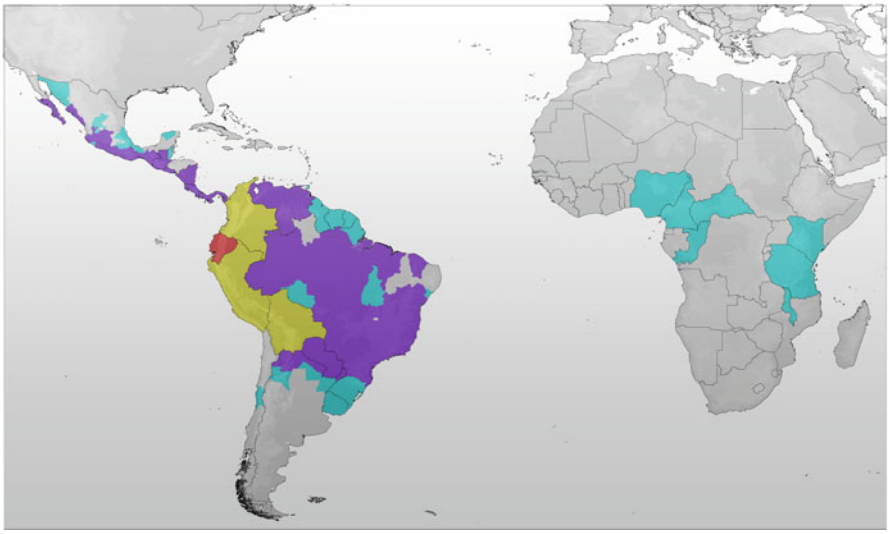

Fig. 5.1 Distribution of the family Caricaceae, except for the worldwide cultivated *Carica papaya* based on data from Badillo (1971). The different colors represent the number of species occurring in each geographic area defined by The International Taxonomic Database Working group (TDWG; http://www.kew.org/gis/tdwg/) as following: (*blue*) one species; (*purple*) 2–5, (*yellow*) 6–9 species and (*red*) >10 species

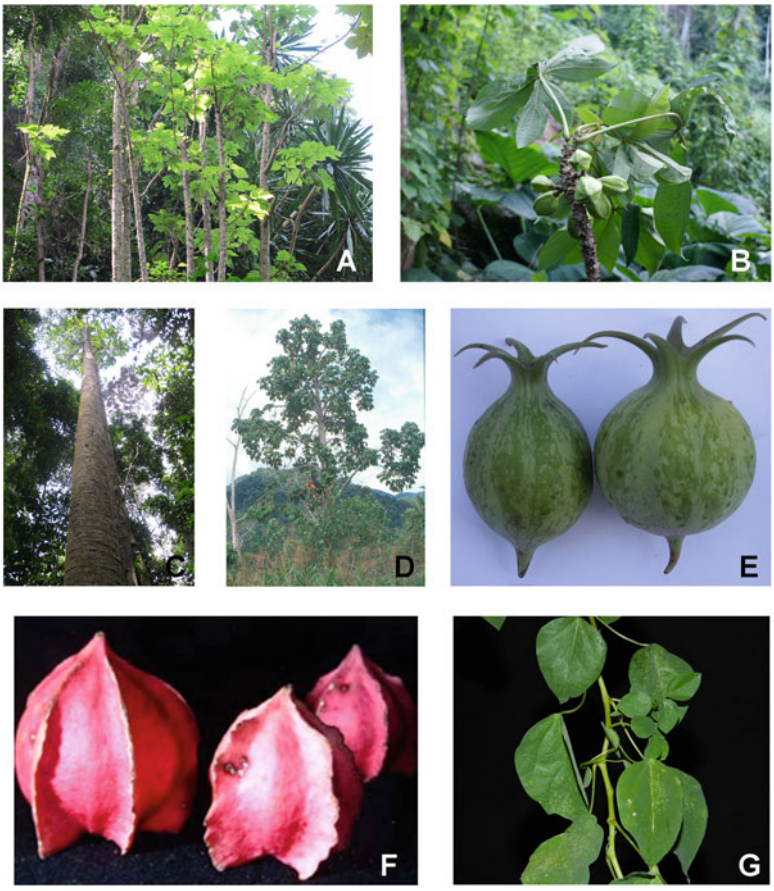

Fig. 5.2 (**a**, **c**) A natural stand and adult individual of *Cylicomorpha parviflora* from Kiangombe hill, Kenya (courtesy of M. Nicholson). (**b**) Fruits of *Cylicomorpha solmsii* from Yaoundé, Cameroon (courtesy of J.P. Gogue). (**d**) *Jacaratia digitata* with fruits (courtesy of X. Scheldeman). (**e**) Fruits of the Mexican herbs *Jarilla heterophylla* (courtesy of J. Lomelí). (**f**) Fruits of *Jarilla chocola* (courtesy of M. Olson). (**g**) Herbaceous stem of *Jarilla heterophylla*

annually (normally in June), with the shoots attaining around 1 m in height (Fig. 5.2e–g; Diaz-Luna and Lomelí Sención 1992). Flowering and fruiting occurs between July and October (Diaz-Luna and Lomelí Sención 1992). Lastly, *Carica papaya*, the only species in the genus *Carica*, is naturalized throughout the Neotropics, as well as cultivated in tropical and subtropical regions worldwide. In the Neotropics, its northern range limit lies in Florida and the southern in Paraguay (Badillo 1971). Truly wild papayas have only been found from Yucatan in Mexico south to Belize and eastern Guatemala and in Costa Rica (Manshardt and Zee 1994; Coppens d'Eeckenbrugge et al. 2007).

Alphonse De Candolle (1864) divided the Caricaceae in three genera, *Papaya, Vasconcellea* (with two sections, *Hemipapaya* and *Euvasconcellea*), and *Jacaratia*, but it was only with the work of Badillo (1971, 1993, 2000) that the classification of the family gradually attained its current form, which matches well with molecular phylogenies (see previous discussion). Studies based on molecular data began in the 1990s (Jobin-Decor et al. 1997; Aradhya et al. 1999) and quickly revealed that species traditionally included in *Carica* section *Vasconcellea* were more closely related to *Jacaratia* than to *C. papaya* (the type species of section *Carica*). In addition, Aradhya et al. (1999) suggested that the progenitor of the Central America Caricaceae (*C. papaya*, *Horovitzia, Jarilla*) might have dispersed across an island chain between South America and North America. Reacting to these molecular findings, Badillo (2000) reinstated *Vasconcellea* as a genus distinct from *Carica*, a decision supported by another molecular study that found *Vasconcellea* and *Jacaratia* forming the sister clade to *C. papaya* (Van Droogenbroeck et al. 2002).

Most species-level molecular-systematic work has concentrated on the highland papayas, *Vasconcellea* (Van Droogenbroeck et al. 2004, 2006; Kyndt et al. 2005a, b). These studies not only elucidated relationship within *Vasconcellea* but also provided evidence for natural hybridization and introgression. Kyndt et al. (2005b) found contradictions between nuclear and chloroplast and mitochondrial data (using amplified fragment length polymorphisms or AFLP, and restriction fragment length polymorphism, PCR-RFLP) and suggested possible hybrids among *Vasconcellea* species as a plausible explanation. They also found intraindividual variation in the biparentally inherited internal transcribed spacer (ITS) sequences of rDNA from a suspected hybrid tree growing in the Botanical Garden in Loja, Ecuador, and sequences from this plant placed differently in nuclear ITS and plastid *psb*A–*trn*H phylogenies (Kyndt et al. 2005a). A later study of nuclear, chloroplast, and mitochondrial markers from five *Vasconcellea* species from a suspected hybrid zone in southern Ecuador suggested two possible pathways for the origin of the hybrid *Vasconcellea* × *heilbornii* (Van Droogenbroeck et al. 2006; for more information see Coppens d'Eeckenbrugge et al. 2013; also, Chap. 4 in this text).

Because of the economic importance of papaya, this species was among the first flowering plants selected for full genome sequencing and genomic studies (Liu et al. 2004; Ming et al. 2008; Wu et al. 2010; Yu et al. 2008; Zhang et al. 2008; VanBuren and Ming 2013; also Chap. 11 in this book). Surprisingly, however, none of the phylogenetic studies prior to our own work (Carvalho and Renner 2012) included all species and genera of the Caricaceae, and the phylogenetic relationships therefore remained insufficiently understood. Most importantly, the Mexican genus *Horovitzia* had never been sequenced for any DNA locus and only one of the two African species. The complete phylogeny of the family (Fig. 5.3) now reveals that the closest relative of papaya is a group of four herbaceous species distributed only in Mexico, Guatemala and El Salvador.

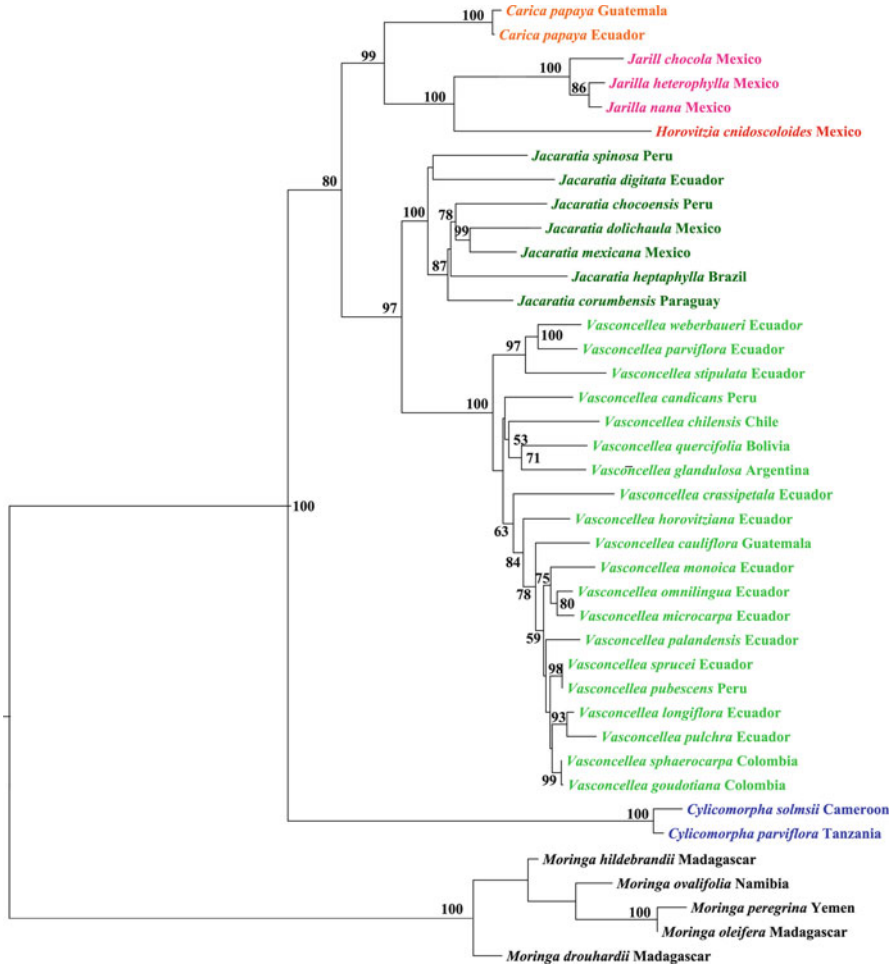

Fig. 5.3 Maximum likelihood tree for 37 accessions representing 35 species of Caricaceae based on 4,711 aligned nucleotides of nuclear (ITS region) and plastid sequences (*trn*L–*trn*F, *rpl*20–*rps*12, *psb*A–*trn*H intergenic spacers, *mat*K and *rbc*L genes). The geographic origin of each accession is shown after the species name. Only bootstrap values higher than 60 are shown (numbers on nodes) (modified with permission of Elsevier from Carvalho and Renner 2012)

A Phylogeny for All Genera and Species of the Caricaceae and Its Implications for the Origin of Papaya in Mexico/Guatemala

Based on 4,711 nucleotides of plastid and nuclear DNA, we generated a maximum likelihood phylogeny that includes 35 species (Fig. 5.3). Because of its composite genome, the analysis does not include the known hybrid *V.*×*heilbornii*. The

phylogeny shows that all genera with more than one species are monophyletic and that *Horovitzia* could be included in *Jarilla* to make the classification more informative (since a generic name is only informative about relationships if the respective genus includes more than one species).

The places of origin and domestication of papaya have been deduced either from centers of early human civilizations or from centers of Caricaceae species diversity. Classical studies on crop domestication proposed Mexico as the place of origin (De Candolle 1883; Solms-Laubach 1889; Vavilov 1940 [1992]), but others suggested northwestern South America because this area is the center of species diversity, with most *Vasconcellea* species occurring only there (Badillo 1971; Prance 1984). Of Vavilov's two centers of crop domestication in the Neotropics, papaya could have been domesticated either in the Peruvian/Bolivian center by the Incas or in the (southern) Mexican center by the Mayas. Since wild papaya has never been found in South America (Manshardt and Zee 1994), domestication by Mayan Indians in the Mesoamerican lowlands appears likely, although there is no direct archaeological evidence because papaya cannot be identified from small cell inclusions as can many other plants (D. Piperno, Smithsonian Tropical Research Institute, Panama; email to S. Renner, 17 October 2010).

The complete phylogeny for the family, combined with ancestral area reconstruction and molecular clock dating (Figs. 5.3 and 5.4; Carvalho and Renner 2012), shows that the closest relatives of *C. papaya* are *Jarilla* and *Horovitzia* genera together with four species occurring in southern Mexico and Guatemala (their names appear in Fig. 5.4), suggesting that the species *C. papaya* also evolved in Central America. A morphological trait supporting this relationship is the unilocular ovary present in all five species (including papaya), while the remaining Caricaceae have 5-locular ovaries (Badillo 1993; F. Carvalho, personal observation).

Ecological or morphological studies of wild Caricaceae other than *Vasconcellea* are extremely scarce, and breeding programs for papaya improvement have focused entirely on hybridization between papaya and highland papayas (Sawant 1958, and many later studies reviewed in Coppens d'Eeckenbrugge et al. 2013; also, Chap. 4 of this text). Given the new phylogenetic findings, crossing experiments of the herbaceous species with papaya might be worthwhile. Fruits and tubers of *Jarilla chocola* are edible and have a papain-like enzyme (Tookey and Gentry 1969); a study on the species' cultivation showed that it does not tolerate freezing and is susceptible to soil parasites (Willingham and White 1976). The other two species of *Jarilla* occur at higher altitudes (Diaz-Luna and Lomelí Sención 1992) and may be more tolerant to cold conditions. The refreshing milk and fruits (up to 18 cm × 9 cm large) of *J. heterophylla* are found in local markets in Mexico (Diaz-Luna and Lomelí Sención 1992, 1997), and a natural hybrid between this species and *J. nana* has been described from an area where they co-occur (Diaz-Luna and Lomelí Sención 1992). *Horovitzia cnidoscoloides* occurs in cloud forest in Oaxaca (at ca. 1,250 m alt.) and may be cold adapted.

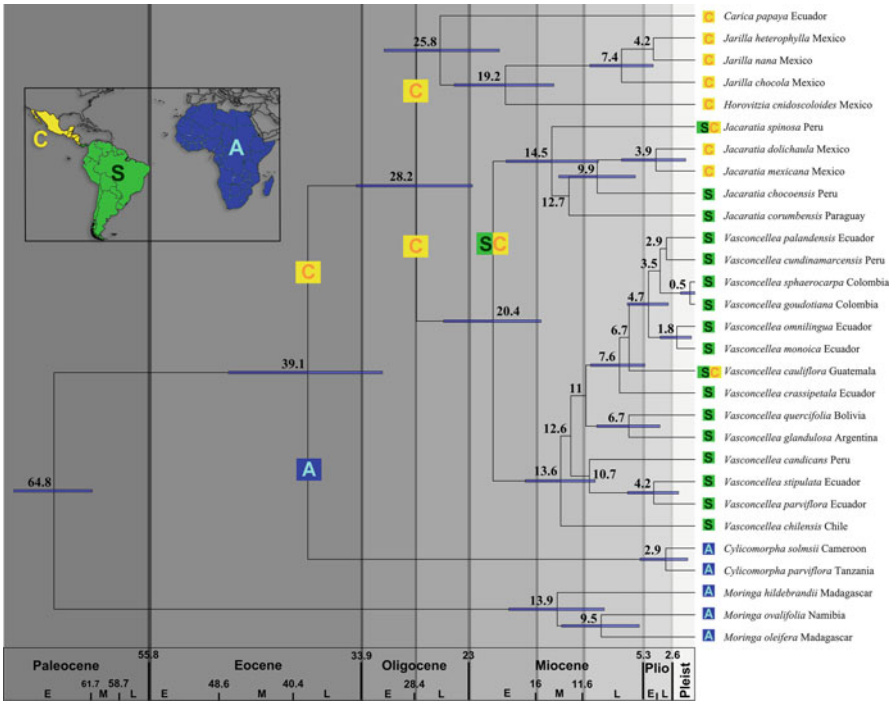

Fig. 5.4 A chronogram for the Caricaceae obtained under a strict clock model applied to two plastid genes (*mat*K and *rbc*L). The blue bars indicate 95 % posterior probability intervals (equivalent to confidence intervals). The geological time scale is in million years and follows Walker et al. (2009). The geographic origin of each accession is shown after the species name. Letters on branches represent the inferred ancestral areas, with *A* meaning Africa, *S* South America, and *C* Central America Mexico/Guatemala. Species with near-identical sequences were excluded from the clock run, since zero-length branches cause problems for molecular clock-based estimation (modified with permission of Elsevier from Carvalho and Renner 2012)

The History of Caricaceae in Africa and Their Eocene Arrival in Central America Following Dispersal Across the Atlantic

The deepest divergence in the Caricaceae is between the two African species (Fig. 5.2a–c) and the Neotropical clade including all other Caricaceae (Fig. 5.3). Molecular clock dating, combined with ancestral area reconstruction, shows that the Caricaceae originated in Africa and that the two African species (*Cylicomorpha solmsii* and *C. parviflora*) shared a most recent common ancestor during the Pliocene, around 2.8 (0.6–5.2) million years ago (Fig. 5.4). At this time, Africa was characterized by extreme climate variability with alternating periods of high moisture levels and extreme aridity (Sepulchre et al. 2006; Trauth et al. 2009). A change from

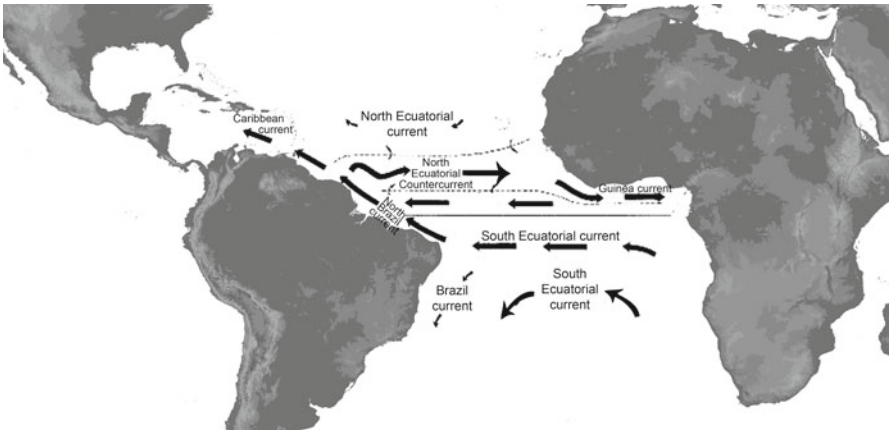

Fig. 5.5 Surface currents in the tropical Atlantic Ocean (modified with permission from Fratantoni et al. 2000. Copyright © American Meteorological Society)

wet to dry conditions occurred during the Late Pliocene, between 4 and 3 Ma (Sepulchre et al. 2006), the time when the West and East African *C. solmsii* and *C. parviflora* are inferred to have diverged from each other (Fig. 5.4). Both species are trees occurring in rainforest or along rainforest margins at 500–1,500 m high (Fig. 5.2a–c). Their modern ranges clearly result from the fragmentation of evergreen tropical forests during the Pliocene, and their divergence time matches the inferred ages of other East African and West African rainforest clades (Carvalho and Renner 2012).

Although the history of the two extant African species is relatively young, the divergence between African and Neotropical species occurred during the Late Eocene around 35 (28.1–43.1) million years ago (Fig. 5.4), long after the separation of South America and Africa at 100–90 million years. Based on statistical area reconstruction, the ancestor of the Neotropical Caricaceae apparently arrived in Central America and then dispersed from there to South America. A conceivable mechanism for the dispersal across the Atlantic is a floating island carried by ocean currents (Fig. 5.5) from the Congo delta via the North Atlantic Equatorial and Caribbean Current (Houle 1999; Fratantoni et al. 2000; Renner 2004; Antoine et al. 2011). Arrival from Africa to Central America has also been inferred for several groups of lizards that apparently were transported on rafting vegetation from the west coast of Northwestern Africa to the West Indies (Carranza et al. 2000; Vidal et al. 2008). Caricaceae have soft, fleshy fruits not suitable for water dispersal, but seeds could have been transported stuck to floating vegetation. Even if transport took several weeks, seeds might not have germinated because germination in the family is slow and erratic, which has been attributed to inhibitors present in the fleshy seed coat (Tokuhisa et al. 2007).

A molecular clock implies that Caricaceae reached South America from Central America between 27 and 19 Ma ago, which matches recent geological evidence

suggesting that the formation of the Isthmus of Panama already began 23–25 Ma ago, earlier than previously thought (Farris et al. 2011). This may have facilitated range expansion from Mexico to Colombia, where a newly established population then began to diversify and gradually to expand the family's range south to Paraguay, Uruguay, and Argentina. Mountain building in the northern Andes first peaked around 23 Ma and again around 12 Ma (Hoorn et al. 2010). It was during this period that the *Vasconcellea/Jacaratia* clade started to diversify (Fig. 5.4; around 19 Ma ago). Today, 18 out of 20 species of *Vasconcellea* occur in the northern Andean region, with 14 species found at altitudes between 750 and 2,500 m (Scheldeman et al. 2007).

The main groups of *Vasconcellea* species seen in our phylogeny (Fig. 5.3) agree with those discussed by Coppens d'Eeckenbrugge et al. (2013) (also, Chap. 4 in this text). A small clade consisting of species from the western Andes (*V. candicans*); the coastal region of Central Chile (*V. chilensis*), Peru, Bolivia, and Brazil (*V. glandulosa*); and a species from southern Peru to the north of Argentina and adjacent Brazil (*V. quercifolia*) is the sister group to a clade comprising all remaining species. The latter either have a narrow distribution in the western Andes in Peru and Ecuador or range from Ecuador to Mexico (*V. cauliflora*). A clade of *V. stipulata*, *V. parviflora*, and *V. weberbaueri* is statistically highly supported and is supposed to have formed the hybrid *V. heilbornii* (Van Droogenbroeck et al. 2006; Coppens d'Eeckenbrugge et al. 2013; also, Chap. 4 in this text). The ancestor of *V. cauliflora*, the only truly Central American species of *Vasconcellea*, originated in South America and reached its present range in Guatemala sometime after 6.3 Ma, at least based on the rough molecular clock estimates shown in Fig. 5.4.

Within *Jacaratia*, *J. spinosa* (a widespread species throughout the Neotropical region) is sister to a clade comprising the remaining species (Fig. 5.3). The two Central American species, *J. dolichaula* and *J. mexicana*, are embedded in a South American clade, implying that they reached Central America from South America (Fig. 5.3). Based on the molecular clock-inferred divergence times, *Jacaratia* diversified during the drier climates and into more open vegetation that predominated during the Late Miocene, 12–7 Ma ago (Pound et al. 2011). This would have favored species adapted to dry, open environments and seasonal forest, such as *Jacaratia corumbensis*, and *J. mexicana.* We are currently working with geo-referenced specimens to provide a better understanding of species distributions and habitat requirements, and the data are available at the cyber monograph of Caricaceae website (http://herbaria.plants.ox.ac.uk/bol/caricaceae).

References

Antoine PO, Marivaux L, Croft DA, Billet G, Ganerod M, Jaramillo C, Martin T, Orliac MJ, Tejada J, Altamirano AJ, Duranthon F, Fanjat G, Rousse S, Gismondi RS (2011) Middle Eocene rodents from Peruvian Amazonia reveal the pattern and timing of caviomorph origins and biogeography. Proc R Soc B Biol Sci 12. doi:10.1098/rspb.2011.1732

Aradhya MK, Manshardt RM, Zee F, Morden CW (1999) A phylogenetic analysis of the genus *Carica* L. (Caricaceae) based on restriction fragment length variation in a cpDNA intergenic spacer region. Genet Resour Crop Evol 46:579–586

Badillo VM (1971) Monografia de la familia *Caricaceae*, Asociación de profesores. Universidad Central de Venezuela, Maracay, 220 pp

Badillo VM (1993) Caricaceae. Segundo esquema. Rev Fac Agron Univ Central Venezuela 43:1–111

Badillo VM (2000) *Vasconcella* St.-Hil. (Caricaceae) con la rehabilitacion de este ultimo. Ernstia 10:74–79

Beilstein MA, Nagalingum NS, Clements MD, Manchester SR, Mathews S (2010) Dated molecular phylogenies indicate a Miocene origin for Arabidopsis thaliana. Proc Natl Acad Sci USA 107:18724–18728

Carranza S, Carranza S, Arnold EN, Mateo JA, López-Jurado LF (2000) Long-distance colonization and radiation in gekkonid lizards, *Tarentola* (Reptilia: Gekkonidae), revealed by mitochondrial DNA sequences. Proc R Soc B Biol Sci 267:637–649

Carvalho F, Renner SS (2012) A dated phylogeny of the papaya family (Caricaceae) reveals the crop's closest relatives and the family's biogeographic history. Mol Phylogenet Evol 65:46–53

Cheek M (2004) *Cylicomorpha solmsii*. In: IUCN 2011. IUCN Red List of threatened species. Version 2011.1. www.iucnredlist.org. Downloaded on 3 Nov 2011

Coppens d'Eeckenbrugge G, Drew R, Kyndt T, Scheldeman X (2013) *Vasconcellea* for papaya improvement. In Ming R, Moore P (eds) Genetics and genomics of papaya. Springer Science+Business Media, New York

Coppens d'Eeckenbrugge G, Restrepo MT, Jiménez D (2007) Morphological and isozyme characterization of common papaya in Costa Rica. Acta Hortic 740:109–120

De Candolle A (1864) Papayaceae. Prodromus Syst Nat Regni Veg 15:413–420

De Candolle A (1883) Origine des plantes cultivées. G Baillière et cie, Paris

Diaz-Luna CL, Lomelí Sención JA (1992) Revisión del género *Jarilla* Rusby (Caricaceae). Acta Bot Mexicana 20:77–99

Diaz-Luna CL, Lomelí Sención JA (1997) Flora de Mexico: Fanerogamas 7(1), Familia Caricaceae. Consejo Nacional de La Flora de Mexico, Mexico, 21 pp

Farris DW, Jaramillo C, Bayona G, Restrepo-Moreno SA, Montes C, Cardona A, Mora A, Speakman RJ, Glascock MD, Valencia V (2011) Fracturing of the Panamanian isthmus during initial collision with South America. Geology 11:1007–1010

Fratantoni DM, Johns WE, Townsend TL, Hurlburt HE (2000) Low-latitude circulation and mass transport pathways in a model of the tropical Atlantic Ocean. J Phys Oceanogr 30:1944–1966

Hoorn C, Wesselingh FP, Ter Steege H, Bermudez MA, Mora A, Sevink J, Sanmartín I, Sanchez–Meseguer A, Anderson CL, Figueiredo JP, Jaramillo C, Riff D, Negri FR, Hooghiemstra H, Lundberg J, Stadler T, Särkinen T, Antonelli A (2010) Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. Science 330:927–931

Houle A (1999) The origin of platyrrhines: an evaluation of the Antarctic scenario and the floating island model. Am J Phys Anthropol 109:541–559

Jobin-Decor MP, Graham GC, Henry RJ, Drew RA (1997) RAPD and isozyme analysis of genetic relationships between *Carica papaya* and wild relatives. Genet Resour Crop Evol 44:471–477

Kyndt T, Van Droogenbroeck B, Romeijn-Peeters E, Romero-Motochi JP, Scheldeman X, Goetghebeur P, Van Damme P, Gheysen G (2005a) Molecular phylogeny and evolution of Caricaceae based on rDNA internal transcribed spacers and chloroplast sequence data. Mol Phylogenet Evol 37:442–459

Kyndt T, Romeijn-Peeters E, Van Droogenbroeck B, Romero-motochi JP, Gheysen G, Goetghebeur P (2005b) Species relationships in the genus *Vasconcellea* (Caricaceae) based on molecular and morphological evidence. Am J Bot 92:1033–1044

Liu Z, Moore PH, Ma H, Ackerman CM, Ragiba M, Yu Q, Pearl HM, Kim MS, Charlton JW, Stiles JI, Zee FT, Paterson AH, Ming R (2004) A primitive Y chromosome in papaya marks incipient sex chromosome evolution. Nature 427:348–352

Lorence DH, Colin RT (1988) *Carica cnidoscoloides* (sp. nov) and sect. *Holostigma* (sect. nov.) of Caricaceae from Southern Mexico. Syst Bot 13:107–110
Manshardt RM, Zee FTP (1994) Papaya germplasm and breeding in Hawaii. Fruit Varieties J 48:146–152
Ming R et al (2008) The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). Nature 452:991–996
Olson ME, Rosell JA (2006) Using heterochrony to detect modularity in the evolution of stem diversity in the plant family Moringaceae. Evolution 60:724–734
Pound MJ, Haywood AM, Salzmann U, Riding JB, Lunt DJ, Hunter SJ (2011) A Tortonian (Late Miocene, 11.61–7.25 Ma) global vegetation reconstruction. Palaeogeogr Palaeoclimatol Palaeoecol 300:29–45
Prance GT (1984) The pejibaye, *Guilielma gasipaes* (H.B.K.) Bailey, and the papaya, *Carica papaya* L. In Stone D (ed) Pre-Columbian plant migration. Papers of the Peabody Museum of Archaeology and Ethnology, vol 76. Harvard University Press, Boston, pp 85–104
Renner S (2004) Plant dispersal across the tropical Atlantic by wind and sea currents. Int J Plant Sci 165:S23–S33
Sawant AC (1958) Crossing relationships in the genus *Carica*. Evolution 12:263–266
Scheldeman X, Willemen L, Coppens d'Eeckenbrugge G, Romeijn-Peeters E, Restrepo MT, Romero Motoche J, Jiménez D, Lobo M, Medina CI, Reyes C, Rodríguez D, Ocampo JA, Damme P, Goetgebeur P (2007) Distribution, diversity and environmental adaptation of highland papayas (Vasconcellea spp.) in tropical and subtropical America. Biodivers Conserv 116:1867–1884
Sepulchre P, Ramstein G, Fluteau F, Schuster M, Tiercelin JJ, Brunet M (2006) Tectonic uplift and Eastern Africa aridification. Science 313:1419–1423
Solms-Laubach (1889) Die Heimat und der Ursprung des kultivierten Melonenbaumes, *Carica papaya* L. Botanische Zeitung 44:709–720
Stevens PF (2001 onwards) Angiosperm phylogeny website. Version 9, June 2008 continuously updated since. http://www.mobot.org/MOBOT/research/APweb/
Tokuhisa D, Cunha D, Dos F, Dias S, Alvarenga EM, Hilst PC, Demuner AJ (2007) Compostos fenólicos inibidores da germinação de sementes de mamão (*Carica papaya* L.). Rev Brasil Sementes 29:161–168
Tookey HL, Gentry HS (1969) Proteinase of *Jarilla chocola*, a relative of papaya. Phytochemistry 8:989–991
Trauth MH, Larrasoaña JC, Mudelsee M (2009) Trends, rhythms and events in Plio-Pleistocene African climate. Quat Sci Rev 28:399–411
Van Droogenbroeck BV, Breyne P, Goetghebeur P, Romeijn-Peeters E, Kyndt T, Gheysen G (2002) AFLP analysis of genetic relationships among papaya and its wild relatives (Caricaceae) from Ecuador. Theor Appl Genet 105:289–297
Van Droogenbroeck BV, Kyndt T, Maertens I, Romeijn-Peeters E, Scheldeman X, Romero-Motochi JP, Van Damme P, Goetghebeur P, Gheysen G (2004) Phylogenetic analysis of the highland papayas (Vasconcellea) and allied genera (Caricaceae) using PCR-RFLP. Theor Appl Genet 108:1473–1486
Van Droogenbroeck BV, Kyndt T, Romeijn-Peeters E, Van Thuyne W, Goetghebeur P, Romero-Motochi JP, Gheysen G (2006) Evidence of natural hybridization and introgression between Vasconcellea species (Caricaceae) from southern Ecuador revealed by chloroplast, mitochondrial and nuclear DNA markers. Ann Bot 97:793–805
VanBuren R, Ming R (2013) Sequencing and assembly of the transgenic papaya genome. In: Ming R, Moore P (eds) Genetics and genomics of papaya. Springer Science+Business Media, New York
Vavilov NI (1992) Origin and geography of cultivated plants. Translated by D Löve. Cambridge University Press, Cambridge (English translation, original published 1940 in Russian)
Vidal N, Azvolinsky A, Cruaud C, Hedges SB (2008) Origin of tropical American burrowing reptiles by transatlantic rafting. Biol Lett 4:115–118
Walker JD, Geissman JW, Compilers (2009) Geologic time scale: geological society of America. Available at http://geology.gsapubs.org/cgi/doi/10.1130/G22967A.1. Accessed 5 July 2011

Willingham BC, White GA (1976) Agronomic evaluation of prospective new crop species. V. *Jarilla chocola* – a proteinase source. Econ Bot 30:189–191

Wu X, Wang J, Na J, Yu Q, Moore RC, Zee F, Huber SC, Ming R (2010) The origin of the non-recombining region of sex chromosomes in Carica and Vasconcellea. Plant J 63:801–810

Yu Q, Navajas-Pérez R, Tong E, Robertson J, Moore PH, Paterson AH (2008) Recent origin of dioecious and gynodioecious Y chromosomes in papaya. Trop Plant Biol 1:49–57

Zhang W, Wang X, Yu Q, Ming R, Jiang J (2008) DNA methylation and heterochromatinization in the male-specific region of the primitive Y chromosome of papaya. Genome Res 18:1938–1943

Part II
Classical and Molecular Genetics and Breeding

Chapter 6
History and Future of the Solo Papaya

Richard Manshardt

Botanical Background

Taxonomy

The papaya (*Carica papaya* L.) is the most important crop species in the Caricaceae, a plant family consisting of five New World genera and one African genus. In 2000, the largest genus, *Carica*, was split to better reflect molecular and morphological differences between a group of 21 predominantly South American species, now assigned to genus *Vasconcellea*, and the single remaining *Carica* species, *C. papaya,* which has a native range restricted to Central America (Badillo 2000, 2001).

Origin and Domestication

The homeland of the typical wild papaya is limited to the region extending from southern Mexico to perhaps as far south as northern Costa Rica (Manshardt and Zee 1994). Wild papayas are dioecious, and the fruits borne on the pistillate plants are small (<100 g) and seedy, with very little edible flesh. During the domestication of wild papayas, human selection was focused on enlarging the fruit size and increasing the succulence of the ovary wall, leading to present day dioecious land races like "Cera" in Veracruz, Mexico, with fruit weight typically in excess of 1 kg. An allele for hermaphroditism, rare in wild populations, was exploited by early domesticators to create genetically stable pure lines by inbreeding, and other mutants with red

R. Manshardt (✉)
Department of Tropical Plant and Soil Sciences, University of Hawaii at Manoa, St. John Plant Science Lab, Rm 102, 3190 Maile Way, Honolulu, HI 96822, USA
e-mail: manshard@hawaii.edu

R. Ming and P.H. Moore (eds.), *Genetics and Genomics of Papaya*, Plant Genetics and Genomics: Crops and Models 10, DOI 10.1007/978-1-4614-8087-7_6,

flesh color were favored over the orange color typical of wild plants. These gave rise to gynodioecious land races like “Roja” of Central America that segregate for hermaphrodites and females and have more elongate fruit shapes. Papayas became common in dooryard gardens because of their short juvenile period of less than 6 months, continuous bearing habit, and plentiful attractive fruits.

Geographic Dispersal

Seeds of wild plants have long viability and marked dormancy, but the latter has been mostly lost in domesticates. Papayas are preadapted to agriculture, since they fill an ecological niche as a colonizing species in areas disturbed by human or other disruptive activities (Bartlett 1937; Lundell 1936). These traits have facilitated dispersal of papaya by human travelers and explorers since 1500 to tropical regions around the world. Papaya probably arrived in SE Asia via the Spanish who operated a lucrative trade route between Acapulco, Mexico, and Manila in the Philippine Islands for 250 years after the discovery of the New World (Hayes 2001). It was rapidly spread to India, Indonesia, Malaysia, and Africa by other European colonial powers, including the Portuguese, Dutch, British, and French. Selection in Asia and Africa has produced numerous local papaya land races and cultivars in those tropical regions. Around the world today, quality of fruit is variable, with many genotypes having low total soluble solids (TSS) and/or objectionable flavor or texture, but large fruit size seems generally to have been a selection priority of value to locally based subsistence agriculture.

Early Introductions in Hawaii

Dioecious and Gynodioecious Land Races

Papaya arrived relatively late in Hawaii. Its introduction is attributed to Don Francisco de Paula Marin, a legendary Spanish adventurer and arms purveyor to Kamehameha I, who was also an avid horticulturist in the early 1800s in pre-colonial Hawaii. While there are no records of earlier contacts, there is a possibility that Spanish galleons straying off the normal trade routes to or from the Orient may have brought papaya seed to Hawaii, as there are legends of large ships and uniformed men disembarking before Captain Cook’s first visit in 1778 (Kane 1996). Regardless of the origin of their conveyors, the first description of papayas in Hawaii is of large-fruited dioecious types, presumably from the western coast of Mexico, followed later by large-fruited hermaphrodites. There was little organized effort to select or improve papayas in Hawaii until early in the era of US territorial governance. Higgins and Holt (1914) made an important contribution in their bulletin titled

"The Papaya in Hawaii." At that time, they stated that dioecious papayas were more common than gynodioecious types and that no true-breeding cultivars existed, in the sense that qualities of the offspring could not be reliably predicted from those of the parental selections. They emphasized the advantages of inbreeding and growing pure lines of gynodioecious papayas in which both segregating sexes would be fruit-bearing, and the quality of the fruit would be predictable. Consequently, as employees of the Hawaii Agricultural Experiment Station, they bent their efforts toward developing inbred lines by self-pollinating selected hermaphrodites. In addition, Higgins and Holt undertook some of the first experiments to test feasibility of shipping to export markets in Portland, Seattle, and Vancouver (in 1914, California markets were already under fruit fly quarantine). They noted that an advantage of shipping hermaphrodite fruits was that the more cylindrical shape of the fruit lent itself to more efficient packaging than the spherical fruits of female plants. However, the gynodioecious papaya germplasm available at the time had two major shortcomings with regard to the fruit criteria necessary for export markets. First, the quality of most early Hawaiian materials was not exceptional, and in fact, it led some critics to refer to them as food fit only for pigs (Crawford 1937). In addition, the fruits were relatively large, judging by photos from their bulletin, making the price per fruit excessively high in West Coast markets.

Arrival of the Solo Papaya

Three years before Higgins and Holt published their work, an event occurred that would have a major impact in subsequent years on the trajectory of the papaya industry in Hawaii. This was the introduction in 1911 of seed of a small papaya from the Caribbean island of Barbados by Gerrit P. Wilder (Storey 1941). Wilder was the son of an important *kama'aina* family that had developed rail transportation for sugarcane processors on the Islands. He spent the majority of his working life in the family business, but he was also an enthusiastic and competent horticulturist and botanist. After his retirement, he devoted himself to his avocation and made many contributions to Pacific science as botanist at the Bishop Museum in Honolulu (University of Hawaii 2012). Probably, none of his other efforts had as large an impact on Hawaii's economy as his collection and introduction of the little fruit that was the progenitor of the solo papaya.

Characteristics of Solo Papayas

Solo is not the name of a specific cultivar; rather, it refers to the general class of export quality, gynodioecious papayas having pear-shaped fruits weighing about 450–675 g, yellow or red flesh color, TSS in the 12–15 % range, and superior flavor characteristics. The name was given by J.E. Higgins (HAES 1920) and was said to derive from HAES personnel of Puerto Rican descent, who differentiated the small-fruited papayas that could be consumed by a single person (solo) from the large

"watermelon" types that could feed a family. Small fruit size was as important as the excellent flavor and texture traits in establishing the solo as the export standard, since it kept the cost of individual fruits, sold on a cost/lb basis, acceptable in mainland US markets.

Germplasm Deficiencies and Genetic Solutions

Sex Segregation

For many people in Hawaii, the solo papaya and its derivatives provided a first experience of a desirable papaya fruit with immediate taste appeal, and it revealed the potential for a new fruit export industry in Hawaii. However, along with the adoption of gynodioecious solo lines as the preferred commercial production model came several significant production problems that required attention from geneticists and breeders. The occurrence of both pistillate and hermaphrodite sexes among seedling progenies of gynodioecious lines, and the inability to distinguish these using purely vegetative characters, led growers to question how to establish a field of seedlings with the maximum number of commercially desirable hermaphrodite plants.

The phenomenon of sex segregation in papaya had been noted by all early researchers, including Higgins and Holt (1914), Wilcox (1916), and Willard Pope, who published "Papaya Culture in Hawaii" (1930). However, it was William B. Storey who first provided experimental evidence to explain papaya sex segregation. Storey was born in Hawaii and was educated at Cornell University. He returned to become a horticulturist at the University of Hawaii (UH). Over a period of ten years from 1936 to 1945, Storey (1938b) and Hofmeyr (1938), who worked independently on the same topic in South Africa, published a series of articles that showed that sex segregation was controlled by a single Mendelian locus with three alleles. The male and hermaphrodite states are determined by different dominant alleles (M^m and M^h, respectively), and these sexes are genetically heterozygous, sharing the locus with the recessive allele (m) that in homozygous condition determines the female state. A sex-linked lethal gene prevents the formation of homozygous M^m/M^m or M^h/M^h or M^m/M^h genotypes. With this genetic model, Storey worked out the expected segregation ratios for various matings of the different sexes. The theoretical sex segregation ratios allowed efficient and predictable establishment of commercial papaya fields based on calculations of the minimum number of seedlings required per planting hole to achieve any desired probability of sex uniformity in the field (Jones and Storey 1941). The standard planting procedure now involves planting each hole with three seedlings from a self-pollinated hermaphrodite parent, followed several months later at flowering time by roguing of females and extra hermaphrodites to yield a field with hermaphrodites in ~96 % of the holes.

Fruit Carpellody and Carpel Abortion

A second production problem that accompanied the commercial exploitation of gynodioecious lines was the sensitivity of many hermaphrodite genotypes to fruit deformity caused by stamen carpellody. This is a genetic proclivity affecting floral development in hermaphrodites, such that stamens become carpel-like and attach to the ovary in irregular or occasionally symmetrical lobes (Storey 1938a, 1941), particularly during the cool, wet season. This causes distorted growth resulting in production of unattractive and unmarketable fruit. The opposite tendency toward abortion of carpels from the ovary in some hermaphrodite genotypes results in fruits with mango or banana shapes or total loss of the ovary leading to unproductive zones in the fruit column. Carpel abortion is usually most pronounced under warm, dry conditions. These environmental influences on fruit morphology are unique to hermaphrodites, but the tendencies are under genetic control, and it is possible to select effectively against these negative characteristics.

Early Improvement Objectives and Accomplishments

An early description of papaya production in the United States, focusing primarily on Florida, listed six named cultivars there, including Hawaiian solo (Traub et al. 1942). Writing in Hawaii at the same time, Storey (1941) echoed the observations of Higgins and Holt a quarter century earlier that there were no true-breeding cultivars in Hawaii, due to the tendency to outcross among types unless grown in isolation or intentionally self-pollinated. Named types were usually descriptive of economically important morphologies, such as the small pyriform-fruited "solo" or large red-fleshed "watermelon" papayas, or they were named after the farmer that produced them. Storey pointed out the desirability of a more uniform and predictable crop and set about developing such by stabilizing the genetics through inbreeding selected hermaphrodites with good fruit qualities and production characteristics. The objectives that he and subsequent breeders have generally identified for improvement included fruit weight of 450–675 g, high TSS in the range of 12–15 %, good flavor and texture, minimal seasonal variation in fruit shape caused by stamen carpellody or carpel abortion, and early flowering leading to fruit production low on the trunk. Working primarily at the Waimanalo Experiment Station on windward Oahu, he released Line 5 in 1948 and Line 8 in 1953. Line 8 is a yellow-fleshed cultivar with flavor characteristics generally acknowledged to be excellent, but fruits have the unusual property of ripening internally before the skin has lost its green color. Consequently, determining the proper time to harvest is problematic, and if removed from the tree at color-break like other papayas, they are too soft for export. It was still grown to a limited extent on Oahu until the early 1990s.

Growth of Export Industry

By the mid-1950s, papaya production reached 4.5 million kg annually and became the largest component of the diversified crops sector (crops other than sugar and pineapple) in terms of gross production, and it was third in value behind coffee and tomatoes (Hawaii Agricultural Extension Service 1958). However, papaya was nearly exclusively marketed locally. The large growth in the industry over the subsequent 25 years was due to development of export markets in North America and Japan. The shift in marketing was made possible by several advances, which overcame technical problems or allowed improvements in logistical capacity.

Fruit Fly Disinfestation

The initial barrier to papaya shipments to California was the quarantine restriction imposed in 1914 to exclude fruit flies. Although color-break papaya fruits are not a major host for fruit flies, compliance with quarantine regulations required a disinfestation procedure. An existing vapor heat treatment was modified for papaya by UH plant physiologists (Jones 1940a), and subsequent experimental shipping and testing in mainland markets led to the first commercial shipments in 1940. That same year, methyl bromide was approved for fruit fly disinfestation (Jones 1940b), followed by ethylene dibromide (EDB) in 1951. EDB treatment became the standard disinfestation method until 1984, due to its relative low cost and freedom from adverse effects on fruit quality.

The "Kapoho" Papaya

Another step in the evolution of the papaya industry toward export markets resulted from the confluence of several factors that effectively moved the major production areas from Oahu to Hawaii Island. Increasing urbanization and rising land values in the 1950s and 1960s, combined with repeated outbreaks of papaya ringspot virus (PRSV), served to make papaya production increasingly problematic on Oahu. The area of the state with cheap land available for agriculture was the Puna District in East Hawaii. The chief agricultural employer of the area, Puna Sugar Company, was in decline through the 1970s and closed in 1982, gifting laid-off workers with 5-acre land parcels and creating a pool of new farmers looking for profitable crops (UH 2006). This region became the new center of the papaya industry. Puna has abundant rainfall, well distributed throughout the year, but it is a volcanically active region with geologically young lava substrate and little soil development. It was not clear that papaya could be grown successfully under Puna conditions, but in fact a cultivar specifically adapted to Puna had been under selection in the region for several decades. Hanichi Masumoto was a farmer with land near Hilo, who requested seed from the Cooperative Extension Service to grow papayas. He was initially

Fig. 6.1 "Kapoho." Its excellent postharvest characteristics and specific adaptation to raw volcanic soils and rainforest conditions made East Hawaii Island the center of production for export markets (courtesy of Brian Sato for Hawaii Papaya Industry Association)

given "pig food" papaya seed, since the extension agent handling the request knew that Masumoto raised pigs, and he assumed that was the intended use for the fruit. Anecdotal information suggests that Masumoto, upon discovering the poor quality of the fruits, made his displeasure abundantly clear and was quickly provided with solo papaya seed, from which he gradually refined his selection. It was initially known as "Masumoto" solo but was renamed "Kapoho" as it became more widely planted in the Puna District (Hamilton and Ito 1986).

"Kapoho" has several characteristics which destined it for success as the standard export cultivar from Hawaii. Fruit quality is excellent, and it is free of carpellody and carpel abortion (Fig. 6.1). "Kapoho" fruits remain firm after harvest, which makes them well suited for packing and shipment to distant export markets. Under the high rainfall conditions in Puna, the fruit size is the preferred 450 kg, but grown anywhere else, the combination of different soil and climatic factors tend to produce very tall, robust trees and fruits that are too small for export. For this reason, Puna is the only area of the state that can benefit from the other desirable qualities of "Kapoho." The combination of land availability, unique environmental conditions, and a papaya cultivar adapted to those conditions conspired to favor the Puna District as the new home of the Hawaii papaya industry after its decline on Oahu during the decades of the 1950s and 1960s (Loudat et al. 1987).

Tourism Connection

Another factor contributing to and enabling the growth of papaya as an export crop was the concurrent development of the tourism industry in Hawaii. In the years

following World War II, increasing numbers of visitors from the mainland and later from Japan travelled to Hawaii for their first encounter with tropical environments. Exposure to papaya in restaurants and in produce sections of local groceries amounted to free advertising for the industry and resulted in a growing demand for the fruit in temperate regions when tourists returned to their homes. Moreover, tourists increasingly arrived in the Islands by air, so that starting in the 1960s, air freight provided an alternative to sea shipment with an accompanying improvement in quality at market destinations. The growth in papaya exports from 1960 to 1986 closely tracked the increase in air traffic from the mainland to Hawaii, and in 1969, the first air shipments to Japan commenced (Loudat et al. 1987).

Cultivar Diversification

"Sunrise" and "Sunset"

New cultivars were developed and introduced in the 1960s and 1970s by breeders at the UH College of Tropical Agriculture and Human Resources (CTAHR). All earlier commercial releases had produced yellow- or orange-fleshed fruits, including the industry standard cultivar "Kapoho." In 1963, Richard A. Hamilton and Philip Ito selected the red-fleshed cultivar "Sunrise" (Fig. 6.2), derived from a cross between

Fig. 6.2 "Sunrise," "Sunset," "SunUp." These cultivars all share pink-tinged flesh with a melting texture. "SunUp" is a genetically engineered version of "Sunset" with resistance to papaya ringspot virus (courtesy of Brian Sato for Hawaii Papaya Industry Association)

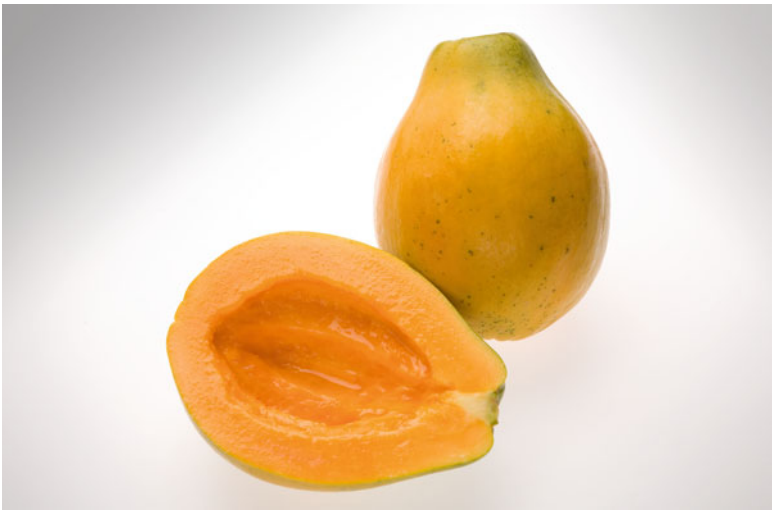

Fig. 6.3 "Waimanalo," "Kamiya." Larger fruited (up to 1 kg) and tolerant to *Phytophthora* root and fruit rots, these closely related cultivars are narrowly adaptated to the windward eastern shore of Oahu (courtesy of Brian Sato for Hawaii Papaya Industry Association)

the red-fleshed Line 9 and a yellow-fleshed solo breeding line called "Kariya" (Hamilton and Ito 1968). "Sunrise" is a high-quality papaya with fruit weights greater than "Kapoho" under the same growing conditions. It is free of carpellody and usually has only a brief sterile period due to carpel abortion in warm weather. "Sunrise" flesh softens to a greater extent than that of "Kapoho" and is not as well suited for export from Hawaii, although it became the basis of Brazil's European export industry in the 1980s and seems to be adaptable to a broader range of environments than "Kapoho" (Hamilton and Ito 1986). Since its release, "Sunrise" has been popular among breeders as a parental line in the development of papaya F1 hybrid cultivars, including "Tainung No. 2" from Taiwan and "Exotica" from Malaysia, because of its high TSS, excellent flavor, red flesh color, and lack of carpellody. A sib line of the same parental cross was released later under the name "Sunset," but its characteristics are very similar to "Sunrise" (Hamilton et al. 1993). "Sunset" has been planted extensively in Brazil and to some extent has replaced "Sunrise" there as the chief export cultivar.

"Waimanalo" and "Kamiya"

About the same time, the experimental line X-77 was released by CTAHR horticulturist Henry Y. Nakasone as the papaya cultivar "Waimanalo" (Fig. 6.3) (Nakasone et al. 1972). This cultivar was derived from a cross made in 1948 between solo

Line 5 and a dwarf line from Florida called "Betty" with the intent to create a high-quality papaya with a precocious, low-bearing habit. "Waimanalo" produces larger yellow-fleshed fruits weighing 675–900 g and of a more spherical shape than "Kapoho" or "Sunrise"/"Sunset," and it is modestly lower bearing than either. It is notable for its relative resistance to the serious fungal pathogen *Phytophthora palmivora* and has served as a parent in breeding programs to improve *Phytophthora* resistance in other lines. On the negative side, it has a rather narrow adaptation to conditions on windward Oahu and tends to be subject to fruit disfigurement due to stamen carpellody in other locations. A selection from "Waimanalo" called "Kamiya" was made by Ken Kamiya, a farmer on Oahu's windward coast. It is very similar to the parent line and has been popular in Honolulu markets.

"Kapoho" × "Sunrise" Hybrid

In 1984, the Environmental Protection Agency disallowed use of EDB for fruit fly disinfestation of papaya, and this initiated a switch from chemical to physical postharvest treatments that evolved through multiple steps over several decades. Almost immediately, the heat treatment substitutes for EDB were found to be inadequate. Mainland exports were interrupted several times in 1987 when live fruit fly larvae were discovered in papaya shipments by California quarantine inspectors, precipitating an urgent reexamination of quarantine protocol. Francis Zee, curator of the newly opened USDA National Clonal Germplasm Repository in Hilo, made the observation that carpels at the style end of the ovary occasionally fail to fuse completely, allowing a route for fruit fly oviposition directly into the central seed cavity, thereby evading the superficial hot water treatment (Zee et al. 1989). In the years leading up to 1990 when effective high-temperature forced-air disinfestation protocols became available, Zee offered hybrids between "Kapoho," which manifested the problem most often, and "Sunrise," which had a more elliptical fruit shape with better carpel fusion, to the industry to reduce the probability of "blossom-end defect." This marked the first use of an F1 hybrid cultivar in Hawaii, and the unnamed hybrid and derived inbred lines were marketed for some years by a local packing company supplying Japan.

Transgenic "SunUp" and "Rainbow"

In 1992, the perennial problem of PRSV arrived at the main production areas in Puna, about 30 years after it had destroyed most of the state's production, at that time, on Oahu. The geographical isolation that had protected production fields in eastern Puna from PRSV had been breached over this time period by the infringement of housing developments and their associated backyard papaya plants. The damage started slowly enough, but by 1998, papaya production in Puna had

dropped by 50 % from levels of the late 1980s and early 1990s (NASS 1999). Research to solve the PRSV problem had been ongoing since the 1970s and 1980s, involving screening papaya germplasm for resistant lines, employing cross-protective mild-symptom virus strains (Mau et al. 1989; Yeh and Gonsalves 1984), and investigating wide crosses with PRSV-resistant wild relatives (Manshardt and Wenslaff 1989; Mekako and Nakasone 1975). None of these approaches proved successful in delivering economical protection or a resistant cultivar. By the mid-1980s, genetic engineering technology had advanced to the point that it was possible to conceive a plan to provide Hawaii's papaya cultivars with PRSV resistance by this approach. Beginning in 1987, a team of scientists under the leadership of a Cornell University virologist, Dennis Gonsalves, provided the needed skills to accomplish the task. Gonsalves, who was born in Hawaii and was familiar with the PRSV problem, identified and isolated the resistance gene from the PRSV genome itself. Jerry Slightom of the Upjohn Company engineered the gene into a functional transformation vector. Maureen Fitch of the UDSA's Sugarcane Research Lab in Aiea, Oahu, developed the tissue cultures for transformation and regeneration of the genetically engineered plants as part of her doctoral research at UH at Mānoa. Confirmation of the efficacy of PRSV resistance under field conditions in Hawaii was the contribution of Richard Manshardt and Stephen Ferreira of UH at Mānoa, aided by cooperating growers Delan and Jenny Perry of Kapoho, Hawaii (Lius et al. 1997; Ferreira et al. 2002). Ironically, "SunUp," the first "transgenic" papaya with successful PRSV resistance (Fitch et al. 1992), was rejected by the papaya industry, because it was the wrong color. Marketing of Hawaiian papayas was based on the yellow flesh color of the standard "Kapoho" cultivar, and "SunUp," a genetically engineered version of the existing "Sunset" cultivar, was pink-fleshed. This impediment was overcome by a conventional sexual cross between "Kapoho" and "SunUp" to yield the yellow-fleshed, PRSV-resistant F1 hybrid named "Rainbow" (Fig. 6.4) (Manshardt 1998). These names were suggested provisionally by UH Horticulture Department chair H.C. Bittenbender as symbolically hopeful for papaya growers after the PRSV "storm," and they stuck, becoming the official cultivar appellations. In addition to its resistance to PRSV, the resulting hybrid was more highly productive and more widely adapted to microclimatic variation than its "Kapoho" parent but also somewhat more susceptible to *Phytophthora* fruit rot, stem canker, and root rot. When released in 1998, "Rainbow" and "SunUp" became the world's first genetically engineered tree fruit cultivars to reach commercial production and the first transgenic cultivars to be released by public institutions in the USA (Gonsalves 1998). "Rainbow" was rapidly adopted by growers in the Puna District, and by 2009, plantings of "Rainbow" accounted for about 75 % of commercial papaya acreage in Hawaii (NASS 2009) (Fig. 6.5). Several other PRSV-resistant papaya hybrids have been produced by conventional crosses with the transgenic cultivars subsequent to their 1998 release. The most important of these, released by USDA plant physiologist Maureen Fitch in 2002, is the cross of "Kamiya" with a transgenic inbred derived from "Rainbow." The resulting hybrid, called "Laie Gold," is a high-quality, yellow-fleshed fruit that is popular with growers along the windward eastern coast of Oahu.

Fig. 6.4 "Rainbow." This F1 hybrid combines the major desirable qualities of "Kapoho" (shipping durability, orange flesh color) with resistance to PRSV from genetically engineered "SunUp." "Rainbow" is adapted to a wider range of environments in Hawaii than "Kapoho" (courtesy of Brian Sato for Hawaii Papaya Industry Association)

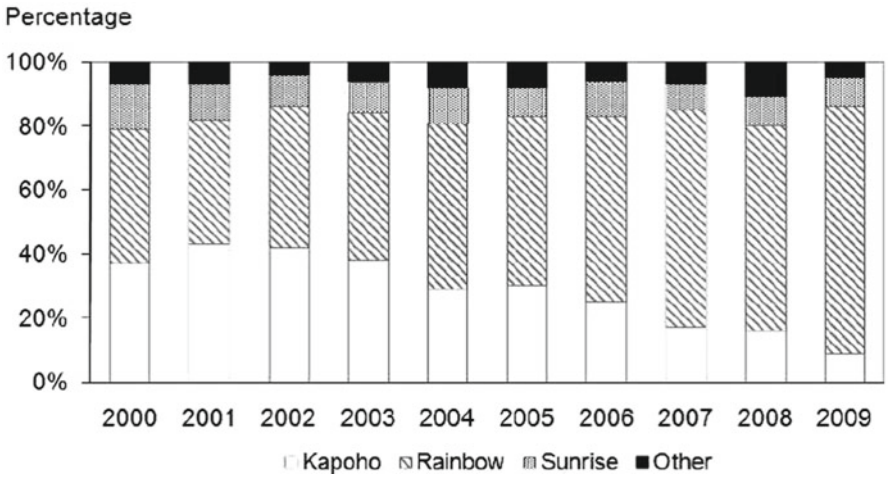

Fig. 6.5 Percentage of total crop area in Hawaii by cultivar from 2000 to 2009, showing the rapid adoption of PRSV-resistant "Rainbow" by growers [from National Agricultural Statistics Service (2009) Hawaii Papayas. Monthly. October 17, 2009. http://www.nass.usda.gov/Statistics_by_State/Hawaii/Publications/Fruits_and_Nuts/papaya.pdf]

Breeding for the Future

Economic Realities

The industry today is leaner than in the last decades of the twentieth century. Market share in the US mainland was lost to international competitors in Central and South America during the drop in Hawaiian production caused by PRSV in the 1990s, and that loss was not regained subsequent to introduction of PRSV-resistant varieties (Figs. 6.6 and 6.7). Imports of the large Mexican "Maradol" papayas into the USA have risen dramatically since 1990 but are less damaging competition for Hawaii than other producers of solo papayas, such as Belize and Brazil. Reductions also occurred in lucrative exports to Japan as a result of Japanese quarantine restrictions against shipments of transgenic papaya cultivars, although these were recently lifted (January 2012). Statistics from the late 1980s indicate that more than 300 papaya farms statewide harvested about 975 ha of papaya, compared to 177 farms and 535 ha in 2009 (NASS 2009). In recent years, Hawaii has produced about 13.5 million kg of fruit annually for fresh consumption, about half the amount produced in the mid- to late 1980s, and annual crop value has averaged about $14 million.

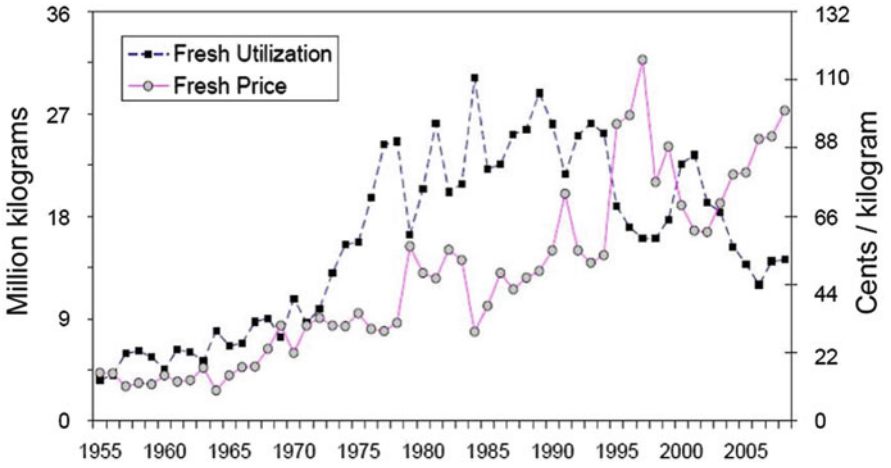

Fig. 6.6 Total utilization and grower price per kilogram of fresh papaya in Hawaii from 1955 through 2008, showing recent declines in production due to disease and foreign competition (from National Agricultural Statistics Service. Hawaii Papayas. Monthly. September 17, 2009. http://www.nass.usda.gov/Statistics_by_State/Hawaii/Publications/Archive/xpap0709.pdf)

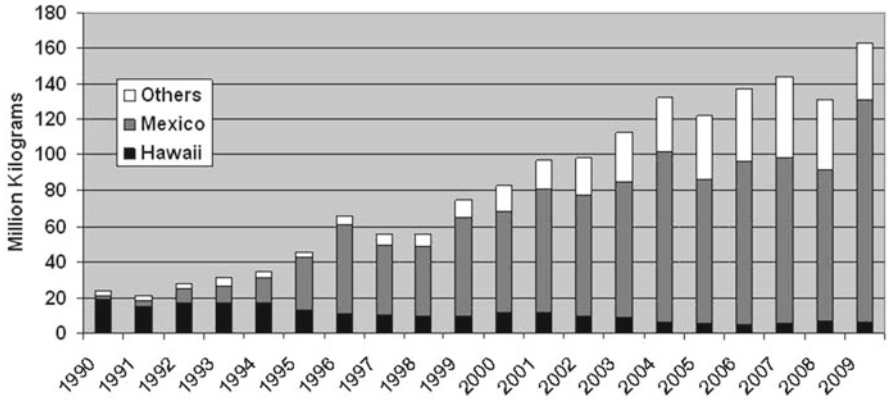

Fig. 6.7 Origin and magnitude of fresh papaya imports into the US mainland from 1990 through 2009, showing the impact of strong foreign competition on Hawaii's share of the US market. "Others" category consists mainly of Belize, Brazil, Guatemala, and Dominican Republic (data from National Agricultural Statistics Service)

Reconfiguring Plant Architecture

In order for Hawaii's papaya growers to remain competitive in world markets, solo cultivars need to achieve a more ideal phenotype for multiyear commercial production. Some of its needs can be met by standard plant breeding objectives, such as improved disease resistance, while others require a more radical restructuring of the plant's genome or energy allocation during development.

More productive genotypes will require that the plant architecture of the solo papaya be moved away from its present position near that of the wild phenotype, toward one that is more suitable for an agricultural setting. Wild papayas in Central America are members of a secondary successional community that rapidly replaces abandoned agricultural plots. A dormant papaya seed bank deposited by birds quickly exploits any opening in the forest canopy and joins the community of early opportunists pioneering disturbed areas. Stands of wild papayas can be rather dense and weedy. Papayas are adapted to these conditions, in that the plant's resources during its early development are channeled into vertical growth enabling it to overtop neighboring competitors. Delayed flowering is also adaptive in the wild to minimize processes detracting from this upward race. Once the plant reaches a sufficient aerial size and root mass, the reproductive mode is initiated, and flowering and fruit development follows continuously. Filling the reproductive sink represented by the developing seeds in the fruit column then becomes the dominant function of the remainder of the plant's life cycle, sometimes to the detriment of maintenance of canopy and root systems, which seem to be secondary sinks. Any resources the plant encounters during the juvenile phase go primarily to increase height and, after flowering, to increase the number of fruits produced at each leaf node. The biological

strategy of the wild papaya is to quickly achieve vegetative dominance and then exhaust its reserves in seed production before lianas or more persistent species overtake it. Long-term survival of individual plants is not an important part of the plan.

This strategy has negative consequences for commercial production in managed landscapes. Spaced plantings and weed control reduce biological competition in agricultural fields, yet the wild-type developmental program emphasizes upward growth, with the result that the first fruit occurs high on the trunk. Contributing to excessive early vertical growth is the practical necessity to plant at each hole a handful of seedlings segregating for sex to ensure the selection of a single hermaphrodite plant at flowering time. Competition among seedlings leads to vertical stretching and that reduces the productive life of the tree, since a commercial field is usually terminated when trees reach a height that makes harvesting impractical. Moreover, the extra seedlings needed to obtain uniform plantings of hermaphrodites must be maintained until flowering, even though they are ultimately removed, resulting in a waste of inputs, labor, and the productive life of the field.

The idiotype that overcomes these difficulties is a true-breeding hermaphrodite papaya that flowers earlier and lower on the trunk. A true-breeding hermaphrodite line of the cultivar Sunrise has been reported from Taiwan (Chan-Tai et al. 2003), and other genetic engineering approaches for creating such are under way, but practically speaking, all available papaya germplasm segregates for sex. Germplasm that flowers precociously does exist but usually faces several shortcomings. Exceptional plants can initiate flowering very early (in their third month), but when these become reproductive, they typically have trouble supporting continuous development of the fruit column. Once maturing fruits start absorbing the resources of the tree, little further root growth can occur. Precocious flowering cuts short normal vegetative development, and the growing fruit column soon outstrips the ability of truncated root and canopy systems to provide, resulting in flower abscission and slowing of apical growth during the 4–6 months period of fruit maturation. Flowering may commence again as fruits are harvested, but at best, this causes gaps in the fruit column and erratic production. Often, such trees lack the photosynthetic reserve capacity to tolerate normal biological or environmental stresses and become essentially annual plants, producing one crop of fruits before succumbing. Any damage in the canopy, such as that caused by black spot (*Asperisporium*), powdery mildew (*Oidium*), or leaf roll mites, leads to an undernourished root system that quickly becomes vulnerable to *Phytophthora* and other soilborne fungi.

Increased fertilizer application can somewhat offset this problem, but it causes another. Fruiting papayas respond to nitrogen by developing multiple fruits at each leaf node and by converting some stamens into supernumerary carpels, a phenomenon called stamen carpellody. On the cymose inflorescence of hermaphrodite trees, the terminal flower develops first and has a greater tendency to stamen carpellody, whereas subterminal flowers tend in the opposite direction toward carpel abortion. Selecting for symmetrical, non-carpellodic fruits at the terminal position of the inflorescence may cause subterminal fruits with “mango” or “banana” shapes resulting from carpel abortion. These are unmarketable and constitute a drain on the plant’s reserves. They also contribute to a closed fruit column in which fruit

compression and disease due to poor air circulation occur. We need germplasm that produces only a single 500–750-g terminal fruit per node and has reduced environmental sensitivity to factors contributing to stamen carpellody and carpel abortion. Keeping fruit size in this range is necessary for the export markets that Hawaii serves, but it also functions to reduce the reproductive sink on the tree, relative to larger-fruited cultivars.

In order to breed a better papaya for multiyear commercial production, we need to redirect the sink during the juvenile phase from vertical growth to root development and trunk girth, encouraging a precocious attainment of the vegetative base necessary to support continuous flowering and fruit set. Resistances to important canopy and root diseases, such as black spot, powdery mildew, and *Phytophthora*, will help ensure that the integrity of the photosynthetic source is maintained. With this base, flowering that commences between the 20th and 25th node (~0.5 m above ground level) may be sustainable for the life of the field. In addition, it will be helpful to reduce the reproductive sink by limiting fruit production to a single 1.0–1.5-lb. terminal fruit at each leaf node. This will reduce destructive competition between developing seeds and root or canopy systems during the commercial life of the tree.

Tapping Hybrid Vigor

To power precocious development, we need to enlist the heterotic vigor of F_1 hybrids. Careful selection of synergistic parents with complementary traits and good combining ability can produce hybrids with rapid seed germination and enhanced vegetative vigor. Compared with current inbred cultivars, hybrids may also have advantage in better adaptation to environmental variability, allowing them to be grown over a wider range of geographic locations. In Hawaii, wider adaptation has been observed in the F_1 hybrid cultivar "Rainbow," relative to its inbred parent "Kapoho," which loses ability to produce fruits of sufficient size outside of the Puna District, the region of its development. Chan (1995) has documented similar hybrid advantages in vegetative vigor and yield over varying environments at M.A.R.D.I. in Malaysia.

True-Breeding Hermaphrodites

The previously mentioned improvements will have maximum effect when seedling trees can be planted singly in the field, avoiding early competition. To make this possible, sex segregation must be eliminated, either through direct manipulation of the genetic loci determining lethality of homozygous hermaphrodites or, until that ambitious goal can be achieved, by linking an easily selected seedling marker, such as anthocyanin pigmentation, with the sex-determining locus, so that growers can identify and discard female seedlings based on the color of their foliage and transplant only non-transgenic hermaphrodites to the field. In 2008, an international

team of researchers based in Hawaii published a draft sequence for the papaya genome of 327 million base pairs (Ming et al. 2008). Hawaii principals in this effort were Ray Ming, formerly of the Hawaii Agriculture Research Center in Aiea, Oahu; Maqsudul Alam of the UH at Manoa genomics center; and Dennis Gonsalves, Director of the USDA Pacific Basin Agriculture Research Center in Hilo, along with an international group of collaborators including a number of faculty and graduate students at UH. This project is informing research in papaya improvement that will favorably impact the industry in the future. The first genome sequence has already benefitted papaya breeding efforts by making available a broad assortment of DNA markers, potentially linked with important economic traits, for marker-assisted selection. In the near future, it may allow dissection of the genetic control of sex in papaya and enable development of homozygous hermaphrodite genotypes that will eliminate the current necessity of roguing females from segregating seed-propagated plantings. In time, the solo papaya genome may be restructured to achieve the early partitioning to root development and single fruit per node architecture described in the preceding sections. Like a book of deep knowledge that reveals its meaning in proportion to our ability to comprehend, the genome sequence will provide guidance in measure with our ability to reinterpret it in the light of new experience.

In 1914, Higgins and Holt wrote: "Excepting the banana, there is no fruit grown in the Hawaiian Islands that means more to the people of this Territory than the papaya, if measured in terms of the comfort and enjoyment furnished to the people as a whole." A century later, their words apply more than ever. In Hawaii, a combination of progressive growers and modern horticultural research moved papaya from "pig feed" status to iconic tropical delicacy. The prosperity of the Hawaiian papaya industry has been due to its success in extending that appeal to other countries and peoples of the world. With markets opening overseas for virus-resistant papayas, and new technologies providing tools to achieve new breeding objectives, there is new hope that solo papayas will remain an important Hawaiian export commodity.

References

Badillo VM (2000) *Carica* L vs *Vasconcella* St Hil (Caricaceae) con la rehabilitacion de este ultimo. Ernstia 10:74–79

Badillo VM (2001) Nota correctiva *Vasconcellea* St Hill y no *Vasconcella* (Caricaceae). Ernstia 11: 75–76

Bartlett HH (1937) Botany of the Maya area. Miscellaneous Paper No. 1:19. Carnegie Institution, Washington, DC

Chan Y-K (1995) Development of F1 hybrids of papaya (Carica papaya L): seed production and performance of F1 hybrids. Ph.D. Thesis. Department of Genetics and Cellular Biology, University of Malaya, Kuala Lumpur

Chan-Tai C, Yen C-R, Chang L-S, Hsiao C-H, Ko T-S (2003) All hermaphrodite progeny are derived by self-pollinating the sunrise papaya mutant. Plant Breed 122:431–434

Crawford DL (1937) Hawaii's crop parade. Advertiser Publishing Co, Honolulu

Ferreira SA, Pitz KY, Manshardt R, Zee F, Fitch M, Gonsalves D (2002) Virus coat protein transgenic papaya provides practical control of Papaya ringspot virus in Hawaii. Plant Disease 86:101–105

Fitch M, Manshardt R, Gonsalves D, Slightom J, Sanford J (1992) Virus resistant papaya plants derived from tissues bombarded with the coat protein gene of papaya ringspot virus. Biotechnology 10:1466–1472

Gonsalves D (1998) Control of papaya ringspot virus in papaya: a case study. Annu Rev Phytopathol 36:415–437

Hamilton RA, Ito P (1968) Sunrise Solo: a different colored solo papaya. Hawaii Agricultural Experiment Station, Circular 69

Hamilton RA, Ito P (1986) Papaya breeding activities in Hawaii. In: Proceedings of the 22nd annual Hawaii Papaya Industry Association conference, 6–8, Honolulu

Hamilton RA, Ito P, Paull R (1993) Sunset Solo papaya. Hawaii Agricultural Experiment Station, Commodity Fact Sheet PA-3(B) Fruit

Hawaii Agricultural Experiment Station (1920) The solo papaya. Ann Rpt 1919

Hawaii Agricultural Extension Service (1958) Statistics of Hawaiian Agriculture 1957. Agric Econ Rpt No. 35

Hayes D (2001) Historical atlas of the North Pacific Ocean: maps of discovery and scientific exploration 1500–2000. Douglas and McIntyre, Vancouver

Higgins JE, Holt VS (1914) The papaya in Hawaii. USDA, Hawaii Agricultural Experiment Station, Bul 32. Washington, DC

Hofmeyr JDJ (1938) Determination of sex in Carica papaya. Farming South Afr 13:332

Jones WW (1940a) The influence of relative humidity on the respiration of papaya at high temperatures. Am Soc Hortic Sci Proc (1939) 37:119–124

Jones WW (1940b) Methyl-bromide fumigation of papaya and tomato. Hawaii Agricultural Experiment Station Cir 17

Jones WW, WB Storey (1941) Propagation and culture of the papaya. In: Papaya production in the Hawaiian Islands. Univ Hi Agric Exp Sta Bul 87, pp 23–31

Kane HK (1996) The Manila Galleons. In: Dye B (ed) Hawaii chronicles: island history from the pages of Honolulu magazine. University of Hawaii Press, Honolulu, pp 25–32

Lius S, Manshardt RM, Fitch MM, Slightom JL, Sanford JC, Gonsalves D (1997) Pathogen-derived resistance provides papaya with effective protection against papaya ringspot virus. Mol Breed 3:161–168

Loudat TA, Roecklein JC, Leung P-S (1987) A method of assessing new crop potential in Hawaii: a case study of the Hawaiian Papaya Industry. University of Hawaii Res Ext Ser 79

Lundell CL (1936) The vegetation of Peten. Carnegie Institution, Washington, DC, pp 194–195

Manshardt R (1998) "UH Rainbow" papaya. University of Hawaii Germplasm G-1

Manshardt RM, Wenslaff TF (1989) Interspecific hybridization of papaya with other *Carica* species. J Am Soc Hortic Sci 114:689–694

Manshardt RM, Zee FT-P (1994) Papaya germplasm and breeding in Hawaii. Fruit Varieties J 48:146–152

Mau R, Gonsalves D, Bautista R (1989) Use of cross protection to control papaya ringspot virus at Waianae. In: Proceedings of the 25th Ann Hawaii Papaya Ind Assoc Conf., pp 77–84

Mekako HU, Nakasone HY (1975) Interspecific hybridization among 6 Carica species. J Am Soc Hortic Sci 100:237–242

Ming R et al (2008) Genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). Nature 452:991–996

Nakasone HY, Crozier JA Jr, Ikehara DK (1972) Evaluation of "Waimanalo" a new papaya strain. Univ Hi Agric Exp Sta Tech Bul No. 79

National Agricultural Statistics Service (1999) Statistics of Hawaiian Agriculture

National Agricultural Statistics Service (2009) Statistics of Hawaiian Agriculture

Pope WT (1930) Papaya Culture in Hawaii. University of Hawaii Agric Exp Sta Bul 61

Storey WB (1938a) The primary flower types of papaya and the fruit types that develop from them. Amer Soc Hortic Sci Proc 35:80–82

Storey WB (1938b) Segregation of sex types in solo papaya and their application to the selection of seed. Amer Soc Hortic Sci Proc 35:83–85
Storey WB (1941) The botany and sex relationships of the papaya. In: Papaya production in the Hawaiian Islands. University of Hawaii Agric Exp Sta Bul 87, pp 5–22
Traub HP, Robinson TR, Stevens HE (1942) Papaya production in the United States. USDA Circular No 633. Washington, DC
University of Hawaii at Manoa Library, Hawaiian Collection (2006) Hawaiian Sugar Planters' Association Plantation Archives. http://www2.Hawaii.edu/~speccoll/p_puna.html. 3 June 2012
University of Hawaii at Manoa, Botany Department (2012) Brief history of G.P. Wilder. http://sites.google.com/site/uhmbotany/Home/employment/wilder-chair/brief-history-of-gp-wilder. 16 Feb 2012
Wilcox EV (1916) Tropical agriculture. D Appleton and Co, New York
Yeh S-D, Gonsalves D (1984) Evaluation of induced mutants of papaya ringspot virus for control by cross protection. Phytopathology 74:1086–1091
Zee FT-P, Nishina MS, Chan HT, Nishijima KA (1989) Blossom end defects and fruit fly disinfestation in papayas following hot water quarantine treatment. Hortscience 24:323–325

Chapter 7
Hawaii's Transgenic Papaya Story 1978–2012: A Personal Account

Dennis Gonsalves

Introduction

In the early 1990s, Hawaii's papaya industry was doing just fine. It had escaped destruction from papaya ringspot virus (PRSV) in the 1950s and early 1960s by relocating from Oahu Island to the Puna district of Hawaii Island which was free of PRSV. In 1991, Puna produced 46 million pounds of the state's 48 million pounds of fresh papaya. However, in May 1992 PRSV was detected in Puna and dramatically changed the course of the papaya industry (Gonsalves 1998).

This review provides my personal interpretation of Hawaii's transgenic papaya story in which I was fortunate to be part of from the very beginning. The aim of this writing is not to simply review the Hawaii papaya story because a number of reviews have been written on this subject (Fermin et al. 2004; Gonsalves et al. 1998, 2004a, b, 2006, 2008a, b; Gonsalves 1998, 2006). Furthermore, it *only* focuses on the Hawaii transgenic papaya, and even at that it is not an in-depth review of the technical aspects of the Hawaii transgenic papaya. Instead, I give my personal and undoubtedly biased view of the conditions, circumstances, people, and other factors that contributed to the making of the Hawaii transgenic papaya story. The story will be described chronologically but the current status of the events is injected into the sections. The last section provides my overall view of the contribution of the Hawaii transgenic papaya story to the world of what I call "translation" biotechnology. That is, getting the product to the market.

D. Gonsalves (✉)
US Department of Agriculture, Agricultural Research Service,
Pacific Basin Agricultural Research Center, 64 Nowelo Street, Hilo, HI 96720, USA
e-mail: dennisgonsal@gmail.com

R. Ming and P.H. Moore (eds.), *Genetics and Genomics of Papaya*, Plant Genetics and Genomics: Crops and Models 10, DOI 10.1007/978-1-4614-8087-7_7,

The Question

In 1978 the late Dr. Richard Hamilton a renowned horticulturist at the University of Hawaii took me to see Dr. William Furtick, the University's Dean of the College of Tropical Agriculture. During the course of our meeting, the Dean asked me: "What would happen if PRSV got into Puna?" My natural reaction was that PRSV would devastate the industry because PRSV is rapidly transmitted by aphids and no resistant germplasm had ever been found in *Carica papaya* (Gonsalves et al. 2008b). His next question was: "What could be done to control PRSV?" I told him that cross protection could be tried because I had been actively experimenting in Florida with Dr. Steve Garnsey on the control of citrus tristeza virus by mild strain cross protection. Cross protection is the phenomenon whereby plants that are systemically infected with a mild strain of a virus are protected against the effects of infection by a more virulent related strain (Gonsalves and Garnsey 1989). Some months after the meeting with the Dean, my lab started characterizing PRSV and trying to control PRSV by cross protection. The important take-away of this paragraph is that proactive research has started to control PRSV even though Puna was free of PRSV. It is clear to me that if this proactive research was not undertaken, timely control of PRSV after it entered Puna in 1992 would have been virtually impossible.

Characterization of PRSV and Cross Protection

We started working on PRSV in 1978 with a team composed of Drs. Mamoru Ishii and Ryoji Namba of the University of Hawaii and my graduate student Shyi-Dong Yeh. The goal was to develop antiserum against PRSV for use in ELISA test (Clark and Adams 1977) and to develop a mild strain of PRSV from Hawaii. Virus purification and detection by ELISA (Gonsalves and Ishii 1980) were accomplished and by 1983, Shyi-Dong Yeh had developed the mild strain of PRSV HA 5-1 via nitrous acid mutation (Yeh and Gonsalves 1984). Cross-protection trials were done extensively in Taiwan and also in Hawaii in efforts led by Ronald Mau (Mau et al. 1989) and Steve Ferreira. Although cross protection afforded some protection in Taiwan and good protection in Hawaii, the effects were not good enough such that the farmers adopted it long term (Yeh and Gonsalves 1994). Importantly, we subsequently used the coat protein gene of the mild strain PRSV HA 5-1 in our transgenic papaya work. Also, it taught us lessons on when to abandon a selected research approach and search for other avenues to control PRSV.

The Pathogen-Derived Resistance Approach

We were indeed fortunate that as we were seeing the limited success of our cross-protection approach for controlling PRSV, two key papers published in the mid-1980s revolutionized research for controlling viruses through the use of transgenic plants.

Roger Beachy's group (Powell-Abel et al. 1986) showed that transgenic tobacco expressing the coat protein gene of tobacco mosaic virus was protected against infection by tobacco mosaic virus. And Sanford and Johnston (1985) published the concept of parasite-derived resistance or now more commonly referred to as pathogen-derived resistance (PDR). PDR is a phenomenon whereby transgenic plants containing genes or sequences of a parasite (in our case, the coat protein gene of PRSV) are protected against detrimental effects of the same or related pathogens.

Development of Transgenic Papaya

This section, up through the deregulation and commercialization of the transgenic papaya, is largely taken from a review I wrote in 1998 (Gonsalves 1998) because it is difficult for me to provide the rationale any better, since the events were fresh in my mind at that time. Our laboratory began utilizing the PDR concept in 1986 by cloning and sequencing the coat protein gene of PRSV HA 5-1 in collaboration with molecular biologist Jerry Slightom who was then employed by "The Upjohn Company." Actually, our initial joint interests were not focused on PRSV of papaya. Instead, they were focused on the important viruses that infect vegetables, including cucumber mosaic virus, watermelon mosaic virus 2, zucchini yellow mosaic virus, and PRSV. The then Asgrow Seed Company subsequently developed commercial squash with resistance to zucchini yellow mosaic virus and watermelon mosaic virus 2 (Fuchs and Gonsalves 1995; Tricoli et al. 1995).

The expertise and reagents obtained from the initial vegetable work were then applied to papaya since our specific goal was to control PRSV. The cross-protection work had been consistently supported through funds from the USDA Section 406 grant program, which focused on agricultural problems of the Pacific region. Fortunately, this program also funded the research to use PDR to control PRSV in papaya in Hawaii. Thus, the original transgenic papaya team was formed and consisted of Richard Manshardt and Maureen Fitch from the University of Hawaii, Slightom from The Upjohn Company, and myself from Cornell University. We were also very fortunate to tap the expertise and services of John Sanford at Cornell University, who had recently coinvented the gene gun.

We followed the prevailing thoughts of the mid-1980s that the coat protein per se was required to confer resistance. Our target gene was the coat protein gene of PRSV HA 5-1, the mild mutant of PRSV HA that had been recently cloned and sequenced for the vegetable work (Quemada et al. 1990). Jerry Slightom developed our expression cassette. Because of various technical difficulties and "our" requirement that the gene be expressed as a protein, the gene was engineered as a chimeric protein containing 17 amino acids of cucumber mosaic virus at the N terminus of the full-length coat protein gene of PRSV HA 5-1. Jerry blended well into the team because he had state-of-the-art expertise in molecular biology while we contributed to biological thinking as we ourselves sought to master the art of molecular biology.

Arguably, the most difficult task for developing the transgenic papaya was tissue culture and transformation since virtually nothing had been published on that subject. This task was taken up in 1987 by Maureen Fitch, an ARS researcher that was a graduate student of Richard Manshardt. The target cultivars for transformation were "Kapoho," "Sunrise," and "Sunset." Numerous efforts by Fitch to develop a papaya regeneration system via organogenesis failed. However, a technique to develop transgenic walnuts by transforming embryogenic cultures had recently been reported. The research moved rapidly once the decision was made to shift to transforming embryogenic tissue. A technique to produce highly embryogenic tissue starting from immature zygotic embryos was developed (Fitch and Manshardt 1990). In 1988–1989, embryogenic tissue were bombarded with tungsten particles coated with DNA of the PRSV HA 5-1 coat protein gene using the gene gun in Sanford's laboratory. Transgenic plants were obtained and were growing in the greenhouse 15 months later (Fitch et al. 1990, 1992).

What were the most important factors that helped to achieve this timely breakthrough in getting transformed papaya? My vote goes to Maureen Fitch with her unending enthusiasm and skills in tissue culture. Another was the emphasis on obtaining a sufficient number of transgenic plants to test for resistance, rather than focusing on refining the details of the transformation technique. We wanted to advance quickly to the stage of testing transgenic papaya to determine whether PDR would work for controlling PRSV in papaya.

Clones of nine R0 transgenic lines, six "Sunset" and three "Kapoho," were sent to Cornell for inoculation tests with PRSV HA, which is the severe parent of the mild strain PRSV HA 5-1. We were lucky, the R0-micropropagated plants of the first line, designated 55-1, that was tested showed excellent resistance to PRSV HA (Fitch et al. 1992). It was almost unbelievable as I watched the inoculated line 55-1 plants grow normally while the inoculated controls showed severe infection. By April 1991, we had identified a line that was resistant to PRSV under greenhouse conditions (Fig. 7.1). The other transgenic lines in the R0 inoculation were not resistant to PRSV.

The team's mentality in going for the "jugular" showed in the subsequent testing of line 55-1. It would play a huge role in the timely deployment of the transgenic papaya. Line 55-1 was female and thus progenies could not be obtained directly from the R0 plants, as would be the case for a hermaphrodite. A two-pronged approach was instituted to move the research ahead aggressively and to determine whether line 55-1 would be resistant to PRSV under field conditions and have suitable horticultural characteristics. First, a decision was made to conduct a field trial using R0 plants instead of waiting for more than a year to obtain Rl plants. And second, seeds were obtained by pollinating female R0 clones of line 55-1 with pollen from nontransgenic "Sunset" under greenhouse conditions at Cornell University and the University of Hawaii, and the resulting R1 plants were screened under greenhouse conditions at Cornell for resistance to PRSV isolates from around world.

Fig. 7.1 Transgenic line 55-1 on left compared to nontransgenic papaya (photo taken: 1991)

Transgenic Rl Plants of Line 55-1 Are Highly Resistant to Hawaii Strains but Largely Susceptible to Strains Outside Hawaii

Research headed by my graduate student Paula Tennant at Cornell University assessed the resistance of Rl plants of line 55-1 against 3 PRSV isolates from Hawaii and 13 isolates from different parts of the world (Tennant et al. 1994). Analysis clearly showed that 50 % of the progenies were transgenic, with the rest nontransgenic; this confirmed that transgenic plants had one insert of the nptll gene and, presumably, the coat protein gene. Inoculation results showed that transgenic Rl plants of line 55-1 were highly resistant to Hawaiian isolates but showed variable levels of resistance (largely susceptible) to non-Hawaiian isolates: for example, plants inoculated with an isolate from Thailand developed severe symptoms with no delay in symptom appearance, whereas isolates from Jamaica infected line 55-1, but symptoms were delayed and attenuated; isolates from Florida and Mexico infected only a percentage of the plants and symptoms were milder. Cross-protection results on nontransgenic plants roughly paralleled those observed for transgenic plants.

That is, protection was complete against the Hawaiian isolates but not against non-Hawaiian isolates. Also, isolates that rapidly infected line 55-1 also rapidly infected cross-protected papaya. For example, papaya cross-protected with PRSV HA 5-1 showed almost no protection against the Thailand isolate, which also rapidly overcame the resistance of line 55-1. In summary, the results clearly showed that Rl plants of line 55-1 would not be effective against all isolates of PRSV, but fortunately line 55-1 was resistant to PRSV isolates in Hawaii. Later work would show that resistance of line 55-1 was via a posttranscriptional RNA-mediated mechanism (Tennant et al. 2001).

Field Trials with R0 Plants Show Line 55-1 Is Effective for Controlling PRSV in Hawaii

In 1991, Manshardt obtained a permit from Animal Plant Health Inspection Service (APHIS) for a field trial at the University of Hawaii's Experiment station at Waimanalo, on Oahu Island. The importance of a field evaluation at an early stage cannot be overemphasized because it allowed us to appraise the resistance and horticultural characteristics of line 55-1, to bulk up seeds by crossing line 55-1 with nontransgenic plant cultivars, and to demonstrate the long-term resistance of a transgenic fruit crop to infection. Plants were set in the field by the end of June 1992. The transgenic papaya showed excellent resistance throughout the 2-year trial (Lius et al. 1997). Line 55-1 plants grew normally and fruit appearance and total soluble solids of about 13 % were within the expected range. By mid-1993, the trial had provided convincing evidence that line 55-1 would be useful for controlling PRSV in Hawaii or at least on Oahu Island.

1992–1998: Invasion and Impact of PRSV in Puna

Detection and Spread of PRSV

The inevitable entry of PRSV into the Puna district on Hawaii Island was discovered during the first week of May in a papaya field in Pahoa, 1–3 miles from the major papaya-growing areas in Puna (Isherwood 1992). I still have a vivid memory of that day. Steve Ferreira and I had just landed at Honolulu airport following a trip to Guam and were met by Steve's technician, Karen Pitz, who told us that apparently PRSV was discovered in Pahoa in the Puna district. We flew to Hilo that day and sure enough we saw many trees in a papaya plantation with advanced PRSV symptoms. Infection had been established in this area for at least several months, as judged by symptoms on the fruits and the fact that many plants were infected in one location. Surveys of the immediate area revealed PRSV in abandoned orchards, as well as in young orchards that were not yet producing fruit. PRSV was poised to

Fig. 7.2 PRSV-infected abandoned papaya fields in Puna, 1994

invade the major papaya-growing areas of Puna, which included Kapoho, Opihikao, Kahuawai, and Kalapana.

The Hawaii Department of Agriculture (HDOA) immediately launched an eradication program. The area was surveyed and infected trees were rouged out. A suggestion to destroy all papaya in that area, however, was not approved by the growers. Nevertheless, an HDOA program to mark trees to be rouged by growers was started in 1992 (Isherwood 1992). The next few months seemed hopeful for containing PRSV. By September 1992, 4,915 trees had been rouged in Pahoa, and the number of infected trees being cut each week had decreased to below 85 (Isherwood 1992).

However, the hope of containment was short-lived. The incidence of PRSV increased dramatically in Kapoho, which was closest to Pahoa, as the program of voluntary cutting of trees was not strictly followed and as farmers experiencing high infection rates abandoned their fields, which created huge reservoirs of inocula for aphids to acquire and spread the virus (Isherwood 1994). By late 1994, nearly all papaya of Kapoho was infected by the virus. In October 1994, the HDOA declared that PRSV was uncontrollable and stopped the practice of marking trees for rouging. In less than 3 years, a third of the Puna papaya area was infected. By 1997, Pohoiki and Kahuawai were completely infected. Kalapana was the last place to become heavily infected. By September 1997, rouging was also discontinued in Kalapana. Five years after the onset of the virus in Pahoa, the entire Puna area was severely affected (Fig. 7.2).

Table 7.1 Fresh papaya production in the state of Hawaii and in the Puna district from 1992 to 2008

Year	Total (×1,000 lbs)	Puna (×1,000 lbs)	Puna (%)
1992[a]	55,800	52,955	95
1993	58,200	55,785	95
1994	56,200	52,525	93
1995	41,900	29,215	93
1996	37,800	34,195	90
1997	35,700	27,810	77
1998[b]	35,600	26,750	75
1999	39,400	25,455	64
2000	50,250	33,950	67
2001	52,000	40,290	77
2002	42,700	35,880	84
2003	40,800	35,735	87
2004	34,100	29,995	87
2005	30,700	26,910	87
2006	26,600	24,090	90
2007	31,200	28,340	90
2008	31,500	28,500	90

Fresh papaya utilization in Hawaii

[a]PRSV first reported in Puna. Note that figures represent Hawaii County but we assume for nearly all years, much of the papaya was produced in the Puna District of Hawaii County

[b]Transgenic seed released

Impact of PRSV on Production and Industry

The spread of PRSV was rapid and dramatic. Equally dramatic was the impact of PRSV on the industry (Gonsalves 2006). In looking at impact, it is worthwhile to note that papaya plants that become infected within the first several months of their growth stage will not produce saleable fruit. Thus, all the expenses for land preparation and caring for the papaya will be lost. Under these severe disease conditions, the difference between resistant transgenic and susceptible nontransgenic papaya is a matter of fruit production versus no fruit production. It is difficult to quantify the number of farmers who tried to grow papaya when the virus was widespread and subsequently abandoned the field even before they could harvest a single fruit.

The overall impact on papaya production in Puna and Hawaii as a whole is depicted in Table 7.1. The papaya production figures of Hawaii bear out the effect that PRSV was having on Hawaii's papaya industry (Gonsalves et al. 2004b). Puna's production of marketable papaya had decreased from 52.9 million pounds in 1992 to 26.7 million pounds in 1998, a decrease of 50 % in a 6-year span. In 1992, Puna produced 95 % of Hawaii's papaya, whereas in 1998 its share of the Hawaii's production was 75 %. Other regions, primarily in the Hamakua district of Hawaii Island, parts of Oahu Island, and Kauai, increased production and were able to pick up some of the production void left by Puna. Even with the establishing of new production areas, Hawaii's overall marketable papaya production had slipped from 55.8 million pounds in 1992 to 35.6 million pounds in 1998.

Importantly, much of the papaya that was being harvested in Puna were infected and not of superior quality, which was the big selling point for the Hawaiian papaya. In fact, the standards of not shipping symptomatic fruit to Japan and the mainland was lifted as a measure to keep the exports of papaya to Japan and mainland at a reasonably good volume. Another impact is on the people associated with the papaya industry. While in 1992 Hawaii had eight packing houses, by 1998 only three packing houses were in operation and these were not running at full capacity as they were in 1992. It was difficult to find noninfected papaya in Hawaii.

1993–1995: Manshardt Develops "SunUp" and "Rainbow" Using Plants in the R0 Field Trial on Oahu

As PRSV was raging in Puna, Manshardt was busy developing a cultivar(s) that might replace the "Kapoho" in Puna. Line 55-1 was brought to homozygosity for the CP gene by first crossing the R0 female line 55-1 with its nontransgenic parent "Sunset" and doing backcrossing until he had a transgenic line 55-1 that was homozygous for the CP gene. The dominant cultivar in Puna in the 1992 was the yellow-fleshed "Kapoho" and unfortunately, we had not successfully developed a PRSV-resistant transgenic "Kapoho" (Fitch et al. 1992). Manshardt crossed the homozygous transgenic line 55-1 with "Kapoho," anticipating the yield of a yellow-fleshed hybrid (because yellow is dominant over red), which might be an acceptable substitute for "Kapoho." The homozygous line 55-1 was later named "UH SunUp," and the F1 hybrid from the cross of the transgenic "UH SunUp," and Kapoho was named "UH Rainbow" (Manshardt 1998) but the names "Rainbow" and "SunUp" are commonly used today (Figs. 7.3 and 7.4). The establishment of the R0 field trial in 1992 and the continual use of the initial R0 field trial by Manshardt cannot be over emphasized in its contribution to the success of the papaya story. However, Manshardt did

Fig. 7.3 "Rainbow"

Fig. 7.4 "SunUp"

not have extensive data on the resistance or horticultural characteristics of these cultivars under field conditions with intense disease pressure. A field trial in Puna was necessary.

1995–1998: Performance in the "Red Zone": Establishment of Transgenic Field Trial in Puna, Deregulation and Commercialization of Transgenic Papaya

The Team Has Entered the Red Zone and Let's See if They Can Score a Touchdown and at the Least Come Up with a Field Goal

American football broadcasters often say this or similar sayings to emphasize that it is great to move the football down the field but the obvious goal is to score. Teams that consistently score after entering the "red zone" usually win championships. In looking back, we had done lots of good research to develop a transgenic papaya line that showed resistance to PRSV in the initial R0 field trial during the critical time that the industry was being severely damaged by PRSV. However, these achievements would simply be academic exercises if we did not get the transgenic papaya to the industry in a timely manner. We had indeed entered what I call the "red zone" of "translational" biology.

Field Trial in Kapoho

By 1994, personnel at the University of Hawaii (e.g., Ferreira and Mau 1994) had recognized that PRSV was out of control in Puna and would significantly reduce the

state's papaya production. To maintain production and eventually reclaim the Puna area for papaya growth, a bold plan was proposed (Ferreira and Mau 1994): move new papaya plantings to areas of Hawaii Island where PRSV had not been detected, completely eradicate papaya and cucurbits from the papaya area in Puna, and place a 1-year moratorium on planting papaya and cucurbits. The hope was that papaya acreage and production would be maintained, and PRSV would be eliminated from the Puna area. The program was anticipated to be completed in 6–8 years. USDA was to provide funding to support the initial phases of the plan.

Coincidentally, Steve Ferreira faxed me a copy of the plan while I was at the meeting on Biosafety Results of GMO crops in Monterrey, California in 1994 (Gonsalves et al. 1994). The next day I gave my talk on assessment on our field trials of transgenic papaya, cucurbits, and tomato and mentioned the dire situation of PRSV in Hawaii. At the meeting, I talked with a person in APHIS and asked: Would it be possible to put a large field trial in the devastated area of Puna? He said that might be possible and to submit an application.

Why take a "risk" in establishing a field trial in the middle of the devastated area at such a stage in the development of the transgenic papaya? Arguments could be marshaled for and against establishing such a field test: line 55-1 had performed very well in field trials on Oahu Island, and the line was resistant to PRSV Panaewa, a greenhouse isolate from Hawaii Island; the industry needed drastic actions to survive; the plan for moving the industry might not succeed; and it would be very difficult to eradicated PRSV from Puna; the risk of PRSV-resistant papaya becoming a weed was not relevant because papaya is not a weed in areas that do not have PRSV; wild relatives of papaya are not grown in Hawaii; and the potential benefits of transgenic papaya far outweighed the risks. Weighing against the field trial were the facts that pollen from the transgenic papaya might contaminate commercial plantings, resulting in the potential sale of commercial fruit with a nonderegulated transgene; preventing pilferage in a trial installed on a farmer's field would be difficult, and thus there might be serious consequences if stolen fruit ended up in commercial markets.

Steve Ferreira joined the transgenic papaya team and led the task of getting the application going and initiating the field test in Kapoho on a farm that was being devastated by PRSV. APHIS approved the field trial in 1995 with the stipulation that (a) the field must be sufficiently isolated from commercial orchards to minimize the chance of transgenic pollen escaping to nontransgenic material outside of the field test, (b) all abandoned trees in the area must be monitored for the introgression of the transgene into fruits of these trees, and (c) all fruits had to be buried on site.

Led by Ferreira and his excellent technician Karen Pitz, the field trial was set up in Kapoho in October 1995 on the property of a farmer who had ceased growing papaya because of PRSV. One part of the trial consisted of replicated blocks to compare virus-resistance performances of "SunUp," "Rainbow," and line 63-1. The latter was a red-flesh transgenic "Sunrise" from initial transformation experiments that were screened in later greenhouse tests, and found to be PRSV resistant (Souza et al. 2005; Tennant et al. 1994, 2005). Additionally, "Kapoho" cross-protected with PRSV HA 5-1, and several PRSV-tolerant lines were tested in the replicated blocks. Another part of the trial was established to simulate commercial conditions.

Fig. 7.5 Solid block of "Rainbow" surrounded by PRSV-infected nontransgenic papaya in 1997

A one-square acre solid block of "Rainbow" was planted adjacent to the replicated blocks. Several rows of nontransgenic "Sunrise" were planted on the perimeter of the replicated and solid blocks. An abandoned papaya field alongside the field plot was used as a primary source of the virus (Ferreira et al. 2002).

The results were dramatic (Fig. 7.5). Resistance of "Rainbow" and "SunUp" were complete, and the yield and quality were excellent (Ferreira et al. 2002). In the solid block, "Rainbow" averaged about 110,000 lbs of marketable fruit per acre/year, whereas nontransgenic plants yield of marketable fruit were down to 50 lbs per acre after 1.5 years in the field. In "Rainbow" it appeared that the papaya industry potentially had a substitute for "Kapoho." Additionally, line 63-1 performed as well as "SunUp" or "Rainbow" in the replicated blocks.

Deregulation and Commercialization of Transgenic Papaya

From the 1990s and until today, deregulation and commercialization have been largely the purview of private companies who stand to benefit financially from developing a product and moving it to the market. However, the papaya industry did not have the infrastructure or finances to do the work. The papaya team took on the efforts to deregulate the transgenic papaya lines 55-1 and 63-1 and thus entered the "red zone."

The tasks of securing deregulation of papaya by APHIS, Food and Drug Administration (FDA), and Environmental Protection Agency (EPA) were taken up

by Richard Manshardt and our lab. APHIS was largely concerned with the potential risk of transgenic papaya on the environment. Two main risks were of heteroencapsidation of the incoming virus with coat protein produced by the transgenic papaya and of recombination of the transgene with incoming viruses. The former might allow nonvectored viruses to become vector transmissible, whereas the latter might result in the creation of novel viruses. A third concern was that the escape of the transgenic genes to wild relatives might make the relatives more weedy or even make papaya more weedy because its resistance to PRSV. The concern was of no consequence because Hawaii does not have papaya relatives in the wild, and papaya is not considered a weed even in areas where there is no PRSV. Transgenic lines 55-1 and 63-1 and their derivatives were deregulated by APHIS in November 1996 (Strating 1996). This action greatly increased the efficiency of the ongoing field trial because fruit no longer had to be buried at the test site, which allowed us to sample and send fruit to various laboratories and to the packinghouse without undue constraints.

The EPA considered the coat protein of PRSV pesticide because it conferred resistance to plant viruses. A pesticide is subjected to tolerance levels in the plant. In the permit application, we petitioned for an exemption from tolerance levels of the coat protein produced by the transgenic plant. We contended that the pesticide (the coat protein) was already present in many fruits consumed by the public. In fact, we had earlier used cross protection (the deliberate infection of papaya with a mild strain of PRSV) to control PRSV. Fruit from these trees was sold to consumers. Furthermore, there is no evidence to date that the coat protein of PRSV or other plant viruses is allergenic or detrimental to human health in any way. Finally, measured amounts of coat protein in transgenic plants were much lower than those of infected plants. An exemption from tolerance to lines 55-1 and 63-1 was granted in August 1997.

The FDA is concerned with food safety of transgenic products. This agency follows a consultative process whereby the investigators submit an application with data and statements corroborating that the product is not harmful to human health. Several aspects of the transgenic papaya were considered: the concentration range of some important vitamins, including vitamin C; the presence of GUS and nptll genes; and whether transgenic papaya had abnormally high concentrations of the naturally occurring benzyl isothiocyanate. This latter compound has been reported in papaya. As part of our analysis we had done some initial southern blot tests of line 55-1 and line 63-1. These tests suggested that line 63-1 had large segments of vector sequences in its genome and thus we elected not to move line 63-1 forward for consultation with FDA. FDA consultation approval was granted for line 55-1 in September 1997.

In the United States, a transgenic product cannot legally be commercialized unless it is fully deregulated and until licenses are obtained for the use of the intellectual property rights for processes or components that are part of the product or that have been used to develop the product. To move things forward, the ownership of the papaya was placed in the hands of the Papaya Administrative Committee (PAC), which consisted of Hawaii papaya growers under a USDA marketing order agreement. The processes in question were the gene gun and parasite-derived

resistance, in particular, coat protein-mediated protection. The components were translational enhancement leader sequences and genes (nptll, GUS, and coat protein). This crucial hurdle involved legal and financial considerations beyond means and expertise of the transgenic papaya team. The tasks were taken up by PAC and its legal counsel Michael Goldman. License agreements were obtained from all parties in April 1998, allowing the commercial cultivation of the papaya or its derivatives in Hawaii only. Fruits can be sold outside Hawaii, provided that the importing state or country allows the importation and sale of transgenic papaya. Fruit derived from the licensed transgenic papaya grown outside of Hawaii cannot be sold commercially.

Distribution of seeds sufficient for 1,000 acres was started on May 1, 2008 after a celebration at the Hilo Hawaiian Hotel. Almost 7 years to the day PRSV was discovered in Puna, efforts in the "red zone" were successful in providing to growers a PRSV-resistant transgenic papaya. What would the impact be?

Impact of the Transgenic Papaya in Hawaii

Puna Papaya Production

The Kapoho field trial showed that the transgenic papaya was resistant to PRSV, and "Rainbow" papaya could serve as adequate replacement for "Kapoho." But would the farmers adopt it? Carol Gonsalves did a study to capture the farmer adoption rate of the transgenic cultivars from the initial seed distribution and up to September 1999 (Gonsalves et al. 2004a, 2007). Personal interviews were conducted with 93 of the 171 farmers who had registered to obtain transgenic papaya seeds in 1998. The data collected also included information on the size and types of farms, farmer attitudes, and demographic information. To the question on why they wanted to plant "Rainbow," 96 % said that it was because of the ability of "Rainbow" to resist PRSV. Interestingly, 71 % of the farmers had received information of the transgenic papaya through their own farmer organizations or through PAC.

Farmer adoption of the transgenic papaya was very high. Adoption was defined as to whether the farmer had planted the seeds, and not based merely on whether the farmer had signed up and obtained seeds. By September 1999, 90 % of the farmers had obtained transgenic seeds and 76 % of them had planted (adopted) the seeds. The survey also showed that farmers were ready to implement the technology; that is, 80 % of the farmers surveyed planted the seeds within 3 months of obtaining them. As noted earlier, Kapoho area of Puna succumbed to PRSV the earliest. By September 1999, 94 % of the farmers in Kapoho who were surveyed had obtained transgenic papaya seeds, and 88 % had planted them, and 29 % had already begun to harvest fruit.

Interestingly, in 2011 Carol Gonsalves again interviewed 16 of the 93 farmers that are still farming and their answers were still the same. They are growing

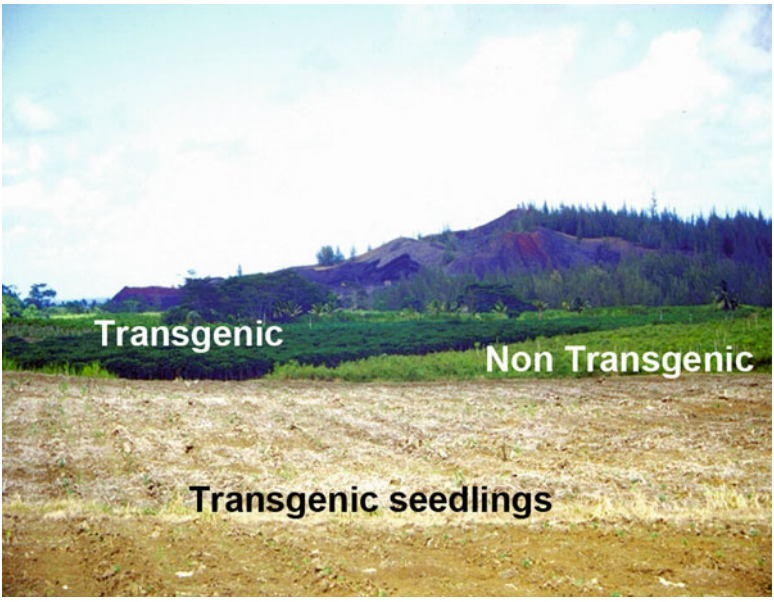

Fig. 7.6 Reclamation of Puna papaya lands with "Rainbow" (photo taken: 1999). Green field is "Rainbow" surrounded by PRSV-infected abandoned papaya, and foreground is newly planted "Rainbow" seedlings that are barely visible

"Rainbow" papaya and a main reason was that it is resistant to PRSV. They still have vivid memories of PRSV and know that PRSV is still around.

As mentioned previously, seeds of the transgenic papaya were released in May 1998, and many growers started planting the seeds (Fig. 7.6). Table 7.1 shows that trend of the marketable papaya production in Puna following the release of the transgenic papaya. The year 2000 was the first full harvest cycle of transgenic papaya in Puna, and the harvest was 33,950 million pounds which was an increase of 7.2 million pounds over the 1998 harvest which would have been all "Kapoho." Clearly, "Rainbow" papaya had reversed the downward trend of papaya production in Puna. However, Puna's part of the state's papaya crop was 67 % showing that other parts of the states had picked up production as mentioned previously. The peak production of Puna papaya since 1998 occurred in 2001 at 40.2 million pounds and gradually declined to 28.5 million pounds in 2008.

Since its release in 1998, the amount of "Rainbow" papaya has steadily increased from 32 % in 2001 to 77 % in 2009, while "Kapoho" has gone from 37 % in 2001 to 9 % in 2009 (Table 7.2). "Kapoho" and "Rainbow" are grown almost exclusively in Puna with very small acreage of "Rainbow" also grown on Oahu Island. Without a doubt, the popularity of "Rainbow" is due to its virus resistance and its good horticultural qualities. "Sunrise" is a nontransgenic papaya that has consistently occupied about 10 % of Hawaii's acreage. The other transgenic cultivars "SunUp"

Table 7.2 Relative percentage of acreage of papaya varieties grown in the state of Hawaii from 2000 to 2009

Variety	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Kapoho	37	43	42	38	29	30	25	17	16	9
Rainbow	32	39	44	46	52	53	58	68	64	77
Sunrise	14	11	10	10	11	9	11	8	9	9
Other[a]	7	7	4	6	8	8	6	7	11	5

[a]Transgenic "SunUp" and other transgenic cultivars derived from "Rainbow" or "SunUp" are included in this category, and thus it is reasonable assume that this percentage refers to transgenic papaya other than "Rainbow"

Fig. 7.7 Background shows "Rainbow" field growing well and foreground shows severely PRSV-infected "Kapoho" field that was cut down due to nonproductivity (photo taken: 2004)

and "Kamiya" are grouped in the "Other" category in Table 7.2 and it is reasonable to assume that this category is almost all composed of transgenic papaya. Taken together, one would surmise that transgenic papaya acreage has gone from 39 % of total acreage in 2000 to 83 % in 2009. Very likely, transgenic papaya acreage in 2012 is about 85 % of the total in the state of Hawaii. This is a testament to the fact that PRSV, although controlled by transgenic papaya, is still around and growers are leery of moving back into nontransgenic papaya because of the potential of losing their crop to PRSV. A graphic example is shown in a "Rainbow" field next to a "Kapoho" field in Puna in 2004 (Fig. 7.7).

Table 7.1 also shows that Hawaii's papaya production has not returned to its 1992 level, when PRSV was first detected in Puna. Perhaps the primary reason is that Hawaii lost a good share of its traditional markets in mainland USA and in Japan due to the drastic decrease of "Kapoho" production in the mid-1990s. And thus other areas

supplied the void. In regard to Japan, the release of the transgenic papaya applied specifically to markets in the USA and did not allow export of "Rainbow" to Japan because it was not deregulated. This scenario was recently summarized in an FAS/USDA Gain report which showed that the export value of Hawaii papaya decreased from $16 million in 1996 to $1 million in 2010 (Sato 2011). Clearly, deregulation of the transgenic papaya in Japan was a potential way to regain market share. This part is detailed later in this review.

More Papaya Cultivars Grown

A common concern has been that the introduction of the transgenic papaya would cut down on the horticultural diversity of papaya cultivars by relying only on growing transgenic crops. In actuality, Hawaii's papaya industry was virtually a monoculture of "Kapoho" before "Rainbow" papaya was released in 1998. In looking at diversity, one has to take into account the nature of PRSV on Hawaii, Maui, and Oahu islands. The virus is still around and will again become severe if susceptible papaya is widely grown again. Thus, in reality we need to look at "PRSV resistance" as a basic requirement to profitably grow papaya on a commercial scale on Hawaii and Oahu Islands. If one accepts that PRSV resistance is a "basic" trait, then the Hawaiian papaya industry has indeed become more diversified since the introduction of the transgenic "PRSV-resistance" trait in papaya. While "SunUp" is a new variety, it essentially is a transgenic sib of "Sunset." However, "Rainbow" is a new variety since it is an F1 hybrid of a cross of between "Kapoho" and "SunUp." The new transgenic cultivar "Laie Gold," which is a hybrid between "Rainbow F2" (selfed Rainbow) and the nontransgenic "Kamiya," also serves a niche market on Oahu Island. My opinion is that all cultivars that are generated for Hawaii need to have the "PRSV-resistance" trait.

Help to Enable the Production of Nontransgenic Papaya in Puna

One might ask the logical question: Why doesn't Hawaii produce only transgenic papaya? It is critical that Hawaii continues to produce nontransgenic papaya to supply the Japan market. Arguably, one of the major contributions that the transgenic papaya has made to the papaya industry is that it helps prolong the economic production of nontransgenic papaya (Gonsalves and Ferreira 2003). This has occurred in several ways. Firstly, the initial large-scale planting of transgenic papaya in established farms, along with the elimination of abandoned virus-infected fields drastically reduced virus inocula. In fact, HDOA instituted a plan in 1999 to ensure the production of nontransgenic papaya in the Kahuawai area of Puna. Kahuawai was isolated from established papaya fields, and the prevailing winds in Kahuawai came

from the ocean which borders the area (Gonsalves and Ferreira 2003). Growers were to monitor for infection and rogue infected plants quickly. Growers who followed the recommended practices were able to economically produce "Kapoho" without major losses from PRSV. Secondly, although definitive experiments have not been carried out, it seems that transgenic papaya can provide a buffer zone to protect nontransgenic papaya that are planted within the confines of the buffer. The reasoning is that viruliferous aphids will feed on transgenic plants and thus be purged of virus before traveling to the nontransgenic plantings within the buffer. This approach also has the advantage of allowing the grower to produce transgenic and nontransgenic papaya in relatively close proximity. Timely elimination of infected trees would need to be practiced to delay large-scale infection of the nontransgenic plants.

Environmental Impact

The transgenic papaya has also had positive environmental impact. Since transgenic papaya can be grown on previous papaya land with no damage from PRSV, this situation has cut down on the amount of new papaya land that has to be cleared. In other words, it has helped to slow down the expansion of the industry into new lands simply to escape the virus. This situation helps environmentally by preserving forest and other lands that might otherwise be cleared for papaya plantings to escape PRSV infection.

Resurgence of Papaya Cultivation on Oahu

As noted earlier, the papaya industry originally was centered on Oahu, but production on that island was largely eliminated by PRSV in the 1950s. The availability of PRSV-resistant papaya provided options for papaya growers on Oahu Island. Prior to the release of transgenic papaya, Oahu growers farmed only small plots of papaya due to the effect of PRSV on production. Growers on Oahu enjoy a niche market, currently growing "Rainbow" and "Laie Gold" papaya for residents in Honolulu and other urban areas of the island. Whereas Oahu grew 50 acres of papaya in 1960, it grew 200 acres of virus-resistant papaya in 2008. This would not have happened without the release and utilization of the virus-resistant transgenic papaya.

Coexistence to Help Continue Production of Nontransgenic Papaya for Japan

Japan for many years had been a lucrative and major export market for "Kapoho." But, Japan had zero tolerance for transgenic papaya. Could both "Rainbow" and "Kapoho" coexist in Puna without contamination?

The following is a description of a program that indeed has served Hawaii well in allowing the shipment of nontransgenic papaya to Japan even though the plants were growing in close proximity to "Rainbow" papaya fields. The main points of the program are at the request of Japanese importers, HDOA developed an Identity Preservation Protocol (IPP) that growers and shippers must adhere to in order to receive an IPP certification letter from HDOA that accompanies the papaya shipment. This is a voluntary program. Papaya shipments with this certification can be distributed in Japan without delay during the time Japanese officials are conducting spot testing to detect contaminating transgenic papaya. In contrast, papaya shipments without this certificate must remain in custody at the port of entry until Japanese officials complete their spot checks for transgenic papaya. Completing the tests may take several days or a week, during which time fruit loses quality and marketability.

Some significant features of the IPP are that the nontransgenic papaya must be harvested from papaya orchards that have been approved by HDOA (Camp 2003). To get approval, every tree in the proposed field must be tested for the transgenic GUS reporter gene that is linked to the virus-resistance gene, and found negative. Trees (nontransgenic) must be separated by at least a 4.5 m papaya-free buffer zone, and new fields to be certified must be planted with papaya seeds that have been produced in approved nontransgenic fields. Tests for detecting transgenic papaya trees in the field are monitored by HDOA and conducted by the applicant who must submit detailed records to HDOA. Before final approval of a field, HDOA will randomly test one fruit from 1 % of papaya trees in the field. If approved by HDOA, fruit from these fields can be harvested. Additionally, the applicant must submit the detailed protocols that will be followed to minimize the chance of contamination of nontransgenic papaya by transgenic papaya. This includes a protocol by the applicant on the random testing of papaya before they are packed for shipment. If the procedures are followed and tests are negative, a letter from HDOA will accompany the shipment stating that the shipment is in compliance with a properly conducted IPP.

This procedure represents a good faith effort by HDOA and applicants to prevent transgenic papaya contamination in shipments of nontransgenic papaya to Japan. It also illustrates meaningful collaboration between Japan and HDOA resulting in continued shipment of nontransgenic papaya to Japan with a minimum of delay once they arrive in Japan, while adhering to the policy that transgenic papaya will not commercially enter Japan until it is deregulated by the Japanese government. These efforts, along with the effectiveness of the transgenic papaya in boosting production of nontransgenic papaya, have allowed Hawaii to maintain significant shipments of the latter to Japan.

Transgene Flow from "Rainbow" to "Kapoho" Under Field Conditions

In the years 2002 to 2009 "Kapoho" and "Rainbow" were actually being grown in adjacent fields and thus represented an excellent opportunity to study transgene flow from "Rainbow" to "Kapoho" under commercial conditions (Fig. 7.8). It should be

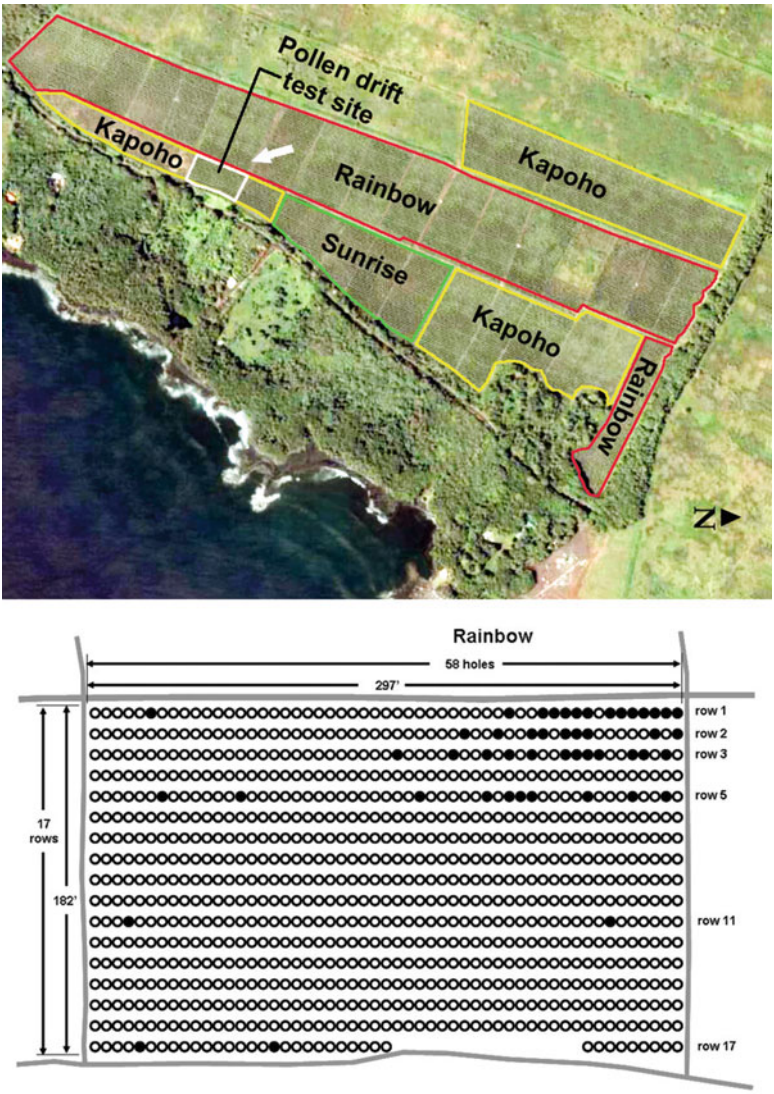

Fig. 7.8 "Rainbow" and "Kapoho" fields growing side by side. Gene flow studies were done in this area. *Black dots* are trees that were samples

noted that a very important feature of Hawaii's papaya industry is that the solo papaya bears only female and hermaphrodite plants and the commercial plantings select for only hermaphrodite plants. Starting in 2004 through 2009, we studied transgene flow of "Rainbow" to "Kapoho" under commercial conditions (Gonsalves et al. 2012).

The results are summarized as follows. In three commercial fields in Kalapana and one in Keaau, transgene flow was not detected in "Kapoho" fields growing adjacent "Rainbow" fields when eight embryos were sampled per fruit of "Kapoho." In a more intense sampling of trees in a "Kapoho" field growing only about 15 feet from "Rainbow" papaya fields in which 90 embryos were analyzed per fruit (roughly 17 % of total seeds in a fruit), the transgene was detected in an average of 0.76 % of embryos tested in fruit from the two rows bordering the "Rainbow" field and none were detected in the fifth row bordering the "Rainbow" field (Fig. 7.8). Using an experimental field condition with high transgenic pollen pressure in which "Kapoho" trees were surrounded by "Rainbow" trees planted about 9 feet away, and 12 embryos were examined per fruit, transgene pollen drift to "Kapoho" averaged 1.3 % of tested embryos. In the same setting with female "Kapoho," 67.4 % of tested embryos were GUS positive. The very low transgene flow to close-by "Kapoho" plantings is likely due to the fact that hermaphrodite trees are used commercially in Hawaii and that these trees are largely self-pollinated before the stigma is exposed to external pollen. In summary, from the practical standpoint, observed transgene pollen drift to hermaphrodite "Kapoho" from neighboring "Rainbow" plants is very low, but as expected the observed pollen drift from "Rainbow" to female "Kapoho" trees is high.

Deregulation and Export of Hawaii's Transgenic Papaya to Canada and Japan

Canada

Canada accounts for 11 % of Hawaii's papaya export market. Canada considers foods derived from GMOs as "novel foods" and import requires review and approval by Health Canada (HC), the government organization responsible for food safety. Health Canada examined the properties of "SunUp" and "Rainbow" papaya and gave approval for food purposes only in January 2003. Labeling of the approved transgenic papaya imported into Canada was not required. The data used for the nutritional assessment of the transformant line 55-1 included fruit composition (total soluble solids, carotenoids, vitamin C, and minerals), which were within the range found in fruit of nontransgenic cultivars grown in Hawaii. In the toxicology assessment, PRSV CP was not considered a "novel" protein due to the history of human consumption of PRSV-infected fruit without adverse health effects (http://www.hc-sc.gc.ca/fn-an/gmf-agm/appro/papaya_e.html).

Japan

Once "Rainbow" was released, it seemed logical that deregulating the papaya in Japan should be the next step to expand the papaya market and to prevent "Rainbow"

from being inadvertently introduced into Japan without approval. Almost immediately after Hawaii's transgenic papaya was released in Hawaii, the PAC asked us investigators to help deregulate the transgenic papaya in Japan. The scientific group of my lab, Manshardt, and Ferreira started work on the deregulation package quite soon after transgenic papaya was released to Hawaii growers in May 1998. Although we would do the work and fill out the petition, the petition would be formally submitted by PAC, and later by HPIA when PAC was discontinued.

Briefly the deregulation process in Japan required getting approval from the Ministry of Agriculture Fisheries and Forestry (MAFF), Ministry of Health Labor and Welfare, and later in 2009 from the Consumer Affairs Agency (CAA). Another aspect is that we scientists did not directly interact with the regulators in Japan. Instead, the petition was submitted to the agencies by consultants in Japan who were selected by PAC and HPIA. The reverse would follow when the Ministry replied to the submitted petition.

MAFF

We started preparations of documents for deregulating the transgenic papaya in Japan shortly after commercialization in the USA. Our first efforts were to get the transgenic papaya approved by MAFF. We used much of the information from our US petition, but very importantly, MAFF personnel themselves actually grew the transgenic papaya under their greenhouse conditions in Japan to assure that our transgenic papaya was essentially as described in the USA. And in December 2000, MAFF gave tentative approval. However, the MAFF case would later be reevaluated in view of the Cartagena protocols, which Japan accepted to follow. After various modifications, the reevaluation was completed and MAFF tentatively approved our petition following the public comment period in May 2010.

MHLW

The group from my laboratory, Steve Ferreira, and Richard Manshardt submitted a formal application to MHLW in 2003. The application followed a protocol composed by MHLW. Over the next 6 years, this initial submission was followed by numerous updated submissions that were aimed at answering questions or comments by the Japanese reviewers. On August 2009, the Food Safety Committee of MHLW tentatively approved our application.

By all standards, approval by MHLW took a long time, and it is reasonable to ask why. In actuality, the Japan ministries normally responded within weeks following our submissions of our original and revised petitions. The bulk of the time was taken up in us doing experiments to satisfy various requests or comments made by the ministries after review of our petitions. The process could have been accelerated if

we had more financial help and scientific manpower. In fact, we had very little funding and scientific manpower to move things along as say, a commercial company would be able to do. But, we expected this situation since we were public sector scientists, and the papaya industry did not have the financial resources to help us. Limited financial help did come from USDA/FAS and from the state of Hawaii in the form of grants. Also, once the MHLW application was tentatively approved in 2009, we really could not do much while the CAA was making its decisions on labeling. Final approval had to await approval of all three agencies.

I would like to acknowledge the key individuals who contributed to the MHLW approval process. They were Steve Ferreira, Karen Pitz, Richard Manshardt, and my post docs (Savarni Tripathi, Jon Suzuki, and Gustavo Fermin) and technicians (Ron Keith and James Carr). Additionally, Gustavo Fermin, while a Cornell University graduate student in my lab, made one of the major contributions in isolating a cDNA clone that contained the complete functional transgene cassette in line 55-1.

Key contributions were also made by the labs of Ray Ming and of Maqs Alam (Shaobin Hou of Alam's lab played a key role) and Takahashi Sugimura of the Maui High Performance Computer Facilities. I'd like to take the liberty to give my view on how the labs got together. Sometime around 2005, Maqs Alam and Tak Sugimura came to see me to discuss their intended work on sequencing the papaya genome. During our discussions, we all realized that their proposed sequencing work could help us characterize the transgene inserts and I could help with the needed biology background. Ray Ming had always expressed to me that he wanted to help in whatever way he could; his lab had already established themselves as leaders on the determination of sex factors in papaya using the "SunUp" as a basic model and had established a BAC library. The three labs cooperated on the sequencing project, and the combination of these efforts resulted in characterizing the three inserts and their border sequences, which was crucial for our Japan deregulation work. The collaborative effort helped each party meet their goals. In fact, a draft genome of "SunUp" was published in *Nature* (Ming et al. 2008), and the deregulation of "Rainbow" in Japan has been accomplished.

Some things we had to determine for the Japan application:

- Number of transgene or vector elements inserts in the genome of the papaya. Indeed line 55-1 has three inserts: the functional CP cassette, a partial *TetA* gene fragment, and a partial *nptII* gene fragment (Suzuki et al. 2008).
- The host sequences that bordered each of the inserts (Suzuki et al. 2008).
- Composition analysis of Rainbow and its nontransgenic counterpart (Tripathi et al. 2011).
- Potential change in composition of fruit at the stage of picking for marketing and at the stage of consumption if the fruit is left to ripen on the tree or if the fruit is picked and then allowed to ripen in storage (Tripathi et al. 2011).
- Detailed studies on the potential of the coat protein gene as an allergen both real and in the context of the border host sequences affecting the allergenicity of the coat protein (Fermin et al. 2011).
- Evidence for safe consumption of the papaya fruit by the public.

Fig. 7.9 Carol Gonsalves and I holding "Rainbow" papaya at the US Embassy in Tokyo, Japan, September 2011

- Detailed history of the transgene constructs and vector elements used to transform the papaya.
- Addressing the potential environmental impact of "Rainbow" papaya if it is grown in Japan either as a commercial crop or as plants that germinate from seeds that are discarded following the consumption of Hawaii grown "Rainbow" papaya that are purchased from markets in Japan.

The CAA was established on September 1, 2009, and the authority for food labeling was transferred from MHLW/MAFF to CAA. Before final deregulation, the labeling requirements for the transgenic papaya had to be worked out with CAA. On August 31, 2011 CAA released the official notification of labeling for Hawaii biotech papaya. This cleared the way for final approval of the Hawaiian papaya. On December 1, 2011 MAFF released the notification that the environmental review of rainbow papaya was completed, and MHLW lifted the sanction to "Rainbow" papaya and released the notification of food safety process being completed which was the go-ahead for the commercial import and distribution of biotech crops for Japanese public. During the first week in September, several boxes of "Rainbow" papaya were sent to the US Embassy in Japan for initial outreach prior to final approval, and Carol and I took part in the outreach efforts (Fig. 7.9). But when would commercial shipments of "Rainbow" be shipped to Japan since markets had not yet been established?

The late Rickie Deniz of Diversified Ag Products Inc in Hilo stepped up to the plate and performed in the "red zone." Rickie came to see me in the middle of

October 2011 regarding possible shipments of "Rainbow" papaya to Costco Japan. In an incredible span of 6 weeks, he worked closely with Jimmy Nakatani of HDOA, and with help from USDA/FAS in Japan, he established a "Rainbow" papaya market with Costco Japan stores. On December 5, Rickie and Jimmy and others accompanied the first commercial shipment of "Rainbow" papaya to Japan. In my book, Rickie Deniz was the ultimate performer of translational biotechnology in the "red zone." Unfortunately, Rickie left this world on March 25, 2012.

A little over 13 years after the commercial release of "Rainbow" papaya in Hawaii, this same papaya was commercially imported into Japan. We have opened the door to the Japan market. Rickie Deniz led the initial effort, but now it remains to be seen if Hawaii's papaya industry will take advantage of the Japan market. They are now in the "red zone."

Current Efforts at Deregulation of Transgenic Papaya in Mainland China

The export market of Hawaiian papaya to mainland China has not been tapped, but this could be a huge market for Hawaii's papaya. The favorable conditions are that China's economy is booming, their standard of living has increased dramatically, and more tourists are coming to Hawaii with direct flights from China.

A collaborative effort between USDA/ARS in Hawaii and University of Hawaii at Manoa is underway to petition China for importation of Hawaii's transgenic papaya for marketing papaya fruit but not for growing in China. The process is somewhat similar to that of Japan. Needless to say, the experience and scientific information we gathered for the successful deregulation of the Hawaiian papaya in Japan will accelerate our efforts for China. Although it took years for us to compile all information for the Japan case, we expect to compile that information for China in a matter of several months. We plan to submit a completed petition to the Ministry of Agriculture in October 2012 for consideration by the National Biosafety Committee in November 2012.

Closing Remarks

The transgenic "Rainbow" and "SunUp" papaya were commercialized in 1998 and are often viewed as a model of translational biotechnology because it really helped to save Hawaii's papaya industry. Over a quarter of century ago when the transgenic papaya approach was started, the future of biotechnology looked extremely bright and predictions were that lots of commercial products would be marketed in the next 15 years. The transgenic commodity crops of corn, soybean, canola, and cotton have been highly successful and currently occupy well over 95 % of the world's transgenic acreage. However, the promise of translational biotechnology with

specialty crops is found desperately wanting as only virus-resistant squash and papaya are grown commercially in the USA. Other specialty crops should have pushed the transgenic papaya off center stage by now. Many have said that the papaya case is unusual and unique. I disagree. It represents the efforts of a small band of scientists simply doing their job to develop and implement a practical control of a virus using current technology. What will the next 25 years of translational biotechnology with specialty crops be like? No doubt, the "science" part of biotechnology will move forward rapidly as it has in the past 25 years. But translational biotechnology in the "red zone" not only relies on science but other characters as will, people's skill, and of course taking advantage of lady luck. Controversy is part of the game and thus, one needs to be willing to deal with it. I hope this review gave some insights on the "inside" story of the Hawaii papaya effort. This time, I will sit by the sideline and see the next 25 years of translational biotechnology unfold in the "red zone."

Acknowledgements I am fortunate to have been associated with many people that contributed to the papaya story. They are too many to name and I undoubtedly missed a number of them in the text of the story. I thank all of you as I complete my involvement in the Hawaii transgenic papaya story and wish good fortunes to those that will continue. I specifically thank Carol my wife who has been with me throughout this entire story. Most importantly, Rickie Deniz, a true friend that I had for such a short time, showed me how to really get things done in the "red zone."

References

Camp III SG (2003) Virus resistant transgenic papaya in Hawaii: a case for technology transfer to lesser developed countries. In: OECD/USAID/ARS conference, Hilo, 22–24 October 2003

Clark MF, Adams AN (1977) Characteristics of the microplate method for enzyme-linked immunosorbent assay for the detection of plant viruses. J Gen Virol 34:475–483

Fermin G, Tennant P, Gonsalves C, Lee D, Gonsalves D (2004) Comparative development and impact of transgenic papayas in Hawaii, Jamaica, and Venezuela. In: Pena L (ed) Transgenic plants: methods and protocols. The Humana Press Inc, Totowa, pp 399–430

Fermin G, Keith RC, Suzuki JY, Ferreira S, Gaskill DA, Pitz KY, Manshardt RM, Gonsalves D, Tripathi S (2011) Allergenicity assessment of the papaya ringspot virus coat protein expressed in transgenic Rainbow papaya. J Agric Food Chem 59(18):10006–10012

Ferreira SA, Mau RFL (1994) A synopsis of the phased plan for eradicating papaya ringspot disease from the Puna district, Island of Hawaii. In: Proceedings of the 30th annual Hawaii Papaya Industry Association conference, Kihei, 23–24 September 1994, pp 7–9

Ferreira SA, Pitz KY, Manshardt R, Zee F, Fitch M, Gonsalves D (2002) Virus coat protein transgenic papaya provides practical control of papaya ringspot virus in Hawaii. Plant Disease 86:101–105

Fitch M, Manshardt R (1990) Somatic embryogenesis and plant regeneration from immature zygotic embryos of papaya (*Carica papaya* L.). Plant Cell Rep 9:320–324

Fitch MMM, Manshardt RM, Gonsalves D, Slightom JL, Sanford JC (1990) Stable transformation of papaya via microprojectile bombardment. Plant Cell Rep 9(4):189–194

Fitch MMM, Manshardt RM, Gonsalves D, Slightom JL, Sanford JC (1992) Virus resistant papaya derived from tissues bombarded with the coat protein gene of papaya ringspot virus. Biotechnology 10:1466–1472

Fuchs M, Gonsalves D (1995) Resistance of transgenic hybrid squash ZW-20 expressing the coat protein genes of zucchini yellow mosaic virus and watermelon mosaic virus 2 to mixed infections by both potyviruses. Biotechnology 13(13):1466–1473

Gonsalves D (1998) Control of papaya ringspot virus in papaya: a case study. Annu Rev Phytopathol 36:415–437

Gonsalves D (2006) Transgenic papaya: development, release, impact, and challenges. Adv Virus Res 67:317–354

Gonsalves D, Ferreira S (2003) Transgenic papaya: a case for managing risks of papaya ringspot virus in Hawaii. Plant Health Progress. doi:10.1094/PHP-2003-1113-03-RV

Gonsalves D, Garnsey SM (1989) Cross protection techniques for control of plant virus diseases in the tropics. Plant Disease 73:592–597

Gonsalves D, Ishii M (1980) Purification and serology of papaya ringspot virus. Phytopathology 70:1028–1032

Gonsalves D, Fuchs M, Klas F, Tennant P (1994) Field assessment of risks when using transgenic papayas, cucurbits, and tomatoes expressing viral coat protein genes. In: Jones DD (ed) Proceedings of 3rd international symposium on the biosafety results of field tests of genetically modified plants and microorganisms. University of California, Division of Agriculture and Natural Sciences, Oakland, 13–16 November 1994, pp 117–127

Gonsalves D, Ferreira S, Manshardt R, Fitch M, Slightom J (1998) Transgenic virus resistant papaya: new hope for control of papaya ringspot virus in Hawaii. APSnet feature story for September 1998 on world wide web. http://www.apsnet.org/education/feature/papaya/Top.htm

Gonsalves C, Lee D, Gonsalves D (2004a) Transgenic virus resistant papaya: The Hawaiian "Rainbow" was rapidly adopted by farmers and is of major importance in Hawaii today. Online. APSnet Feature, American Phytopathological Society (August–September). http://www.apsnet.org/online/feature/rainbow

Gonsalves D, Gonsalves C, Ferreira S, Pitz K, Fitch M, Manshardt R, Slightom J (2004b) Transgenic virus resistant papaya: from hope to reality for controlling of papaya ringspot virus in Hawaii. Online. APSnet Feature, American Phytopathological Society (August–September). http://www.apsnet.org/online/feature/ringspot/ (July 2004)

Gonsalves D, Vegas A, Prasartsee V, Drew R, Suzuk J, Tripathi S (2006) Developing papaya to control papaya ringspot virus by transgenic resistance, intergeneric hybridization, and tolerance breeding. Plant Breed Rev 26:35–78

Gonsalves C, Lee DR, Gonsalves D (2007) The adoption of genetically modified papaya in Hawaii and its implications for developing countries. J Dev Stud 43:177–191

Gonsalves D, Ferreira S, Suzuki J, Tripathi S (2008a) Papaya. In: Kole C, Hall TC (eds) A compendium of transgenic crop plants: tropical and subtropical fruits and nuts, vol 5. Blackwell, Oxford, pp 131–162

Gonsalves D, Suzuki J, Tripathi S, Ferreira S (2008b) Papaya ringspot virus (Potyviridae). In: Mahy B, Van Regenmortel M (eds) Encyclopedia of virology, 5 vols (vol 4), 3rd edn. Elsevier, Oxford, pp 1–8

Gonsalves D, Gonsalves C, Carr J, Tripathi S, Matsumoto T, Suzuki J, Ferreira S, Pitz K (2012) Assaying for pollen drift from transgenic "Rainbow" to nontransgenic "Kapoho" papaya under commercial and experimental field conditions in Hawaii. Trop Plant Biol 5(1):53–160

Isherwood MOJ (1992) Papaya ringspot virus in Puna: chronology of events and eradication program update. In: Proceedings of the 28th annual Hawaii Papaya Industry Association conference, 25–26 September 1992, Honolulu, pp 7–9

Isherwood MOJ (1994) Status of papaya ringspot virus program. In: Proceedings of the 30th annual Hawaii Papaya Industry Association conference, 23–24 September 1994, Kihei, pp 1–3

Lius S, Manshardt RM, Fitch MMM, Slightom JL, Sanford JC, Gonsalves D (1997) Pathogen-derived resistance provides papaya with effective protection against papaya ringspot virus. Mol Breed 3:161–168

Manshardt RM (1998) "UH Rainbow" papaya. University of Hawaii College of Tropical Agriculture and Human Resources Germplasm G-1, 2 pp

Mau RFL, Gonsalves D, Bautista R (1989) Proceedings of the 25th annual Papaya Industry Association conference, Hilo

Ming R, Hou S, Feng Y, Yu Q, Dionne-Laporte A, Saw JH et al (2008) A draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). Nature 452:991–996

Powell-Abel P, Nelson RS, De B, Hoffmann N, Rogers SG, Fraley RT, Beachy RN (1986) Delay of disease development in transgenic plants that express the tobacco mosaic virus coat protein gene. Science 232(4751):738–743

Quemada H, L'Hostis B, Gonsalves D, Reardon IM, Heinrikson R, Hieber EL, Sieu LC, Slightom JL (1990) The nucleotide sequences of the 3′-terminal regions of papaya ringspot virus strains w and p. J Gen Virol 71(1):203–210

Sanford JC, Johnston SA (1985) The concept of parasite-derived resistance – deriving resistance genes from the parasite's own genome. J Theor Biol 113:395–405

Sato S (2011) Japan approved GM papaya. USDA Foreign Agricultural Service Gain Report number JA1048

Souza MT Jr, Tennant PF, Gonsalves D (2005) Influence of coat protein transgene copy number on resistance in transgenic line 63–1 against papaya ringspot virus isolates. Hortscience 40(7): 2083–2087

Strating A (1996) Availability of determination of nonregulated status for papaya lines genetically engineered for virus resistance. Fed Regist 61(48663)

Suzuki JY, Tripathi S, Fermin GA, Jan F-J, Hou S, Saw JH, Ackerman CM, Yu Q, Schatz MC, Pitz KY, Yepes M, Fitch MMM, Slightom JL, Ferreira SA, Salzberg SL, Alam M, Ming R, Moore PH, Gonsalves D (2008) Characterization of insertion sites in Rainbow papaya, the first commercialized transgenic papaya fruit crop. Trop Plant Biol 1:293–309

Tennant PF, Gonsalves C, Ling KS, Fitch M, Manshardt R, Slightom JL, Gonsalves D (1994) Differential protection against papaya ringspot virus isolates in coat protein gene transgenic papaya and classically cross-protected papaya. Phytopathology 84(11):1359–1366

Tennant P, Fermin G, Fitch MM, Manshardt RM, Slightom JL, Gonsalves D (2001) Papaya ringspot virus resistance of transgenic Rainbow and SunUp is affected by gene dosage, plant development, and coat protein homology. Eur J Plant Pathol 107(6):645–653

Tennant PF, Souza M, Fitch MM, Manshardt RM, Slightom JL, Gonsalves D (2005) Line 63–1, a new virus-resistant transgenic papaya. Hortscience 40:1196–1199

Tricoli DM, Carney KJ, Russell PF, McMaster JR, Groff DW, Hadden KC, Himmel PT, Hubbard JP, Boeshore ML, Quemada HD (1995) Field evaluation of transgenic squash containing single or multiple virus coat protein gene constructs for resistance to cucumber mosaic virus, watermelon mosaic virus 2, and zucchini yellow mosaic virus. Biotechnology 13(13):1458–1465

Tripathi S, Suzuki JY, Carr JB, McQuate GT, Ferreira SA, Manshardt RM, Pitz KY, Wall MM, Gonsalves D (2011) Nutritional composition of Rainbow papaya, the first commercialized transgenic fruit crop. J Food Compos Anal 24(2):140–147

Yeh SD, Gonsalves D (1984) Evaluation of induced mutants of papaya ringspot virus for control by cross protection. Phytopathology 74(9):1086–1091

Yeh S-D, Gonsalves D (1994) Practices and perspective of control of papaya ringspot virus by cross protection. In: Harris KF (ed) Advances in disease vector research, vol 10. Springer, New York, pp 237–257

Chapter 8
Molecular Genetic Mapping of Papaya

Jennifer Han and Ray Ming

Introduction: Genetic Mapping Techniques

High-density genetic maps are indispensable tools in marker-assisted selection, QTL mapping, map-based cloning, comparative genomics, and integration of physical, cytomolecular, and genetic maps (Martin et al. 1993; Tanksley et al. 1995; Klein et al. 2000; Paterson et al. 2000; Draye et al. 2001). Low-resolution genetic linkage maps have been reported for numerous plant species, but high-resolution maps, which are the most informative and necessary for genomic analysis, were at one time limited to the model species Arabidopsis (Peters et al. 2001) and major crop species such as maize (Davis et al. 1999), rice (Harushima et al. 1998), tomato, potato (Tanksley et al. 1992; Haanstra et al. 1999), wheat (Boyko et al. 2002), soybean (Keim et al. 1997), and rapeseed (Lombard and Delourme 2001). Increased availability of genomic sequences and decreased sequencing costs have made it practical to develop high-density genetic maps for non-model species such as ryegrass (Bert et al. 1999), non-domesticated *Gossypium* (Becerra Lopez-Lavalle et al. 2011), cacao tree (Risterucci et al. 2000), and water yam (Mignouna et al. 2002).

Many types of morphological and molecular markers are used for genetic mapping. Historically, readily identified morphological markers were used to develop classical genetic linkage maps. Continued improvements in molecular techniques, increased genome sequence availability, and the development of software capable of mapping hundreds of markers in a single run have made it possible to use molecular markers for mapping and allow for efficient high-density genetic mapping. The major types of molecular markers used for genetic mapping include restriction fragment length polymorphisms (RFLP) (Grodzieker et al. 1974), randomly amplified polymorphic DNA (RAPD) markers (Williams et al. 1990), amplified fragment

J. Han • R. Ming (✉)
Department of Plant Biology, University of Illinois at Urbana-Champaign, 1201 W. Gregory Drive, 148 ERML, MC-051, Urbana, IL 61801, USA
e-mail: rming@life.uiuc.edu

R. Ming and P.H. Moore (eds.), *Genetics and Genomics of Papaya*, Plant Genetics and Genomics: Crops and Models 10, DOI 10.1007/978-1-4614-8087-7_8,

length polymorphism (AFLP) markers (Vos et al. 1995), and simple sequence repeats (SSR) (Morgante and Olivieri 1993).

Although extensively used for genetic mapping when the techniques first became available, RFLP and RAPD markers are now obsolete. The RFLP technique involves fragmenting genomic DNA with restriction enzymes, size separation of the fragments, and hybridizing these fragments with DNA probes. Because this technique relies on hybridization, probes are limited to low-copy regions due to the tendency of repeat regions to hybridize to multiple areas. Low-copy regions are achieved by digesting the genomic DNA with a methylation-sensitive restriction enzyme such as *Pst*I. Such digestions enrich the library for single or low-copy regions. RAPD markers, unlike RFLP markers, are not limited to low-copy regions; rather, they are randomly distributed throughout the genome. The RAPD technique is able to detect more polymorphisms compared to RFLP, making it useful for many domesticated crop species with a narrow genetic base, such as papaya (Stiles et al. 1993) and wheat (Chao et al. 1989). RAPD technology is based on PCR amplification of random genomic regions using short ten-base primers. As the RAPD technique is based on random PCR amplification, it is influenced by the concentration of PCR elements, template concentration and quality, and PCR cycling conditions and is therefore difficult to reproduce. Additionally, mismatches between primer and template can produce results that are difficult to interpret. RAPD markers are also less informative than RFLP markers since RAPD markers are typically dominant in nature, rather than the codominant type seen in RFLP studies. Dominant markers are inherently less informative than codominant markers because the heterozygote genotype is masked by the homozygous dominant genotype, which leads to inflated genetic map distances.

Similar to RAPD markers, AFLP markers are anonymous, randomly distributed throughout the genome, and are typically dominant. The anonymous nature of the markers complicates aligning genetic maps with sequence data, but it has the advantage in that no prior sequence knowledge is required. This is beneficial for studying non-model organisms in which limited genomic information is available. On the other hand, SSR markers are sequence based; biased toward low-copy, gene-rich areas; and are usually codominant. Because SSR markers are sequence based, they are ideal for integrating genetic map with physical maps and whole genome sequences, but the development of SSR markers requires prior sequence knowledge. High reproducibility, sensitivity, relatively low cost, and high information content have made these two techniques the current most common techniques used for genetic mapping.

Papaya Genetic Maps

Papaya and the model species *Arabidopsis thaliana* and *Brassica rapa* are both in the order Brassicales. Their respective families, Caricaceae and Brassicaceae, diverged from a common ancestor approximately 72 million years ago (Wikstrom et al. 2001), making papaya an ideal out-group for comparative genomic studies

Table 8.1 Characteristics of five genetic maps of papaya from 1939 to 2009

Type and quantity of markers												
Flower color	Stem color	Sex type	Fruit flesh color	Ringspot virus coat protein	RAPD	AFLP	SSR	No. of loci	Total cM mapped	No. of LG	Average distance (cM)	Source
1	1	1	–	–	–	–	–	3	41	n/a	20.5	Hofmeyr (1939)
–	–	1	–	–	61	–	–	62	999	11	19.6	Sondur et al. (1996)
–	–	1	1	1	–	1,498	–	1,501	3,294	12	2.2	Ma et al. (2004)
–	–	–	1	–	–	–	706	707	1,069	12	1.5	Chen et al. (2007)
–	–	–	1	–	–	277	712	990	945	14	0.9	Blas et al. (2009)

Marker type is the type of either genetic marker or phenotype used to create the genetic map. No. of loci refers to the total number of loci used to produce the map. Total cM mapped is the total cM distance mapped for all linkage groups. No. of LG is the total number of linkage groups generated. Average distance (cM) is the mean distance between adjacent markers

within Brassicaceae. Papaya is an ideal candidate for the development of high-density genetic maps. It has a relatively small genome size of 372 Mb (Arumuganathan and Earle 1991) and a small number of chromosomes ($2n = 18$). It flowers year round, yields ripe fruit in as short as 9 months after planting, produces high amounts of seeds (approximately 1,000 seeds per fruit), and is readily transformed (Fitch et al. 1992), making it an ideal model species to study flower and fruit formation and fruit quality in perennial fruiting trees. Botanically, papaya is not a tree but an herbaceous shrub because it does not have a cambium for forming secondary wood. However, the size, perennial nature, and monopodal form of the papaya plant make it useful to talk of it as a tree. Papaya is a trioecious species with separate male, female, and hermaphrodite trees. A pair of nascent sex chromosomes determines sex; female plants have two X chromosomes; male plants have an XY chromosome pair; and hermaphrodite plants have an X and slightly different Y, Y^h, chromosome pair. The recently evolved sex chromosomes provide a unique opportunity to study the dynamic changes in the early stages of sex chromosome evolution.

The first genetic map for papaya explored three readily identifiable traits: sex type, flower color, and stem color. It had been previously reported that flower color, yellow vs. white, was linked with sex type (Hofmeyr 1938), whereas stem color, purple vs. non-purple, was most likely inherited independently of sex type. This first map spanned a distance of 41 cM and showed that stem color is linked with flower color with 17.3 cM between the loci and distantly linked with sex type at 41 cM away (Hofmeyr 1939). The low resolution of this map made it unable for use in predicting sex type (Table 8.1).

The next genetic map was not developed until almost 60 years later. Unrelated and morphologically distinct lines were selected as parents to maximize polymorphisms in the F_2 mapping population. In spite of these precautions, the extent of morphological differences between the parents did not correlate well with the minimal molecular differences seen in the parents. This map was based on a total of 61 RAPD markers and one morphological marker, sex type (Sondur et al. 1996). The 62 markers mapped to 11 linkage groups (LGs), spanning a total of 999.3 cM, with an average distance of 19.6 cM between adjacent markers (Table 8.1). The dominant nature of RAPD markers resulted in inflated map distances between markers as the heterozygote genotype was masked. Nevertheless, the sex type locus, *Sex1*, was mapped for the first time to LG 1 with eight additional markers. No suppression of recombination was found on any of the linkage groups, including the linkage group containing the sex locus. This is most likely due to the low marker density, and therefore low resolution, of this RAPD map. While recombination was found throughout LG 1, the *Sex1* locus displayed segregation that deviated from expected Mendelian ratios. The two flanking markers, OPT12 and OPT1C, both located approximately 7 cM away, segregated in the expected 3:1 ratio for dominant markers, whereas *Sex1* segregated two hermaphrodite to one female.

The third genetic map was the first high-density map. It was based on 1,498 AFLP markers, papaya ringspot virus coat protein marker (PRSVCO), sex type, and fruit flesh color (Ma et al. 2004). The 1,501 markers mapped to 12 linkage groups spanning 3294.2 cM with a mean distance of 2.2 cM between markers (Table 8.1). Another 269 AFLP markers were mapped to minor linkage groups or remained unlinked. This map was a vast improvement over the previous linkage map, which had an average distance of 19.6 cM between markers. The linkage group numbers were assigned in descending order based on the genetic length of each group, with the exception that LG 1 was named on the basis that it contained the sex locus.

The region surrounding the sex type locus on LG 1, henceforth referred to as the sex-determining region (SDR), displayed severe suppression of recombination. Of the 342 markers mapped to LG 1, 225 (66 %) co-segregated with sex in the expected 2:1 ratio seen in YY sex-lethal systems. Storey (1953) proposed that the sex determination gene and the genes responsible for the characteristics associated with specific sex types, such as peduncle length, flower morphology, and fruit shape, are closely linked. This complex of genes lay in a region where crossing-over is inhibited, thereby causing the entire region to act as a single unit. The suppression of recombination observed at the SDR provided support for this hypothesis. The high percentage of co-segregating markers on LG 1 also suggests considerable sequence divergence between the X and Y^h chromosomes in the SDR.

DNA methylation is a prominent feature of plant genomes. Genome methylation was assessed by comparing the distribution of *Pst*I to *Eco*RI markers. If the genome is not methylated, and if *Pst*I and *Eco*RI restriction sites are randomly distributed throughout the genome, then the frequency of *Pst*I and *Eco*RI markers on the linkage groups should be equivalent. Lower than expected frequency of *Pst*I markers suggests high levels of DNA methylation (hypermethylation), whereas higher than expected frequency of *Pst*I markers suggests low levels of methylation

(hypomethylation). The observed frequency of *Pst*I markers deviated significantly from the expected equal distribution on all linkage groups. There was a significantly low frequency of *Pst*I markers compared to *Eco*RI markers at the SDR on LG 1, suggesting hypermethylation and therefore a low gene density in this region. Hofmeyr proposed over 40 years ago that the SDR contained few active genes (Hofmeyr 1967). The high methylation levels support this hypothesis. The high methylation and low gene paucity in the SDR was later confirmed using fluorescence in situ hybridization, immunofluorescence assays, and detailed BAC analysis (Liu et al. 2004; Yu et al. 2007; Zhang et al. 2008). A total of 59 marker clusters were found distributed throughout all 12 linkage groups. Clusters are typically associated with regions in which recombination is suppressed. It was previously shown that the heterochromatic nature of centromeres leads to marker clusters on genetic maps (Alonso-Blanco et al. 1998; Copenhaver et al. 1999; Vuylsteke et al. 1999). While some of the marker clusters found are likely indicative of centromere location, the disproportionate number of clusters (59) to chromosome number (9) makes it difficult to ascribe specific clusters with centromere location.

Other notable features were the mapping of the fruit flesh color and the papaya ringspot virus coat protein marker. Yellow fruit flesh color is associated with fruit firmness, whereas red fruit flesh color is associated with shorter shelf life, sweeter taste, and better flavor. The fruit flesh color and PRSVCO markers mapped to opposite ends of LG 7. Mapping the gene of interest and finding markers associated with it is the first step toward a directed breeding program based on marker-assisted selection.

While generation of the high-density AFLP genetic map provides a significant step toward using genetic maps for genome analysis, the anonymous nature of AFLP markers makes it difficult to integrate genetic and physical maps. Attempts were made to anchor the AFLP marker data to genome sequence using plate, row, column, and diagonal pools of BAC DNA from the papaya BAC library (Ming et al. 2001), but these attempts resulted in high percentages of false positives (Q. Yu, P.H. Moore, R. Ming, unpublished results). The development of a sequence-based genetic map was necessary for anchoring the genetic map with the physical map and assigning the linkage groups to individual chromosomes. It was toward this end that a sequence-based high-density genetic map was made.

An F_2 population with 54 plants was used to map 706 simple sequence repeat (SSR) markers and the morphological marker fruit flesh color. The SSR markers were generated from the 13.7× coverage SunUp hermaphrodite BAC end sequences and the SunUp female whole genome shotgun sequence. The map produced 12 linkage groups (nine major and three minor linkage groups), comprising a total of 1068.9 cM with a mean of 1.5 cM between markers (Table 8.1) (Chen et al. 2007). The nine major linkage groups correspond to the nine chromosomes of papaya, but the minor linkage groups had insufficient data to map to one of the chromosomes.

Although only 707 markers were used to generate the SSR map, compared to 1,501 markers for the AFLP map, the genetic distance of the SSR map was condensed as SSR markers are inherently more informative. Unlike AFLP markers, SSR markers are usually codominant, allowing for the identification of heterozygotes in a

mapping population. Another factor contributing to decreased marker distance is the genetic distance between the parents used to form the mapping F_2 population. The AFLP genetic map was generated using a cross between cultivars Kapoho (female)×SunUp (pollen donor) (Ma et al. 2004). While these parents are morphologically distinct, they were derived from the same gene pool and therefore genetically very similar. The SSR genetic map was based on a mapping population generated from a cross between AU9 (female)×SunUp (pollen donor), which are more genetically distant from each other than are Kapoho and SunUp. Parents that are more genetically similar display an increase in recombination rates leading to increased genetic distances, whereas parents that are genetically more distant exhibit a decrease in recombination resulting in smaller genetic distances. Another technique used by Chen et al. (2007) to minimize inflating genetic distances was the removal of markers missing more than five data points. Missing data points can inflate genetic distances.

There were 11 sex co-segregating SSR markers mapped to the SDR. SSR marker P3K2981YC0 mapped 1 cM away from the SDR, which is attributed to two missing data points for this marker. LG 1 contains 78 SSR markers, 11 (14 %) of which co-segregate with sex type. Again, as seen on the AFLP map, segregation distortion was apparent in the SDR on LG 1. The major region of distortion was made up of 51 of the 78 markers found on LG 1. The 51 markers include the 11 that segregated 2:1 instead of the expected Mendelian 3:1 ratio. This 2:1 ratio is the same seen on the AFLP genetic map (Ma et al. 2004). While the remaining 27 markers do not segregate 2:1, they do deviate from the expected 3:1. The 11 sex co-segregating markers compose 14 % of the total number of markers found on LG 1. This is a dramatic decrease compared to the previous map that found 66 % of all markers on LG 1 co-segregating with sex type (Ma et al. 2004). This is most likely due to the nature of the markers used in both studies. SSR markers are found in genic regions, while AFLP markers are randomly distributed throughout the genome, regardless of gene content. The SDR is known to be gene poor, accounting for the decrease in percentage of markers that co-segregate with sex (Yu et al. 2007).

Another major region of segregation distortion was found near the end of LG 6, contained 24 markers (22.6 %), and spanned an 8 cM region. The SSR markers in this region are codominant and should segregate 1:2:1, but instead there was a bias toward the homozygous SunUp genotype and a deficiency of the heterozygote genotype. Unfortunately there are no annotated genes in this region; however, Chen et al. (2007) hypothesized that either the genes in this co-segregating region are associated with YY genotype abortion or possibly the function of the genes located here are advantageous when in the homozygous state. The other minor regions of segregation distortion are likely due to genetic or physical characteristics on the chromosomes. This phenomenon of localized sites of low recombination is likely the basis for the inability for integrating the three minor linkage groups with the nine major linkage groups.

Similar to the previous high-density genetic map, the SSR markers were not homogeneously distributed throughout the linkage groups. Clusters of markers were found on each major linkage group. These clusters most likely correspond to

centromere location, but over 20 clusters were found, again making it difficult to attribute specific clusters to centromere location. Gaps greater than 5 cM between markers were also found on all nine major linkage groups.

To integrate the minor linkage groups with the major linkage groups, Blas et al. (2009) enriched the SSR map first reported by Chen et al. (2007) with AFLP markers. The same mapping population used to generate the SSR map was used for this study. In total, 712 SSR, 277 AFLP, and the morphological marker fruit flesh color were used. These markers mapped to nine major and five minor linkage groups spanning a total of 945.2 cM with a mean of 0.9 cM between adjacent markers (Table 8.1). As the same SSR markers were used, the same linkage group designation was used between the two maps.

Contrary to expected, the addition of AFLP markers to the SSR map decreased the overall mapped genetic distance by 123.4 cM (11.5 %). The decrease in map size is attributed to tighter linkage associations between markers. Typically the addition of AFLP markers increases total map length due to the dominant nature of the markers. The AFLP markers also reduced the number of gaps found on linkage groups to 27 gaps and rearranged marker order on eight of the nine linkage groups. It also incorporated six SSR markers that previously had been unlinked to any linkage groups. Two additional minor linkage groups were created; LG 13 consists of entirely AFLP markers, and LG 14 contains 2 AFLP markers and three SSR markers previously mapped to the end of LG 1.

Recombination suppression at the SDR on LG 1 was more apparent on the combined AFLP and SSR map than on the earlier SSR map. In this combined map, 38 markers co-segregated with sex type: 11 SSR and 27 AFLP markers. These 38 markers account for 32 % of all markers mapped to LG 1 compared to 14 % of markers found on the SSR map (Chen et al. 2007). This phenomenon can be explained by the high repetitive sequence content, high methylation, low gene content, and the type of markers used (Liu et al. 2004; Yu et al. 2007, 2008; Ming et al. 2008; Zhang et al. 2008). The high percentage of sex co-segregating AFLP markers is indicative of the high polymorphism rate found in the SDR and further validated the hypothesis of rapid divergence in the SDR between the X and Y chromosomes (Yu et al. 2008).

The SSR genetic map was aligned with the papaya whole genome sequence. Linkage groups 1, 2, and 4 aligned with significantly less genome sequence per cM when compared to the linkage group average (Ming et al. 2008). The increased levels of recombination on LGs 1, 2, and 4 compared to the other linkage groups were thought to contribute to LGs 1, 2, and 4 aligning with less sequence. Addition of AFLP markers to the SSR map decreased the size of the LGs 1, 2, and 4 by 42 %, 33 %, and 5 %, respectively, suggesting the previous linkage group size was inflated due to technical reasons rather than actual increased recombination rates. Additionally, when the SSR genetic map was integrated with the physical map and whole genome sequences, there appeared to be an inversion of marker order on LG 8 when compared against sequence data (Ming et al. 2008; Yu et al. 2009). Enrichment of the SSR generated LG 8 with 40 AFLP markers resulted in a substantial inversion of marker order in a 26.3 cM recombinationally suppressed pericentromeric region.

The new marker order better correlates with genome sequence data. These data show that combining SSR data with AFLP data leads to an improvement in quality of the papaya genetic map. It has been shown in other species that combining different molecular marker types leads to higher accuracy and increased genome coverage (Castiglioni et al. 1999; Peng et al. 2000; Verde et al. 2005).

Integration of Genetic Map with Physical Map, Genome Sequence, and Cytological Data

Integrated genetic and physical maps have many applications including estimation of physical distances between genetic markers and assembling whole genome shotgun sequences. In turn, assembled genome sequences can be used to more precisely map physical distances between genetic markers. There are minor errors within each approach, but the combined genetic and physical map with assembled whole genome sequences provides a way to correct these. Genetic maps can also be used to anchor linkage groups to specific chromosomes using fluorescence in situ hybridization (FISH). Comparisons between cytological data and genetic and physical maps can reveal chromosomal regions that exhibit recombination suppression and areas of increased recombination.

The SSR genetic map developed by Chen et al. (2007) was used to integrate the physical map and the whole genome shotgun sequences. Assuming a genome size of 372 Mb, the genetic map covers a total of 63 % of the whole genome shotgun sequences and 72 % of the physical map (Yu et al. 2009). The integrated map revealed regions of recombination suppression which was most apparent in the SDR on LG 1. The SSR markers used in the genetic map were derived from BAC end sequences (BES) (153 markers), shotgun sequence reads (466 marker), and from assembled shotgun contigs (87 markers) (Chen et al. 2007). Direct anchoring of BES-derived SSR markers to the FingerPrinted Contigs (FPC) anchored 132 Mb (35.5 % of papaya genome sequence) to the genetic map. The remaining markers were aligned to genome sequence using a BLASTN search. The order and orientation of FPC contigs were verified using the shotgun sequence scaffolds.

The integrated map was used to compare genetic distance against physical distance. The ratio between genetic and physical distance was not homogeneous between the linkage groups. Recombination suppression was evident near the center of all major linkage groups, which is indicative of the centromere location. The largest region of recombination suppression spans a distance of about 8 Mb and corresponds to the SDR on LG 1. Using FISH, the SDR was physically mapped to the middle of chromosome 1 (Yu et al. 2007). The regions flanking the SDR gradually increased rates of recombination moving outward from the SDR (Yu et al. 2009). There is a sharp elevation of recombination rates to about seven times the genome average approximately 10 Mb from the SDR. Regions of increased recombination rates were also apparent at one end of LGs 2, 3, 4, and 5 which correspond to telomeric regions.

The combination of FPC, whole genome shotgun sequence scaffolds, chromosome-specific markers, and FISH was used to assign and orientate the linkage group generated by Chen et al. (2007) to specific chromosomes, including the three minor linkage groups (Wai et al. 2010; Zhang et al. 2010). Specifically, LG 10 was mapped to the same chromosome with LG 8 and so are LGs 11 and 9 and LGs 12 and 7 (Fig. 8.1a–c). The modified linkage groups developed by Blas et al. (2009) were able to assign linkage groups to chromosomes because the same SSR markers were used in the genetic maps developed by Blas et al. and Chen et al.

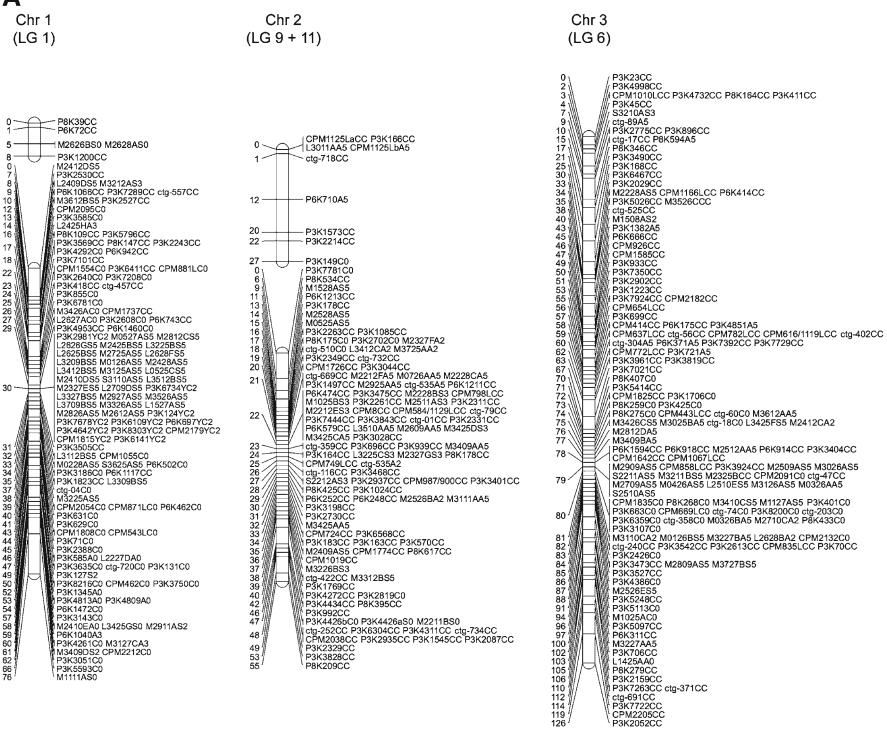

Fig. 8.1 (**a**–**c**) Linkage groups developed by Blas et al. (2009) with chromosome number assignment. Minor linkage groups are orientated against the major linkage groups using chromosome-specific markers and FISH. The numerical scale to the left of each linkage group represents cumulative map distances in centimorgans. Marker names are indicated to the right of each linkage group. SSR markers begin with CPM, ctg, P3K, P6K, or P8K. AFLP markers begin with L, M, or S. For both marker types, a Y at the third to last position specifies sex co-segregating markers. The second to last character represents the parent from which the marker was derived (A, AU9; S, SunUp; C, codominant). The last character provides information about segregation pattern using a χ^2 test (0, does not fit any expected segregation ratio; 2, sex-linked codominant in 2:1 ratio; 3, dominant fitting 3:1 ratio; 5, dominant fitting both 2:1 and 3:1 ratios; C, codominant fitting 1:2:1 ratio). The sex locus is mapped to LG 1 or the X/Y chromosome (chromosome 1). Fruit flesh color is mapped to LG 5 or chromosome 5 [modified with permission from Blas et al. (2009). Copyright © Canadian Science Publishing or its licensors]

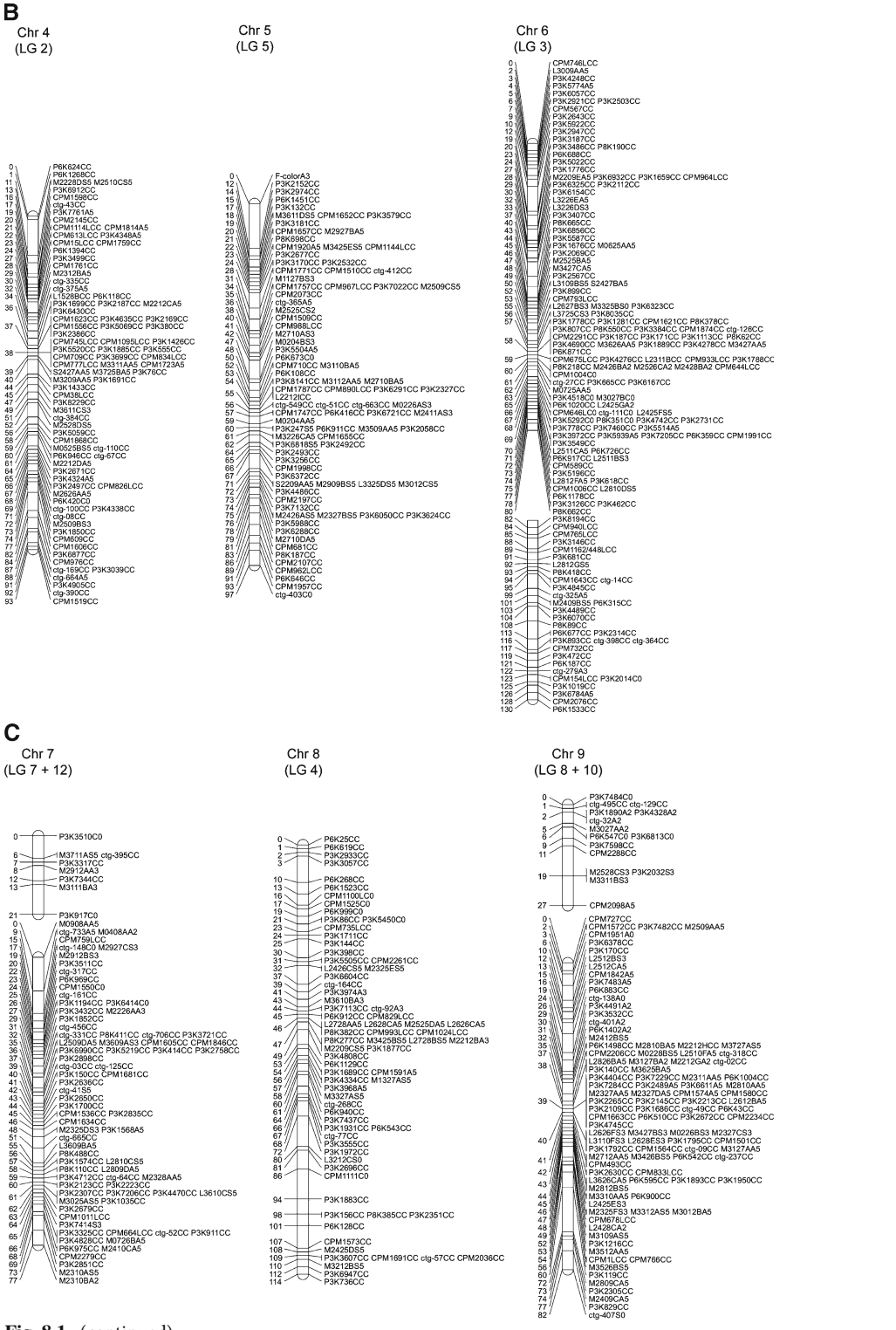

Fig. 8.1 (continued)

Concluding Remarks

The development of high-density papaya genetic maps has revealed much about papaya chromosome characteristics, including centromere locations, recombination hot spots, sequence divergence, and methylation patterns. The integration of the high-density genetic map with genome sequences will prove to be an indispensable tool for identifying candidate genes of agronomic importance. RAPD markers have already been used to map QTLs for plant height, stem diameter, and flowering time (Sondur et al. 1995). SCAR markers were developed from RAPD markers that were able to accurately identify female plants from hermaphrodites and males (Parasni et al. 2000; Deputy et al. 2002; Urasaki et al. 2002). These are the first steps toward using genetic maps for QTL analysis, marker-assisted selection, and map-based cloning.

References

Alonso-Blanco C, Peeters AJM, Koornneef M, Lister C, Dean C et al (1998) Development of an AFLP based linkage map of Ler, Col and Cvi Arabidopsis thaliana ecotypes and construction of a Ler/Cvi recombinant inbred line population. Plant J 14:259–271

Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. Plant Mol Biol Rep 93:208–219

Becerra Lopez-Lavalle LA, Matheson B, Brubaker CL (2011) A genetic map of an Australian wild Gossypium C genome and assignment of homoeologies with tetraploid cultivated cotton. Genome 54(9):779–794

Bert PF, Charme G, Sourdille P, Hayward MD (1999) A high-density molecular map for ryegrass (Lolium perenne). Theor Appl Genet 99:445–452

Blas AL, Yu Q, Che C, Veatch O, Moore PH, Paull RE, Ming R (2009) Enrichment of a papaya high-density genetic map with AFLP markers. Genome 52:716–725

Boyko E, Kalendar R, Korzun V, Fellers J, Korol A et al (2002) A high-density cytogenetic map of the Aegilops tauschii genome incorporating retrotransposons and defense-related genes: insights into cereal chromosome structure and function. Plant Mol Biol 48:767–790

Castiglioni P, Ajmone-Marsan P, van Wijk R, Motto M (1999) AFLP markers in a molecular linkage map of maize: codominant scoring and linkage group distribution. Theor Appl Genet 99(3–4):425–431

Chao S, Sharp PJ, Worland AJ, Warham EJ, Koebner RMD et al (1989) RFLP-based genetic maps of wheat homoeologous group 7 chromosomes. Theor Appl Genet 78:495–504

Chen C, Yu Q, Hou S, Li Y, Eustice M et al (2007) Construction of a sequence-tagged high-density genetic map of papaya for comparative structural and evolutionary genomics in Brassicales. Genetics 177(4):2481–2491

Copenhaver GP, Nickel K, Kuromori T, Benito MI, Kaul S et al (1999) Genetic definition and sequence analysis of Arabidopsis centromeres. Science 286:2468–2474

Davis GL, McMullen MD, Baysdorfer C, Musket T, Grant D et al (1999) A maize map standard with sequenced core markers, grass genome reference points and 932 expressed sequence tagged sites (ESTs) in a 1736-locus map. Genetics 152:1137–1172

Deputy JC, Ming R, Ma H, Liu Z, Fitch MMM et al (2002) Molecular markers for sex determination in papaya (Carica papaya L.). Theor Appl Genet 106:107–111

Draye X, Lin YR, Qian XY, Bowers JE, Burow GB et al (2001) Toward integration of comparative genetic, physical, diversity, and cytomolecular maps for grasses and grains, using the sorghum genome as a foundation. Plant Physiol 125:1325–1341

Fitch MMM, Manshardt RM, Gonsalves D, Slightom JL (1992) Virus resistant papaya derived from tissue bombarded with the coat protein gene of the papaya ringspot virus. Nat Biotechnol 10:1466–1472

Grodzieker T, Williams J, Sharp P, Sambrook J (1974) Physical mapping of temperature sensitive mutations of adenovirus. Cold Spring Harb Symp Quant Biol 39:439–446

Haanstra J, Wye C, Verbakel H, Meijer-Dekens F, van den Berg P et al (1999) An integrated high-density RFLP-AFLP map of tomato based on two Lycopersicon esculentum × L. pennellii F_2 populations. Theor Appl Genet 99:254–271

Harushima Y, Yano M, Shomura A, Sato M, Shimano T et al (1998) A high-density rice genetic linkage map with 2275 markers using a single F2 population. Genetics 148:479–494

Hofmeyr JDJ (1938) Genetical studies of *Carica papaya* L. South Afr J Sci 35:300–304

Hofmeyr JDJ (1939) Sex-linked inheritance in *Carica papaya* L. South Afr J Sci 36:283–285

Hofmeyr JDJ (1967) Some genetic breeding aspects of Carica papaya L. Agron Trop 17:345–351

Keim P, Schupp JM, Travis SE, Clayton K, Zhu T et al (1997) A high-density soybean genetic map based on AFLP markers. Crop Sci 37:537–543

Klein PE, Klein RR, Cartinhour SW, Ulanch PE, Dong J et al (2000) A high-throughput AFLP-based method for constructing integrated genetic and physical maps: progress toward a sorghum genome map. Genome Res 10:789–807

Liu Z, Moore PH, Ma H, Ackerman CM, Ragiba M et al (2004) A primitive Y chromosome in papaya marks incipient sex chromosome evolution. Nature 427:348–352

Lombard V, Delourme R (2001) A consensus linkage map for rapeseed (*Brassica napus* L.): construction and integration of three individual maps from DH populations. Theor Appl Genet 103:491–507

Ma H, Moore PH, Liu Z, Kim MS, Yu Q et al (2004) High-density linkage mapping revealed suppression of recombination at the sex determination locus in papaya. Genetics 166:419–436

Martin G, Brommonschenkel SH, Chunwongse J, Frary A, Ganal M et al (1993) Map-based cloning of a protein kinase gene conferring disease resistance in tomato. Science 26(2): 1432–11436

Mignouna HD, Mank RA, Ellis N, Van Den Bosch N, Asiedu R et al (2002) A genetic linkage map of water yam (Dioscorea alata L.) based on AFLP markers and QTL analysis for anthracnose resistance. Theor Appl Genet 105:726–735

Ming R, Moore PH, Zee F, Abbey CA, Ma H, Paterson AH (2001) Construction and characterization of a papaya BAC library as a foundation for molecular dissection of a tree-fruit genome. Theor Appl Genet 102:892–899

Ming R, Hou S, Feng Y, Yu Q, Dionne-Laporte A et al (2008) The draft genome of the transgenic tropical fruit tree papaya (Carica papaya Linnaeus). Nature 452:991–996

Morgante M, Olivieri AM (1993) PCR-amplified microsatellites as markers in plant genetics. Plant J 1:175–182

Parasni AS, Gupta VS, Tamhankar SA, Ranjekar PK (2000) A highly reliable sex diagnostic PCR assay for mass screening of papaya seedlings. Mol Breed 6:337–344

Paterson A, Bowers JE, Burow MD, Draye X, Elsik CG et al (2000) Comparative genomics of plant chromosomes. Plant Cell 12:1523–1539

Peng J, Korol AB, Fahim T, Roder MS, Ronin YI, Li YC, Nevo E (2000) Molecular genetic maps in wild emmer wheat, Triticum dicoccoides: genome-wide coverage, massive negative interference, and putative quasi-linkage. Genome Res 10(10):1509–1531

Peters JL, Constandt H, Neyt P, Cnops G, Zethof J et al (2001) A physical amplified fragment-length polymorphism map of Arabidopsis. Plant Physiol 127:1579–1589

Risterucci AM, Grivet L, N'Goran JAK, Pieretti I, Flament MH, Lanaud C (2000) A high-density linkage map of Theobroma cacao L. Theor Appl Genet 101:948–955

Sondur SN, Manshardt RM, Stiles JI (1995) Genetics of growth rate and flowering time in papaya (Carica papaya L.). J Quant Trait Loci. http://probe.nalusda.gov:8000/otherdocs/jqtl/1995-04/jqtl8r2.html

Sondur SN, Manshardt RM, Stiles JI (1996) A genetic linkage map of papaya based on randomly amplified polymorphic DNA markers. Theor Appl Genet 93:547–553

Stiles JI, Lemme C, Sondur S, Morshidi MB, Manshardt R (1993) Using randomly amplified polymorphic DNA for evaluating genetic relationships among papaya cultivars. Theor Appl Genet 85:697–701

Storey WB (1953) Genetics of papaya. J Hered 44:70–78

Tanksley SD, Ganal MW, Prince JP, de Vicente MC, Bonierbale MW et al (1992) High density molecular linkage maps of the tomato and potato genomes. Genetics 132:1141–1160

Tanksley SD, Ganal MW, Martin GB (1995) Chromosome landing: a paradigm for map-based gene cloning in plants with large genomes. Trends Genet 11:63–68

Urasaki N, Tokumoto M, Tarora K, Ban Y, Kayano T et al (2002) A male and hermaphrodite specific RAPD marker for papaya (Carica papaya L). Theor Appl Genet 104:281–285

Verde I, Lauria M, Dettori MT, Vendramin E, Balconi C, Micali S et al (2005) Microsatellite and AFLP markers in the Prunus persica [L. (Batsch)] P. ferganensis BC1 linkage map: saturation and coverage improvement. Theor Appl Genet 111(6):1013–1021

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T et al (1995) AFLP: a new tool for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Vuylsteke M, Mank R, Antonise R, Bastiaans E, Senior ML et al (1999) Two high-density AFLP linkage maps of Zea mays L.: analysis of distribution of AFLP markers. Theor Appl Genet 99:921–935

Wai CM, Ming R, Moore PH, Paull RE, Yu Q (2010) Development of chromosome-specific cytogenetic markers and merging of linkage fragments in papaya. Trop Plant Biol 3:171–181

Wikstrom N, Savolainen V, Chase MW (2001) Evolution of the angiosperms: calibrating the family tree. Proc Biol Sci 268:2211–2220

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531–6653

Yu Q, Hou S, Hobza R, Feltus FA, Wang X et al (2007) Chromosomal location and gene paucity of the male specific region on papaya Y chromosome. Mol Genet Genomics 278:177–185

Yu Q, Hou S, Feltus FA, Jones MR, Murray JE, Veatch O et al (2008) Low X/Y divergence in four pairs of papaya sex-linked genes. Plant J 53(1):124–132

Yu Q, Tong E, Skelton RL, Bowers JE, Jones MR et al (2009) A physical map of the papaya genome with integrated genetic map and genome sequence. BMC Genomics 10:371

Zhang W, Wang X, Yu Q, Ming R, Jiang J (2008) DNA methylation and heterochromatinization in the male-specific region of the primitive Y chromosome of papaya. Genome Res 18:1938–1943

Zhang W, Wai CM, Ming R, Yu Q, Jian J (2010) Integration of genetic and cytological maps and development of a pachytene chromosome-based karyotype in papaya. Trop Plant Biol 3:155–170

Chapter 9
Molecular Cytogenetics of Papaya

Wenli Zhang and Jiming Jiang

Abbreviations

BAC	Bacterial artificial chromosome
Chr	Chromosome
DAPI	4′,6-Diamidino-2-phenylindole
FISH	Fluorescence in situ hybridization
gDNA	Genomic DNA
Kbp	Kilo base pairs
LG	Linkage group
Mbp	Mega base pairs
MI	Metaphase I
MSY	Male-specific region of the Y chromosome
5mC	5-Methylcytosine
NORs	Nucleolar organizers
rDNA	Ribosomal DNA
SSR	Simple sequence repeat

Introduction

Papaya (*Carica papaya* L.) is an economically important fruit tree grown in tropical and subtropical regions worldwide. It has a relatively small diploid genome of 372 Mbp with nine pairs of chromosomes (Arumuganathan and Earle 1991). Papaya is trioecious with three sex forms, female, male, and hermaphrodite that are

W. Zhang • J. Jiang (✉)
Department of Horticulture, University of Wisconsin-Madison, 1575 Linden Dr., Madison, WI 53706, USA
e-mail: jjiang1@wisc.edu

R. Ming and P.H. Moore (eds.), *Genetics and Genomics of Papaya*, Plant Genetics and Genomics: Crops and Models 10, DOI 10.1007/978-1-4614-8087-7_9,

controlled by an X/Y chromosome system. Most importantly, significant amounts of genomic resources have been developed in recent years in papaya, including two high-density genetic linkage maps (Ma et al. 2004; Chen et al. 2007), BAC libraries generated from both male and female plants of papaya variety SunUp and from male plant of variety AU9 (Ming et al. 2001; Gschwend et al. 2011), and a whole genome sequence for the transgenic papaya variety SunUp (Ming et al. 2008). Those resources have widely been used for gaining insights into papaya genomic organization at the molecular level. Due to its small diploid genome size, fast growth, and transformability and containing a nascent sex chromosome, papaya is becoming an ideal genomic system to work with and serves as a model plant for studying sex chromosome evolution and sex determination (Liu et al. 2004; Ming et al. 2008).

Molecular cytogenetic techniques can be used to delineate genetic marker order, define contig order and gap size, map heterochromatic chromosome regions as well as heterochromatin–euchromatin boundaries, and reveal molecular genetic events such as deletions, reversions, and genome rearrangements happening at the chromosome level in a straightforward way. Molecular cytogenetic analysis has been applied in guiding genome-related projects in several plant species, including *Arabidopsis thaliana* (Fransz et al. 2000; Stupar et al. 2001), rice (*Oryza sativa*) (Feng et al. 2002; Sasaki et al. 2002), tomato (Koo et al. 2008; Szinay et al. 2008), and potato (Milbourne et al. 2009; Xu et al. 2011). In addition, molecular cytogenetic tools have been used in comparative genomics studies among plant species in *Brassicaceae* (Lysak et al. 2003; Jackson et al. 2000; Mandakova and Lysak 2008), *Solanaceae* (Tang et al. 2008; Iovene et al. 2008; Lou et al. 2010), and *Poaceae* (Zwick et al. 1998; Amarillo and Bass 2007; Draye et al. 2001).

Classical cytogenetics revealed that the papaya genome contains nine pairs of chromosomes (Heilborn 1921), but due to the lack of differences in size and morphology, the papaya chromosomes cannot be distinguished from each other (Kumar et al. 1945; Storey 1953; Datta 1971). Thus, contributions of conventional cytogenetics to papaya genome research are very limited. With the availability and wealth of genetic and genomic resources, molecular cytogenetic techniques have already been successfully applied in cytological characterizations of the papaya genome, including high-resolution karyotyping, heterochromatin distribution, and especially in deciphering cytological features of papaya sex chromosomes. Here, recent advances in the molecular cytogenetics of the papaya genome are summarized, and prospective applications in papaya comparative genomics are briefly discussed.

The Karyotype of Papaya

Conventional Karyotyping

Every species can potentially be cytologically identified based on its unique karyotype, which mainly includes number, size, and shape of the chromosomes as well as the distribution pattern of heterochromatin along the chromosomes. Heilborn (1921)

was the first to report Papaya to be a diploid with nine pairs of chromosomes, and this basic chromosome number remained consistent in other studied species belonging to the Caricaceae family (Darlington and Ammal 1945; Badillo 1971). Morphologically, papaya metaphase chromosomes look similar to each other and are quite small in size between 1.0 and 4.25 μm (Datta 1971). With improvements in cytological preparation techniques, papaya premetaphase chromosomes exhibited well-defined primary and secondary constrictions and distinct chromatin condensation patterns, permitting the identification of minor morphological differences between papaya chromosomes (Araújo et al. 2010). For example, according to the relative position of the primary constriction, seven of the nine chromosome pairs (Chr.1, 2, 3, 4, 5, 7, and 8) appear metacentric, whereas the remaining two pairs are submetacentric (Chr. 6 and 9). Overall, due to the small sizes of papaya chromosomes, conventional karyotyping can provide very limited information for cytological characterization and identification of papaya chromosomes.

Development of Chromosome-Specific Cytogenetic Markers and Integration of Genetic and Cytological Maps in Papaya

Unambiguous identification of individual chromosomes is the most important factor for successful cytogenetic investigations in both animal and plant species. FISH-based chromosome identification has been successfully applied in karyotyping and integration of genetic and cytological maps in a number of plant species (Dong et al. 2000; Cheng et al. 2001, 2002; Kulikova et al. 2001; Howell et al. 2002; Pedrosa et al. 2002; Kim et al. 2005; Zhang et al. 2005; Ren et al. 2009; Wang et al. 2008; Iovene et al. 2011).

In papaya, through screening of a BAC library using SSR markers anchored to individual genetic linkage group, Wai et al. (2010) isolated a set of 12 BACs that are specific to each of the 12 papaya linkage groups (LGs). These BACs were then used as FISH probes for mitotic and meiotic chromosome identification and integration of genetic and chromosome maps (Wai et al. 2010; Zhang et al. 2010). According to the locations of the FISH signals, the 12 LGs were assigned to the nine pairs of mitotic and meiotic chromosomes. For example, at mitotic chromosomes, the FISH signal derived from BAC clone 99D21, which was isolated using SSR marker CPM 2098 from LG10, co-localized with the BAC 35I09, which was selected from SSR marker CPM766 from LG8, indicating that LG8 and LG10 are associated with the same chromosome. Similarly, LG9 and LG11, LG7, and LG12 were assigned to the same chromosome (Wai et al. 2010). The X/Y sex chromosome was designated as chromosome 1 and the other eight chromosomes were aligned in descending order of the pachytene chromosome lengths, thus integrating the 12 linkage groups with the nine pachytene chromosomes (Zhang et al. 2010). For example, linkage groups 2 and 3 correspond to chromosomes 4 and 6, respectively, and LG7/LG12, LG8/LG110, and LG9/LG11 were accordingly mapped to chromosome 7, 9, and 2, respectively.

FISH-Based Karyotyping

A total of 15 best papaya pachytene chromosome spreads prepared from "SunUp" were used to develop a high-resolution karyotype of the papaya genome. Individual pachytene chromosomes in each cell were identified using chromosome-specific BAC clones. The measurements of absolute and relative lengths of all pachytene chromosomes revealed that chromosome 2 was the longest one, with a length of about 58.08 ± 11.84 μm, covering approximately 12.8 % of the total length of all nine chromosomes combined. Chromosome 9 was the shortest, about 41.67 ± 6.34 μm, accounting for approximately 9.2 % of the combined total chromosome length. Based on the putative centromere position assigned on each pachytene chromosome, the two arms of chromosome 2, which was merged from LG9 and LG11, have a similar size, while the sizes of the long arms from the other eight chromosomes are unambiguously longer than their corresponding short arms (Zhang et al. 2010).

Using 45S and 5S ribosomal RNA genes (rDNA) as FISH probes, bicolor FISH revealed that one pair of 45S rDNA FISH signals was located in a medial position close to the primary constriction of a metaphase chromosome, and four strong 5S rDNA FISH signals were observed at interstitial positions of two pairs of chromosomes in *C. papaya* cultivars, "Solo" and "Maradol" (Costa et al. 2008) and "SunUp" (Wai et al. 2010). In "SunUp," the 45S rDNA FISH signal was mapped to chromosome 4 (Zhang et al. 2010). High-resolution pachytene FISH revealed a total of at least 13 5S rDNA loci on chromosomes 1, 3, 5, 8, and 9 in "SunUp" (Fig. 9.1), suggesting that significant amplification of the 5S rRNA genes occurred in the papaya genome (Zhang et al. 2010). Variations in the sizes of these 5S rDNA loci were visible based on the sizes and intensities of the FISH signals, for example, chromosome 8 had two 5S rDNA domains with different sizes, chromosome 5 contained two major and two small sites, and chromosome 3 possessed four small sites. Differential distribution of 5S rDNA was surprisingly detected in two of the four heterochromatic knobs (knob 2 and 4) within the male-specific region of the Y chromosome (MSY) (Fig. 9.1) (Zhang et al. 2008). Thus, the accumulation of heterochromatin within the MSY was partially caused by 5S rDNA amplification (Zhang et al. 2010).

Distribution of Heterochromatin in the Papaya Genome

Heterochromatin is cytologically defined as the portion of chromosomal domains that remain condensed throughout the whole cell cycle. Heterochromatin, which at the molecular level is mainly composed of various types of repetitive DNA sequences, is frequently associated with telomeric, centromeric, and pericentromeric regions, as well as the nucleolar organizers in plant genomes (Topp and Dawe 2006; Copenhaver et al. 1999). In papaya, heterochromatin, represented by domains stained brightly by 4′,6-diamidino-2-phenylindole (DAPI) on pachytene chromosomes, accounts for approximately 17 % of the papaya genome (Ming et al. 2008), indicating that the majority of the papaya genome is largely euchromatic. Zhang et al. (2010) mapped the overall distribution patterns of the heterochromatic

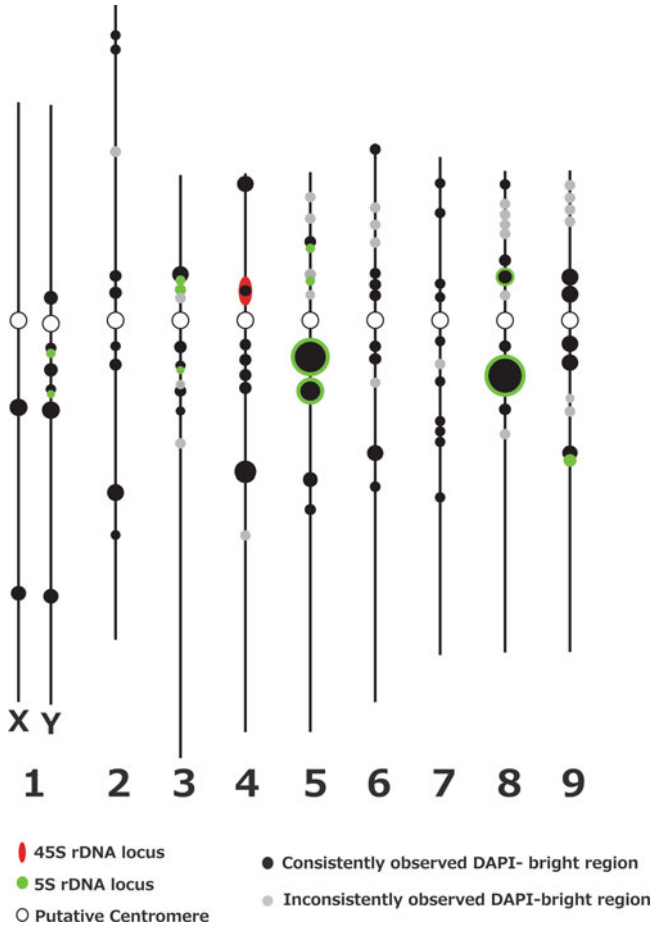

Fig. 9.1 Ideogram illustrating the cytological distribution of DAPI-bright regions on meiotic pachytene chromosomes of papaya. The ideogram is generalized on the basis of observations of 15 pachytene cells. Consistently and inconsistently observed DAPI-bright regions in different cells are indicated by *solid* and *grey circles*, respectively. Putative centromeres are indicated by *open circles*. The 45S rDNA locus is marked by the *red oval* on the short arm of chromosome 4. The 5S rDNA loci distributed in chromosome 3, 5, 8, 9, and the Y chromosomes are marked by the *green circles*; the size of each *green circle* represents the relative amount of the 5S rDNA based on the size and intensity of the FISH signal. The relative length and arm ratio of each chromosome are drawn on basis of the data published by Zhang et al. (2010) [modified with kind permission of Springer Science+Business Media from Zhang et al. (2010)]

regions at each pachytene chromosomes using DAPI staining. The majority of the heterochromatic domains are located at pericentromeric regions on the pachytene chromosomes (Fig. 9.1).

Compared to the amount of heterochromatin found on other pachytene chromosomes, the X chromosome represents the most euchromatic chromosome in the papaya genome. The X chromosome contains only two heterochromatic knobs that

shared with the Y chromosome, with one of the knobs located in the pericentromere and the other at the distal part of the long arm. Papaya pachytene chromosomes contain many knob-like heterochromatin regions, with the two most dominant regions being detected in the pericentromeric domains of chromosomes 5 and 8, which are also associated with 5S rDNA. Four knob-like structures are associated MSY, and these knobs are not present in its counterpart of chromosome X (Zhang et al. 2008) (Fig. 9.1). Two of these four MSY-specific knobs contain 5S rDNA related sequences (Zhang et al. 2010) (Fig. 9.1).

Locations of Centromeres of Papaya Chromosomes

Morphologically, the centromere appears as a constriction on a condensed mitotic chromosome. However, the centromere-associated constriction is often not consistently visible on pachytene chromosomes in some plant species. In well-spread papaya meiotic cells, the primary constrictions of papaya pachytene chromosomes were found to be consistently under-stained by DAPI compared to the pericentromeric heterochromatin. Zhang et al. (2010) predicted the centromere positions of all pachytene chromosomes based on this unique DAPI staining pattern (Fig. 9.1). This prediction, however, needs to be further verified by FISH mapping using centromere-specific sequences that have not been isolated in papaya.

FISH signal derived from the papaya MSY-specific BAC clone 54H01 was located close to the primary constriction of the metaphase Y^h chromosome and the apex of the V-shaped anaphase chromosome, suggesting that the MSY is located close to the centromere (Yu et al. 2007). At pachytene Y^h chromosome, five distinct knobs with bright DAPI staining, named from K1 to K5, respectively, were observed in the MSY, and BACs 12I03, 99O03, 95B12, 52H15, and 85B24 were specifically mapped to K1 to K5, respectively (Zhang et al. 2008). The centromeric position of the Y^h chromosome was accurately mapped by FISH using those five BAC clones on MI bivalent chromosomes. The FISH signals derived from BAC 52H15 were found at the most poleward position of the MI bivalent chromosome compared with the FISH signal from any one of the other four BAC clones, indicating that the centromere of the Y^h chromosome either is likely located at K4 or is immediately adjacent to this knob (Fig. 9.1) (Zhang et al. 2008).

Cytological Structure of the Papaya Y Chromosome

The XY chromosome system is a classical sex determination system associated mammalian species. This system has also been reported in a number of model dioecious plant species, including *Asparagus officinalis*, *Silene latifolia*, *Rumex acetosa*, sorrel, kiwifruit, spinach, hop, and *Marchantia polymorpha* (Jamilena et al. 2008; Negrutiu et al. 2001; Matsunaga and Kawano 2001; Ming et al. 2011). X/Y chromosomes are considered to be derived from a pair of autosomes. Most X/Y chromosomes show cytologically detectable differences in the heterogametic sex. The ancient Y

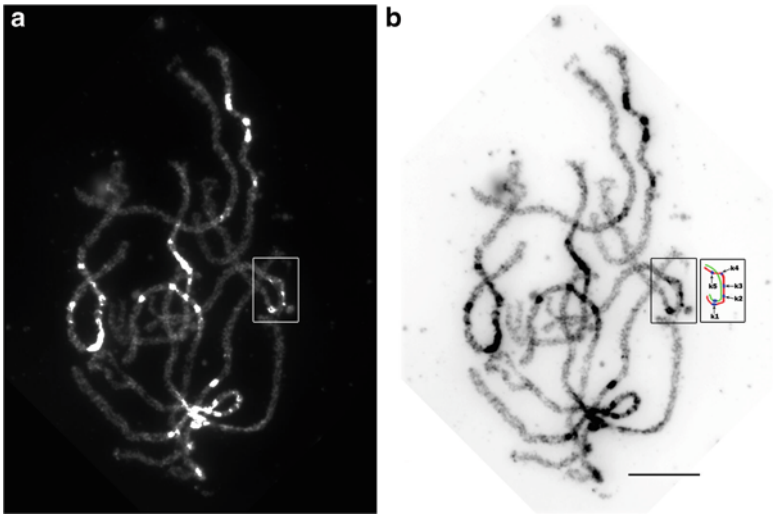

Fig. 9.2 MSY-specific heterochromatin distribution in the papaya sex chromosome. (**a**) A complete pachytene cell stained by DAPI. The MSY containing five knobs with bright DAPI staining is marked by the *rectangle*. (**b**) A black-white image was converted from the same pachytene cell. The diagram in the *right rectangle* shows the (*green*) X and (*red*) Y^h chromosomes and the five (*blue*) knobs. Bars, 10 μm

chromosome in the liverwort *Marchantia polymorpha* is small and largely heterochromatic (Yamato et al. 2007), whereas Y chromosomes in some dioecious plants are larger in size than the other chromosomes in male genomes. Accumulation of transposable elements and satellite DNA has played an important role in the divergence and size enlargement of Y chromosomes (Hobza et al. 2006; Negrutiu et al. 2001; Vyskot and Hobza 2004; Kubat et al. 2008; Charlesworth 2002; Navajas-Pérez et al. 2009; Matsunaga 2006; Jamilena et al. 2008).

Preliminary cytological analysis showed that papaya has a homomorphic pair of X/Y chromosomes (Vyskot and Hobza 2004; Wai et al. 2010; Yu et al. 2007, 2008a, b). The male-specific region of the Y^h chromosome was previously mapped to near the middle of the genetic linkage group (LG1) (Ma et al. 2004), and it was estimated to account for approximately 10 % of the Y chromosome (Liu et al. 2004). The MSY was then physically mapped to the vicinity of the centromere at mitotic and meiotic chromosomes using MSY BAC clones as FISH probes (Yu et al. 2008b; Zhang et al. 2008). Many MSY-specific BACs generated strong FISH signals on Y^h chromosome very faint signals on the corresponding X chromosome, indicating a high level divergence of the DNA sequences located in MSY and its corresponding X chromosome region.

Pachytene chromosome-based cytogenetic analysis revealed that the papaya MSY covers the centromeric region and spanned approximately 13 % of the Y^h chromosome (Zhang et al. 2008). Five distinct knobs specific to the MSY, named from K1 to K5, respectively, were clearly visible on the XY^h pachytene bivalent. Knob 1 was shared by both the X and Y^h chromosomes, the other four knobs were only observed in MSY. The size of the five knobs was variable (Fig. 9.2a, b).

In the majority of meiotic cells observed, knob 1 was the largest, and knob 2 was the smallest. A protruded spot was reproducibly observed around K4 on the XY^h pachytene bivalent, and a twist of paring between the X and Y^h chromosomes was cytological visible between K4 and K5 (Fig. 9.2a, b), suggesting that significantly more DNA accumulated at the MSY around the K4 region than the corresponding domain on the X chromosome (Zhang et al. 2008). The DNA sequences located the five MSY knobs are also significantly more methylated based on immunofluorescence assays using antibodies against 5-methylcytosine (Zhang et al. 2008). Thus, cytological analysis revealed that the DNA sequences within the heterochromatic knobs of MSY are highly divergent and heavily methylated compared with the sequences in the corresponding X chromosomal domains. These results suggest that DNA methylation and heterochromatinization are possible mechanisms responsible for the early stage evolution of sex chromosome.

Outlook

Papaya is a member of the Caricaceae, which consists of six genera and collectively encompasses 35 species. Within the family, except for *Vasconcellea monoica*, *V. cundinamarcensis*, and *C. papaya* that are monoecious and trioecious species, respectively, all other Caricaceae species are regarded as dioecious (Badillo 1971, 1993, 2000). The incipient sex chromosome in papaya makes an exceptionally promising model system for studies of origin and evolution of sex chromosomes. Physical mapping and genomic sequencing of the MSY and the corresponding region of the X chromosomes are ongoing and the data are expected to be available to the public soon. Using the genomic resources developed in papaya variety "SunUp" as a reference, future comparative cytological studies of papaya and its wild relatives (both monoecious and dioecious) will provide molecular cytological evidences to clarify the mechanisms that led to the evolution of sex chromosomes and dioecy. Centromere-specific DNA sequence can possibly be cloned in the future, which will provide a clue of a possible coevolution of centromere and MSY. Thus, papaya is and will continue to be an excellent model species for comparative genomic studies in both molecular cytogenetics and molecular biology.

References

Amarillo FI, Bass HW (2007) A transgenomic cytogenetic sorghum (Sorghum propinquum) bacterial artificial chromosome fluorescence in situ hybridization map of maize (Zeamays L.) pachytene chromosome 9, evidence for regions of genome hyperexpansion. Genetics 177:1509–1526

Araújo FS, Carvalho CR, Clarindo WR (2010) Genome size, base composition and karyotype of Carica papaya L. Nucleus 53:25–31

Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. Plant Mol Biol Rep 9:208–218

Badillo VM (1971) Monografia de la familia Caricaceae. Publicada por la Asociacion de Profesores. Univ Centr Venezuela, Maracay

Badillo VM (1993) Caricaceae. Segundo esquema. Rev Fac Agron Univ Centr Venezuela 43: 1–111

Badillo VM (2000) *Carica* L. vs *Vasconcella* St. Hil. (Caricaceae): con la rehabilitación de este último. Ernstia 10:74–79

Charlesworth D (2002) Plant sex determination and sex chromosomes. Heredity 88:94–101

Chen CX, Yu QY, Hou SB, Li YJ, Eustice M et al (2007) Construction of a sequence-tagged high-density genetic map of papaya for comparative structural and evolutionary genomics in brassicales. Genetics 177:2481–2491

Cheng Z, Buell CR, Wing RA, Gu M, Jiang J (2001) Toward a cytological characterization of the rice genome. Genome Res 11:2133–2141

Cheng ZK, Dong FG, Langdon T, Ouyang S, Buell CR et al (2002) Functional rice centromeres are marked by a satellite repeat and a centromere-specific retrotransposon. Plant Cell 14:1691–1704

Copenhaver GP, Nickel K, Kuromori T, Benito MI, Kaul S et al (1999) Genetic definition and sequence analysis of *Arabidopsis* centromeres. Science 286:2468–2474

Costa FR, Pereira TNS, Hodnett GL, Pereira MG, Stelly DM (2008) Fluorescent in situ hybridization of 18S and 5S rDNA in papaya (*Carica papaya l.*) and wild relatives. Caryologia 61:411–416

Darlington CD, Ammal EKJ (1945) Chromosome atlas of cultivated plants. George Allen and Unwin Ltd, London

Datta P (1971) Chromosomal biotypes of Carica papaya L. Cytologia 36:555–562

Dong F, Song J, Naess SK, Helgeson JP, Gebhardt C et al (2000) Development and applications of a set of chromosome-specific cytogenetic DNA markers in potato. Theor Appl Genet 101:1001–1007

Draye X, Lin YR, Qian XY, Bowers JE, Burow GB et al (2001) Toward integration of comparative genetic, physical, diversity, and cytomolecular maps for grasses and grains, using the sorghum genome as a foundation. Plant Physiol 125:1325–1341

Feng Q, Zhang YJ, Hao P, Wang SY, Fu G et al (2002) Sequence and analysis of rice chromosome 4. Nature 420:316–320

Fransz PF, Armstrong A, de Jong JH, Parnell LD, van Drunen C et al (2000) Integrated cytogenetic map of chromosome arm 4S of A. thaliana: structural organization of heterochromatic knob and centromere region. Cell 100:367–376

Gschwend AR, Yu QY, Moore P, Saski C, Chen CX et al (2011) Construction of papaya male and female BAC libraries and application in physical mapping of the sex chromosomes. J Biomed Biotechnol. Epub 2011 July 5. doi:10.1155/2011/929472

Heilborn O (1921) Taxonomical and cytological studies on cultivated Ecuadorian species of Carica. Ark Bot 17:1–16

Hobza R, Lengerova M, Svoboda JL, Kubekova H, Kejnovsky E et al (2006) An accumulation of tandem DNA repeats on the Y chromosome in *Silene latifolia* during early stages of sex chromosome evolution. Chromosoma 115:376–382

Howell EC, Barker GC, Jones GH, Kearsey MJ, King GJ et al (2002) Integration of the cytogenetic and genetic linkage maps of *Brassica oleracea*. Genetics 161:1225–1234

Iovene M, Wielgus SM, Simon PW, Buell R, Jiang JM (2008) Chromatin structure and physical mapping of chromosome 6 of potato and comparative analyses with tomato. Genetics 180:1307–1317

Iovene M, Cavagnaro PF, Senalik D, Buell CR, Jiang JM et al (2011) Comparative FISH mapping of *Daucus* species (Apiaceae family). Chromosome Res 19:493–506

Jackson SA, Cheng ZK, Wang ML, Goodman HM, Jiang JM (2000) Comparative fluorescence in situ hybridization mapping of a 431-kb Arabidopsis thaliana bacterial artificial chromosome contig reveals the role of chromosomal duplications in the expansion of the Brassica rapa genome. Genetics 156:833–838

Jamilena M, Mariotti B, Manzano S (2008) Plant sex chromosomes: molecular structure and function. Cytogenet Genome Res 120:255–264

Kim JS, Klein PE, Klein RR, Price HJ, Mullet JE et al (2005) Chromosome identification and nomenclature of *Sorghum bicolor*. Genetics 169:1169–1173

Koo DH, Jo SH, Bang JW, Park HM, Lee SH et al (2008) Integration of cytogenetic and genetic linkage maps unveils the physical architecture of tomato chromosome 2. Genetics 179: 1211–1220

Kubat Z, Hobaz R, Vyskot B, Kejnovsky E (2008) Microsatellite accumulation on the Y chromosome in *Silene latifolia*. Genome 51:350–356

Kulikova O, Gualtieri G, Geurts R, Kim DJ, Cook D et al (2001) Integration of the FISH pachytene and genetic maps of *Medicago truncatula*. Plant J 27:49–58

Kumar LSS, Abraham A, Srinivasan VK (1945) The cytology of Carica papaya Linn. Indian J Agric Sci 15:242–253

Liu ZY, Moore PH, Ma H, Ackerman CM, Ragiba M et al (2004) A primitive Y chromosome in papaya marks incipient sex chromosome evolution. Nature 427:348–352

Lou QF, Iovene M, Spooner DM, Buell CR, Jiang JM (2010) Evolution of chromosome 6 of *Solanum* species revealed by comparative fluorescence in situ hybridization mapping. Chromosoma 119:435–442

Lysak MA, Pecinka A, Schubert I (2003) Recent progress in chromosome painting of Arabidopsis and related species. Chromosome Res 11:195–204

Ma H, Moore PH, Liu ZY, Kim MS, Yu QY et al (2004) High-density linkage mapping revealed suppression of recombination at the sex determination locus in papaya. Genetics 166:419–436

Mandakova T, Lysak MA (2008) Chromosomal phylogeny and karyotype evolution in x = 7 crucifer species (Brassicaceae). Plant Cell 20:2559–2570

Matsunaga S (2006) Sex chromosome-linked genes in plants. Genes Genet Syst 81:219–226

Matsunaga S, Kawano S (2001) Sex determination by sex chromosomes in dioecious plants. Plant Biol 3:481–488

Milbourne D, Orjeda G, Sagredo B, Tang XM (2009) Sequencing the potato genome: outline and first results to come from the elucidation of the sequence of the world's third most important food crop. Am J Potato Res 86:417–429

Ming R, Moore PH, Zee FT, Abbey CA, Ma H et al (2001) Construction and characterization of a papaya BAC library as a foundation for molecular dissection of a tree-fruit genome. Theor Appl Genet 102:892–899

Ming R, Hou SB, Feng Y, Yu QY, Dionne-Laporte A et al (2008) The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya Linnaeus*). Nature 452:991–996

Ming R, Bendahmane A, Renner SS (2011) Sex chromosomes in land plants. Annu Rev Plant Biol 62:485–514

Navajas-Pérez R, Schwarzacher T, Rejón MR, Garrido-Ramos MA (2009) Molecular cytogenetic characterization of Rumex papillaris, a dioecious plant with an XX/XY1Y2 sex chromosome system. Genetica 135:87–93

Negrutiu I, Vyskot B, Barbacar N, Georgiev S, Moneger F (2001) Dioecious plants. A key to the early events of sex chromosome evolution. Plant Physiol 127:1418–1424

Pedrosa A, Sandal N, Stougaard J, Schweizer D, Bachmair A (2002) Chromosomal map of the model legume *Lotus japonicus*. Genetics 161:1661–1672

Ren Y, Zhang ZH, Liu JH, Staub JE, Han YH et al (2009) An integrated genetic and cytogenetic map of the cucumber genome. PLoS One 4:e5795

Sasaki T, Matsumoto T, Yamamoto K, Sakata K, Baba T et al (2002) The genome sequence and structure of rice chromosome 1. Nature 420:312–316

Storey WB (1953) Genetics of papaya. J Hered 44:70–78

Stupar RM, Lilly JW, Town CD, Cheng ZK, Kaul S et al (2001) Complex mtDNA constitutes an approximate 620-kb insertion on *Arabidopsis thaliana* chromosome 2: Implication of potential sequencing errors caused by large-unit repeats. Proc Natl Acad Sci USA 98:5099–5103

Szinay D, Chang SB, Khrustaleva L, Peters S, Schijlen E et al (2008) High-resolution chromosome mapping of BACs using multi-colour FISH and pooled-BAC FISH as a backbone for sequencing tomato chromosome 6. Plant J 56:627–637

Tang XM, Szinay D, Lang C, Ramanna MS, van der Vossen EA et al (2008) Cross-species bacterial artificial chromosome-fluorescence in situ hybridization painting of the tomato and potato chromosome 6 reveals undescribed chromosomal rearrangements. Genetics 180:1319–1328

Topp CN, Dawe RK (2006) Reinterpreting pericentromeric heterochromatin. Curr Opin Plant Biol 9:647–653

Vyskot B, Hobza R (2004) Gender in plant: sex chromosomes are emerging from the fog. Trends Genet 20:432–438

Wai CM, Ming R, Moore PH, Paul RE, Yu QY (2010) Development of chromosome-specific cytogenetic markers and merging of linkage fragments in papaya. Trop Plant Biol 3:171–181

Wang K, Guan B, Guo WZ, Zhou BL, Hu Y et al (2008) Completely distinguishing individual A-genome chromosomes and their Karyotyping analysis by multiple bacterial artificial chromosome-fluorescence in situ hybridization. Genetics 178:1117–1122

Xu X, Pan SK, Cheng SF, Zhang B, Mu DS et al (2011) Genome sequence and analysis of the tuber crop potato. Nature 475:189–195

Yamato KT, Ishizaki K, Fujisawa M, Okada S, Nakayama S, Fujishita M et al (2007) Gene organization of the liverwort Y chromosome reveals distinct sex chromosome evolution in a haploid system. Proc Natl Acad Sci USA 104:6472–6477

Yu QY, Hou SB, Hobza R, Feltus FA, Wang X et al (2007) Chromosomal location and gene paucity of the male specific region on papaya Y chromosome. Mol Genet Genomics 278:177–185

Yu QY, Hou S, Feltus FA, Jones MR, Murray J et al (2008a) Low X/Y divergence in four pairs of papaya sex-liked genes. Plant J 53:124–132

Yu QY, Navajas-Pérez R, Tong E, Robertson J, Moore PH et al (2008b) Recent origin of dioecious and gynodioecious Y chromosomes in papaya. Trop Plant Biol 1:49–57

Zhang D, Yang QY, Bao WD, Zhang Y, Han B et al (2005) Molecular cytogenetic characterization of the *Antirrhinum majus* genome. Genetics 169:325–335

Zhang WL, Wang XE, Yu QY, Ming R, Jiang JM (2008) DNA methylation and heterochromatinization in the male-specific region of the primitive Y chromosome of papaya. Genome Res 18:1938–1943

Zhang WL, Wai CM, Ming R, Yu QY, Jiang JM (2010) Integration of genetic and cytological maps and development of a pachytene chromosome-based karyotype in papaya. Trop Plant Biol 3:166–170

Zwick MS, Islam-Faridi MN, Czeschin DC Jr, Wing RA, Hart GE et al (1998) Physical mapping of the liguleless linkage group in Sorghum bicolor using rice RFLP-selected sorghum BACs. Genetics 148:1983–1992

Chapter 10
Physical Map of Papaya Genome

Qingyi Yu

Introduction

The high-throughput and low-cost next-generation sequencing technologies make genome sequencing possible for non-model species. The typical approaches of genome sequencing include either "BAC-by-BAC" or "whole-genome shotgun sequencing" approaches. A high-quality physical map plays essential roles in both sequencing strategies (The Arabidopsis Genome Initiative 2000; Schnable et al. 2009; The Tomato Genome Consortium 2012; Yu et al. 2002). The key step in "BAC-by-BAC" sequencing projects is to identify BAC clones on the "minimal tiling path," which heavily relies on an accurate BAC-based physical map. In whole-genome shotgun sequencing projects, a physical map, in combination with BAC end sequences (BES), provides anchor points to build scaffolds, especially for large repeat regions. Physical maps provide a valuable genomic resource not only for genome sequence assembly but also for studying genome structure and organization and positional cloning of genes associating with economically and agronomically important traits. In addition, physical maps can be used for comparative genomics studies of closely related species to enhance our understanding of the evolutionary process.

Genetic and physical maps are constructed using independent methods and are not directly related to each other. Various methods have been developed to integrate the genetic and physical maps, including (1) genetically mapping BAC-anchored markers (Yu et al. 2009; Troggio et al. 2007; Mun et al. 2006), (2) BAC pooling with PCR screening (Klein et al. 2000), and (3) hybridization of overgo probes (Yüksel et al. 2005) (Fig. 10.1a–c).

Among these three methods, genetically mapping BAC-anchored markers is the most commonly used. Genetic and physical maps are independent genomic

Q. Yu (✉)
Texas A&M AgriLife Research Center at Dallas, Texas A&M University System, 17360 Coit Road, Dallas, TX 75252, USA
e-mail: qyu@ag.tamu.edu

R. Ming and P.H. Moore (eds.), *Genetics and Genomics of Papaya*, Plant Genetics and Genomics: Crops and Models 10, DOI 10.1007/978-1-4614-8087-7_10,

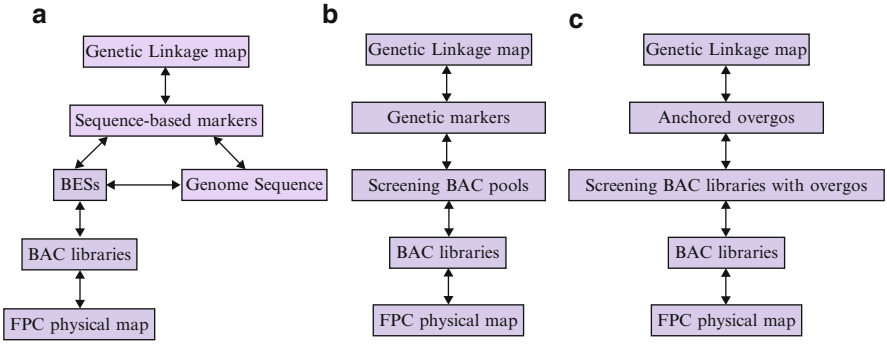

Fig. 10.1 The three major methods for integrating genetic and physical maps. (**a**) Genetically mapping BAC-anchored markers. (**b**) BAC pooling with PCR screening. (**c**) Hybridization of overgo probes

Table 10.1 Comparison of the three papaya genetic maps

	AFLP-based map	SSR-based map	SSR-AFLP map
Population	Kapoho × SunUp	AU9 × SunUp	AU9 × SunUp
Marker types	AFLP	SSR	SSR and AFLP
No. of markers	1,501	707	990
Length of map (cM)	3,294.2	1,069.9	945.2
No. of linkage groups	12	12	14
Reference	Ma et al. (2004)	Chen et al. (2007)	Blas et al. (2009)

resources and complementary to one another. The integrated map could minimize errors and enhance the quality of these resources. The integrated maps with BES allow assignment of genome sequence contigs on linkage groups. The integrated genomic resources provide a simplified and efficient interface for data mining and bioinformatics analysis.

Significant progress has been made in recent years in developing genomic resources to expedite genetic research for improvement of papaya. Three bacterial artificial chromosome (BAC) libraries, one for each sex type, were constructed, providing the foundation for studying papaya genome structure and organization (Ming et al. 2001; Gschwend et al. 2011). BAC ends of the entire hermaphrodite papaya BAC library were sequenced (Lai et al. 2006; Ming et al. 2008; Yu et al. 2009), providing the first glimpse of the sequence composition of the papaya genome. Three high-density genetic linkage maps have been constructed (Ma et al. 2004; Chen et al. 2007; Blas et al. 2009) (Table 10.1), providing essential tools for comparative genomic analysis, marker-assisted selection, and genomic dissection of complex traits. The first high-density genetic map of papaya was constructed with 1,501 markers, including 1,498 amplified fragment length polymorphism (AFLP) markers, 2 morphological markers, and one transgenic marker (Ma et al. 2004). Due to the anonymous nature of AFLP markers, this map cannot be used to

integrate the genetic and physical maps and is not suitable for aligning papaya genome sequence to linkage groups. To overcome this limitation, highly informative sequence-based simple sequence repeat (SSR) markers were used to construct the second high-density genetic map. This sequence-tagged genetic map contains 707 markers including 706 sequence-based SSR markers and one morphological marker (Chen et al. 2007). The SSR markers were developed from either BES or whole-genome shotgun sequence reads (Chen et al. 2007); and this map has contributed to the integration of genetic and physical maps and the genome sequences (Yu et al. 2009). A BAC-based physical map of papaya was constructed using the high-information-content fingerprinting method (Yu et al. 2009). The draft genome sequence of papaya was published in 2008 and contains 271 Mb of contig sequence and 370 Mb of scaffolds with embedded gaps (Ming et al. 2008). The physical map was integrated with the genetic map and genome sequence through mapping micro-satellite markers derived from BES and the whole-genome shotgun sequences (Yu et al. 2009). This integrated map provides essential information for studying the papaya genome structure and comparative genomics.

Papaya BAC Libraries

Bacterial artificial chromosome (BAC) libraries, containing large genomic fragments with deep coverage, have served as valuable research tools for physical mapping, positional cloning, and genome sequencing. The first papaya BAC library was constructed from a hermaphrodite plant of inbred cultivar "SunUp" using *Hin*dIII partial digestion (Ming et al. 2001) (Table 10.2). This BAC library consists of 39,168 clones with an average insert size of 132 kb, providing 13.7× papaya genome equivalents (Ming et al. 2001). This BAC library has been used for cloning flower development (Yu et al. 2005, 2008; Ackerman et al. 2008) and fruit flesh-color genes (Skelton et al. 2006; Blas et al. 2010) and characterization of transgenic insertions in genetically modified papaya (Suzuki et al. 2008). The entire library was end sequenced (Lai et al. 2006; Yu et al. 2009). After trimming the vector sequences and removing the clones with cross-well contamination, a total of 67,179 BES remained with an average length of 666 bp. The total length of the cleaned BES is 44,725,370 bp, accounting for 12 % of the papaya genome. The BES played essential roles in integration of genetic and physical maps and in assembling the genome sequence.

Table 10.2 The characteristics of the three papaya BAC libraries

	Hermaphrodite BAC lib.	Female BAC lib.	Male BAC lib.
Genotype	SunUp, hermaphrodite	SunUp, female	AU9, male
Restriction enzyme	*Hin*dIII	*Bst*YI	*Eco*RI
No. of clones	39,168	36,864	55,296
Average insert size (kb)	132	104	101
Genome equivalents	13.7	10.3	15.0

Two additional papaya BAC libraries, one for "SunUp" female and the other for "AU9" male plants (Table 10.2), were constructed to assist physical mapping of the male-specific Y chromosome (MSY) region and its corresponding region on the X chromosome (Gschwend et al. 2011). These two independent BAC libraries used two different restriction enzymes, *Eco*RI and *Bst*YI, to increase representative genome coverage. The female BAC library consists of 36,864 clones with an average insert size of 104 kb, providing 10.3× genome equivalents (Gschwend et al. 2011). The male BAC library consists of 55,296 clones with an average insert size of 101 kb, providing 15.0× genome equivalents (Gschwend et al. 2011). The three papaya BAC libraries of three sex types provided essential genomic resources for physical mapping of the hermaphrodite-specific region of the Y^h chromosome (HSY), the male-specific region of the male Y chromosome (MSY), and the corresponding region on the X chromosome (Na et al. 2012).

BAC Fingerprinting and Contig Assembly

The basic strategy for construction of a physical map is to identify overlapping clones based on shared restriction fragments or DNA markers. Various methods have been developed to create physical maps. Briefly, these methods can be grouped into either PCR/hybridization-based or DNA fingerprinting-based approaches. Compared with PCR/hybridization-based methods, fingerprinting-based methods are rapid and better suited to map at a whole-genome level. The high-information-content fingerprinting (HICF) method has proved to be the most effective method owing to its high resolution and high throughput (Luo et al. 2003).

The entire library of the hermaphrodite papaya BAC library was fingerprinted using the HICF method. After the 1st round of fingerprinting, the failed clones were repeated twice to obtain successful fingerprints for 38,522 BAC clones (98.4 % of the BAC library). After excluding the clones with no or small inserts (referred as less than 20 true fragments after editing), and cross-contamination (70 % or higher shared fragments for neighboring clones), the remaining 30,824 fingerprints (78.7 % of the BAC library) were subjected to contig assembly using the FPC (FingerPrinted Contigs) program (Soderlund et al. 1997).

A total of 26,466 BAC clones were assembled into 963 contigs, while 4,358 clones remained as singletons. The average number of fragments (bands) of each clone was 69.4. Considering the clones with no or small inserts were excluded from contig assembly and the average insert size of the BAC clones used for contig assembly should be higher than 132 kb, the BAC clones mapped on contigs (26,466 BAC clones) represented at least 9.4× genome equivalents of the papaya genome. On average, each contig contains 27.5 BAC clones. The longest contig contains 1,571 consensus bands, which is about 0.7 % of the total length of the FPC contigs.

Overgo and Single-Copy Probe Hybridization

Overgo probes consist of two complementary oligonucleotides that can anneal at the 3′-end to generate a double-stranded probe with labeled deoxynucleotides (Ross et al. 1999). Overgo probes are designed to bind to conserved regions between two genomes at regular intervals. The overgo hybridization system was developed for large-scale BAC library screening in a high-throughput format to anchor EST sequences, or genetic markers, on physical maps. Overgo hybridization has been successfully used for integration of genetic and physical maps in cotton, cacao and peanut (Lin et al. 2010; Saski et al. 2011; Yüksel et al. 2005).

Papaya and *Arabidopsis* belong to the order Brassicales and shared the last common ancestor approximately 72 MYA (Wikström et al. 2001). Unlike *Arabidopsis*, papaya has not undergone recent whole-genome duplication events (Ming et al. 2008; Lyons et al. 2008a, b; Tang et al. 2008). To facilitate comparative and evolutionary genomic study within Brassicales, a total of 2,277 overgo probes representing conserved sequences in *Arabidopsis* and genetically mapped Brassica loci were tested against 36,864 papaya BACs. A total of 1,329 overgos (58 % of overgos designed) detected positive BACs in papaya. After eliminating low-quality data, the remaining 1,181 overgos were anchored on the papaya FPC map. Among these overgos, 756 (64.0 %) hit single contigs; the average number of clones per overgo is 6.0.

In addition to overgo probes, 16 probes designed from single-copy loci of papaya were used to screen the papaya BAC library. A total of 153 positive BAC clones were obtained. These 16 single-copy probes were anchored on the FPC map. All the probes except papaya *Pistillata* (*PI*) gene hit single contigs. The probe of papaya *Pistillata* (*PI*) gene hit two contigs, ctg1350 and ctg577. Ctg577 contains only three clones and all of them are PI positive. Thus, ctg577 should be part of the ctg1350, but the high stringency setting failed to merge them into a single contig.

These anchored overgos and single-copy genes further improved the quality of the physical map and provide direct links among papaya, *Arabidopsis,* and *Brassica* genomes for comparative genomic research. These anchored overgos and single-copy genes could help reveal synteny and rearrangements in target regions of these genomes, particularly in these recently duplicated genomes of *Arabidopsis* and *Brassica*.

Integration of Physical Map, Genetic Map, and Genome Sequence

Physical map, genetic map, and genome sequence are independent genomic resources. Among these resources, physical map forms an intermediate between genetic map and genome sequence. Integration of physical map, genome sequence, and genetic map could help to correct errors and enhance the quality of each of these resources that are complementary to one another. Integration of these three sets of

Table 10.3 Summary of the integrated genetic and physical map and genome sequence

Number of clones with successful fingerprints	30,824
Number of clones on FPC contigs	26,466
Number of FPC contigs	963
Number of BAC end sequences	67,179
Number of overgos anchored on the FPC contigs	1,181
Mapped SSR markers	706
FPC contigs anchored on the genetic linkage map	535
Scaffolds anchored on the genetic linkage map	255
Genome coverage of the FPC physical map	95.8 %

genomic information provides a valuable resource for comparative genomics, particularly for analyses of Brassicales (Freeling et al. 2008).

The papaya genetic and physical maps were integrated by mapping microsatellite markers derived from BES and whole-genome shotgun sequences (Yu et al. 2009). The sequence-tagged high-density genetic map of papaya was constructed by mapping 706 microsatellite markers (Chen et al. 2007). The resulting genetic map consists of nine major and three minor linkage groups. The three minor linkage groups, LGs 10–12, were merged with major linkage groups, LGs 8, 9, and 7, respectively, based on fluorescence in situ hybridization (FISH) (Wai et al. 2010).

Among the 706 mapped SSR markers, 153 were derived from BES, 466 were designed from shotgun sequence reads, and 87 were developed from assembled shotgun contigs. The BES-derived SSRs directly anchored the associated FPC contigs to the genetic map. Using the 153 SSRs derived from BESs, 122 FPC contigs containing 46,475 consensus bands and 97 shotgun scaffolds covering 132 Mb (35.5 % of the papaya genome) were anchored to the genetic map. The initial integration was extended by aligning BESs to the draft genome sequences. The remaining 553 microsatellites were positioned on shotgun scaffolds by blastn. The BAC clones with paired ends covering the mapped microsatellites were used as bait to search the FPC physical map to find the corresponding FPC contigs. The order and orientation of the FPC contigs were validated by aligning associated BESs on the shotgun scaffolds.

In the final integrated map, 55.6 % FPC contigs (535 FPC contigs) and 81.2 % BAC clones (21,371 BACs) were anchored on the genetic linkage map. A total of 255 shotgun scaffolds covering 233 Mb were anchored to the genetic map. Overall, 63 % of the papaya genome sequences were placed on the genetic map. The summary of the integrated map is shown in Table 10.3.

Integration of Genetic and Cytogenetic Maps

Papaya chromosomes are relatively small and have similar size and morphology at metaphase stage (Storey 1953). Thus, it is difficult to identify individual papaya chromosomes using traditional karyotyping methods, such as banding pattern, arm ratio, morphology, and size. A set of chromosome-specific cytogenetic markers

Table 10.4 The physical size of papaya pachytene chromosomes and the length of their corresponding linkage groups

Chromosome	Total length (μm)	Relative length (%)	Linkage group	Total length (cM)
X/Y	55.70 ± 8.78	12.30 ± 2.74	LG1	145.0
2	58.08 ± 11.84	12.82 ± 3.70	LG9 + LG11	91.3
3	53.49 ± 8.28	11.81 ± 2.59	LG6	100.2
4	51.69 ± 8.74	11.41 ± 2.73	LG2	138.8
5	51.59 ± 9.71	11.39 ± 3.03	LG5	103.6
6	51.31 ± 10.81	11.33 ± 3.38	LG3	132.4
7	47.57 ± 9.11	10.50 ± 2.85	LG7 + LG12	117.8
8	41.78 ± 8.48	9.23 ± 2.65	LG4	120.6
9	41.67 ± 6.34	9.20 ± 1.98	LG8 + LG10	118.9

were developed to assist papaya chromosome identification and cytological analysis (Wai et al. 2010). The chromosome-specific cytogenetic markers were derived from the integrated genetic and physical maps. A total of 104 BAC clones containing SSR markers distributed on 12 linkage groups were tested for signals on papaya chromosomes using FISH. Fifty-four BACs producing repeatable and unambiguous FISH signals were selected as chromosome-specific markers (Wai et al. 2010). This set of chromosome-specific markers served as a foundation for papaya chromosome karyotyping and assigning linkage groups to chromosomes.

The papaya SSR genetic map consists of nine major linkage groups (LGs 1–9) and three minor linkage groups (LGs 10–12) (Chen et al. 2007). The nine major linkage groups were assigned to each individual chromosome by hybridizing LG-specific markers on papaya chromosomes (Wai et al. 2010; Zhang et al. 2010). The three minor linkage groups, LGs 10–12, were assigned to the major linkage groups by simultaneously labeling cytogenetic markers derived from the minor linkage groups with the LG-specific markers developed from the nine major linkage groups. Using this strategy, LGs 10–12 were assigned to the major LGs 8, 9, and 7, respectively (Wai et al. 2010).

A pachytene chromosome-based karyotype of papaya was developed with the assistance of chromosome-specific cytogenetic markers (Zhang et al. 2010). The chromosomes were numbered based on their length except X/Y chromosome pair (Table 10.4). In general, the papaya chromosomes have similar sizes as they were observed at metaphase stage (Table 10.4). The longest chromosome is about 58.08 μm and accounts for 12.8 % of the total length of papaya chromosomes, while the shortest chromosome is about 41.67 μm and accounts for 9.20 % of the total length of papaya chromosomes. The difference between the longest and shortest chromosomes is about 28 %.

The papaya genome is composed largely of euchromatin distributed mainly on chromosome arms (Ming et al. 2008; Zhang et al. 2010). Approximately 83 % of the papaya genome is euchromatic, and 17 % is heterochromatic (Ming et al. 2008). The DAPI brightly stained heterochromatin is concentrated in the centromeric and pericentromeric regions (Zhang et al. 2010), which may explain the recombination suppression found in the middle of most linkage groups (Chen et al. 2007).

The nucleolus organizing regions (NOR), one of the major chromosome structure features, were anchored on the papaya genetic linkage map by FISH (Wai et al. 2010; Zhang et al. 2010). The 5S rDNA loci were located on chromosome 3, 5, 8, 9, and the Y chromosome (Zhang et al. 2010). The strongest signals of 5S rDNA were found at interstitial positions of chromosomes 5 and 8 (Wai et al. 2010; Zhang et al. 2010). The 25S rDNA was located at constriction sites of chromosome 4 (Wai et al. 2010; Zhang et al. 2010).

Genome Coverage

The genome coverage of the fingerprinted physical map was evaluated by estimating the average consensus band size and excluding the organelle genome contamination. BAC clones containing papaya chloroplast genome sequences were identified by hybridizing sorghum chloroplast *ropB* and *trnK* probes against papaya BAC grids (Yu et al. 2009). By searching the BES using the papaya chloroplast genome sequence as a query, additional BAC clones containing the papaya chloroplast genome were identified. The positive clones were scanned against the papaya FPC map, and FPC contigs 972 and 426 were identified as containing papaya chloroplast genome sequence.

The BAC clones containing papaya mitochondrial genome sequences were identified by searching the papaya BES database. A total of 356 BAC clones were identified with at least one end sharing over 95 % identity with the papaya mitochondrial genome sequence. Among them, 144 BACs were verified with both ends containing the papaya mitochondrial genome sequences. FPC contigs 867, 524, 1,172, 553, and 47 were identified as containing the papaya mitochondrial genome fragments.

Lengths of fingerprinted contigs are analyzed as consensus band (CB) units. To estimate the genome coverage of the fingerprinted physical map, the average band size was determined on 22 randomly selected, non-overlapping contigs (Yu et al. 2009). Based on the length of these 22 FPC contigs of 7,503 CB units and their physical size of 12,256,987 bp in the shotgun assembly, the average band size of the papaya FPC map was estimated at about 1.6 kb.

The total length of the papaya FPC map is 224,354 CB units. After excluding the FPC contigs containing papaya organelle genome fragments, the total length of the remaining contigs was 222,808 CB units and approximately equal to 356.5 Mb. Considering the genome size of papaya at 372 Mb (Arumuganathan and Earle 1991), the papaya physical map covers about 95.8 % of the papaya genome, and 72.4 % of papaya genome (269.15 Mb in total length) was anchored to the genetic map.

Genetic Recombination

Genetic maps are constructed by using recombination frequencies between markers. Thus, the distance between two loci on a genetic map reflects their recombination frequency, instead of an actual physical distance. On the other hand, a physical map

is developed using restriction enzyme fingerprinting data which reflect the relative physical location of overlapped clones. The integrated genetic and physical map makes it possible to reveal the relationship between physical distance and genetic distance at the genome level.

The ratio between genetic and physical distance on the integrated map varies among chromosomes. Most recombination suppressed regions were observed near the center of linkage groups, while most recombination hot spots were found toward the ends of linkage groups. DAPI (4′,6-diamidino-2-phenylindole)-stained papaya chromosomes revealed highly condensed heterochromatin knobs located in the centromeric and pericentromeric regions of all nine pairs of chromosomes (Ming et al. 2008). Apparent lengths of the condensed heterochromatin knobs varied among the chromosomes (Ming et al. 2008). Consistent with papaya chromosome structure, recombination suppression regions were found near the center of all the major linkage groups, suggesting that those regions might be centromeric. Higher recombination rates observed toward the telomeric regions could reflect these regions favored by natural selection to contain beneficial alleles.

Extensively suppression of recombination was observed at the male-specific Y chromosome region (MSY) on LG1. The size of the suppressed region was estimated at about 8–9 Mb based on the integrated genetic and physical map. Based on pachytene FISH images, the physical location of the MSY is in the middle of the Y chromosome (Zhang et al. 2008), whereas the MSY was located on the upper half of LG1 on the genetic map (Chen et al. 2007). This discrepancy in locating the MSY is likely caused by the lower recombination rate on one arm of the Y chromosome (Chen et al. 2007; Wai et al. 2012). Also, pachytene FISH images consistently revealed denser heterochromatic regions on one of the Y chromosome arms (Ming et al. 2008). Suppression of recombination has spread to adjacent regions of the MSY, but recombination rates recovered gradually and then increased dramatically at about 10 Mb from the MSY.

Map-Based Cloning of the Major Gene Controlling Fruit Flesh Color in Papaya

Carotenoid antioxidants, which are also called provitamin A, are precursors of vitamin A that carries out a number of functions in the human body including serving as antioxidants, precursors of steroid hormones involved in growth and differentiation, and production of visual proteins (Paiva and Russell 1999; Krinsky 1998; Bartley and Scolnik 1995). Papaya is one of the top fruits that are rich in vitamin A and vitamin C (Liebman 1992). Promoting papaya consumption in underdeveloping and developing tropical and subtropical countries will significantly improve health of poor people as golden rice projected (Dawe et al. 2002). In papaya, the flesh color of the fruit is considered as a quality trait that is often targeted for papaya improvement because it correlates with nutritional value and is linked to shelf life of the fruit (Yamamoto 1964). Two fruit flesh colors, yellow and red, of papaya fruit are present

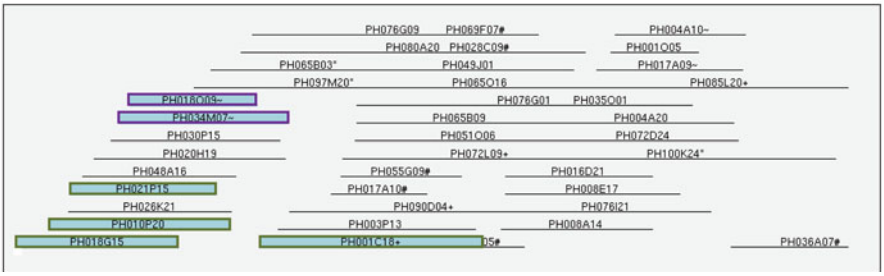

Fig. 10.2 FPC Ctg 962 of the papaya BAC-based physical map. The *purple color box* highlights the BAC clones identified by the cDNA probe of the tomato *CYC-B* gene. The *green color box* highlights the BAC clones identified by the genetic marker linked with *flesh color*

in the papaya germplasm. Red-fleshed papaya is preferred by some consumers, but it is associated with a faster rate of softening thus a shorter shelf life.

The fruit flesh color of papaya is controlled by a single gene with yellow color as dominant (Hofmeyr 1938; Storey 1969). It has been previously documented that yellow-fleshed fruits have high amounts of the carotenoid β-cryptoxanthin and ζ-carotene but lack lycopene, while red-fleshed fruits have high amounts of lycopene along with β-cryptoxanthin, β-carotene, and ζ-carotene (Karrer and Jucker 1950; Yamamoto 1964; Chandrika et al. 2003). The red color of papaya fruit is due to the accumulation of lycopene; the yellow color is the result of converting lycopene to β-carotene and β-cryptoxanthin (Yamamoto 1964).

The high-density AFLP genetic map of papaya was constructed using 54F2 plants derived from cultivars Kapoho and SunUp (Ma et al. 2004). The map consists of 1,501 markers, including 1,498 amplified fragment length polymorphism (AFLP) markers, the papaya ringspot virus coat protein marker, morphological sex type, and fruit flesh color (Ma et al. 2004). The gene controlling the fruit flesh color was placed on linkage group 7 where it is flanked by two AFLP markers at recombination distances 3.4 and 3.7 cM (Ma et al. 2004). These closely linked AFLP markers were converted to sequence-characterized amplified region (SCAR) markers and used as probes to screen the papaya BAC library (Blas et al. 2010). Meanwhile, candidate gene approaches were carried out to clone the gene controlling the fruit flesh color. A tomato *CYC-b* cDNA was used as a probe to screen the papaya BAC library. Five BACs were identified by the SCAR marker and two by the tomato *CYC-b* cDNA probe. These seven positive BACs were used as seed BACs to search into papaya BAC-based physical map. Six of the positive BACs were mapped to a single fingerprint contig FPC-962 of the papaya physical map (Fig. 10.2). Co-localization of the fruit flesh-color-linked SCAR marker and the *CYC-b* cDNA probe on a single contig indicated that the papaya *CYC-b* orthologous gene was a strong candidate for controlling fruit flesh color. The papaya draft genome sequence integrated with FPC-962 was used to identify the papaya homologous gene of tomato *CYC-b* (named as *cp CYC-b*).

The two alleles, the dominant yellow-fleshed allele and recessive red-fleshed allele, were fully sequenced (Blas et al. 2010). A 2 bp insertion was identified within

the red-fleshed *CpCYC-b* coding region, which produces a frameshift mutation resulting in a premature stop codon and a truncated coding region in red-fleshed *CpCYC-b* allele (Blas et al. 2010). The function of *CpCYC-b* alleles was tested by transformation into lycopene-accumulating *E. coli* strain (Cunningham et al. 1994). The gene product of the full-length yellow-fleshed *CpCYC-b* was able to mediate conversion of lycopene (red) to β-carotene (yellow) (Blas et al. 2010). While the red-fleshed *CpCYC-b* lost the lycopene β-cyclase activity as indicated by the lack of color change (Blas et al. 2010).

A SCAR marker located at 580 bp away from *CpCYC-b* gene was developed to support marker-assisted selection in papaya breeding programs. Using this marker, a simple PCR-based screening test using agarose gels can identify individuals in a segregating population with the desired fruit flesh color (Blas et al. 2010). Due to the extremely high recombination rate in the surrounding region, this simple PCR test will be able to identify fruit flesh-color genes with approximately 98 % certainty (Blas et al. 2010). A more precise test could be done at a higher cost and lower throughput using acrylamide sequencing gels by targeting the two nucleotides insertion of the red flesh *CpCYC-b* (Blas et al. 2010).

Perspective

The three major genomic resources, genetic map, physical map, and genome sequence, complement one another and correct errors from each individual source. The integrated information enhances the capacity for map-based cloning, identification of underlying genes controlling quantitative traits in papaya, and comparative genomics studies.

Although the draft genome sequence of papaya covers 92 % of the genes (Ming et al. 2008), most quantitative trait loci (QTLs) controlling economically and agronomically important traits in papaya have not yet been characterized, including those controlling sugar content, fruit size, shape, weight, and complex biotic and abiotic stress reactions. The integrated map in combination with genome sequence will expedite the mapping and cloning of target genes to facilitate papaya breeding through marker-assisted selection.

The current draft sequence of the papaya genome represents about 75 % of the papaya genome (Ming et al. 2008). With rapid progress in next-generation sequencing technology, finishing the papaya genome may be carried out in the foreseeable future. The papaya BAC-based physical map integrated with the genome sequences and genetic map will be an essential resource for closing gaps of any particular genomic region under investigation. The BAC clones located at the ends of the scaffolds can be used as baits to search the FPC physical map. For example, the gap between scaffolds 6 and 129 of the draft genome can be filled by FPC Ctg 355 (Fig. 10.3). The BAC clones on the minimal tilling path can be sequenced to fill the targeted gaps. Moreover, the integrated genetic and physical map could also provide a framework to guide the genome sequencing of related species.

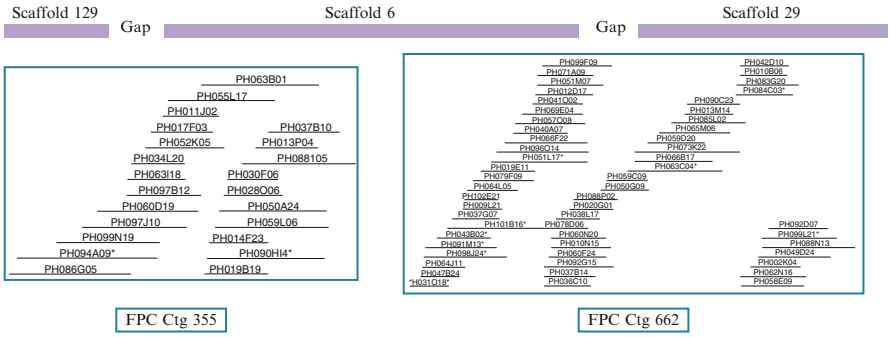

Fig. 10.3 With assistance from BAC end sequences and BAC-based physical map, the gap between shotgun scaffolds 129 and 6 was filled by FPC Ctg 355, and the gap between shotgun scaffolds 6 and 29 was filled by FPC Ctg 662

The integrated genetic and physical map can serve as a reference to study gene and genome evolution and to reveal the genetic base of unique ecological adaptations. A total of 1,181 overgos representing conserved sequences of *Arabidopsis* and genetically mapped *Brassica* loci were anchored on the integrated genetic and physical map and the draft genome sequence of papaya. These overgos provide direct links among papaya, *Arabidopsis,* and *Brassica* genomes for comparative genomic research among species within the order Brassicales. The overgos were designed from single-copy genes and sequences of *Arabidopsis* and *Brassica*. Along with the FPC contigs and WGS sequences, overgo markers could help identify synteny and rearrangements in target regions of these genomes, particularly in these recently duplicated genomes of *Arabidopsis* and *Brassica*.

References

Ackerman CM, Yu Q, Kim S, Paull RE, Moore PH, Ming R (2008) Differential expression of B-class MADS-box genes in trioecious papaya flowers. Planta 227:741–753

Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. Plant Mol Biol Rep 9:208–218

Bartley GE, Scolnik PA (1995) Plant carotenoids: pigments for photoprotection, visual attraction, and human health. Plant Cell 7:1027–1038

Blas AL, Yu Q, Chen C, Veatch O, Moore PH, Paull RE, Ming R (2009) Enrichment of a papaya high-density genetic map with AFLP markers. Genome 52:716–725

Blas AL, Ming R, Liu Z, Veatch OJ, Paull RE, Moore PH, Yu Q (2010) Cloning of papaya chromoplast specific lycopene β-cyclase, CpCYC-b, controlling fruit flesh color reveals conserved microsynteny and a recombination hotspot. Plant Physiol 152:2013–2022

Chandrika UG, Jansz ER, Wickramasinghe SMDN, Warnasuriya ND (2003) Carotenoids in yellow- and red-fleshed papaya (*Carica papaya* L). J Sci Food Agric 83:1279–1282

Chen C, Yu Q, Hou S, Li Y, Eustice M, Skelton RL, Veatch O, Herds R, Diebold L, Saw J, Feng Y, Bynum L, Wang L, Moore PH, Paull RE, Alam M, Ming R (2007) Construction of a sequence-tagged high-density genetic map of papaya for comparative structural and evolutionary genomics in brassicales. Genetics 177:2481–2491

Cunningham FX Jr, Sun Z, Chamovitz D, Hirschberg J, Gantt E (1994) Molecular structure and enzymatic function of lycopene cyclase from the cyanobacterium Synechococcus sp strain PCC7942. Plant Cell 6:1107–1121

Dawe D, Robertson R, Unnevehr L (2002) Golden rice: what role could it play in alleviation of vitamin A deficiency? Food Policy 27:541–560

Freeling M, Lyons E, Pedersen B, Alam M, Ming R, Lisch D (2008) Many or most genes in *Arabidopsis* transposed after the origin of the order Brassicales. Genome Res 18:1924–1937

Gschwend AR, Yu Q, Moore P, Saski C, Chen C, Wang J, Na J-K, Ming R (2011) Construction of papaya male and female BAC libraries and application in physical mapping of the sex chromosomes. J Biomed Biotechnol 2011:929472

Hofmeyr JDJ (1938) Genetical studies of Carica papaya L.I. The inheritance and relation of sex and certain plant characteristics. II Sex reversal and sex forms. South African Department of Agriculture and Science Bulletin, no. 187

Initiative TAG (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796–815

Karrer P, Jucker E (1950) Carotenoids. Elsevier, New York

Klein PE, Klein RR, Cartinhour SW, Ulanch PE, Dong J, Obert JA, Morishige DT, Schlueter SD, Childs KL, Ale M, Mullet JE (2000) A high-throughput AFLP-based method for constructing integrated genetic and physical maps: progress toward a sorghum genome map. Genome Res 10:789–807

Krinsky NI (1998) Overview of lycopene, carotenoids, and disease prevention. Exp Biol Med 218:95–97

Lai CWJ, Yu Q, Hou S, Skelton RL, Jones MR, Lewis KLT, Murray J, Eustice M, Guan P, Agbayani R, Moore PH, Ming R, Presting GG (2006) Analysis of papaya BAC end sequences reveals first insights into the organization of a fruit tree genome. Mol Genet Genomics 276:1–12

Liebman B (1992) Nutritional aspects of fruit. Nutr Action Newslett (May 10–11)

Lin L, Pierce GJ, Bowers JE, Estill JC, Compton RO, Rainville LK, Kim C, Lemke C, Rong J, Tang H, Wang X, Braidotti M, Chen AH, Chicola K, Collura K, Epps E, Golser W, Grover C, Ingles J, Karunakaran S, Kudrna D, Olive J, Tabassum N, Um E, Wissotski M, Yu Y, Zuccolo A, ur Rahman M, Peterson DG, Wing RA, Wendel JF, Paterson AH (2010) A draft physical map of a D-genome cotton species (Gossypium raimondii). BMC Genomics 11:395

Luo MC, Thomas C, You FM, Hsiao J, Ouyang S, Buell CR, Malandro M, McGuire PE, Anderson OD, Dvorak J (2003) High-throughput finger-printing of bacterial artificial chromosomes using the SNaP-shot labeling kit and sizing of restriction fragments by capillary electrophoresis. Genomics 82:378–389

Lyons E, Pedersen B, Kane J, Alam M, Ming R, Tang H, Wang X, Bowers J, Paterson A, Lisch D, Freeling M (2008a) Finding and comparing syntenic regions among Arabidopsis and the outgroups papaya, poplar, and grape: CoGe with Rosids. Plant Physiol 148:1772–1781

Lyons E, Pedersen B, Kane J, Freeling M (2008b) The value of nonmodel genomes and an example using SynMap within CoGe to dissect the hexaploidy that predates the Rosids. Trop Plant Biol 1:181–190

Ma H, Moore PH, Liu Z, Kim MS, Yu Q, Fitch MMM, Sekioka T, Paterson AH, Ming R (2004) High-density linkage mapping revealed suppression of recombination at the sex determination locus in papaya. Genetics 166:419–436

Ming R, Moore PH, Zee F, Abbey CA, Ma H, Paterson AH (2001) Construction and characterization of a papaya BAC library as a foundation for molecular dissection of a tree-fruit genome. Theor Appl Genet 102:892–899

Ming R, Hou S, Feng Y, Yu Q, Dionne-Laporte A, Saw JH, Senin P, Wang W, Ly BV, Lewis KLT, Salzberg SL, Feng L, Jones MR, Skelton RL, Murray JE, Chen C, Qian W, Shen J, Du P, Eustice M, Tong E, Tang H, Lyons E, Paull RE, Michael TP, Wall K, Rice D, Albert H, Wang M-L, Zhu YJ, Schatz M, Nagarajan N, Agbayani R, Guan P, Blas A, Wai CM, Ackerman CM, Ren Y, Liu C, Wang J, Wang J, Na J-K, Shakirov EV, Haas B, Thimmapuram J, Nelson D, Wang X, Bowers JE, Gschwend AR, Delcher AL, Singh R, Suzuki JY, Tripathi S, Neupane K, Wei H, Irikura B, Paidi M, Jiang N, Zhang W, Presting G, Windsor A, Navajas-Pérez R, Torres MJ, Alex FF, Porter B, Li Y, Burroughs AM, Luo M-C, Liu L, Christopher DA, Mount SM,

Moore PH, Sugimura T, Jiang J, Schuler MA, Friedman V, Mitchell-Olds T, Shippen DE, de Pamphilis CW, Palmer JD, Freeling M, Paterson AH, Gonsalves D, Wang L, Alam M (2008) The draft genome of the transgenic tropical fruit tree papaya (Carica papaya Linnaeus). Nature 452:991–996

Mun J-H, Kim D-J, Choi H-K, Gish J, Debellé F, Mudge J, Denny R, Endré G, Saurat O, Dudez A, Kiss GB, Roe B, Young ND, Cook DR (2006) Distribution of microsatellites in the genome of *Medicago truncatula*: a resource of genetic markers that integrate genetic and physical maps. Genetics 172:2541–2555

Na J-K, Wang J, Murray JE, Gschwend AR, Zhang W, Yu Q, Navajas-Pérez R, Feltus AF, Chen C, Kubat Z, Moore PH, Jiang J, Paterson AH, Ming R (2012) Construction of physical maps for the sex-specific regions of papaya sex chromosomes. BMC Genomics 13:176

Paiva SAR, Russell RM (1999) ß-Carotene and other carotenoids as antioxidants. J Am Coll Nutr 18:426–433

Ross M, LaBrie T, McPherson S, Stanton VP (1999) Screening large-insert libraries by hybridization. In: Haines JL, Korf BR, Morton CC, Seidman JG, Smith DR (eds) Current protocols in human genetics. Wiley, New York, pp 5.6.1–5.6.32

Saski CA, Feltus FA, Staton ME, Blackmon BP, Ficklin SP, Kuhn DN, Schnell RJ, Shapiro H, Motamayor JC (2011) A genetically anchored physical framework for *Theobroma cacao* cv. Matina 1–6. BMC Genomics 12:413

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, Minx P, Reily AD, Courtney L, Kruchowski SS, Tomlinson C, Strong C, Delehaunty K, Fronick C, Courtney B, Rock SM, Belter E, Du F, Kim K, Abbott RM, Cotton M, Levy A, Marchetto P, Ochoa K, Jackson SM, Gillam B, Chen W, Yan L, Higginbotham J, Cardenas M, Waligorski J, Applebaum E, Phelps L, Falcone J, Kanchi K, Thane T, Scimone A, Thane N, Henke J, Wang T, Ruppert J, Shah N, Rotter K, Hodges J, Ingenthron E, Cordes M, Kohlberg S, Sgro J, Delgado B, Mead K, Chinwalla A, Leonard S, Crouse K, Collura K, Kudrna D, Currie J, He R, Angelova A, Rajasekar S, Mueller T, Lomeli R, Scara G, Ko A, Delaney K, Wissotski M, Lopez G, Campos D, Braidotti M, Ashley E, Golser W, Kim H, Lee S, Lin J, Dujmic Z, Kim W, Talag J, Zuccolo A, Fan C, Sebastian A, Kramer M, Spiegel L, Nascimento L, Zutavern T, Miller B, Ambroise C, Muller S, Spooner W, Narechania A, Ren L, Wei S, Kumari S, Faga B, Levy MJ, McMahan L, Van Buren P, Vaughn MW, Ying K, Yeh CT, Emrich SJ, Jia Y, Kalyanaraman A, Hsia AP, Barbazuk WB, Baucom RS, Brutnell TP, Carpita NC, Chaparro C, Chia JM, Deragon JM, Estill JC, Fu Y, Jeddeloh JA, Han Y, Lee H, Li P, Lisch DR, Liu S, Liu Z, Nagel DH, McCann MC, SanMiguel P, Myers AM, Nettleton D, Nguyen J, Penning BW, Ponnala L, Schneider KL, Schwartz DC, Sharma A, Soderlund C, Springer NM, Sun Q, Wang H, Waterman M, Westerman R, Wolfgruber TK, Yang L, Yu Y, Zhang L, Zhou S, Zhu Q, Bennetzen JL, Dawe RK, Jiang J, Jiang N, Presting GG, Wessler SR, Aluru S, Martienssen RA, Clifton SW, McCombie WR, Wing RA, Wilson RK (2009) The B73 maize genome: complexity, diversity, and dynamics. Science 326:1112–1115

Skelton RL, Yu Q, Srinivasan R, Manshardt R, Moore PH, Ming R (2006) The expression of the gene for Lycopene β-cyclase is elevated in leaves and flowers and down-regulated in both yellow- and red-fleshed papaya fruits. Cell Res 16:731–739

Soderlund C, Longden I, Mott R (1997) FPC: a system for building contigs from restriction fingerprinted clones. Comput Appl Biosci 13:523–535

Storey WB (1953) Genetics of papaya. J Hered 44:70–78

Storey WB (1969) Papaya. In: Ferwerda FP, Wit F (eds) Outlines of perennial crop breeding in the tropics. Veenman, Wageningen, pp 389–408

Suzuki JY, Tripathi S, Fermín GA, Jan F-J, Hou S, Saw JH, Ackerman CM, Yu Q, Schatz MC, Pitz KY, Yépes M, Fitch MMM, Manshardt RM, Slightom JL, Ferreira SA, Salzberg SL, Alam M, Ming R, Moore PH, Gonsalves D (2008) Characterization of insertion sites in rainbow papaya, the first commercialized transgenic fruit crop. Trop Plant Biol 1:293–309

Tang H, Bowers JE, Wang X, Ming R, Alam M, Paterson AH (2008) Synteny and collinearity in plant genomes. Science 320:486–488

The Tomato Genome Consortium (2012) The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485:635–641

Troggio M, Malacarne G, Coppola G, Segala C, Cartwright D, Pindo M, Stefanini M, Mank R, Moroldo M (2007) A dense SNP-based genetic linkage map of grapevine (*Vitis vinifera* L.) anchoring Pinot Noir BAC contigs. Genetic 176:2637–2650

Wai CM, Moore PH, Paull RE, Ming R, Yu Q (2012) An integrated cytogenetic and physical map reveals unevenly distributed recombination spots along the papaya sex chromosomes. Chromosome Research 20:753–767

Wai CM, Ming R, Moore PH, Paull RE, Yu Q (2010) Development of chromosome-specific cytogenetic markers and merging of linkage fragments in papaya. Trop Plant Biol 3:171–181

Wikström N, Savolainen V, Chase MW (2001) Evolution of the angiosperms: calibrating the family tree. Proc R Soc Lond B 268:2211–2220

Yamamoto HY (1964) Comparison of the carotenoids in yellow- and red-fleshed carica papayas. Nature 201:1049–1050

Yu J, Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X, Cao M, Liu J, Sun J, Tang J, Chen Y, Huang X, Lin W, Ye C, Tong W, Cong L, Geng J, Han Y, Li L, Li W, Hu G, Huang X, Li W, Li J, Liu Z, Li L, Liu J, Qi Q, Liu J, Li L, Li T, Wang X, Lu H, Wu T, Zhu M, Ni P, Han H, Dong W, Ren X, Feng X, Cui P, Li X, Wang H, Xu X, Zhai W, Xu Z, Zhang J, He S, Zhang J, Xu J, Zhang K, Zheng X, Dong J, Zeng W, Tao L, Ye J, Tan J, Ren X, Chen X, He J, Liu D, Tian W, Tian C, Xia H, Bao Q, Li G, Gao H, Cao T, Wang J, Zhao W, Li P, Chen W, Wang X, Zhang Y, Hu J, Wang J, Liu S, Yang J, Zhang G, Xiong Y, Li Z, Mao L, Zhou C, Zhu Z, Chen R, Hao B, Zheng W, Chen S, Guo W, Li G, Liu S, Tao M, Wang J, Zhu L, Yuan L, Yang H (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science 296:79–92

Yu Q, Moore PH, Albert HH, Roader AHK, Ming R (2005) Cloning and characterization of a *FLORICAULA/LEAFY* ortholog, *PFL*, in polygamous papaya. Cell Res 15:576–584

Yu Q, Steiger D, Kramer EM, Moore PH, Ming R (2008) Floral MADS-box genes in trioecious papaya: characterization of *AG* and *AP1* subfamily genes revealed a sex-type-specific gene. Trop Plant Biol 1:97–107

Yu Q, Tong E, Skelton RL, Bowers JE, Jones MR, Murray JE, Hou S, Guan P, Acob RA, Luo MC, Moore PH, Alam M, Paterson AH, Ming R (2009) A physical map of the papaya genome with integrated genetic map and genome sequence. BMC Genomics 10:371

Yüksel B, Bowers J, Estill J, Goff L, Lemke C, Paterson A (2005) Exploratory integration of peanut genetic and physical maps and possible contributions from *Arabidopsis*. Theor Appl Genet 111:87–94

Zhang W, Wang X, Yu Q, Ming R, Jiang J (2008) DNA methylation and heterochromatinization in the male-specific region of the primitive Y chromosome of papaya. Genome Res 18:1938–1943

Zhang W, Wai CM, Ming R, Yu Q, Jiang J (2010) Integration of genetic and cytological maps and development of a pachytene chromosome-based karyotype in papaya. Trop Plant Biol 3:166–170

Part III
The Papaya Genome

Chapter 11
Sequencing and Assembly of the Transgenic Papaya Genome

Robert VanBuren and Ray Ming

Introduction

Fruit tree genomics and genetics are notoriously difficult to study because of the long generation times, residual heterozygosity within a genome, limited production of offspring, and large mature size associated with most fruit trees. Many commercial fruit trees take years to flower from seed, rendering slow advancements in both classical plant breeding and modern molecular genetics. New cultivars can take decades to develop, and progress on quantitative trait loci (QTL) mapping, construction of genetic maps for genome sequencing projects, and generating mutant lines are often hindered due to the long generation time. These problems are further complicated by the fact that some trees such as mango, avocado, and lychee produce one seed per fruit, making it difficult to obtain the large number of progeny needed for phenotyping traits and genetic mapping. The large mature size of most fruit trees makes growing and maintaining populations for genetic studies laborious and cost prohibitive.

Often, fruit tree genome sequencing projects run into problems associated with residual within genome heterozygosity. Maize, rice, and Arabidopsis have undergone countless generations of self-crossing and, as a result, are highly inbred with little heterozygosity (Schnable et al. 2009; The Arabidopsis Genome Initiative 2000; Jaillon et al. 2007). The lack of sequence polymorphisms makes whole-genome shotgun approaches feasible. However, grafting and other forms of vegetative propagation have maintained the natural heterozygosity in most fruit trees, making it difficult to accurately sequence the genomes using shotgun approaches. Heterozygosity between the haplotypes can pull apart contigs leading to an erroneous

R. VanBuren • R. Ming (✉)
Department of Plant Biology, University of Illinois at Urbana-Champaign, 1201 W. Gregory Drive, 148 ERML, MC-051, Urbana, IL 61801, USA
e-mail: rming@life.uiuc.edu

R. Ming and P.H. Moore (eds.), *Genetics and Genomics of Papaya*, Plant Genetics and Genomics: Crops and Models 10, DOI 10.1007/978-1-4614-8087-7_11,

assembly. Together, these problems have complicated and hindered the study of fruit tree genomics and genetics. Papaya, however, is not subjected to these problems and is thus an excellent model for studying fruit tree genomics. At the time of its completion, papaya was the first fruit tree genome to be sequenced (Ming et al. 2008).

Papaya as a Model Fruit Tree

Papaya has a number of characteristics that make it a suitable model for fruit tree genomics. Papaya is a perennial tree that produces an abundance of fruit year round. The generation time of papaya is remarkably short in comparison to other fruit trees, with a juvenile phase of 3–8 months, and ripe fruits produced 9–12 months after planting. Well-pollinated fruits can contain upwards of 1,000 seeds, and an individual tree can produce hundreds of fruits a year. This bountiful fruit and seed production leads to an almost unlimited number of progeny for variety development and genetic studies. Papaya trees are relatively small requiring as little as 5 m^2 per tree, with a density of 1,200–2,000 plants per hectare. The trioecious nature of papaya coupled with large, easily removable anthers and carpels makes controlled crosses feasible. Hermaphrodite plants are mostly selfing, but male and female plants are incapable of self-fertilization, and are easily crossed. Individual plants can be screened in multiple environments, as papaya is easily cloned from cuttings. Genetics and genomic studies in papaya are further facilitated by a well-established transformation system. Sequencing the papaya genome was feasible partly due to the small genome size and lack of within genome heterozygosity. Papaya is diploid, with nine pairs of chromosomes and a relatively small haploid genome size of 371 Mbp. In comparison, banana, plum, and avocado genomes are estimated to be over twice as large at 875, 883, and 883 Mbp, respectively (Arumuganathan and Earle 1991). The passion fruit genome is over six times as large and is estimated to be 2,191 Mbp. The SunUp variety of papaya was chosen for sequencing, as its progenitor Sunset had undergone over 25 generations of inbreeding and, as a result, has a low level of residual heterozygosity (Storey 1969). These characteristics made sequencing and assembly of the papaya genome possible and have also made papaya an excellent model for studying fruit tree genomics.

Additional reasons for sequencing the papaya genome were its agricultural importance, incipient sex chromosomes, and successful transgenic applications. Papaya is the most nutritious of the 34 commonly consumed fruits based on its high levels of vitamins A and C, folate, potassium, niacin, riboflavin, riboflavin, calcium, and fiber; see Chap. 20 in this text for details of papaya's nutritional value. Papaya consumption has also been suggested for preventing vitamin A deficiency, which causes childhood blindness in tropical and subtropical developing countries (Chandrika et al. 2003). Discovering genes related to fruit development and nutrition could facilitate the further improvement of papaya or enrichment of other fruits. Papaya is also the primary source of papain, a cysteine protease primarily used to

tenderize red meat, chill-proof beer, and to treat external hard tissues such as scars and warts. Identifying genes related to papain biosynthesis could significantly increase its abundance in papaya fruit tissues. Sequencing the papaya genome also revealed a number of economically important genes involved in disease resistance, lignin biosynthesis, and starch accumulation.

The draft genome formed the starting point for studying papaya sex chromosome evolution. Unlike most plants, sex in papaya is controlled by a pair of nascent XY sex chromosomes (Liu et al. 2004). Males and hermaphrodites have two slightly different Y chromosomes, which share 98.8 % sequence homology, and are denoted as Y for males and Yh for hermaphrodites. Female papaya have a XX genotype, and males and hermaphrodites have XY or XYh genotype, respectively, with any combination of YY, YhY, or YhYh chromosomes resulting in embryo abortion. The sex chromosomes recombine normally at the pseudo-autosomal region which spans approximately 87 % of the chromosome length. The remaining 13 % is recombinationally suppressed, signifying the sex-determination regions. The hermaphrodite-specific or male-specific region of the Y chromosome (HSY or MSY, respectively) is 8.1 Mbp, and its X counterpart is 3.5 Mbp (Wang et al. 2012). The draft genome was sequenced from a female SunUp plant to avoid assembly errors in the heterogametic X and Y regions. As a result, the Y sequence is completely absent in the draft genome. The draft genome, however, spans the pseudo-autosomal region on chromosome 1 and parts of the X-specific region, which formed the basis of sequencing and mapping the papaya sex chromosomes. The short evolutionary history of the papaya sex chromosomes has yielded enormous insights into the initiation and degeneration of sex chromosomes.

Sequencing and Assembly of the Papaya Draft Genome

The SunUp variety of papaya was chosen for the papaya genome sequencing project. SunUp is a recently developed cultivar that originated from the transformation of Sunset, a variety with low residual heterozygosity and 25 generations of inbreeding (Storey 1969). SunUp was genetically engineered to produce the papaya ringspot virus (PRSV) coat protein (Lius et al. 1997). The PRSV protein provided resistance to the papaya ringspot virus that devastated the Hawaiian papaya industry in the 1990s (see Chap. 7 in this text for details). Sequencing a transgenic variety provided conclusive evidence of the transgene insertion site, and the copy number. A female plant was chosen for sequencing to avoid the complication of assembling the heterogametic X and Y chromosomes.

A whole shotgun sequencing approach was used with integration of BAC-end sequences and physical and genetic maps to sequence the papaya genome (Ming et al. 2008). The draft genome was sequenced before the advent of next-generation sequencing technologies, using small-insert libraries and Sanger sequencing. A total of 2.8 million paired-end shotgun reads were generated from total papaya genomic

DNA. The total genomic DNA was sheared and cloned in two separate libraries. Low-quality reads and organelle-based sequences were removed from the assembly, leaving 1.55 million high-quality plasmid-based reads. Cumulatively, these reads represent roughly three times coverage, meaning, on average; each base pair in the genome would be sequenced three times. The high-quality reads were assembled to generate 47,483 contiguous, gap-free sequences. The N50 statistic of the contig sequences is 11 kb. N50 refers to the minimum contig length where half of the base pairs in the genome assembly are represented. In papaya, this signifies that half of the assembled genome is in pieces equal to or larger than 11 kb. Within genome heterozygosity of the assembled contigs is low, at an estimation of 0.06 %. The total assembled contigs represent 75 % of papaya genome and roughly 90 % of the euchromatic regions. Contigs derived from the whole-genome shotgun reads were assembled into scaffold sequences using the physical map and the BAC-end sequences.

Physical Mapping

A library of 39,168 BAC clones was generated to physically map the papaya genome and to anchor genome sequences. The average insert size for the BAC library is 132 kb, and the entire library represents 13.7× genome equivalents (Ming et al. 2001). The female BAC library was made with two separate ligations, to ensure the desired insert size and to remove digestion bias. All of the BACs in the hermaphrodite BAC library were fingerprinted using five restriction enzymes and assembled into contigs based on their fingerprint profiles. BAC clones with no inserts, highly repetitive sequences, and incomplete restriction digests were removed from the analysis, leaving 30,824 fingerprints, representing 11 times genome equivalents. These BACs were assembled into 963 contigs, with each contig containing on average, 27.5 clones. BAC-end sequencing (BES) was conducted on all of the BACs, resulting in 67,179 high-quality sequences. Reactions that failed the first round were resequenced, and low-quality sequences and empty BAC clones were removed. The BES had a total combined length of 45 Mbp, or 12 % of the papaya genome. The BES was used with the draft assembly to create scaffolds. Contigs from the assembly that matched individual BAC ends or BAC ends within fingerprint contigs were assembled together with imbedded gaps. Since the relative size of each BAC insert was known (based on the fingerprint map), paired-end sequences were used to estimate gap sizes in scaffold sequences. 32,397 BES were integrated into the shotgun genome assembly, to extend the scaffold sequences (Ming et al. 2008; Yu et al. 2009). Overall, the scaffold sequences covered 370 Mbp, with an N50 of 1 Mbp. These scaffold sequences are reasonably large, but difficult to draw conclusions from, as they weren't ordered or anchored to linkage groups or chromosomes. Thus, a genetic map with 706 SSR markers was used to map the scaffolds to linkage groups and, subsequently, chromosomes.

Genetic Mapping and Scaffold Anchoring

Short sequence repeats (SSR) were identified in the BES and the assembled genome reads to construct the high-density genetic map for integrating the physical map and shotgun assembly (Chen et al. 2007). SSRs are highly polymorphic, dispersed evenly throughout the genome, and are almost always at single loci. They can also be quickly and accurately scored using PCR and agarose gels. 11,976 SSR markers were identified in the BES and in the assembled genome reads. The candidate SSR markers were amplified via PCR with 8,763 (73 %) of markers successfully amplified. Of the amplified markers, 1,167 showed polymorphisms between the parents in the mapping population, and 886 SSR markers were used to score the offspring.

The F2 mapping population used to create the high-density genetic map arose from a cross between the dioecious cultivar AU9 (with male and female plants) and the gynodioecious SunUp (with hermaphrodite and female plants). Fifty-four plants were scored for each of the 886 SSR markers. The high-density genetic map consisted of 706 SSR markers and 1 morphological marker for flesh color. The markers were mapped to nine major linkage groups and three minor linkage groups. The nine major linkage groups correlate to the nine chromosomes in papaya, and the three minor groups represent chromosome arms or regions that failed to link with the other major linkage groups. These three linkage groups were eventually anchored to the chromosomes and merged with the major linkage groups using fluorescence in situ hybridization (FISH) (Ming et al. 2008; Zhang et al. 2010).

The high-density genetic map was used to anchor the shotgun scaffolds and the fingerprint contigs from the physical map to the nine major and three minor linkage groups (Ming et al. 2008; Yu et al. 2009). About 153 of the markers were derived from BES and anchored 122 BAC fingerprint contigs to the genetic map (Fig. 11.1). Ninety-seven scaffolds encompassed these fingerprint contigs, representing a total size of 132 Mb or 35 % of the genome. The remaining 553 markers were identified in the scaffold sequences and position the scaffolds within the 12 linkage groups. Overall, 233 Mb (63 %) of the papaya genome sequences were anchored to the linkage groups (Fig. 11.1). The shotgun scaffolds and fingerprint contigs joined linkage groups 8 and 10, and linkage groups 9 and 11 were merged based on FISH (Ming et al. 2008). Linkage groups 7 and 12 were also merged using BAC clones and FISH (Zhang et al. 2010).

The sequence coverage of each linkage group is uneven. Linkage groups 9 and 11 (chromosome 2) had 90 % coverage by anchored shotgun scaffolds and nearly 100 % coverage by BAC fingerprint clones. The sex chromosomes are confined to linkage group 1 and had the lowest physical coverage, with only 43 % of the estimated length represented by anchored shotgun scaffolds and BAC fingerprint clones (Yu et al. 2009). This is due in part to the sex-specific regions, which account for 13 % of the overall sequence length, but the remaining gaps are difficult to explain. The papaya draft genome contained X-specific sequences, but the Y region is completely absent from the assembly.

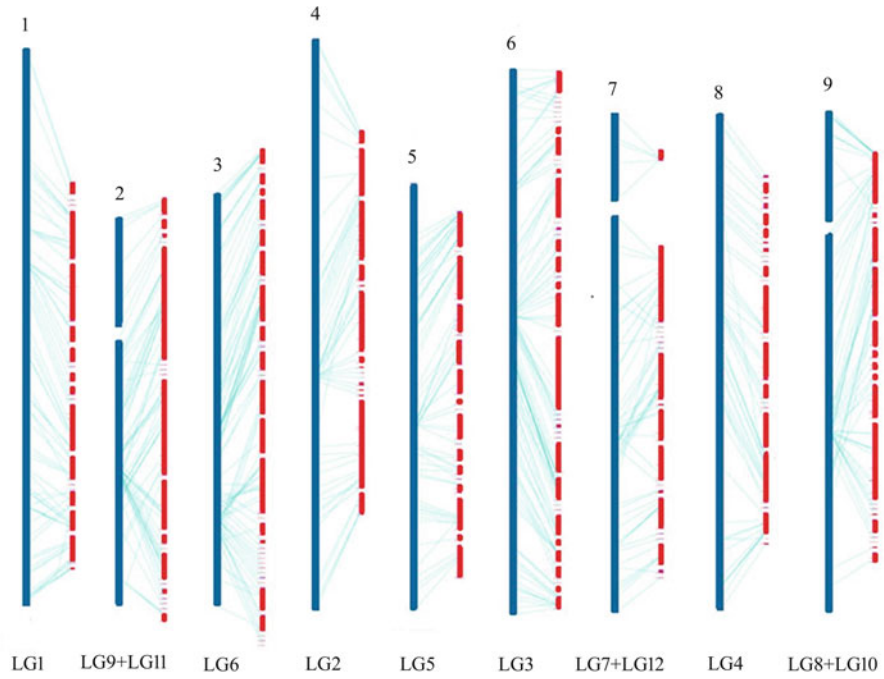

Fig. 11.1 The assembled papaya genome. SSR markers from the 12 linkage groups (*blue*) were used to anchor scaffold sequences (*red*). *Light-blue* lines correlate to SSR marker positions on the genetic map and scaffold sequences. A cytogenetic map was used to assign chromosome numbers to the linkage groups (top numbers). Significant gaps exist in several of the chromosome, most notable on chromosome 1, the sex chromosomes [modified with permission of Macmillan Publishers Ltd from Ming et al. (2008)]

FISH was also used to assign the linkage groups to each of the nine chromosomes. The original 12 linkage groups were numbered based on size but were not assigned to individual chromosomes. The cytological map was created using BAC probes specific to each of the 12 linkage groups (Zhang et al. 2010). The BAC probes were hybridized with meiotic pachytene chromosomes and yielded excellent signals. Linkage group 1 was assigned to chromosome 1 (the second largest chromosome), and the remaining chromosomes were ordered based on size. Linkage groups are integrated with the cytogenetic map based on BAC hybridization.

EST Sequencing and Gene Prediction

The papaya genome is largely euchromatic, as revealed by 4′,6-diamidino-2-phenylindole (DAPI) staining at the pachytene stage of meiosis. DAPI binds to the characteristic AT-rich sequences found in highly condensed heterochromatic regions.

Heavily stained bivalents were identified in highly condensed knob structures at the centromeric and pericentromeric regions, accounting for 17 % of the genome. Most of the heterochromatic sequences remain unassembled, as evident by the lack of centromere- and telomere-specific repeats in the draft assembly.

Annotating the papaya genome began with large-scale EST sequencing. Initially, five cDNA libraries were constructed from floral tissue. Flower ESTs were sequenced in the hope of identifying the key sex-determination genes, or sex-specific gene expression. Three libraries were created from pre-meiosis (<4 mm) male, female, and hermaphrodite flower buds, and two were generated from mature hermaphrodite and female flowers. cDNA clones from the five libraries were sequenced from the 5′ end using the Sanger method. EST reads were trimmed for vector contamination and quality with a minimum usable length of 200 nucleotides. Overall, 31,652 high-quality sequences were generated, with an average length of 486 nucleotides. Repeat sequences were annotated, masked, and removed from the assembly. The remaining ESTs were clustered and assembled. Assembly parameters were stringent, with a minimum identity of 95 % and overlap greater than 30 bp to avoid erroneous assembly of closely related genes in the same gene family. A total of 8,572 unigenes were assembled from the floral EST libraries (Ming et al., unpublished data).

Representative samples from diverse tissue types ensured the capture and sequencing of genes with tissue-specific and developmental-specific expression. Additional EST resources came from a normalized, subtractive cDNA library constructed from pooled RNA samples. Equimolar ratios of RNA were pooled from roots, stems, leaves, callus tissue, seeds, flowers, and three stages of fruit development. Over 50,000 cDNA clones were sequenced from this library. Trimmed, masked, and filtered ESTs were assembled with the floral unigenes to obtain a final set of 16,362 unigenes. ESTs were annotated using BLASTx and the non-redundant protein database on NCBI, with an e value cutoff of 1×10^{-5}. Approximately 83 % of the ESTs have homology to known proteins, with the remaining representing papaya-specific genes and actively transcribed transposable elements. The papaya EST database likely failed to capture low-abundance transcripts, but it served as a valuable starting point for annotating the papaya genome.

92.1 % of the EST sequences were anchored to the papaya draft genome sequence. This percentage is similar the theoretical coverage of 95 % expected by the three times whole-genome shotgun sequence as well as the estimated 90 % coverage of the euchromatic regions. The remaining EST sequences likely lie within the unassembled heterochromatic regions. EST sequences were also used to estimate the average intron size. The average intron size in papaya is 479 bp, significantly larger than the average intron sizes in grape and Arabidopsis of 213 and 165 bp, respectively (Jaillon et al. 2007). Comparison of 4,403 orthologous gene pairs between Arabidopsis, rice, and papaya indicated that papaya contains significantly fewer introns than Arabidopsis, and slightly fewer than rice. The increased intron size in papaya is likely due to excessive repeat accumulation in intragenic regions.

Ab initio gene prediction software was coupled with EST splicing site and protein structure data to generate a reference set of papaya gene models. ESTs are a valuable resource for identifying transcripts, but they are rarely exhaustive, as some

Table 11.1 Comparison of the papaya genome to other sequenced plant genomes

	Genome size (Mb)[a]	Number of genes[a]	Transposable elements (% of genome)[a]
Papaya	370	28,027	52
Maize	2,300	32,450	85
Apple	742	57,386	42.4
Poplar	480	45,654	35
Grape	475	33,514	21.5
Rice	420	40,577	39.5
Sorghum	730	34,496	62
Cucumber	367	26,682	14.8
Arabidopsis	125	27,228	18.5

[a]Values were obtained from Velasco et al. (2010), Paterson et al. (2009), Jaillon et al. (2007), Tuskan et al. (2006), Schnable et al. (2009), The Arabidopsis Genome Initiative (2000), Huang et al. (2009), Schmutz et al. (2010), and International Rice Genome Sequencing Project (2005)

genes have specific expression patterns or extremely low abundance not captured by EST libraries. Five gene prediction programs were coupled with EST data to annotate the papaya genome. Repetitive elements were masked in the draft genome to prevent annotation of active transposable elements. The papaya-specific repeat database as well as preexisting plant repeat databases were used to mask the draft assembly (Nagarajan et al. 2008). Gene prediction software identified 27,950 gene models. 20,067 (72 %) of the gene models have homology to known proteins in the nonredundant protein database from NCBI. Of those, 9,642 (48 %) have EST support. The 7,971 genes without homology to known proteins have an average length of 307 bp (in comparison to the 1,102 average of gene models with homology), and only 647 (8.1 %) have EST support. Most of these predicted genes are likely false-positives. Excluding short ESTs without support brings the predicted gene number to 24,746 gene models. The predicted gene number is lower than most other sequenced plant genomes to date. The predicted gene number is 25 % less than Arabidopsis (The Arabidopsis Genome Initiative 2000; Hanada et al. 2007), 46 % less than poplar (Tuskan et al. 2006), 47 % less than soybean (Schmutz et al. 2010), and 58 % less than apple (Velasco et al. 2010) (Table 11.1). The gene model estimates in papaya are likely inflated, as single genes may be fragmented due to gaps and sequencing errors, causing them to be counted as several separate genes. This is demonstrated by the initially inflated gene number estimates in the rice draft genome (Table 11.1).

A global protein analysis was conducted with five sets of angiosperm gene models to identify gene conservation and gene family expansion and reduction. The papaya gene models were clustered with inferred models from Arabidopsis, poplar, grape, and rice to generate a non-redundant set of 208,901 protein sequences (Wall et al. 2008). The nonredundant proteins clustered into 39,706 distinct tribes with 11,851 tribes containing two or more genes. Gene ontology terms were assigned to each tribe using GO SLIM annotations provided by TAIR. Tribes with gene models from each of the five sequenced plant genomes are likely evolutionarily conserved and are part of a minimum set of genes required for all angiosperms. The smallest observed number of conserved genes in each tribe suggests that the minimum

angiosperm gene number of 13,311. Tribes that contain genes from only one species represent either a species-specific gene family or families with significant diversion. In papaya 5,669 tribes were papaya-specific, and 5,314 of the papaya-specific tribes were singletons. Of the papaya-specific tribes, only 14 % had EST support, suggesting that most papaya-specific tribes may be incorrectly annotated or represent false-positives.

Among the gene tribes, 67 major families of transcription factors (TF) were annotated, with over 6,000 members in papaya and Arabidopsis. Most of the TF families had fewer members in papaya when compared to Arabidopsis. The WRKY superfamily, for instance, has 213 members in Arabidopsis and only 66 members in papaya. Many WRKY genes are involved in plant defense, abiotic stress responses, trichome morphology, and a number of other plant processes (Zhang and Wang 2005). The low number of WRKY genes is reflective of the reduced number of disease resistance genes in the papaya genome. The NAC superfamily is an important, highly conserved family with fewer members in papaya (95 vs 136). The NAC family is one of the largest plant-specific families, and members have diverse functions ranging from development to defense and abiotic stress responses (Olsen et al. 2005). Although the majority of TF families have fewer members in papaya, several families including the MADS-box family have significantly more gene members.

Major Findings

The MADS-Box Family in Papaya

The molecular mechanisms governing flower development are of tremendous interest agriculturally. Understanding the pathways that trigger flowering and the genes that control floral organ identity can be useful for the artificial induction of flowering and the homeotic conversion of flower organs. Many of the genes controlling flower development belong to the MADS-box protein family. MADS-box proteins are found in most eukaryotic organisms and are for the most part, involved in signal transduction and development. MADS-box proteins are named for their highly conserved 58 amino acid DNA-binding motif. The diversification of MADS-box genes has been linked to the evolution of novel characteristics and development patterns in eukaryotes (Gramzow et al. 2010). The papaya genome shows a dramatic increase in MADS-box proteins in comparison to other sequenced plant genomes.

MADS-box genes are especially abundant in plants, displaying a diverse range of developmental functions. In Arabidopsis, they are divided into five distinct groups based on the phylogenetic relationship of their MADS-box domains: Mα, Mβ, Mγ, Mδ, and MIKC^c (Parenicova et al. 2003). A duplication event before the divergence of plants and animals gave rise to two evolutionarily distinct lineages of MADS-box proteins (Alvarez-Buylla et al. 2000). Mα, Mβ, and Mγ belong to lineage I, and Mδ and MIKC^c belong to lineage II. MIKC^c is the largest clade in

Arabidopsis, and it contains most of the MADS-box genes involved in floral and reproductive development. Type II MADS-box proteins are functionally well characterized and conserved between species. In stark contrast, the function of most type I genes remains largely elusive. Type I genes typically have weak expression signals and were initially annotated solely from genomic studies. Recent work, however, has shed light onto the function of type I MADS-box genes, demonstrating roles in female gametophyte, seed, and ovule development. Arabidopsis T-DNA lines with insertions in AGL23 showed arrest in female gametophyte development and albino seeds that are unable to give rise to viable plants (Colombo et al. 2008). Mutants in a second type I gene, AGL61, developed ovules without fused polar nuclei and aberrant cell morphology (Bemer et al. 2008). AGL37 is a Mγ family member shown to be involved in seed development (Kohler et al. 2003). Protein homology suggests that other type I MADS-box genes also play roles in female gametophyte development. Papaya has a dramatic increase in type I MADS-box genes.

Most transcription factor families in papaya have fewer members than the other sequenced plant genomes. The MADS-box family is an exception, however, as papaya has significantly more genes in this family than any other sequenced plant genome. Overall, papaya contains 171 MADS-box genes. 145 MADS-box proteins belong to the type I lineage and 25 belong to the type II lineage. Papaya has a dramatic expansion in type I genes in comparison to Arabidopsis, poplar, and apple, which contain 94, 50, and 56 members, respectively. The largest expansion is seen in the Mα clade. This expansion is intriguing, as it is the only annotated type I gene that functions in female gametophyte development. Perhaps duplication and diversification of type I MADS-box genes played a role in papayas adaptation as a trioecious species. In poplar, the dramatic expansion of type II MIKC^c genes has been suggested to aid in the adaptive requirements needed for dioecious trees (Leseberg et al. 2006). Papaya has fewer Mβ genes than Arabidopsis or poplar. Grape, however, has no MADS-box genes that cluster in the Mβ subclade, suggesting the role of this clade is less evolutionarily conserved (Ming et al. 2008). Type I MADS-box genes in Arabidopsis and rice show a significant ratio of pseudogenes to functional genes, suggesting a high birth and death rate. The functionality of MADS-box genes in papaya has not been tested.

Although papaya has significantly fewer type II MADS-box genes, functions of the remaining genes are likely conserved. Six MIKC^c have been annotated in papaya and show similar function and expression patters to their homologs in Arabidopsis (Ackerman et al. 2008; Yu et al. 2008). CpFUL, a homolog of *FRUITFUL* (*FUL*) in the *APELLATA1* (*AP1*) subfamily, was reported to have expression in papaya flowers and leaves (Yu et al. 2008). *AP1* is an A class gene in Arabidopsis, and *FUL* is an *AP1* subfamily gene that controls silique elongation and development. The *paleoAP3* paralogs, *CpTM6-1* and *CpTM6-2*, and the *PI* ortholog *CpPI* were shown to have petal-specific expression in all three sex types and stamen-specific expression in males and hermaphrodites (Ackerman et al. 2008). Sex-specific expression was also seen in the *AGAMOUS* (*AG*) subfamily genes *CpSTK* and *CpPLE*. *CpSTK* is homologs of *SEEDSTICK* (*STK*), a D class

gene that plays a role in ovule and seed development. *CpSTK* has carpel-specific expression in female and hermaphrodite flowers and no expression in males. *CpPLE* transcripts were found to be stamen- and carpel-specific, suggesting that *CpPLE* may be the sole C class gene in papaya. A papaya C class gene homolog to the Arabidopsis AG has not been identified (Ackerman et al. 2008).

Repeat Composition

Genome size is exceptionally variable in plants, and repetitive elements play a major role in genome expansion. Repeats are abundant in the papaya genome, representing 52 % of the draft genome. Much of the repeats in the papaya are papaya specific, as preexisting repeat databases masked only 14 % of the genome, suggesting a lower than expected repeat composition (Nagarajan et al. 2008; Ming et al. 2008). Instead, de novo repeat annotation was used to accurately gauge repeat composition. Papaya-specific repeats were identified from the draft sequence using a suite of repeat identification software. The annotated repeats identified 889 repeat families. This repeat database also serves as a valuable resource for annotating other sequenced plant genomes. 43.4 % of the papaya genome is homologous to identifiable repetitive elements, and an additional 8.5 % represent unannotated repeats. The diversity of repetitive elements in the papaya genome is astounding, but the vast majority (40 % of the genome) are retrotransposons. 69.5 % of the retrotransposons are *Ty3-gypsy* elements and 13.5 % are *Ty1-copia* elements. Retrotransposons are particularly abundant in most plant genomes, at least those that have been sequenced, because of their mechanism of transposition. MITEs, MuDR, and other types of DNA transposons are excised before translocation. Although transposons frequently move around the genome, the copy number stays the same. Retrotransposons amplify themselves using an RNA intermediate, increasing in copy number and rapidly expanding the genome. The wheat genome is almost 90 % retrotransposons, and in maize, approximately 70 % of the genome is composed of retrotransposons. The percentage of repeats in papaya is likely a low estimate, as much of the unsequenced genome fragments represent repetitive regions. DNA transposons have an extremely low abundance in the papaya genome, representing less than 0.2 % of the total sequence length. This percentage is lower than other plant genomes.

Tandem repeats are widely distributed throughout the papaya genome and represent 1.3 % of the total sequence. Tandem repetitive elements have individual repeating units ranging in size from 1 to 2,000 bp and copy numbers as high as 1,000. Over 57,000 distinct tandem repeat loci were annotated in the papaya draft genome, encompassing a total of 4.8 Mbp. The average repeat length was 79 bp, and the average copy number is 7.3. Minisatellites and satellites (>6 bp repeats) make up the bulk of terminal repeats. Surprisingly, very few tandem repeats characteristic of telomeric and centromeric regions were found. Telomere tracts are relatively long (10 kb) in papaya as revealed by terminal restriction fragment analysis (Skakirov and Shippen, unpublished data). Four clusters of telomeric repeat

motifs (TTTTAGGG) were found, but no centromeric repeat motifs were detected. This finding accurately portrays the unfinished state of the draft genome.

Organelle genome integration is a common feature of plant nuclear genomes, and both chloroplast- and mitochondria-derived sequences are abundant in the papaya genome. Organelle genomes were once comparable in size to the genomes of cyanobacteria, but major gene losses in organelle genomes during the two billion years since their inception have significantly reduced their size. Nuclear genomes are the recipients of these lost organelle sequences and have subsequently expanded drastically in both size and metabolic function (Timmis et al. 2004). The Arabidopsis genome, for instance, contains a large 620 kb mitochondrial DNA insertion and 17 chloroplast insertions totaling 11 kb. Organelle-to-nucleus transfers are still occurring, and a number of integration events are prominent in the papaya genome. The papaya genome contains 786 kb of interspersed chloroplast sequences and 858 kb mitochondria sequences. Overall the chloroplast sequences represent 0.28 % of the genome, and mitochondria represent 0.3 % of the papaya genome (Ming et al. 2008). Sequences from nearly all of the chloroplast genome have been found in the papaya genome, but only around half of the mitochondria genome is present in the nuclear genome.

Synteny Analysis and Genome-Wide Duplications

Polyploidization is among the most important force driving genome evolution. Newly duplicated genes can undergo subfunctionalization or neofunctionalization to take on new, adaptive roles. Increased gene copy number can help shield organisms against deleterious mutations, and enforced homologous recombination in allopolyploids helps to maintain heterosis. An estimated 15 % of angiosperm and 30 % of fern speciation events have been accompanied by an increase in ploidy (Wood et al. 2009). A paleopolyploidization event predates the monocot-eudicot divergence and likely facilitated the enormous radiation of plant species (Bowers et al. 2003). Polyploidization is apparent in most crop plants, as well as many eudicots, cereals, and a range of other plant families (Bowers et al. 2003; Paterson et al. 2004).

Papaya is no exception and also been subjected to polyploidization events in the course of its evolution. Alignment and subalignment of the papaya genome to itself revealed an ancient genome triplication event. 25 % of the sequence from the 200 largest scaffolds demonstrates triplication patterns. This pattern is shared by poplar and a number of other sequenced plant genomes, suggesting an ancient triplication event near the origin of angiosperms (Ming et al. 2008; Cui et al. 2006). Aside from the early triplication event, papaya has not undergone any recent genome duplication events. This is supported by the L-shaped distribution of intra-EST correspondence. A recent duplication event would produce a curve with a large peak representative of paralogous gene pairs. Alignment of collinear regions from papaya and Arabidopsis indicates several genome duplication events in Arabidopsis. The 200 largest scaffold sequences from papaya (representing 247 Mb) were aligned to the

Arabidopsis genome. 121 collinear blocks were observed, ranging from 0.16 Mb and 19 genes to 1.36 Mbp and 181 genes. 26 collinear regions corresponded to one segment in Arabidopsis, 41 to 2 segments, 21 to 3, 30 to 4, and 3 to more than 4 (Ming et al. 2008).

A significant portion of the collinear regions in papaya matched three or four regions in Arabidopsis. This suggests that two genome duplication events occurred after the divergence (72 MYA) of papaya from the Arabidopsis lineage. A significant portion of the collinear regions align to one Arabidopsis segment, further proving that papaya has not had recent genome duplication. The most recent Arabidopsis genome duplication "alpha" was thought to effect a subset of the *Brassicales*, supporting this finding (Bowers et al. 2003). However, phylogenetic dating of the more ancient "beta" duplication suggests an event predating the *Arabidopsis–Carica* divergence. Phylogenetic dating and collinearity results are contradictory, and this incongruity is still a mystery. Although the papaya genome has not undergone any recent genome duplication events, it is not the best representative of the ancestral angiosperm genome. Changes in gene composition and genome structure are rapid in papaya, making it a poor model. Like papaya, grape has only experienced the ancient triplication event shared by all angiosperms, and it has a much slower rate of structural evolution (Tang et al. 2008). Thus, grape is currently the best representation of the ancient angiosperm genome.

Characterization of Transgene Insertions in SunUp

Papaya is the first transgenic organism to have its genome sequenced. Genetic engineering via Agrobacterium-mediated transformation and particle bombardment both produce relatively random integration of the transgene into the host genome. The variable integration site significantly affects expression of the transgenes and can potentially lead to gene disruption. Southern blot hybridization is useful for assessing transgene copy number but gives no indication of the physical location of the transgene insert. Furthermore, small, unintended inserts can be difficult to distinguish from full-length inserts in restriction fragment banding patterns. The SunUp cultivar was developed from an R_0 Sunset line that was transformed via particle bombardment to contain a transgene for the coat protein of the papaya ringspot virus (PRSV). Segregation patterns revealed consistent co-segregation of PRSV resistance and plasmid-derived β-glucuronidase (*uidA*) and neophosphotransferase (*nptII*) genes, indicating that the transgene was intact (Lius et al. 1997). However, the integration site, unintended inserts, and stability of the insert were unknown. Sequencing the papaya genome identified all of the plasmid-derived sequences and their exact genome positions.

All of the functional transgenes were found in a single, 9,789 bp insert. The stable insert contained intact copies of the PRSV coat protein, *uidA*, and *nptII* genes. Two unintended inserts were found in the draft genome. The first was a nonfunctional *nptII* sequence comprising 290 bp of the 3′ end. The second insert was a

1,533 bp fragment composed of a truncated *tetA* gene and flanking vector backbone. The flanking genomic regions of the PRSV transgene insertions reveal sequence preferences for integration. Interestingly, five of the six sequences flanking the transgene inserts were chloroplast-derived nuclear sequences. Agrobacterium-mediated and biolistic approaches typically integrate into AT-rich regions (Sawasaki et al. 1998). Chloroplast sequences are notoriously AT-rich making chloroplast sequences prime candidates for integrations. Furthermore, four of the flanking sequences were at junctions matching topoisomerase I recognition sites. These recognition sites are associated with double-stranded breaking points in the DNA and likely represent sites easily opened for transgene insertion. Identification of the precise transgenic modifications in SunUp should serve to dissipate fears of genetically modified food crops and possibly lower the barriers restricting its production currently in place in several countries.

Concluding Remarks and Future Prospects

The genetic and genomic resources of papaya have advanced rapidly in the past decade, fostering significant advances in comparative plant genomics, papaya improvement, and sex chromosome evolution. Papayas' small genome, low heterozygosity, short generation time, and abundant progeny production have, in part, facilitated these advances. A whole-genome shotgun approach was applied, along with a high-density linkage map and physical map to generate the papaya draft genome. A large collection of ESTs were anchored to this assembly and used as a foundation for annotating the genome. The resulting gene models revealed a number of insights into fruit nutrition and development, as well as disease resistance. The newly annotated genes and ESTs provide powerful insight into crop improvement and cloning of desirable traits. The papaya genome is free of recent genome duplications, making it a useful resource for studying angiosperm evolution. The draft genome and genetic map also revealed several co-segregating markers and sequence regions corresponding to the incipient sex chromosomes in papaya. These markers formed the foundation for physical mapping, sequencing, and annotating the sex chromosomes.

Although the draft genome encompasses 91 % of the ESTs and 92 % of the genetic markers, a significant proportion of the sequence remains unassembled (roughly 25 %), making it difficult to conduct comparative genomics analyses. It is likely that genes and entire regions are missing entirely from the assembly, invalidating certain findings. Furthermore, the low sequence coverage has likely led to assembly errors, resulting in erroneous scaffold arrangements, indels, and base-calling errors. Completion of the papaya draft genome will alleviate these problems and greatly improve the usefulness of the papaya genome to the research community. Refinement of the genetic map will anchor more scaffold sequences and bridge gaps in the linkage groups. The finished papaya genome will also result in the annotation of additional genes missing from the original assembly. Resequencing projects of plants in the Caricaceae family will greatly expedite the discovery of genes

for crop improvement. *Vasconcellea cauliflora*, for instance, is naturally resistant to PRSV, and *V. cundinamarcensis* could provide genes for cold tolerance (Manshardt and Wenslaff 1989). Most of the 35 species in the Caricaceae family are dioecious, and one is monoecious, making them useful species for comparative genomics of sex chromosome evolution in papaya.

References

Ackerman CM, Yu Q, Kim S, Paull RE, Moore PH, Ming R (2008) B-class MADS-box genes in trioecious papaya: two TM6 paralogs, CpTM6-1 and CpTM6-2, and a PI ortholog CpPI. Planta 227:741–753

Alvarez-Buylla ER, Pelaz S, Liljegren SJ, Gold SE, Burgeff C, Ditta GS, Ribas dePouplana L, Martinez-Castill L, Yanofsky MF (2000) An ancestral MADS-box gene duplication occurred before the divergence of plants and animals. Proc Natl Acad Sci USA 97:5328–5333

Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. Plant Mol Biol Rep 9:208–218

Bemer M, Wolters-Arts M, Grossniklaus U, Angenent GC (2008) The MADS domain protein DIANA acts together with AGAMOUS-LIKE80 to specify the central cell in *Arabidopsis* ovules. Plant Cell 20:2088–2101

Bowers JE, Chapman BA, Rong JK, Paterson AH (2003) Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. Nature 422:433–438

Chandrika UG, Jansz ER, Wickramasinghem SMDN, Warnasuriya ND (2003) Carotenoids in yellow- and red-fleshed papaya (*Carica papaya* L). J Sci Food Agric 83:1279–1282

Chen C, Yu Q, Hou S, Li Y, Eustice M, Skelton RL et al (2007) Construction of a sequence-tagged high density genetic map of papaya for comparative structural and evolutionary genomics in Brassicales. Genetics 177:2481–2491

Colombo M, Masiero S, Vanzulli S, Lardelli P, Kater MM, Colombo L (2008) *AGL23*, a type I MADS-box gene that controls female gametophyte and embryo development in *Arabidopsis*. Plant J 54:1037–1048

Cui L, Wall PK, Leebens-Mack JH, Lindsay BG, Soltis DE et al (2006) Widespread genome duplications throughout the history of flowering plants. Genome Res 16:738–749

Gramzow L, Ritz MS, Theissen G (2010) On the origin of MADS-domain transcription factors. Trends Genet 26:149–153

Hanada K, Zhang X, Borevitz JO, Li WH, Shiu SH (2007) A large number of novel coding small open reading frames in the intergenic regions of the *Arabidopsis thaliana* genome are transcribed and/or under purifying selection. Genome Res 17:632–640

Huang S, Li R, Zhang Z, Li L, Gu X et al (2009) The genome of the cucumber, Cucumis sativus L. Nat Genet 41:1275–1281

International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. Nature 436:793–800

Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A et al (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449:463–465

Kohler C, Hennig L, Spillane C, Pien S, Gruissem W, Grossniklaus U (2003) The Polycomb group protein MEDEA regulates seed development by controlling expression of the MADS-box gene PHERES1. Genes Dev 17:1540–1553

Leseberg CH, Li A, Kang H, Duvall M, Mao L (2006) Genome-wide analysis of the MADS-box gene family in Populus trichocarpa. Gene 378:84–94

Liu Z, Moore PH, Ma H, Ackerman CM, Makandar R, Yu Q, Pearl HM, Kim MS, Charlton JW, Stiles JI, Zee FT, Paterson AH, Ming R (2004) A primitive Y chromosome in papaya marks incipient sex chromosome evolution. Nature 427:348–352

Lius S, Manshardt RM, Fitch MMM, Slightom JL, Sanford JC et al (1997) Pathogen-derived resistance provides papaya with effective protection against papaya ringspot virus. Mol Breed 3:161–168

Manshardt RM, Wenslaff TF (1989) Interspecific hybridization of papaya with other Carica species. J Am Soc Hortic Sci 114:689–694

Ming R, Moore P, Zee F, Abbey C, Ma H, Paterson AH (2001) A papaya BAC library as a foundation for molecular dissection of a tree-fruit genome. Theor Appl Genet 102:892–899

Ming R, Hou S, Feng Y, Yu Q et al (2008) The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya Linnaeus*). Nature 452:991–996

Nagarajan N, Navajas-Pérez R, Pop M, Alam M, Ming R, Paterson AH, Salzberg SL (2008) Genome-wide analysis of repetitive elements in papaya. Trop Plant Biology 1:191–201

Olsen A, Ernst H, Leggio LL, Skriver K (2005) NAC transcription factors: structurally distinct, functionally diverse. Trends Plant Sci 10:79–87

Parenicova L, de Folter S, Kieffer M, Horner DS, Favalli C, Busscher J, Cook HE, Ingram RM, Kater MM, Davies B, Angenent GC, Colombo L (2003) Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in Arabidopsis: new openings to the MADS world. Plant Cell 15:1538–1551

Paterson AH, Bowers JE, Chapman BA (2004) Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. Proc Natl Acad Sci USA 101:9903–9908

Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J et al (2009) The Sorghum bicolor genome and the diversification of grasses. Nature 457:551–556

Sawasaki Y, Inomata K, Yoshida K (1998) Trans-kingdom conjugation between Agrobacterium tumefaciens and Saccharomyces cerevisiae, a bacterium and a yeast. Plant Cell Physiol 37:103–106

Schmutz J, Cannon S, Schlueter J, Ma J et al (2010) Genome sequence of the palaeopolyploid soybean. Nature 463:178–183

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S et al (2009) The B73 maize genome: complexity, diversity, and dynamics. Science 326:1112–1115

Storey WB (1969) Papaya. In: Ferwerda FP, Wit F (eds) Outlines of perennial crop breeding in the tropics. H Veenman & Zonen, Wageningen, pp 389–408

Tang H, Bowers JE, Wang X, Ming R, Alam M, Paterson AH (2008) Synteny and collinearity in plant genomes. Science 320:486–488

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796–815

Timmis JN, Ayliffe MA, Huang CY, Martin W (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat Genet Rev 5:123–135

Tuskan GA et al (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr & Gray). Science 313:1596–1604

Velasco R, Zharkikh A, Affourtit J, Dhingra A et al (2010) The genome of the domesticated apple (*Malus*×*domestica* Borkh.). Nat Genet 42:833–839

Wall PK, Leebens-Mack J, Muller K, Field D, Altman N, dePamphilis CW (2008) PlantTribes: a gene and gene family resource for comparative genomics in plants. Nucleic Acids Res 36:970–976

Wang J, Na JK, Yu Q, Gschwend AR, Han J et al (2012) Sequencing papaya X and Yh chromosomes reveals molecular basis of incipient sex chromosome evolution. Proc Natl Acad Sci USA 109:13710–13715

Wood TE, Takebayashi N, Barker MS, Mayrose I, Greenspoon PB, Rieseberg LH (2009) The frequency of polyploidy speciation in vascular. Proc Natl Acad Sci USA 106:13875–13879

Yu Q, Steiger D, Kramer E, Moore P, Ming R (2008) Floral MADS-Box genes in trioecious papaya: characterization of AG and AP1 subfamily genes revealed a sex-type-specific gene. Trop Plant Biol 1:97–107

Yu Q, Tong E, Skelton RL, Bowers JE, Jones MR, Murray JE et al (2009) A physical map of the papaya genome with integrated genetic map and genome sequence. BMC Genomics 10:371–382

Zhang YJ, Wang LJ (2005) The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants. BMC Evol Biol 5:1–12

Zhang W, Wai CM, Ming R, Yu Q, Jiang J (2010) Integration of genetic and cytological maps and development of a pachytene chromosome-based karyotype in papaya. Trop Plant Biol 3:166–170

Chapter 12
Syntenic Sequence Conservation Between and Within Papaya Genes

Eric Lyons and Haibao Tang

Introduction

2008 welcomed the publication of the papaya genome (Ming et al. 2008). At the time, it was the fifth angiosperm sequenced following *Arabidopsis* (*Arabidopsis thaliana* Col-0) in 2000 (The Arabidopsis Genome Initiative 2000), rice (*Oryza sativa japonica* and *indica*) in 2002 (Goff et al. 2000; Yu et al. 2002), poplar in 2006 (Tuskan et al. 2006), and grapevine (*Vitis vinifera*) in 2007 (Jaillon et al. 2007). In addition, the bryophyte moss, *Physcomitrella patens* had been published the prior month (Rensing et al. 2008). Of these genomes, papaya was unique for two major reasons. First, it was the first transgenic organism to have its genome sequenced, and second, it was of rather low coverage (three times fold coverage) which resulted in a relatively poor assembly. At around the same time, high-throughput sequencing technology started to become readily accessible and soon came a plethora of additional plant genomes. Four years later, there are approximately 30 plant genomes publicly available, but those utilizing only high-throughput sequencing technologies also suffer from poor genome assemblies due to the lack of repeat resolution using only short sequence reads. The lessons learned from papaya for comparative genomics have proved particularly useful for these relatively easy to obtain genome assemblies. This chapter will focus on the methods used to identify and compare syntenic sequences within the papaya genome and among related eurosids, starting from whole-genome comparisons and drilling down to identify short, evolutionary conserved noncoding sequences (CNS) neighboring syntenic gene sets.

E. Lyons (✉)
School of Plant Sciences, University of Arizona, Bio5 Institute,
1657 E. Helen St, Tucson, AZ 85721, USA
e-mail: ericlyons@email.arizona.edu

H. Tang
Center for Plant Genomics and Biotechnology, Fujian Agriculture and Forestry University,
Fuzhou 350002, China

R. Ming and P.H. Moore (eds.), *Genetics and Genomics of Papaya*, Plant Genetics and Genomics: Crops and Models 10, DOI 10.1007/978-1-4614-8087-7_12,

The primary set of tools for performing these analyses will come from the publicly accessible Web-based comparative genomics platform, CoGe: http://genomevolution.org (Lyons and Freeling 2008). Links for recreating the presented example analyses will be present throughout the chapter, and readers are encouraged to repeat the analyses online.

All angiosperms sequenced today share a genomic evolutionary history rife with paleopolyploidy events (Tang et al. 2008; Jiao et al. 2011). As observed in modern plants, polyploidy events contemporaneously create a duplicate copy of every chromosome and their underlying genetic features. Such polyploids are created via two general mechanisms: autopolyploidy, where a single parental genome is duplicated, and allopolyploidy, where multiple diverged diploid parental genomes combine. While autopolyploidy is feasible and has a fair number of occurrences in nature, most ancient polyploids are thought to be allopolyploids in nature due to the increase in allelic genetic variable that multiple parent genomes bring to a progeny (Osborn et al. 2003). Such variability may create a variety of beneficial phenotypes similar to those seen due to heterosis (Comai 2000), and the increase in the diversity of genetic elements may help drive long-term morphological complexity and adaptation to new environments (Crow and Wagner 2006). However, such massive genomic events are not with perilous consequences. Massive genomic restructuring often happens after allopolyploids which have a high likelihood of reducing subsequent viable and competitive offspring (Gaeta and Pires 2010), and such events leading to stable lineages over evolutionary time are indeed rare (Mayrose et al. 2011).

Paleopolyploidy events are inferred by identifying syntenic regions within a genome with synteny being defined as genomic regions sharing a common ancestry. If a substantial portion of a genome is covered by intragenomic syntenic regions, then the genome is likely to be derived from a polyploid ancestor. Intragenomic syntenic regions are inferred by identifying a collinear series of homologous genes within a genome, and a similar technique may also be used to infer syntenic regions between genomes (aka orthologous regions). Additionally, the evolutionary distance of such syntenic genes may be calculated using a neutral evolutionary model on a presumably neutral marker, such as the wobble position in codons (Blanc et al. 2003). If the evolutionary distance of population of syntenic genes appears to come from the same statistical distribution, they may be inferred to have been derived from the same evolutionary event, such as polyploidy (Lynch and Conery 2000). One problem with detecting paleopolyploidy events is the decay of the collinear signal used to infer synteny over evolutionary time. While genomes may undergo a variety of changes such as translocations, inversions, and chromosomal fusions and fissions, gene loss is the major factor obfuscating ancient whole-genome duplications. Following polyploidy, genomes undergo a diploidization process whereby many, but not necessarily all, of the duplicated genes are lost from one of the homeologous syntenic regions (or more, depending on the ploidy number). This process of gene loss following polyploidy is known as fractionation (Langham et al. 2004), and the more time that has past, the more genes are likely to be lost. As we will show in the upcoming exercises, gene losses within the duplicated regions occur in

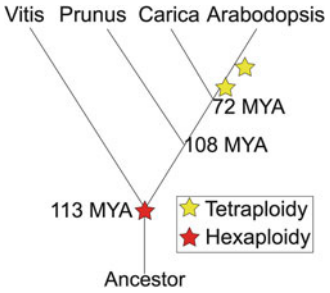

Fig. 12.1 Eurosid phylogeny and polyploidy. 72 MYA (Couvreur et al. 2010). 108 MYA (Magallón and Castillo 2009). 113 MYA (Magallón and Castillo 2009)

a reciprocal, intercalated pattern, which became evident when an unduplicated outgroup is used. Compounding this gene loss problem are sequential polyploidy events. Each new polyploidy event followed by fractionation makes inferring older polyploidies more difficult as the signal of collinear gene order become weaker due to fewer collinear genes retained in duplicate.

Papaya is a member of the eudicots, and its evolutionary history, as with all eudicots, contains a paleohexaploidy event. This event was first inferred by careful analysis of the *Arabidopsis* genome and was termed "gamma" event (Bowers et al. 2003), but neither its hexaploid nature nor evolutionary distribution was discernible until the sequencing of the grapevine genome. This was due to the *Arabidopsis* lineage having undergone two subsequent tetraploidy events (alpha and beta events) (Bowers et al. 2003), while the grapevine genome had none (Jaillon et al. 2007). Papaya's genome was remarkable due to the fact that like the grapevine genome, it too had no subsequent polyploidy events (Ming et al. 2008). Indeed, its timely release confirmed the discovery in the grapevine genome and helped to resolve a competing interpretation at the time (Velasco et al. 2007). The evolutionary relationship of these lineages is shown in Fig. 12.1. Also marked on this figure are the paleopolyploidy events in these lineages and their approximate divergence ages (Magallón and Castillo 2009; Couvreur et al. 2010).

Whole-Genome Comparison of the Papaya Genome

The papaya genome assembly contains over 3,000 contigs, and of the ~331 million nucleotides in the assembled genome, nearly 30 % are the ambiguous nucleotide "N." In comparison, the theoretical value for unsequenced portions based on Lander–Waterman theory (Lander and Waterman 1988) is only 5 % (e^{-3}). Fortunately, the majority of the missing sequences are likely made up of genomic repeat elements such as transposons, and the majority of genic regions of the genome have been sequenced (Ming et al. 2008). While this presents some challenges for comparative genomic methodologies, there are tools that can facilitate the analytical process and other sequenced genomes to which papaya may be compared to leverage

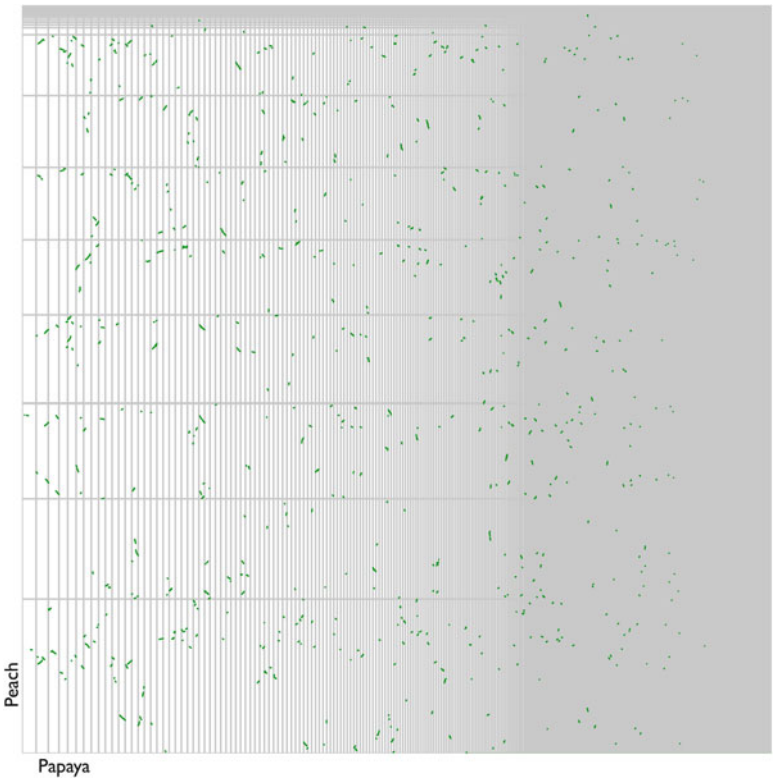

Fig. 12.2 Syntenic dotplot of papaya versus peach (generated using http://genomevolution.org)

the information those genomes contain. The standard method for identifying syntenic regions within or between genomes starts with either the sequence for all the protein-coding genes or the entire genome. These are compared to one another to identify putative homologous genes or sequences, usually using a BLAST-like algorithm such as BlastN (Altschul et al. 1990) or BlastZ (Schwartz et al. 2003). These results may be filtered to condense tandem duplicated genes or remove dispersed repetitive sequences such as transposons. Next, these data are computationally processed to identify collinear sets of homologous genes or sequences; DAGChainer (Haas et al. 2004) is an excellent algorithm for doing that. These sets of collinear homologs are used to infer syntenic regions of genomes. There are various forms for visualizing whole-genome comparisons (Nielsen et al. 2010), and, of those, dotplots are excellent for comparing two genomes.

Figure 12.2 shows a syntenic dotplot between papaya and peach. Along each axis of a syntenic dotplot is the linear representation of a genome: *x*-axis is papaya and *y*-axis is peach. Horizontal and vertical gray lines separate chromosomes or contigs from one another (the gray rectangle on the right side of the dotplot is due to

the many very small contigs of papaya thus dense contig breaks). For each syntenic gene pair identified between the two genomes, a green dot is drawn. A series of these green dots makes a line, defining which regions of these genomes are syntenic with respect to one another. By examining the papaya–peach syntenic dotplot closely, there are many regions of each genome that are syntenic to multiple regions. This is due to these genomes' shared paleohexaploid ancestry (Fig. 12.1). The peach genome is picked as a comparator genome to papaya because it is relatively closely related to papaya, and its genome has not undergone any subsequent polyploidy events since the eurosid paleohexaploidy (Fig. 12.1).

To Generate a Dotplot Between Papaya and Peach Using CoGe

1. Go to CoGe's homepage. Quick link: http://genomevolution.org.
2. Select the tool SynMap from either the list under "CoGe's Entrance Tools" or the "Applications" menu located in the upper-right part of the Web page. Quick link: http://genomevolution.org/CoGe/SynMap.pl.
3. From SynMap, search for the name "papaya" in the "Organism 1 Search" box. You will be presented with a list of organisms whose names match "papaya." From that list, select "Carica papaya."
4. Under these organisms will be a box labeled "Genome:" followed by a drop-down menu. Select the latest version of the genome. If a masked genome is available, you may select it. Masked genomes have been analyzed for repeat sequences and have those sequences converted to "X" therefore ignored in the subsequent analyses. When comparing large genomes with many repetitive elements (e.g., maize), selecting masked genomes may result in faster analysis times.
5. Search for the name "peach" in the "Organism 2 Search" box. Select the organism labeled "Prunus persica (peach)" from the list of returned organisms and select the most recent version of the peach genome.
6. SynMap quick link with genomes selected: http://genomevolution.org/r/487v.
7. Press the red button labeled "Generate SynMap" to run the analysis. The analysis is a multistep process and may take some time to complete. Results from previous comparisons are cached by the system and reused when appropriate in order to cut down on the analysis time. Quick link: http://genomevolution.org/r/470z.
8. Note that below the dotplot, SynMap also provides links to all the data files used to generate the results. Below this is a green button labeled: "Regenerate this analysis" followed by a link. If you copy and use this link, it will load SynMap with all the necessary genomes and parameters to regenerate the exact analysis performed. This is what is used to generate the quick links for these SynMap tutorials.
9. For more information on SynMap, please see its online manual: http://genomevolution.org/r/488j.

Syntenic Path Assembly

Due to the large number of contigs comprising the papaya genome, each identified syntenic region is quite small. One advantage of having a relatively contiguous comparator genome such as peach, which has not undergone any subsequent polyploidy events, is that the contigs of papaya may be arranged and oriented with respect to the structure of the peach genome. This process, known as a "syntenic path assembly" (Lyons et al. 2011), creates a pseudo-assembly of the papaya genome and can help visually identify syntenic regions.

Figure 12.3 shows a syntenic dotplot of papaya (*x*-axis) and peach (*y*-axis) using the syntenic path assembly. The contigs of the papaya genome have been ordered and oriented relative to the peach genome. With this type of visualization, it is relatively straightforward to see that the majority of the peach genome is covered by syntenic regions from the papaya genome. This forms the general pattern of a green line starting in the lower-left corner of the dotplot and extending to the upper-right corner. Also, there are additional, but weaker, syntenic lines that do not fall along

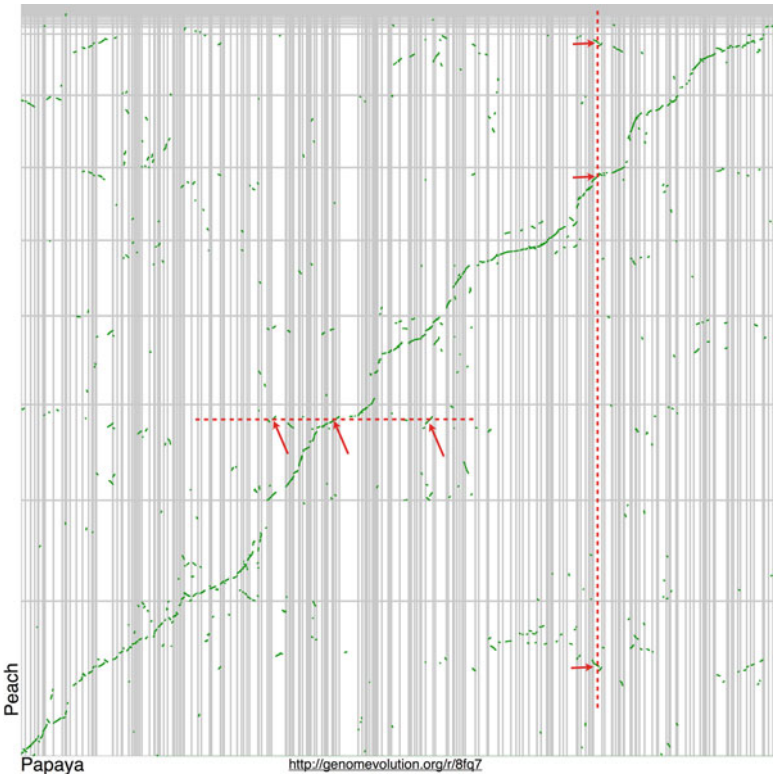

Fig. 12.3 Syntenic path assembly of papaya versus peach (generated using http://genomevolution.org)

this ~45° line. These are presumably derived from their shared eurosid paleohexaploidy event. By examining the dotplot carefully, there are several occurrences where for a given genomic region in one organism, there are three syntenic regions present in the other genome (Fig. 12.3; red dashed lines and arrows).

To Generate a Syntenic Path Assembly Between Papaya and Peach Using CoGe

1. Start with the results from a prior SynMap analysis. Quick link: http://genomevolution.org/r/470z.
2. Select the "Display Options" tab. This tab contains various options to modify the visualization of the syntenic dotplot.
3. Select the checkbox next to the option "Order contigs by best syntenic path."
4. Rerun the analysis by pressing "Generate SynMap." Quick link: http://genomevolution.org/r/4710.
5. When the syntenic dotplot is returned using the best syntenic path, you will notice that there will be a large gray region of papaya contigs off to the edge of the dotplot. These are all the contigs, ordered by size, which did not have any syntenic match to the peach genome. These contigs may be removed from the visualization by selecting the checkbox next to "Don't show contigs without synteny" located next to "Order contigs by best syntenic path." This option is only visible after the syntenic path assembly option has been selected. Quick link: http://genomevolution.org/r/487y.
6. SynMap will dynamically scale the size of the dotplot depending on the number of chromosomes/contigs that are being compared. To force the dotplot to conform to a specific width, you may enter a pixel size next to the option "Master image width (0 == dynamic)." Quick link: http://genomevolution.org/r/4880.

Measuring Evolutionary Distance of Syntenic Gene Pairs

One problem with the syntenic dotplot visualization in Fig. 12.3 is differentiating orthologous gene pairs derived from the divergence of the peach and papaya lineages versus out-paralogous syntenic gene pairs derived from the paleohexaploidy event. These evolutionary relationships are shown in Fig. 12.4, and for an excellent review, please read Koonin (2005).

One method to differentiate orthologous and out-paralogous syntenic gene pairs between papaya and peach is by measuring their evolutionary distance. Orthologous genes between papaya and peach will be younger with respect to one another than out-paralogous genes. In order to calculate an evolutionary distance metric, there needs to be a proxy measurement for evolutionary time. Since genomic sequences change over time, these changes may be used as such a proxy. However, the sequences

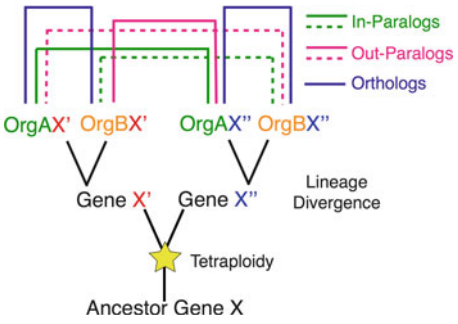

Fig. 12.4 Evolutionary relationships of genes

that are changing should be selectively neutral and therefore are changing due to random mutations alone over the passage of time. As such, sequences under positive selection (selected to change) and under purifying selection (selected not to change) need to be avoided. The wobble nucleotide of a codon may change without affecting the encoded amino acid, and such synonymous genomic changes are often used to measure evolutionary distance between protein-coding sequences. This metric is called the synonymous mutation value (Ks), and the program CodeML (Yang 2007) is frequently used to calculate it for a pair of protein-coding sequences.

Figure 12.5a shows a histogram of the $\log_{10}$ Ks values calculated by CodeML for each syntenic gene pair identified between papaya and peach. Lower Ks values are on the left side of the histogram and signify a smaller evolutionary divergence time between the gene pairs. A color palette has been overlaid on these values to accentuate the differences, and those colors are propagated to their corresponding dots in the syntenic dotplot (Fig. 12.5b). Now it is readily discernible that the majority of the dots falling on the diagonal are mostly green (lower Ks values) while the others dots are mostly red/purple (higher Ks values). The green dots identify gene pairs that are evolutionarily closer than red/purple dots, representing syntenic genes derived from the divergence of these lineages (orthologs) and their shared paleo-hexaploid ancestry (out-paralogs), respectively. Do note that some green dots fall off of the diagonal (Fig. 12.5b; red arrows) and show genomic rearrangements between these genomes (such as inversion) or misassemblies of one of the genomes (Schnable and Lyons 2011).

To Color Dots Based in a Syntenic Dotplot Based on the Evolutionary Distance of Syntenic Gene Pairs Using Synonymous Substitution Values

1. Start with the results of a prior SynMap analysis. Quick link: http://genomevolution.org/r/4880.
2. Select the "Analysis Options" tab. This tab contains various options to modify the parameters used by SynMap's analytical pipeline.

3. Next to the line "CodeML. Calculate syntenic CDS pairs and color dots:" select "Synonymous (Ks)." This option will enable SynMap to calculate the Ks value for every syntenic gene pair. This is a multistep process and may take a while to complete for all identified syntenic gene pairs. However, the results from these calculations are cached and reused when appropriate. Quick link: http://genomevolution.org/r/488p.
4. There is the option to select among various color schemes that may be overlaid on the histogram of Ks values and syntenic dotplot. In addition, you may specify cutoff values to be displayed on the histogram. By specifying cutoff value, and selecting an appropriate set of colors (for instance, using a color palette that has more resolution), one can customize and highlight specific syntenic regions of interest. Quick link: http://genomevolution.org/r/479m.

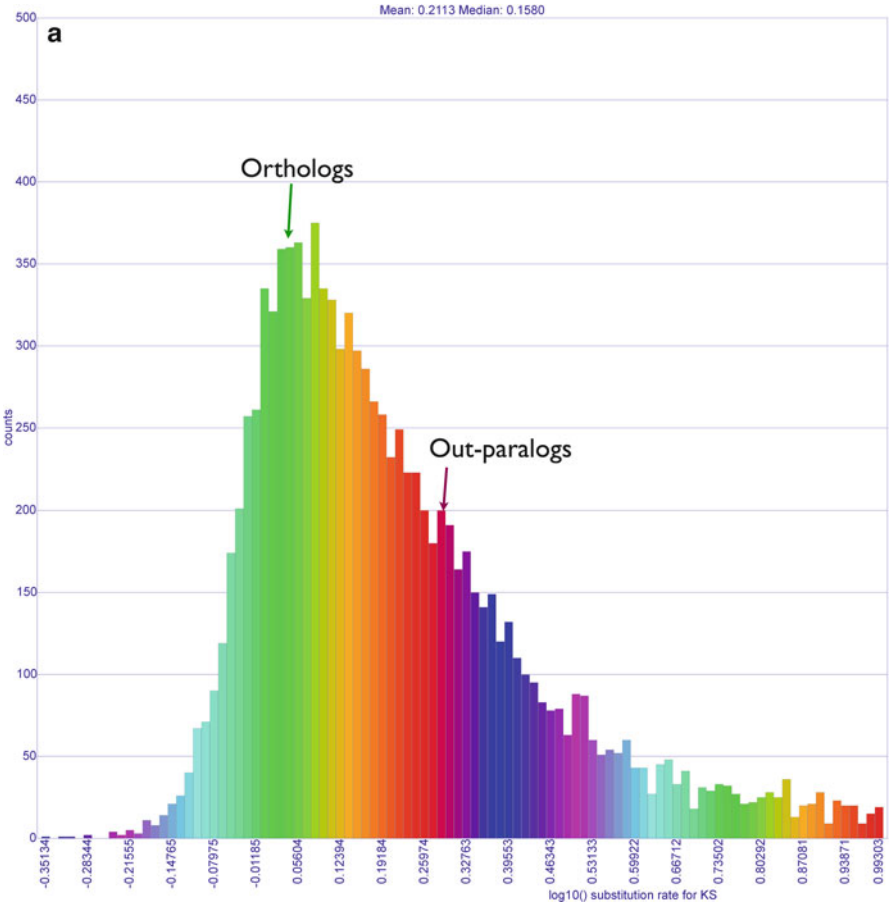

Fig. 12.5 Syntenic path assembly of papaya versus peach. Syntenic gene pairs colored by evolutionary distance (Ks values) (generated using http://genomevolution.org)

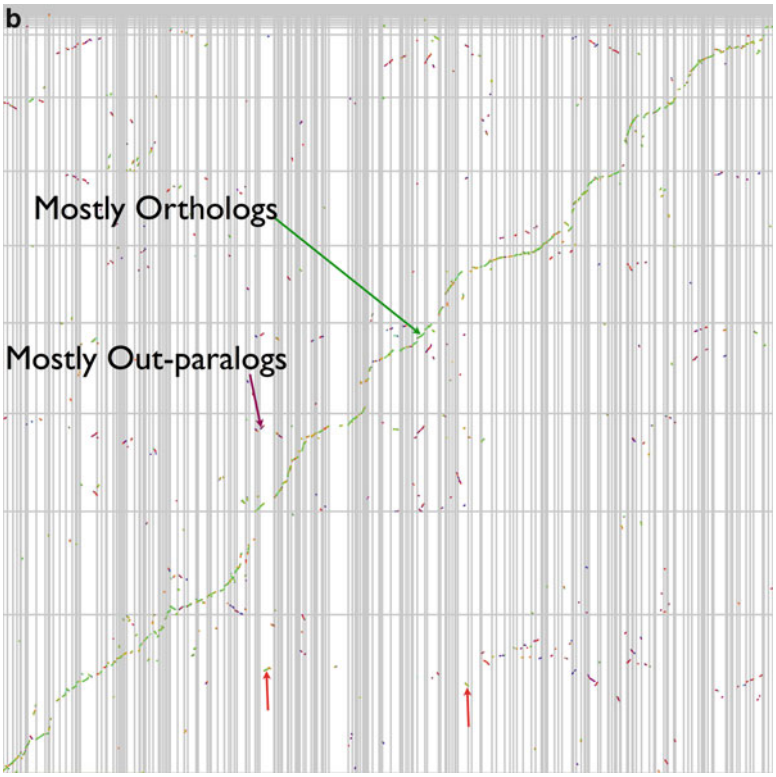

Fig. 12.5 (continued)

High-Resolution Analysis of Syntenic Regions

While syntenic dotplots are useful for pairwise whole-genome comparisons, they lack fine-grain detail, such as gene models, and do not permit tracking small regions of genomic change. Figure 12.6 shows a high-resolution analysis of orthologous syntenic regions between papaya and peach using a parallel line coordinates visualization scheme. This visualization shows two genomic regions: a portion of papaya "supercontig_10" on the top and a portion of peach "scaffold_3" on the bottom. Each genomic region panel has a coordinate bar across the top that designates the number of underling nucleotides encompassing the view. In the middle of each panel is a dash line that separates the forward and reverse strands of DNA comprising the region. Above and below the dashed line are drawn various glyphs representing genomic features such as gene models where gray arrows represent the full extent of a gene, blue arrows represent the portion of a gene which is transcribed to mRNA, and green arrows represent the portion of a gene which is translated into protein (other glyphs represent other genomic features such as rRNAs, pseudogenes, and transposons. For a full list of these please, see http://genomevolution.org/r/3uum).

Fig. 12.6 Evidence of synteny. Collinear arrangement of homologous genes (generated using http://genomevolution.org)

The background of the genomic panels is colored orange to represent sequence comprised of the ambiguous nucleotide "N." "N"s are used to indicate unknown sequence and gaps in the assembly. As is apparent in this graphic, there are gaps in the sequence of both regions, though papaya has some rather large gaps and is an artifact of the low degree of sequence coverage for the papaya genome, compared to the peach genome.

These two genomic regions have been compared to one another using the local alignment algorithm BlastZ/LastZ. BlastZ is excellent for finding large blocks of similar sequences and is more suited for comparing diverged organisms. All such similar sequences identified between these two regions are identified with a pink block located above or below the gene models if the sequences are in the same or opposite orientation, respectively. Most of these regions overlap genomic sequence which codes for protein. This is to be expected as protein-coding sequence is under purifying selection (albeit with the exception of synonymous substitutions). Many of these regions of sequence similarity have a line connecting them between the regions of papaya and peach. This creates a collinear series of lines connecting homologous genes between these regions and is the hallmark of synteny. It is important to note that this pattern of synteny means that these genes are in the same relative genomic region as the ancestral genomic region in common between both these lineages. However, there are many genes unique to one region or the other. These may be due to natural evolution in the position of genes within a genome, gene loss through ongoing fractionation from the paleohexaploidy, gene transpositions, or simply missing from the papaya or peach assemblies (i.e., in the gap of an unsequenced region). These cases may be investigated further by adding in additional syntenic genomic regions from different organisms.

Figure 12.7 shows the same analysis as from Fig. 12.6 with the addition of the orthologous syntenic region from the grapevine genome (bottom panel, mostly purple). The grapevine lineage diverged from the progenitor lineage of papaya and peach and has not had any subsequent whole-genome duplication event (Fig. 12.1). As such, grapevine is an ideal outgroup to papaya and peach as any syntenic gene found in grape between either papaya or peach must be ancestral. In this analysis, all genomic regions have been compared against one another, and each pairwise comparison has its own track of blocks for regions of sequence similarity. However, lines have only been drawn connecting syntenic genes between papaya and peach, and peach and grape. Lines have not been drawn connecting syntenic genes between

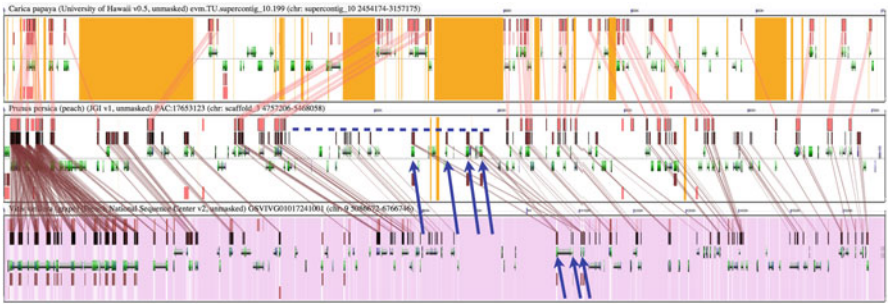

Fig. 12.7 Evidence of synteny. Collinear arrangement of homologous genes with two syntenic regions (generated using http://genomevolution.org)

papaya and grape in order to keep the image more readable but may be done online: http://genomevolution.org/r/479f. The background of the grape genome is purple because all nonprotein-coding sequences have been masked and ignored in this analysis. This region was masked in order to highlight the conservation of gene order. Noncoding sequences often contain additional repetitive sequence such as simple sequence repeats. In this particular experiment, the noncoding repeats can make interpreting the results more difficult as the collinear pattern of homologous genes is obfuscated by many other lines connecting regions of sequence similarity.

By having the orthologous syntenic region of grape in this analysis, genes in peach that overlap a gap in the papaya genome may be investigated in order to determine if orthologs are likely to be present in papaya (Fig. 12.7, dashed blue line). In addition, by comparing the syntenic region of peach and grape regions that overlap a gap in papaya, oddities in their genomic structure may be identified that may shed light as to why this region failed to assemble. The majority of the genes in the peach region are syntenic with genes in the grape region. In addition, there is a tandem gene cluster on one end of the region that is identified by having a set of sequences that all match the same set of sequence in the other region (Fig. 12.7, blue arrows).

To Generate a High Resolution of Syntenic Genomic Regions Starting with a Syntenic Dotplot in CoGe

1. Start with a syntenic dotplot between papaya and peach generated by SynMap. Using the syntenic path assembly option and coloring syntenic gene pairs by their evolutionary distance will help identify orthologous syntenic regions. Quick link: http://genomevolution.org/r/479m.
2. On the syntenic dotplot, click on one of the contig-scaffold squares that contains orthologous syntenic gene pairs (Fig. 12.5b, blue arrow). This will create zoomed-in syntenic dotplot of just that chromosome/contig/scaffold comparison.

Note if a region is less than five pixels wide, you cannot zoom in on it and will need to regenerate the master dotplot as a larger image.

3. Mouse over dots on the zoomed-in syntenic dotplot. A box will appear in the top-left corner of the window and will contain the names and coordinates of gene pairs represented by those dots. At the same time, the crosshairs will turn red. You may mouse over non-syntenic gene pairs (gray dots) or syntenic gene pairs (colored dots).
4. Mouse over a syntenic gene pair and make sure the crosshairs turn red. This also indicates that there is a link to GEvo, CoGe's tool for high-resolution comparison of genomic regions. Click on a syntenic gene pair to launch GEvo. Quick link: http://genomevolution.org/r/48z9.
5. When GEvo loads, it will have two sequence submission boxes, each one loaded with one of the genes on which was clicked. By default, GEvo will load specifying 50,000 nucleotide (nt) bases of additional sequence upstream and downstream of each gene to be compared. Under the "Algorithm" tab, different sequence comparison algorithms may be selected. By default, LastZ is selected which is excellent for finding large blocks of conserved sequence. To run GEvo, press the red button labeled "Run GEvo Analysis!"
6. A status update of the analysis will appear above the area used to configure the analysis. When the analysis is complete, the results will be deposited in this area. The results are interactive, and by clicking on various items in the genomic regions, information boxes will appear and transparent wedges (or lines) drawn connecting regions of sequence similarity. For information on how to use this viewer, see http://genomevolution.org/r/48z4. Of note, all of the files generated by the analysis are available for download along with a log file of the analytical process. In addition, a link is created to regenerate the analysis as configured under the heading "GEvo links." This was used to create the links in this tutorial.
7. GEvo may be easily reconfigured to analyze more sequences by either specifying more sequences to the left and right of the selected genes in the sequence submission boxes or by applying more sequence to all of these text boxes. Analyze 500,000 nt upstream and downstream of both regions by typing "500000" in the box next to "Apply distance to all CoGe submissions" located at the bottom of the sequence submission box. Quick link: http://genomevolution.org/r/48z8.
8. The extent of the regions may be reduced by dragging the slider bars located at the ends of each sequence panel visualized in the results. This will help users zoom in to specific features of interest. Drag these bars to analyze the same amount of sequence as shown in Fig. 12.6. Quick link: http://genomevolution.org/r/48za.
9. There are several ways to identify orthologous genes in different genomes with CoGe. One additional way is to use CoGeBlast, CoGe's tool for searching any number of genomes within its system. To submit a sequence to BLAST from GEvo, click on a gene to get its annotation information to appear in an info box. (e.g., find the papaya gene "evm.TU.supercontig_10.199"). Next click on the

link to CoGeBlast appearing in that info box. Quick link: http://genomevolution.org/r/48zc.

10. When CoGeBlast loads, the gene's sequence will be automatically entered in its sequence submission box. Search for the grapevine genome by typing "grape" in the "Organism Name" search box. From the list of organisms whose names match "grape," select "*V. vinifera* (grape)" and press "+ Add" to add its genome to be blasted. Run CoGeBlast by pressing the red "Run CoGeBlast" button.
11. The results for CoGeBlast appear above the configuration options. On the left is a graphical overview of the hits on individual chromosomes, and on the right is a list of the individual hits. In the hits list, there is a column called "Closest Genomic Feature" which shows the name of the closest gene to the BLAST hit. To evaluate each BLAST hit, click on the BLAST hit's HSP#. This will open an info box with a graphical representation of the BLAST hit in the context of the query sequence and the matching genomic region. For more information about CoGeBlast, see http://genomevolution.org/r/48zd.
12. Using the pre-configured CoGeBlast link, the top hit from CoGeBlast will be the orthologous syntenic gene in grape. Copy this grape gene name from the "Closest Genomic Feature" column. Return to the previous GEvo analysis and add a new sequence submission box by selecting the "Sequence Submission" tab and pressing "+ Add Sequence." In the "Name:" box for this sequence submission box, paste in the grape gene name and make sure there are at least 1,000,000 nucleotides selected upstream and downstream of the gene. In order to make the analysis run faster, dynamically mask the nonprotein-coding sequences in the grape region by pressing on the button "Sequence 3 Options" in Sequence 3's sequence submission box. Select "Non-CDS" from the drop-down menu located next to "Mask Sequence." When configured, rerun the analysis by pressing "Run GEvo Analysis." Quick link: http://genomevolution.org/r/48zi.
13. Adjust the extent of the analyzed regions until the results look similar to those of Fig. 12.7.

Gene Loss Following Polyploidy: Fractionation

While papaya and *Arabidopsis* are in the same order of plants, Brassicales, *Arabidopsis* has had two sequential tetraploidy events since their lineages diverged (Fig. 12.1). After each of these events, *Arabidopsis*' genome underwent extensive gene loss. Together, these two tetraploidy events followed by fractionation created four orthologous syntenic regions in *Arabidopsis* to each region in papaya. However, each syntenic *Arabidopsis* region will only contain a subset of all the ancestral genes present in the papaya region.

Figure 12.8 shows a high-resolution view of syntenic regions among papaya and *Arabidopsis*. Peach may be included as an outgroup region to track common ancestral genes whose extant relatives may be in the unsequenced gaps of papaya (http://genomevolution.org/r/48wj). This view shows that papaya contains the

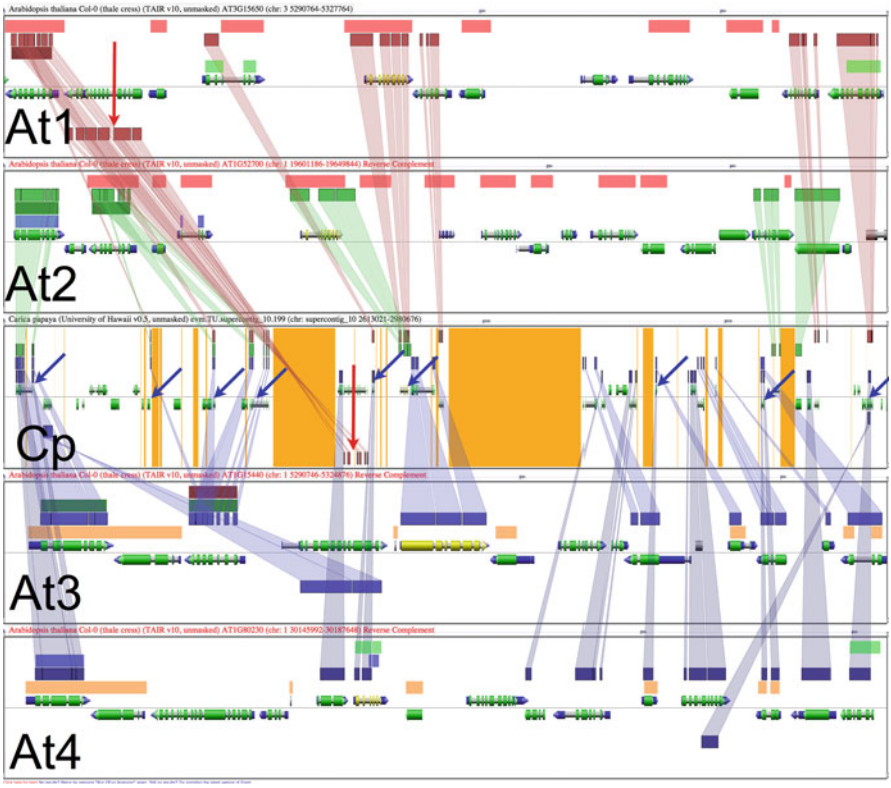

Fig. 12.8 Fractionation of gene content in *Arabidopsis* (generated using http://genomevolution.org)

majority of the gene content of each individual syntenic *Arabidopsis* region, but the gene content of the four *Arabidopsis* regions must be summed in order to represent the full gene complement in papaya. Also, the gene content of *Arabidopsis* is not fully fractionated; there are several ancestral genes that are retained in multiple copies in *Arabidopsis* (alpha or beta duplicates). These are evidenced by multiple syntenic orthologs in different *Arabidopsis* regions matching the same papaya gene (Fig. 12.8, blue arrows).

While each *Arabidopsis* region shows synteny through a collinear arrangement of homologous genes, the gene content of the four syntenic *Arabidopsis* regions is intercalated with respect to one another. In addition, the *Arabidopsis* regions fall into two pairs of syntenic regions where each member of a pair is more closely related to the other member than to the regions in the other pair. These patterns are expected given the sequential nature of these tetraploidy events.

An additional benefit of comparing multiple syntenic genomic regions in high resolution is the identification of potentially missed annotations. The red arrow in Fig. 12.8 points to a set of small regions of sequence similarity between papaya and one of the *Arabidopsis* regions. The gene model in *Arabidopsis* that overlaps

these regions does not match the gene model in papaya. Given that *Arabidopsis* is one of the best annotated plant genomes available, this may be an erroneous gene model in papaya.

How to Find and Compare Orthologous Sequences Between Papaya and Arabidopsis

1. Using the previous GEvo analysis, click on a papaya gene. Quick link: http://genomevolution.org/r/48zi.
2. Click on the link to SynFind in the annotation box for the papaya gene. SynFind takes a gene in one organism and find all syntenic regions in any set of genomes. Quick link: http://genomevolution.org/r/4906.
3. SynFind will load with this papaya gene specified. Use it to identify syntenic regions in the *A. thaliana* genome by search for "*Arabidopsis*" in the "Organism name" search box. For the list of organisms returned, select "*A. thaliana* Col-0." All genomes matching this organism will be presented with the most recent version listed first. Press the "+ Add" button to add the most recent genome to the list of genomes to be searched. Quick link: http://genomevolution.org/r/4904.
4. Run SynFind by pressing the red "Run SynFind" button.
5. When the results are returned, there will be a list of the syntenic genes or regions matching the region from where the query gene was derived (papaya, in this case). If a syntenic gene is identified, the matching gene name will be listed, otherwise a "proxy" position (i.e., its expected position based on synteny) will be listed. Above the list is a link to GEvo with the syntenic regions preloaded; below this list is a link to regenerate the analysis. Follow the link to GEvo to compare these four orthologous syntenic regions of *Arabidopsis* to the papaya region. Quick link: http://genomevolution.org/r/4903.
6. When linked from SynFind, GEvo will automatically start running its analysis. It is also configured to only use papaya's region as a reference sequence. This means that other regions will be compared to it, and this option can be changed for any sequence by clicking on the "sequence option" menu for a given sequence submission box. The sequence submission boxes may be rearranged by dragging them around relative to one another. This will rearrange their stacking order in the GEvo graphic. Try to move the papaya region to have two *Arabidopsis* regions above and two *Arabidopsis* regions below as shown in Fig. 12.8. Quick link: http://genomevolution.org/r/47ra.

Conserved Noncoding Sequences

Besides protein-coding sequences, there are additional sequences in a genome under purifying selection. In plants, these conserved noncoding sequences (CNS) are often found close to genes and are likely involved with the regulation of those

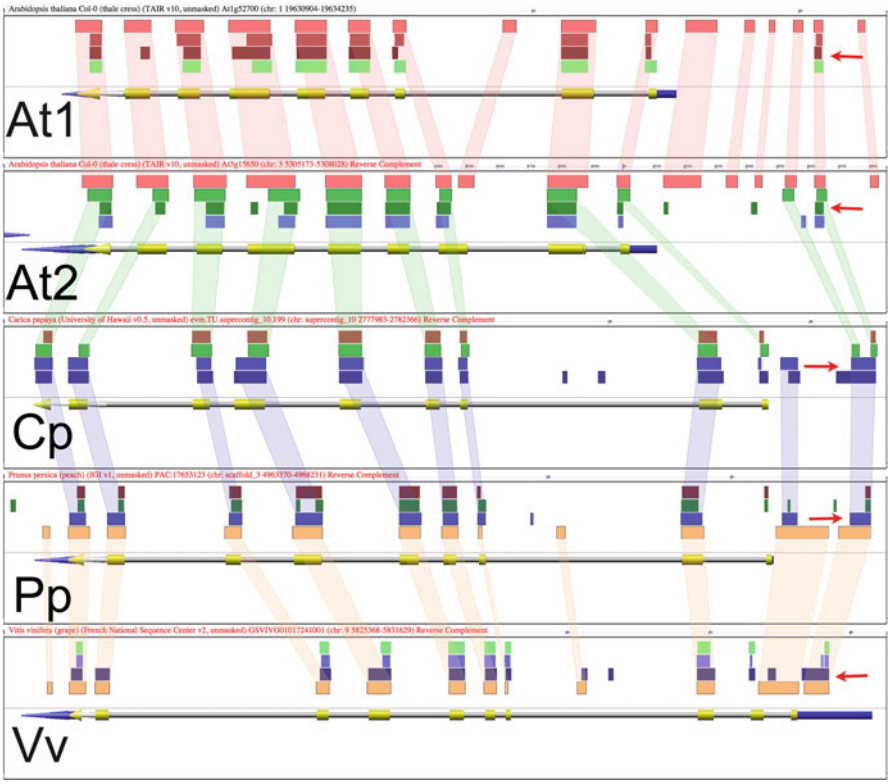

Fig. 12.9 Conserved noncoding sequence (generated using http://genomevolution.org)

genes through *cis*- and *trans*-acting DNA-binding factors (Freeling and Subramaniam 2009). CNSs may be detected by comparing syntenic gene sets. However, since these CNSs are often short, detecting them requires a sensitive alignment algorithm such as BlastN and often with a smaller word size. However, using a sensitive algorithm like small word size BlastN presents two major problems. First, it is likely not scalable across whole-genome comparisons due to large amounts of seeds with a smaller word size. Second, the matches are not specific and can contain a large portion of spurious hits. Therefore, a dedicated CNS detection pipeline first finds syntenic gene pairs and then zooms into the surroundings of those gene pairs to extract CNSs (Freeling and Subramaniam 2009).

Figure 12.9 shows the comparison of a syntenic set of genes among papaya, *Arabidopsis*, peach, and grape. There are two copies of the gene in *Arabidopsis* that have been retained from its additional polyploidy events. The protein-coding sequences of these genes are shown as yellow blocks in the gene model. Between various pairwise comparisons, there are additional conserved sequences that are also collinear and do no overlap protein-coding sequences. The *Arabidopsis* gene pair has an extensive set of CNSs located 5′ of the gene, and the peach-grape

comparison has one in the 3′ UTR of their genes. Of interest in one set of CNSs presents 5′ of all the genes (red arrows). If a scientist is interested in the regulation of this gene, this sequence would be a prime candidate to investigate.

How to Identify Conserved Noncoding Sequences Among a Set of Orthologous Syntenic Genes

1. Start with a GEvo analysis centered on an orthologous syntenic gene set. Quick link: http://genomevolution.org/r/490k.
2. Change the sequence comparison algorithm to BlastN by selecting the "Algorithm" tab and selecting "BlastN: Small Regions" from the drop-down menu next to "Alignment Algorithm."
3. Change the amount of sequences to 1,000 nt upstream and downstream of the gene pairs by selecting the "sequence submission" tab and typing "1000" in the box next to "Apply distance to all CoGe submissions". Quick link: http://genomevolution.org/r/490l.
4. Examine the regions of sequence similarity. There will be many more spurious hits due to short repetitive sequences and simple sequence repeats. CNSs are detected by identifying collinear hits that do not overlap coding sequence. These may be characterized in more detail by clicking on them. This will draw a transparent wedge to the region it matches in another genomic region and open an info box that contains information about the BLAST hit. In the info box, there will be a link to get more information about the BLAST hit including the complete sequence and alignment. Adjust the slider bars in the analysis to include only the putative CNSs. Quick link: http://genomevolution.org/r/490m.

Conclusion

While the papaya genome may currently be under-sequenced and has a relatively poor assembly, its genome still contains much useful information, both for papaya researchers and other plant biologists alike. By leveraging comparison to other genomes, syntenic regions may be identified and the evolutionary history of the genome and its genes understood. Papaya is a member of the Brassicales, sharing a common ancestor with *Arabidopsis* and other crucifers, which serves as an excellent outgroup for this economically and scientifically important clade. The fact that papaya did not experience subsequent genome duplication events since the eurosid paleohexaploidy event adds further value to the phylogenetic importance of this unique genome. The lack of recent duplication events also made papaya genome relatively compact, a ~20 % reduction in gene numbers compared to *Arabidopsis*. In summary, much can be learned by comparing the crucifers to papaya and then from papaya to other rosids and eudicots.

This chapter provides a walk-through of the computational techniques to compare whole-genome sequences and drill down to identify small evolutionarily conserved elements between the matching genes. All the examples are described using the CoGe suite of tools. With the integration of well-tested tools and a wide collection of genomes, plant biologists are well equipped to compare organisms and correlate genomic changes with striking morphological differences, within the same order, genus, or even between closely related species. Throughout this chapter, we have given much focus on the exercise of extracting syntenic blocks, then syntenic gene pairs, and finally small regulatory elements. The resolution and flexibility offered by this multitiered approach is critical to the study of genomic changes and occasionally separating assembly artifacts (as shown to be particularly important in the three times sequencing coverage of the papaya genome). The per-base sequencing cost of a genome has been dropped by three orders of magnitude since the publication of the initial papaya genome (http://www.genome.gov/sequencingcosts/). As more genomes are sequenced, including resequencing of additional strains of papaya, these comparative techniques will be ever more important for identifying and characterizing important functional features of these genomes.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Blanc G, Hokamp K, Wolfe KH (2003) A recent polyploidy superimposed on older large-scale duplications in the Arabidopsis genome. Genome Res 13:137–144

Bowers JE, Chapman BA, Rong J, Paterson AH (2003) Unravelling angiosperm genome evolution by phylogenetic analysis of polyploidization events. Nature 422:433–438

Comai L (2000) Genetic and epigenetic interactions in allopolyploid plants. Plant Mol Biol 43:387–399

Couvreur TLP et al (2010) Molecular phylogenetics, temporal diversification, and principles of evolution in the mustard family (Brassicaceae). Mol Biol Evol 27:55–71

Crow KD, Wagner GP (2006) What is the role of genome duplication in the evolution of complexity and diversity? Mol Biol Evol 23:887–892

Freeling M, Subramaniam S (2009) Conserved noncoding sequences (CNSs) in higher plants. Curr Opin Plant Biol 12:126–132

Gaeta RT, Pires JC (2010) Homoeologous recombination in allopolyploids: the polyploid ratchet. New Phytol 186:18–28

Goff SA et al (2000) A draft sequence of the rice genome (Oryza sativa L. ssp. japonica). Science 296:92–100

Haas BJ, Delcher AL, Wortman JR, Salzberg SL (2004) DAGchainer: a tool for mining segmental genome duplications and synteny. Bioinformatics 20:3643–3646

Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A et al (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449:463–467

Jiao Y et al (2011) Ancestral polyploidy in seed plants and angiosperms. Nature 473:97–100

Koonin EV (2005) Orthologs, paralogs, and evolutionary genomics. Annu Rev Genet 39:309–338

Lander ES, Waterman MS (1988) Genomic mapping by fingerprinting random clones: a mathematical analysis. Genomics 2:231–239

Langham RJ et al (2004) Genomic duplication, fractionation and the origin of regulatory novelty. Genetics 166:935–945
Lynch M, Conery JS (2000) The evolutionary fate and consequences of duplicate genes. Science 290:1151–1155
Lyons E, Freeling M (2008) How to usefully compare homologous plant genes and chromosomes as DNA sequences. Plant J 53:661–673
Lyons E, Freeling M, Kustu S, Inwood B (2011) Using genomic sequencing for classical genetics in E. coli K12. PLoS One 6(2):e16717
Magallón S, Castillo A (2009) Angiosperm diversification through time. Am J Bot 96:349–365
Mayrose I et al (2011) Recently formed polyploid plants diversify at lower rates. Science 333:1257
Ming R et al (2008) The draft genome of the transgenic tropical fruit tree papaya (Carica papaya Linnaeus). Nature 452:991–996
Nielsen CB, Cantor M, Dubchak I, Gordon D, Wang T (2010) Visualizing genomes: techniques and challenges. Nat Methods 7:S5–S15
Osborn TC et al (2003) Understanding mechanisms of novel gene expression in polyploids. Trends Genet 19:141–147
Rensing SA et al (2008) The physcomitrella genome reveals evolutionary insights into the conquest of land by plants. Science 319:64–69
Schnable JC, Lyons E (2011) Comparative genomics with maize and other grasses: from genes to genomes! Maydica 56(1/2):183–200
Schwartz S et al (2003) Human-mouse alignments with BLASTZ. Genome Res 13:103–107
Tang H et al (2008) Synteny and collinearity in plant genomes. Science 320:486–488
The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 408:796–815
Tuskan GA et al (2006) The genome of black cottonwood, Populus trichocarpa (Torr. & Gray). Science 313:1596–1604
Velasco R et al (2007) A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. PLoS One 2:e1326
Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol 24:1586–1591
Yu J et al (2002) A draft sequence of the rice genome (Oryza sativa L. ssp. indica). Science 296:79–92

Chapter 13
Papaya Repeat Database

Niranjan Nagarajan and Rafael Navajas-Pérez

Introduction

Thomas (1971) first suggested that the lack of correlation between genome size and structural complexity is mainly due to the accumulation of repetitive sequences by coining the term C-value paradox. Since then, genomes have been proved to actively expand by means of several mechanisms including polyploidization, transposition, and duplication. Today, it is well known that animal, among them mice and humans, and plant genomes, including such agriculturally important plants as rice, corn, or wheat, have acquired a repertoire of repetitive elements accounting for the vast majority of nuclear DNA in many cases (Kubis et al. 1998).

Three main classes of repetitive sequences are considered: transposable elements (TEs), tandem repeats (TRs), and high copy number genes. On the one hand, TEs constitute the most abundant component of many plant genomes, ranging from 40 up to 80 % of total genomic DNA (Bennetzen et al. 2005). TEs can be further divided into DNA-mediated class II transposons and RNA-mediated class I retrotransposons. DNA transposons were first described by Barbara McClintock as genetic elements capable of transposing to different chromosomal locations in maize plants (1950) and today are known to constitute an important family of TEs in plants (Jiang et al. 2003). The most common TEs in plants though are LTR retrotransposons (Novikov et al. 2012); non-LTR retrotransposons, while numerous, remain mostly inactive and under regulation of the host genome (Cheng and Ling 2006). On the other hand, TRs are main constituents of centromeric, telomeric,

N. Nagarajan
Department of Computational and Systems Biology, Genome Institute of Singapore, Singapore 138672, Singapore

R. Navajas-Pérez (✉)
Facultad de Ciencias, Departamento de Genética, Universidad de Granada, Campus de Fuentenueva s/n, Granada 18071, Spain
e-mail: rnavajas@ugr.es

R. Ming and P.H. Moore (eds.), *Genetics and Genomics of Papaya*, Plant Genetics and Genomics: Crops and Models 10, DOI 10.1007/978-1-4614-8087-7_13,

and subtelomeric regions of many eukaryotes, comprising hundreds to thousands of tandemly arrayed monomeric repeats (Ugarkovic and Plohl 2002). These repeats also appear at interspersed positions and in low-recombining regions, such as sex chromosomes or B chromosomes (Camacho et al. 2000; Navajas-Pérez 2012). This type of sequences can account for a large portion of genomic DNA (Saini et al. 2008). The third class is constituted by high copy number genes. Some molecular data suggest that a great number of plant genes belong to gene families ranging in size from a few members to hundreds (Martienssen and Irish 1999).

Apart from the role of constitutive heterochromatin, traditionally linked to a major architectonic function necessary for cell division (Yunis and Yasmineh 1971), repetitive elements—in the best of cases—have been considered dispensable if not junk or selfish DNA with no function at all (Ohno 1972; Orgel and Crick 1980). This lack of function contrasts with their prevalence in the genomes. In fact, an increasing number of pieces of evidence are changing the whole picture posing that repetitive sequences would play important roles in different biological aspects. It is now evident, for example, that repetitive sequences have been crucial drivers of genome evolution significantly contributing to the expansion of the genomes and consequently shaping contemporary chromosome organization through events of chromosome rearrangements due to interactions between scattered repeats (Fedoroff 2000). Indeed, up to 70 % of flowering plants have evolved through a polyploid ancestor in their lineages so both whole-genome and segmental duplications are common and key events in plant genome evolution (Wang et al. 2012). Also, variations in repeats content are thought to influence the determination of continuous phenotypic characters (Meagher and Vassiliadis 2005; Gemayel et al. 2010) or be related to the response to environmental cues (Schmidt and Anderson 2006). More recently, the implication of repetitive elements in gene regulation has been demonstrated (Thornburg et al. 2006; Lunyak et al. 2007; Román et al. 2011) suggesting they might be fundamental for the creation of new genes and sophisticated regulatory network systems. Thus, the study of repetitive sequence elements is essential to understand the nature and consequences of genome size variation between different species and for studying the large-scale organization and evolution of plant genomes.

Finally, it is worth mentioning the implication of repetitive sequences (mainly satellite DNAs and retrotransposons) in the emergence of sex chromosomes. Many Y chromosomes have more abundant heterochromatin derived from repetitive sequences compared with X chromosomes and autosomes. Accumulation of repetitive sequences contributes to generate gene deserts found in the Y chromosomes, Y-chromosome chromatin expansion, and chromosome breaks and rearrangements and may well be a key factor in the generation of differences in morphology and size observed between X and Y chromosomes that ultimately prevent the recombination in the sex-determining region. This has been proved by cytogenetic analyses and more recent genome-based projects in both plant and animal genomes (Matsunaga 2009; Navajas-Pérez 2012). Also theories on the implication of TEs in the origin of sex reproduction have been put forward (Arkhipova 2005).

Due to their high rate of change, repetitive elements have been used to detect polymorphisms in diverse type of biological analyses. The variation of

minisatellites repeat copy led to the DNA profiling method for general use in human genetic analysis (Jeffreys et al. 1985). Soon, with the advent of PCR, hypervariable and ubiquitous microsatellite markers became widely used in genome mapping and population analysis and genotyping (Ellegren 2004). Also, satellite DNA has helped to clarify phylogenetic relationships among related species by checking the presence/absence status (Navajas-Pérez 2012) or by analyzing the rates of change (Robles et al. 2004) and to understand the dynamics of repetitive elements in the genomes (Navajas-Pérez et al. 2009a). Other repetitive DNA as rRNA sequences has been traditionally used in phylogeny (Pace 2009). TEs have been used to examine genome structure and composition as well as to successfully elucidate evolutionary relationships (Ray 2007).

In these grounds, several databases devoted to store, curate, and classify repetitive DNA have been developed recently: as for satellite repeats (Macas et al. 2002), tandem repeats (Navajas-Pérez and Paterson 2009), or transposable elements (Llorens et al. 2011; Bousios et al. 2012). As the genomic information increases a pleiad of methods for mining, detection and further analysis of repetitive DNA are arising (Benson 1999; Jurka 2003; Navajas-Pérez et al. 2007).

Papaya, because of its position in the tree of life sharing a common ancestor with *Arabidopsis* about 72 million years ago, and with the existence of an incipient pair of sex chromosomes is a promising genomic model. In the past decade many genomic resources have been generated, as a draft whole-genome sequence, an integrated genetic and physical map including sex-determining region, three BAC libraries, and a large collection of ESTs (Ming et al. 2008; Na et al. 2012; Wang et al. 2013; also Chap. 17 in this text). This offers a good opportunity to characterize the papaya repeatome among many other issues. In fact, coupled with this development, a large collection of SSR and AFLP markers comprising sex-specific markers have been characterized (Ma et al. 2004; Chen et al. 2007), and a papaya repeat database has been generated (Nagarajan et al. 2008). In this chapter, we highlight the most relevant information regarding this matter.

Transposable Elements

The papaya repeatome is dominated by TEs, comprising of 52 % of the genome and ~93 % of the repeatome as described in Nagarajan et al. (2008). This is almost certainly a conservative estimate as repeat elements and TEs, in particular, are hard to assemble from whole-genome shotgun sequencing data (Nagarajan and Pop 2009). In addition, the vast majority of identified TEs in the genome are papaya-specific (71 %) and unidentifiable using consensus sequences for other plant repeats, underscoring the rapid divergence of TE families in plant genomes (Nagarajan et al. 2008). Using de novo repeat finders and manual curation, a custom library of TE families was constructed for the papaya genome, providing a curated database of 889 papaya TE families that serve as a resource for annotation of newly sequenced plant genomes (ftp://ftp.cbcb.umd.edu/pub/data/CPR-DB).

Table 13.1 Summary of TE content of papaya WGS and sex chromosomes

Class	Element	Percentage of sequence (%)		
		MSY	X	WGS
(I Retrotransposons)	Ty1/copia	4.6	5.8	5.5
	Ty3/gypsy	47.1	35.1	27.8
	LINE	0.6	1.1	1.1
	SINE	0	0	<0.01
	Other	11.4		8.4
		64.3	49.9	42.8
(II Transposons)	CACTA/En-Sp	<0.01	0	0.01
	MuDR-IS905	0	0	<0.01
	Tc1-IS630-Pogo	0	0	<0.01
	Other	0	0	<0.01
		0.1	0.1	0.01
Unclassified	Unknown	13.4	9.4	8.72
Total		77.8	59.3	51.62

A wide representation of known common types of TEs were found in the papaya genome, with retrotransposons (40 % of the genome) being the dominant class and *Ty3–gypsy* (27.8 %) being the dominant type (71 % of these are papaya specific). *Ty1–copia* (5.5 %) and LINE (1 %) retrotransposons, as well as CACTA-like DNA transposons (0.1 %), were the other major identifiable types. A significant fraction of the TE matches were either unknown retrotransposons (8.4 %) or unannotated families (8.5 %), highlighting the need for further characterization of these repeat families. In particular, the observed lack of known DNA transposons (0.2 % of the genome) compared to other plant genomes could be due to the presence of unannotated papaya-specific DNA transposon families (Table 13.1; Nagarajan et al. 2008).

In agreement with earlier observations, the TE content in the papaya genome is intermediate between the much smaller *Arabidopsis* genome (Arabidopsis Genome Initiative 2001) (14 % TE content) and the much larger maize genome (Messing et al. 2004) (58 % TE content), but as a function of the genome size, it is relatively repeat rich (Nagarajan et al. 2008). The high TE content of the papaya genome serves to explain the observation that it has a smaller gene repertoire than the *Arabidopsis* genome despite having a genome that is three times the size (Ming et al. 2008). The expansion of most TE families in the papaya genome is presumably ancient, with most TE matches being inactive fossils that have diverged substantially from their consensus. However, for several families (papaya-specific, often *Ty3–gypsy* elements), dozens of nearly perfect copies can be found in the papaya genome, some with EST matches, suggesting that some elements may still be active (Nagarajan et al. 2008).

A striking feature of TEs in the papaya genome is the similarity with the rice genome despite their divergence on the species tree. As reported in Nagarajan et al. (2008), 57 % of matches to retrotransposons and 81 % of matches to DNA transposons among TIGR plant repeats (ftp://ftp.tigr.org/pub/data/TIGR_Plant_Repeats) were to rice repeats. Phylogenetic analysis of *Ty1–copia* and *Ty3–gypsy* elements (Fig. 13.1)

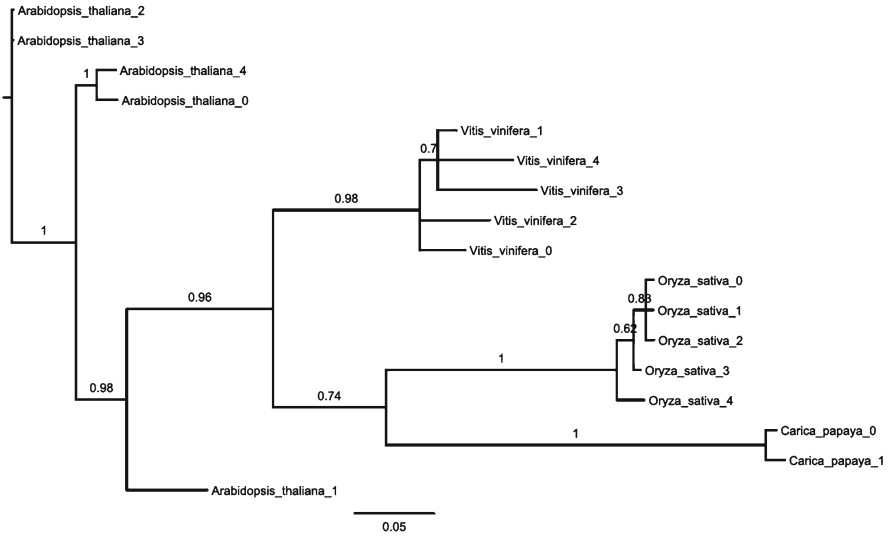

Fig. 13.1 Phylogenetic analysis of plant genome sequences matching the Ty3–gypsy retrotransposon sequence ATGP5A_I in bases 3,700–4,100 (the five best matches for each species were included in the phylogenetic analysis) [modified with kind permission of Springer Science + Business Media from Nagarajan et al. (2008)]

also revealed a similar pattern where papaya sequences tended to cluster with rice sequences. It was also observed that the ratio of *Ty3–gypsy* to *Ty1–copia* elements in the papaya genome was closer to the 2:1 ratio of the rice genome than to the 1:1 of *Arabidopsis* and maize genomes. Taken together these pieces of evidence may suggest a horizontal mode of transfer for introduction of these retrotransposons into the papaya genome.

Tandem Repeats

The existence of 277.4-Mb whole-genome shotgun sequences (WGS) of papaya allows an in silico exploration for TRs. Repeat motifs between 1 and 2,000 bp were analyzed and classified according size into micro- (1–6 bp), mini- (7–100 bp) and satellite (>100 bp) tandemly arrayed sequences, as described in Nagarajan et al. (2008). According to this approach, a total of 414,681 class I (≥20 bp) repeats were characterized in 57,360 loci (spanning a total of 4.8 Mbps, representing 1.3 % of the total genome size). The analysis revealed an average repetitive-unit length of 79 bp and a copy number average of 7.23 (ranging from 1.8 to 969.3 copies). The average AT content was 72 %, slightly higher than the average AT content of the genome (65 %). Tandem repeats are randomly distributed in the papaya genome, and there is no correlation between tandem repeat number and gene density. This supports the observation that papaya genome is mostly euchromatic (Ming et al. 2008).

In terms of physical quantity of DNA, microsatellites represent a 0.19 % of the total papaya genome size, minisatellites a 0.68 %, and satellite DNAs a 0.43 %. Following the same approach, Navajas-Pérez and Paterson (2009) found similar abundance of tandem repeats in angiosperms assemblies, 0.19, 0.83, and 0.5 % on average for micro-, mini-, and satellite DNAs, respectively. Punctual quantifications of TRs in other species reveal that these sequences frequently constitute a large portion of the genomes (Lim et al. 2005; Saini et al. 2008). However, although all papaya TR sequences may not be covered, most known repetitive elements are found to be reasonably well represented.

Despite their low percentage, microsatellites represent the class with the highest number of tandem repeat copies in papaya. Dinucleotides are the best represented with ~180,000 units, the most common being (TA/AT)n and (AG/TC)n along with long A/T stretches. TTC/AAG, AAT/TTA trinucleotides and their multimeric variants (with up to 969.3 repeats in a single locus), and pentanucleotides are also common repeats in papaya. Previous reports based on genomic library screenings and mining of repeated DNA databases have demonstrated the same for a great number of plant species (revised in Navajas-Pérez and Paterson 2009).

Longer TRs are normally specific to a related group of species due to their rapid evolutionary change rate (Miklos 1985). Thus, only a small portion of TRs were annotated. Those sequences fell into DNA binding, pseudogenes, or TE-like categories (Table 13.2) (Navajas-Pérez and Paterson 2009). It can be argued that these sequences could be somehow involved in gene regulation/inactivation or evolved through a TE intermediate. These findings agree with the recent tendency to consider repeat DNAs functional, instead of simply junk or parasitic elements.

There is a general tendency in the distribution of repeat-unit sizes in papaya tandem repeats to sequences between 9 and 50 bp, which account for a high number of copies as well as for the maximum number of variants and loci (Nagarajan et al. 2008). This agrees with data on papaya for perfect SSRs from Wang et al. (2008) who found that the 20-bp repeats were the most common repeats in class I, followed by 24-bp repeats with insignificant variance between EST and WGS or BES sequence data and with data from Navajas-Pérez and Paterson (2009) who found the abundance of repeats in the range 9–30 bp in eight WGS of plants from different sources. This might suggest that structural features such as monomer length could play a role in tandem repeat preservation and evolution (Ugarkovic and Plohl 2002).

Perfect SSR Sampling

Due to the reproducibility of their amplifications and the possibility to better detect polymorphisms among individuals, perfect SSRs are preferred for fine-scale mapping, population analysis, and genotyping. It is important to note that the term perfect repeat is used to denote repeats that do not contain insertions, deletions, and/or mismatches with respect to their basic repetitive motif. In this context, an additional mining has been performed in papaya to detect two types of perfect SSRs: class I or

Table 13.2 Mini- and satellite-DNA BLAST hits summary in *Arabidopsis thaliana* TAIR7 release [modified with kind permission of Springer Science+Business Media from Navajas-Pérez and Paterson (2009)]

Annotation	Arabidopsis	Papaya	Poplar	Grapevine	Rice
Unclassified proteins	4,256	63	71	56	50
Transposable elements, viral, and plasmid proteins	1,047	426	14	9	6
Metabolism	283	14	5	16	12
Cell rescue, defense, and virulence	180	1	6	3	4
Classification not yet clear-cut	179	1	3	6	2
Protein synthesis	88	0	2	8	5
Cellular transport, transport facilitation, and transport routes	80	1	1	2	2
Transcription	75	7	5	6	2
Cellular communication/signal transduction mechanism	58	0	2	1	0
Protein fate	51	2	4	2	3
Subcellular localization	50	1	2	0	3
Biogenesis of cellular components	44	2	0	0	0
Cell cycle and DNA processing	13	2	0	1	1
Energy	12	1	1	1	0
Development (systemic)	10	0	0	1	0
Protein with binding function or cofactor requirement (structural or catalytic)	8	0	0	0	2
Cell fate	8	0	0	0	0
Systemic interaction with environment	3	0	0	0	0
Interaction with environment	2	0	0	0	0
Storage protein	1	0	1	1	0
Regulation of metabolism and protein function	1	0	0	0	0
Total	6,449	521	117	113	92

SSRs ≥20 bp and class II, less variable SSRs between 12 and 20 bp. Following this method, a total of 371,710 perfect SSRs were identified in the papaya genome, of which 32,164 (8.7 %) and 339,546 (91.3 %) belonged to class I and class II SSRs, respectively. The density was of one per 8.6 kb for class I and one per 0.8 kb for class II on average. Thus, according to this approach class II SSRs was substantially more abundant than class I SSRs on a genome-wide scale (Wang et al. 2008).

The same procedure was used to scan 51.2-Mb bacterial artificial chromosome (BAC) end sequences (BES) (Ming et al. 2001) and 13.4-Mb expressed sequence tag (EST) sequences (Ming et al. 2008). A total of 49,738 SSRs were identified from BES, including 3,581 (7.2 %) class I and 61,394 (92.8 %) class II SSRs with densities of one per 14.3 and 1.1 kb, respectively, while 10,688 SSRs with 94.2 % class II and 5.8 % class I were gathered from EST sequences.

The highly mutable nature of SSRs makes them potentially powerful markers for analyzing genetic polymorphisms between closely related genotypes. Around 11,000 primer pairs have been developed by different authors (Santos et al. 2003;

Pérez et al. 2006; Eustice et al. 2008; Wang et al. 2008; Ramos et al. 2011) from different sources (BES, EST, and WGS) for the amplification and polymorphism of class I SSRs in papaya. This batch of primers was tested on four selected genomic DNA samples, including the parents of an F2 mapping population, an "AU9" female and "SunUp" hermaphrodite, and two pooled DNA samples containing either ten female or ten hermaphrodite F2 plants, as described in Wang et al. (2008), and contributed to integrate the WGS data with EST and BES sequences to construct a high-density marker map. This complete set of SSR markers throughout the genome will assist diverse genetic studies in papaya and related species. For example, some of these SSR markers have been used to analyze polymorphisms in tropical accessions of papaya and their cross-amplification with *Vasconcellea* species (Pérez et al. 2006), and others have been used for marker-assisted selection in backcross programs (Ramos et al. 2011).

Gene Families

Despite containing fewer genes overall compared to the *Arabidopsis* genome, the papaya genome has several gene families with increased copy number (Ming et al. 2008). These gene families, identified by a gene "tribe" analysis by comparison with *Arabidopsis*, poplar, grape, and rice genes, highlight the role of gene family expansion in papaya tree and fruit development. In particular, compared to *Arabidopsis*, the papaya genome is marked by an increase in certain families of transcription factors (e.g., RWP-RK), resistance genes (NBS-LRR), lignin synthesis genes, starch-associated genes, and those involved in volatile development (Ming et al. 2008).

While the papaya genome lacks signatures of recent genome duplication, a significant fraction of the genes (>2 %, representing 3 % of the papaya genome) are present in a large number of copies (>20; Nagarajan et al. 2008). Many of the most abundant genes are, not surprisingly, similar to those found in TEs (with matches to integrases and polyproteins). However, a number of them also represent non-TE-associated functions including MADS-box transcription factors, zinc-finger proteins, topoisomerases, and serine/threonine phosphatases (Nagarajan et al. 2008) and could be under strong selection in the papaya genome.

Telomeres

Telomeres are highly conserved structures that maintain chromosome integrity by stabilizing chromosome termini. Telomeric DNA is made up of relatively short arrays of a 7-bp long TG-rich sequence added by a telomerase enzyme. This solves the capping and replication issue at the ends of a DNA double helix (Watson and Riha 2010). The first eukaryotic telomere sequence, TTTAGGG, was cloned for

Arabidopsis thaliana, and found to be present in most higher plants, except for plants of order Asparagales that harbor human-type telomere repeat, TTAGGG (de la Herrán et al. 2005), and plants from several genera of the Solanaceae family (Sykorova et al. 2003). There is no obvious correlation between telomere size and phylogenetic relationships, and the length of telomeric DNA widely varies among plant taxa, ranging from 0.3 kb in green algae (Petracek et al. 1990) to 100 kb in tobacco (Fajkus et al. 1995). Telomeres in *A. thaliana* are 2.5 kb long on average (Richards and Ausubel 1988). Papaya telomeres belong to the *Arabidopsis* type (Nagarajan et al. 2008), and their size ranges from 25 kb to well over 50 kb (Shakirov et al. 2008).

Telomere microsatellite-like repeats are separated from the rest of the genomic DNA by a transitional sequence or subtelomere. Subtelomere does not necessarily participate in telomere function but can facilitate meiotic pairing or protect terminal genes against the loss and gain processes at the chromosome ends (Kipling 1995). Subtelomeric or telomere-associated sequences (TAS), in addition to location and the ineffectiveness for sequence homogenization (Contento et al. 2005), have a similar organization in many plants (Ganal et al. 1991), being frequently constituted by species-specific long tandem repeats, transposons, and degenerate variants of (TTTAGGG)n motifs (Riethman et al. 2005; Navajas-Pérez et al. 2009b). Multiple copia- and gypsy-like retrotransposons and different DNA transposons occupy subtelomeric regions in papaya as well as tracts of microsatellite-like repeats including the vertebrate motif in a small copy number (Ming et al. 2008; Nagarajan et al. 2008). In addition, inspection of subtelomeric regions indicated that nine of them share 0.5–1.5 kb of nearly identical DNA sequence immediately adjacent to terminal telomeric repeats (Ming et al. 2008). Also, Navajas-Pérez and Paterson (2009) found some repeats in papaya WGS showing homology with the telomere-like 500 repeat of *A. thaliana*, all of these, typical features of a TAS. Notably, the organization of subtelomeric DNA in papaya contrasts sharply with *Arabidopsis* subtelomeres, which consist of unique sequence on eight out of ten chromosome arms.

Centromeres

In eukaryotes, centromeres are often composed of cytologically distinctive heterochromatin and are associated with long arrays of satellite DNA (Kipling 1995). This highly repetitive nature makes centromeres difficult for sequencing and fine-scale genetic mapping (Navajas-Pérez and Paterson 2009). Notwithstanding the difficulties, centromere-specific repetitive DNA sequences have been isolated yet from several plant species, including *Brassica napus* (Harrison and Heslop-Harrison 1995), *A. thaliana* (Martinez-Zapater et al. 1986), *Oryza sativa* (Wang et al. 1995), or *Sorghum bicolor* (Miller et al. 1998).

Although the papaya genome is largely euchromatic, highly condensed heterochromatin knobs exist on most chromosomes' centromeric and pericentromeric regions representing an estimated 30–35 % of the genomic DNA. Five BACs that

mapped onto centromeric region have been analyzed up to date. Sequence analysis showed that all of these BACs lack known centromere-specific sequences. BACs contained 19.7 % of known repetitive sequences based on a RepeatMasker search, including 115 gypsy retroelements—which are a typical feature of the pericentromeric region of plant chromosomes—2 copia retroelements, 85 simple repeats, 457 low complexity repeats, one DNA transposon, and one small RNA (Yu et al. 2007). However, a large portion of this heterochromatic DNA was probably not covered neither by the draft genome sequence nor in BAC libraries, and its nature remains understudied (Ming et al. 2008).

It is noteworthy that papaya male-specific region of Y chromosome (MSY) mapped close to the centromere of the Y chromosome (Yu et al. 2007). Fine mapping showed that the centromere of the Y chromosome is either directly associated with knob 4 or is immediately adjacent to either side of this knob, a region showing more divergence between X and Y than the rest of the MSY. It may indicate that the first sex-determining gene of papaya was possibly located within the centromeric region where recombination is severely or completely suppressed. Thus, accumulation of genes related to male functions near the centromere would have favored and triggered the establishment and expansion of the MSY region. Natural selection of such genes may result in a selective advantage to recombination suppression between these genes and the sex-determining region on the proto-sex chromosome (Charlesworth et al. 2005).

Sex-Chromosome Repeatome

Early evolved plant sex chromosomes like those from papaya have given rise to many studies in recent years which have proved chromosomal rearrangements and repetitive DNA accumulation crucial events in sex-chromosome evolution (Sola-Campoy et al. 2012). Sex chromosomes are thought to have evolved from a standard autosomal chromosome pair as a consequence of a rarely recombining region containing genes involved in sex determination (Ming et al. 2011). That progressive suppression of recombination is the ultimate consequence of the accumulation of diverse repetitive sequences, such as mobile elements and satellite DNAs, that consequently gives rise to Y-chromosome degeneration. This may further inhibit recombination between X and Y chromosomes and ensure the maintenance of dimorphic sex chromosomes, while conferring them with exceptional evolutionary features.

In papaya, the lack of recombination might have been caused by the proximity of MSY to the centromere as mentioned before together with two large-scale inversions, followed by numerous additional chromosomal rearrangements (Wang et al. 2012). These data come from a recently constructed physical map of the MSY region and its X counterpart by chromosome walking and sequenced bacterial artificial chromosomes—BACs (Wang et al. 2012; Gschwend et al. 2011). Thus, papaya constitutes the first complete sequencing of a plant Y-specific region together with

its X counterpart. This offers a good opportunity to gain insights into structural organization and composition of plant-sex chromosomes. According to this analysis, papaya MSY encompasses 8.2 Mb, more than twice as large as the corresponding 3.5-Mb female region. As predicted by the model of sex-chromosome evolution, the male-specific region expanded by massively accumulation of repeated DNA, representing 83 %, while the corresponding X region included 70 % of such repeats (Na et al. 2012). In any case, both are much higher than the papaya genome-wide average of 56–58 % (Ming et al. 2008; Nagarajan et al. 2008).

A more detailed analysis revealed that among all interspersed repeats in this region, the retroelements are the most significantly accumulated repeats with 64 % in the MSY and 50 % in the corresponding X region. They are the principal responsible of the larger size of the MSY accounting for nearly 99 % of all identifiable interspersed repeats. DNA transposons were also found but in a minor extension, representing only 0.1 % of the MSY region (Table 13.1). 80.2 % of the younger inversion sequence is already repetitive and at least 80.7 % of second inversion too. This would support the predicted early accumulation of transposable elements in the initial stage of sex-chromosome evolution after recombination stops (Charlesworth et al. 2005). As for the unannotated sequences, a total of 36 new repeats were identified. However, only 21 of them—20 from the MSY and 1 from the X—had no match to papaya genome sequences and then were regarded as potentially sex-specific repeats. Interestingly, all these MSY-specific repeats mapped within two regions where the MSY explosion occurred, suggesting their role in the origin of sex chromosomes (Na et al. 2012). Tandem repeats content has been estimated in 3.1 % for the X region and 3.8 % for the MSY (Na et al. 2012).

Finally, it is remarkable that papaya X counterpart also presented higher repetitive content than the genome-wide average. This has been found in other organisms (Bergero et al. 2007). In papaya, could partly be explained considering the pericentromeric location of sex-determining regions of the X and Y chromosomes (Gschwend et al. 2012). Also, due to the lack of recombination between X and Y chromosomes in males and hermaphrodites, the X region would have a lower effective population size than the autosomes and then a reduced efficacy of purifying selection redounding on a higher accumulation of repeat DNA (Wang et al. 2012).

Methods

Sources

The papaya 277.4-Mb WGS from a "SunUp" female plant (Ming et al. 2008), a total of 51.2-Mb BES from a hermaphrodite BAC library (Ming et al. 2001), and 13.4-Mb EST sequences (Ming et al. 2008) were used for repeats mining.

Annotation of TEs

TEs were annotated using RepeatMasker (http://www.repeatmasker.org) and a non-redundant database combining plant repeats from Repbase (Jurka 2003), CPR-DB (ftp://ftp.cbcb.umd.edu/pub/data/CPR-DB), and TIGR (ftp://ftp.tigr.org/pub/data/TIGR_Plant_Repeats). CPR-DB was constructed by applying the de novo methods RepeatScout (Price et al. 2005) and PILLAR (Edgar and Myers 2005) to the complete set of contigs from the papaya genome. Repeat families were annotated using a combination of manual curation (786 repeat families; N. Jiang, personal communication) and BLAST searches against NR and PTREP (http://wheat.pw.usda.gov/ITMI/Repeats).

Tandem Repeats Detection

Tandem repeats were detected by using the Tandem Repeats Finder software (Benson 1999). Repeat units between 1 and 2,000 bp were analyzed, and only repeats arrayed in tandems >25 bp were considered. Repeats were classified into micro- (1–6 bp), mini- (7–100 bp), and satellite (>100 bp) tandemly arrayed sequences. A nonredundant set of sequences was constructed using the program cd-hit-est, as implemented in the package CD-HIT (Li and Godzik 2006), at the 85 % similarity level. For annotations, the nonredundant sequences were BLASTed with the Arabidopsis TAIR 7 release (Poole 2007) and the hits classified according to the MIPS functional catalogue database (http://mips.gsf.de). Perl scripts were written to automate the process. A perl program Simple Sequence Repeat Identification Tool (SSRIT) available at http://www.gramene.org (Temnykh et al. 2001) was used for perfect SSR automated mining according to Wang et al. (2008).

Analysis of High Copy Number Genes

Annotated genes (DNA and protein sequences) in the papaya genome were BLASTed against the whole-genome sequence to find significant matches (*E*-value < 1e−20), and searches against the NR protein database (NCBI, January 2008) were used to find plant homologs.

Data Access and Retrieval

The sequences and annotations in the papaya repeat database are available via FTP downloads at ftp://ftp.cbcb.umd.edu/pub/data/CPR-DB. The sets of novel TE sequence in papaya (annotated and unannotated) are presented as multi-fasta files in

a format convenient for use with RepeatMasker. For tandem repeats, redundant and nonredundant databases as well as a consensus sequence list are available in multi-fasta files. A file including annotations is also provided. High copy number papaya transcripts and protein sequences are also available as annotated multi-fasta files. Further details can be found in the README file accompanying the database. Also a comprehensive information of 11,000 perfect SSR marker surveyed can be found in Santos et al. (2003), Pérez et al. (2006), Eustice et al. (2008), Wang et al. (2008), and Ramos et al. (2011).

References

Arabidopsis Genome Initiative (2001) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 408:796–815

Arkhipova IR (2005) Mobile genetic elements and sexual reproduction. Cytogenet Genome Res 110(1–4):372–382

Bennetzen JL, Ma J, Devos KM (2005) Mechanisms of recent genome size variation in flowering plants. Ann Bot (Lond) 95:127–132

Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res 27:573–580

Bergero R, Forrest A, Kamau E, Charlesworth D (2007) Evolutionary strata on the X chromosomes of the dioecious plant *Silene latifolia*: evidence from new sex-linked genes. Genetics 175:1945–1954

Bousios A, Minga E, Kalitsou N, Pantermali M, Tsaballa A, Darzentas N (2012) MASiVEdb: the sirevirus plant retrotransposon database. BMC Genomics 13(1):158

Camacho JP, Sharbel TF, Beukeboom LW (2000) B-chromosome evolution. Philos Trans R Soc Lond B Biol Sci 355:163–178

Charlesworth D, Charlesworth B, Marais G (2005) Steps in the evolution of heteromorphic sex chromosomes. Heredity 95(2):118–128

Chen C, Yu Q, Hou S, Li Y, Eustice M, Skelton RL, Veatch O, Herdes RE, Diebold L, Saw J, Feng Y, Qian W, Bynum L, Wang L, Moore PH, Paull RE, Alam M, Ming R (2007) Construction of a sequence-tagged high-density genetic map of papaya for comparative structural and evolutionary genomics in Brassicales. Genetics 177(4):2481–2491

Cheng XD, Ling HQ (2006) Non-LTR retrotransposons: LINEs and SINEs in plant genome. Yichuan 28:731–736

Contento A, Heslop-Harrison JS, Schwarzacher T (2005) Diversity of a major repetitive DNA sequence in diploid and polyploid *Triticeae*. Cytogenet Genome Res 109:34–42

de la Herrán R, Cuñado N, Navajas-Pérez R, Santos JL, Ruiz Rejón C, Garrido-Ramos MA, Ruiz Rejón M (2005) The controversial telomeres of lily plants. Cytogenet Genome Res 109(1–3):144–147

Edgar RC, Myers EW (2005) PILER: identification and classification of genomic repeats. Bioinformatics 21(Suppl 1):i152–i158

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. Nat Rev Genet 5(6):435–445

Eustice M, Yu Q, Lai C, Hou S, Thimmapuram J, Liu L, Alam M, Moore P, Presting G, Ming R (2008) Development and application of microsatellite markers for genomic analysis of papaya. Tree Genet Genomes 4:333–341

Fajkus J, Kovarik A, Kralovics R, Bezdek M (1995) Organization of telomeric and subtelomeric chromatin in the higher plant *Nicotiana tabacum*. Mol Gen Genet 247:633–638

Fedoroff N (2000) Transposons and genome evolution in plants. Proc Natl Acad Sci USA 97(13):7002–7007

Ganal MW, Lapitan NL, Tanksley SD (1991) Macrostructure of the tomato telomeres. Plant Cell 3:87–94

Gemayel R, Vinces MD, Legendre M, Verstrepen KJ (2010) Variable tandem repeats accelerate evolution of coding and regulatory sequences. Annu Rev Genet 44:445–477

Gschwend AR, Yu Q, Moore P, Saski C, Chen C, Wang J, Na JK, Ming R (2011) Construction of papaya male and female BAC libraries and application in physical mapping of the sex chromosomes. J Biomed Biotechnol 2011:929472

Gschwend AR, Yu Q, Tong EJ, Zeng F, Han J, VanBuren R, Aryal R, Charlesworth D, Moore PH, Paterson AH, Ming R (2012) Rapid divergence and expansion of the X chromosome in papaya. Proc Natl Acad Sci USA. 109(34):13716–13721

Harrison GE, Heslop-Harrison JS (1995) Centromeric repetitive DNA sequences in the genus *Brassica*. Theor Appl Genet 90:157–165

Jeffreys AJ, Wilson V, Thein SL (1985) Hypervariable "minisatellite" regions in human DNA. Nature 314:67–73

Jiang N, Bao Z, Zhang X, Hirochika H, Eddy SR, McCouch SR, Wessler SR (2003) An active DNA transposon family in rice. Nature 421(6919):163–167

Jurka J (2003) Repetitive DNA: detection, annotation, and analysis. In: Krawetz SA, Womble DD (eds) Introduction to bioinformatics: a theoretical and practical approach. Humana Press, Totowa

Kipling D (1995) The telomere. Oxford University Press, Oxford

Kubis SE, Schmidt T, Heslop-Harrison JS (1998) Repetitive DNA elements as a major component of plant genomes. Ann Bot (Lond) 82:45–55

Li W, Godzik A (2006) CD-HIT: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22:1658–1659

Lim KB, de Jong H, Yang TJ, Park JY, Kwon SJ, Kim JS, Lim MH, Kim JA, Jin M, Jin YM, Kim SH, Lim YP, Bang JW, Kim HI, Park BS (2005) Characterization of rDNAs and tandem repeats in the heterochromatin of *Brassica rapa*. Mol Cells 19:436–444

Llorens C, Futami R, Covelli L, Domínguez-Escribá L, Viu JM, Tamarit D, Aguilar-Rodríguez J, Vicente-Ripolles M, Fuster G, Bernet GP, Maumus F, Muñoz-Pomer A, Sempere JM, Latorre A, Moya A (2011) The gypsy database (GyDB) of mobile genetic elements: release 2.0. Nucleic Acids Res 39 (database issue):D70–D74

Lunyak VV, Prefontaine GG, Núñez E, Cramer T, Ju BG, Ohgi KA, Hutt K, Roy R, García-Díaz A, Zhu X, Yung Y, Montoliu L, Glass CK, Rosenfeld MG (2007) Developmentally regulated activation of a SINE B2 repeat as a domain boundary in organogenesis. Science 317(5835):248–251

Ma H, Moore PH, Liu Z, Kim MS, Yu Q, Fitch MM, Sekioka T, Paterson AH, Ming R (2004) High-density linkage mapping revealed suppression of recombination at the sex determination locus in papaya. Genetics 166(1):419–436

Macas J, Mészáros T, Nouzová M (2002) PlantSat: a specialized database for plant satellite repeats. Bioinformatics 18:28–35

Martienssen R, Irish V (1999) Copying out our ABCs: the role of gene redundancy in interpreting genetic hierarchies. Trends Genet 15(11):435–437

Martinez-Zapater JM, Estelle MA, Somerville CR (1986) A highly repeated DNA sequence in *Arabidopsis thaliana*. Mol Gen Genet 204:417–423

Matsunaga S (2009) Junk DNA promotes sex chromosome evolution. Heredity 102:525–526

McClintock B (1950) The origin and behavior of mutable loci in maize. Proc Natl Acad Sci USA 36:344–355

Meagher TR, Vassiliadis C (2005) Phenotypic impacts of repetitive DNA in flowering plants. New Phytol 168:71–80

Messing J, Bharti AK, Karlowski WM, Gundlach H, Kim HR, Yu Y, Wei F, Fuks G, Soderlund CA, Mayer KF, Wing RA (2004) Sequence composition and genome organization of maize. Proc Natl Acad Sci USA 101:14349–14354

Miklos GL (1985) Localited highly repetitive DNA sequences in vertebrate and invertebrate genomes. In: McIntryre JR (ed) Molecular evolutionary genetics. Plenum, New York

Miller JT, Jackson SA, Nasuda S, Gill BS, Wing RA, Jiang J (1998) Cloning and characterization of a centromere specific DNA element from Sorghum bicolor. Theor Appl Genet 96:832–839

Ming R, Moore PH, Zee F, Abbey CA, Ma H, Paterson AH (2001) Construction and characterization of a papaya BAC library as a foundation for molecular dissection of a tree-fruit genome. Theor Appl Genet 102:892–899

Ming R, Hou S, Feng Y et al (2008) The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). Nature 452:991–996

Ming R, Bendahmane A, Renner SS (2011) Sex chromosomes in land plants. Annu Rev Plant Biol 62:485–514

Na JK, Wang J, Murray JE, Gschwend AR, Zhang W, Yu Q, Navajas-Pérez R, Feltus FA, Chen C, Kubat Z, Moore PH, Jiang J, Paterson AH, Ming R (2012) Construction of physical maps for the sex-specific regions of papaya sex chromosomes. BMC Genomics 13:176

Nagarajan N, Pop M (2009) Parametric complexity of sequence assembly: theory and applications to next generation sequencing. J Comput Biol 16(7):897–908

Nagarajan N, Navajas-Pérez R, Pop M, Alam M, Ming R, Paterson AH, Salzberg SL (2008) Genome-wide analysis of repetitive elements in papaya. Trop Plant Biol 1(3–4):191–201

Navajas-Pérez R (2012) The genus *Rumex*: a plant model to study sex-chromosomes evolution. In: Navajas-Pérez R (ed) New insights on plant sex chromosomes, 1st edn. Nova, Hauppauge

Navajas-Pérez R, Paterson AH (2009) Patterns of tandem repetition in plant whole genome assemblies. Mol Gen Genomics 281:579–590

Navajas-Pérez R, Rubio-Escudero C, Aznarte JL, Ruiz Rejón M, Garrido-Ramos MA (2007) satDNA Analyzer: a computing tool for satellite-DNA evolutionary analysis. Bioinformatics 23(6):767–768

Navajas-Pérez R, Quesada del Bosque ME, Garrido-Ramos MA (2009a) Effect of location, organization, and repeat copy number in satellite-DNA evolution. Mol Gen Genomics 282: 395–406

Navajas-Pérez R, Schwarzacher T, Ruiz Rejón M, Garrido-Ramos MA (2009b) Characterization of RUSI, a telomere-associated satellite-DNA, in the genus *Rumex* (Polygonaceae). Cytogenet Genome Res 124(1):81–89

Novikov A, Smyshlyaev G, Novikova O (2012) Evolutionary history of LTR retrotransposon chromodomains in plants. Int J Plant Genomics 2012:874743

Ohno S (1972) So much "junk" DNA in our genome. Brookhaven Symp Biol 23:366–370

Orgel LE, Crick FH (1980) Selfish DNA: the ultimate parasite. Nature 284:604–607

Pace NR (2009) Mapping the tree of life: progress and prospects. Microbiol Mol Biol Rev 73(4):565–576

Pérez OJ, Dambier D, Ollitrault P, Coppens DG et al (2006) Microsatellite markers in *Carica papaya* L.: isolation, characterization and transferability to *Vasconcellea* species. Mol Ecol Notes 6:212–217

Petracek ME, Lefebvre PA, Silflow CD, Berman J (1990) *Chlamydomonas* telomere sequences are A+T rich but contain three consecutive G-C base pairs. Proc Natl Acad Sci USA 87: 8222–8226

Poole RL (2007) The TAIR database. Methods Mol Biol 406:179–212

Price AL, Jones NC, Pevzner PA (2005) De novo identification of repeat families in large genomes. Bioinformatics 21:351–358

Ramos HC, Pereira MG, Silva FF, Gonçalves LS, Pinto FO, de Souza Filho GA, Pereira TS (2011) Genetic characterization of papaya plants (*Carica papaya* L.) derived from the first backcross generation. Genet Mol Res 10(1):393–403

Ray DA (2007) SINEs of progress: mobile element applications to molecular ecology. Mol Ecol 16(1):19–33

Richards EJ, Ausubel FM (1988) Isolation of a higher eukaryotic telomere from *Arabidopsis thaliana*. Cell 53:127–136

Riethman H, Ambrosini A, Paul S (2005) Human subtelomere structure and variation. Chromosome Res 13:505–515

Robles F, De La Herrán R, Ludwig A, Ruiz Rejón C, Ruiz Rejón M, Garrido-Ramos MA (2004) Evolution of ancient satellite DNAs in sturgeon genomes. Gene 338:133–142

Román AC, González-Rico FJ, Moltó E, Hernando H, Neto A, Vicente-García C, Ballestar E, Gómez-Skarmeta JL, Vavrova-Anderson J, White RJ, Montoliu L, Fernández-Salguero PM (2011) Dioxin receptor and SLUG transcription factors regulate the insulator activity of B1 SINE retrotransposons via an RNA polymerase switch. Genome Res 21(3):422–432

Saini N, Shultz J, Lightfoot DA (2008) Re-annotation of the physical map of Glycine max for polyploid-like regions by BAC end sequence driven whole genome shotgun read assembly. BMC Genomics 9:323

Santos SC, Ruggiero C, Silva CLSP, Lemos GM (2003) A microsatellite library for *Carica papaya* L. cv Sunrise Solo. Rev Bras Frutic 25:263–267

Schmidt AL, Anderson LM (2006) Repetitive DNA elements as mediators of genomic change in response to environmental cues. Biol Rev Camb Philos Soc 81(4):531–543

Shakirov EV, Salzberg SL, Alam M, Shippen DE (2008) Analysis of *Carica papaya* telomeres and telomere-associated proteins: insights into the evolution of telomere maintenance in Brassicales. Trop Plant Biol 1(3–4):202–215

Sola-Campoy PJ, de la Herrán R, Ruiz Rejón C, Navajas-Pérez R (2012) Plant sex-chromosomes evolution. In: Navajas-Pérez R (ed) New insights on plant sex chromosomes, 1st edn. Nova, Hauppauge

Sykorova E, Lim KY, Chase MW, Knapp S, Leitch IJ, Leitch AR, Fajkus J (2003) The absence of *Arabidopsis*-type telomeres in *Cestrum* and closely related genera *Vestia* and *Sessea* (*Solanaceae*): first evidence from eudicots. Plant J 34:283–291

Temnykh S, DeClerck G, Lukashova A, Lipovich L, Cartinhour S, McCouch S (2001) Computational and experimental analysis of microsatellites in rice (Oryza sativa L.): frequency, length variation, transposon associations, and genetic marker potential. Genome Res 11: 1441–1452

Thomas CA Jr (1971) The genetic organization of chromosomes. Annu Rev Genet 5:237–256

Thornburg BG, Gotea V, Makalowski W (2006) Transposable elements as a significant source of transcription regulating signals. Gene 365:104–110

Ugarkovic D, Plohl M (2002) Variation in satellite DNA profiles, causes and effects. EMBO J 21:5955–5959

Wang ZX, Kurata N, Saji S, Katayose Y, Minobe Y (1995) A chromosome 5-specific repetitive DNA sequence in rice (*Oryza sativa* L.). Theor Appl Genet 90:907–913

Wang J, Chen C, Na JK, Yu Q, Hou S, Paull RE, Moore PH, Alam M, Ming R (2008) Genome-wide comparative analyses of microsatellites in papaya. Trop Plant Biol 1(3–4):278–292

Wang Y, Wang X, Paterson AH (2012) Genome and gene duplications and gene expression divergence: a view from plants. Ann N Y Acad Sci 1256(1):1–14

Wang J, Na JK, Yu Q, Gschwend AR, Han J, Zeng F, Aryal R, VanBuren R, Murray JE, Zhang W, Navajas-Pérez R, Feltus FA, Lemke C, Tong EJ, Chen C, Wai CM, Singh R, Wang ML, Min XJ, Alam M, Charlesworth D, Moore PH, Jiang J, Paterson AH, Ming R (2012) Sequencing papaya X and Yh chromosomes reveals molecular basis of incipient sex chromosome evolution. Proc Natl Acad Sci USA. 109(34):13710– 13715

Wang J, Na J-K, Ming R (2013) Physical mapping of papaya sex chromosomes. In: Ming R, Moore P (eds) Genetics and genomics of papaya. Springer Science + Business Media, New York

Watson JM, Riha K (2010) Comparative biology of telomeres: where plants stand. FEBS Lett 584(17):3752–3759

Yu Q, Hou S, Hobza R, Feltus FA, Wang X, Jin W, Skelton RL, Blas A, Lemke C, Saw JH, Moore PH, Alam M, Jiang J, Paterson AH, Vyskot B, Ming R (2007) Chromosomal location and gene paucity of the male specific region on papaya Y chromosome. Mol Genet Genomics 278(2):177–185

Yunis JJ, Yasmineh WG (1971) Heterochromatin, satellite DNA, and cell function. Science 174(4015):1200–1209

Chapter 14
Genomics of Papaya Fruit Development and Ripening

Robert E. Paull, Pingfang Wu, and Nancy J. Chen

Introduction

The papaya fruit is a fleshy berry and fruit growth follows a single sigmoid growth curve (Zhou and Paull 2001). During development, all tissues in the gynoecium less than 1 mm are meristematic (Roth and Clausnitzer 1972). Later the outer layer of the epidermis increases in size, while the subepidermal layer continues to divide both anticlinally and periclinally. The central parenchyma of the pericarp increases in size and divides with the placenta forming opposite the marginal vascular bundles. This meristematic activity lasts 28–42 days and determines final fruit size. Fruit growth shows two major phases. The first lasts about 80 days post-anthesis, with a large increase in dry weight occurring just before fruit maturity. Fruit development takes 150–164 days that is extended another 14–21 days in Hawaii in the colder months (Paull and Chen 1983; Qiu et al. 1995). Mesocarp growth parallels seed and total fruit growth.

Papaya fruit shape is a sex-linked character and ranges from spherical to ovoid in female flowers to long, cylindrical, or pyriform (pear shaped) in hermaphrodite flowers (Table 14.1). The fruit is normally composed of five carpels united to form a central ovarian cavity that is lined with the placenta carrying numerous black seeds. Placentation is parietal with the seeds attached by 0.5–1 mm stalks. The ovarian cavity is larger in female fruit than hermaphrodite. The shape of the cavity at the transverse cut ranges from star shape with five to seven furrows to smooth and circular (Chan and Paull 2007).

Papaya ripening is climacteric with the rise in ethylene production occurring at the same time as the respiratory rise (Paull and Chen 1983). Respiration is a critical factor in fruit growth and development, especially as it relates to fruit ripening.

R.E. Paull (✉) • P. Wu • N.J. Chen
Department of Tropical Plant and Soil Sciences, University of Hawaii at Manoa, 3190 Maile Way, Honolulu, HI 96822, USA
e-mail: paull@hawaii.edu

R. Ming and P.H. Moore (eds.), *Genetics and Genomics of Papaya*, Plant Genetics and Genomics: Crops and Models 10, DOI 10.1007/978-1-4614-8087-7_14,

Table 14.1 Range of papaya fruit genetic characteristics [modified with permission from Ito et al. (1990)]

	Range	
(A) Fruit characteristics		
Fruit size	++++	100 g to 12 kg
Total soluble solids	+++	Thai (8–11 %) < Solo (12–18 %)
Fruit taste	+	Perfumey (Hortus Gold) ↔ off-flavor
Fruit shape	++	Round, pear shape (Solo), oval with pointed stylar end
Fruit shape uniformity		Inbreed lines are uniform
Seed cavity	++	Star shaped (Solo) ↔ round cavity (Kaek Dum)
Skin spots "freckles"	+++	None Malay Yellow < "Waikane Solo" < Solo
(B) Ripening		
Days to full skin color	++	
Days to edible condition	++	Sunrise < Kapoho
Flesh texture of ripe fruit	++	Soft water soaked to firm Hortus Gold < Ostrem < Sunrise < Kapoho

+ indicates extent of range from small (+) to large (++++)

Fruits are divided into two broad groups based on the role of ethylene in the ripening process and its relationship to the respiratory pattern. Climacteric fruits (e.g., banana, papaya, peach, tomato) demonstrate a peak in respiration and ethylene production during ripening, including autostimulatory (System 2) ethylene production (Biale 1964). The timing of the autostimulatory ethylene peak (Lelievre et al. 1997; Barry et al. 2000) and of other ripening events (skin color changes, carotenoid synthesis, flavor development, softening) varies widely between species and cultivars (Biale 1964; Burg and Burg 1965; Bruinsma and Paull 1984; Gussman et al. 1993). Non-climacteric fruit shows a gradual decline in the respiration rate and shows no marked peak in ethylene production and gradual change in other ripening parameters.

Fruit Development

Fruit Growth

The plant growth regulator abscisic acid (ABA) plays a crucial role in the plant's adaptation to stress and in seed maturation and dormancy, fruit ripening, and senescence (Zeevart and Creelman 1988). ABA and strigolactone are products from the cleavage of carotenoids at one of its double bonds. Carotenoid cleavage is carried out by two groups of related enzymes. One group uses multiple carotenoid substrates and is referred to as carotenoid cleavage dioxygenase (CCD). The other group has C_{40}-9-*cis*-epoxycarotenoid as the preferred substrate and referred to as 9-*cis*-epoxycarotenoid dioxygenase (NCED). The NCEDs are thought to be involved in ABA synthesis while the role of CCDs is less clear. Some evidences suggest that CCDs have a role in lateral shoot growth (Auldridge et al. 2006b;

Azarkan et al. 2006), possibly by cleavage of carotenoids to strigolactone or a related compound (Gomez-Roldan et al. 2008).

Two *Arabidopsis* genes *CCD7/MAX3* and *CCD8/MAX4* control lateral shoot growth (Booker et al. 2004; Auldridge et al. 2006a; Gomez-Roldan et al. 2008). The citrus homolog gene (*CsNCED1*) likely plays a role in ABA synthesis in leaves and fruits (Rodrigo et al. 2006). ABA has also been given a role in the regulation of citrus fruit skin coloration (Rodrigo et al. 2003). Citrus are non-climacteric fruit and normally respond to ethylene by induced fruit coloration (Alferez and Zacarias 1999). Papaya is a climacteric fruit though fruit mesocarp carotenoid development and fruit degreening are ethylene independent (Manenoi et al. 2007). ABA via CCD and/or NCED may therefore play a role in fruit mesocarp carotenoid development and skin degreening. One CCD (*CpCCD1*) and two NCEDs (*CpNCED1*, *CpNCED2*) are predicted in papaya (Paull et al. 2008). A partial open reading frame (ORF) was found for another possible *Cp*C*CD* that had 67 % identity to *Arabidopsis* CCD1. A papaya expressed sequence tag (EST) is not found for *CpCCD1* or *CpNCED1*, and a chloroplast transit peptide was not found for *CpNCED1*. The *Arabidopsis* carotenoid cleavage family has five putative NCEDs and four putative CCDs (Auldridge et al. 2006b), and tomato has two: *LeCCD1A* and *LeCCD1B* (Schwab et al. 2008). *CpCCD1* had 69 % identity with *ArabidopsisAtNCED4*; *CpNCED1* has 73 % identity with *ArabidopsisAtNCED5*. *CpNCED2* had 77 % identity with *AtNECED3*. The *CpCCD* and the *CpNCEDs* were all predicted to have a carotenoid oxygenase motif.

Fruit Growth

The plasticity in papaya size from 0.1 to 12 kg (Nakasone and Paull 1998) provides a unique opportunity to study fruit development control (Table 14.1). Tomato fruit can similarly vary from small berries (~2 g) to large fruit (1,000 g) (Lippman and Tanksley 2001) with a single gene *ORFX* (QTL loci fw 2.2) being responsible for 30 % of the difference (Frary et al. 2000; Bartley and Ishida 2003; Cong and Tanksley 2006). *ORFX* appears to act at or near the plasma membrane β subunit of CKII kinase (Cong and Tanksley 2006) and exerts its control through early fruit cell division (Lippman and Tanksley 2001). One homolog to *ORFX* was predicted in papaya (*CpORFX*) with 63 % identity with *ORFX* from *Solanum pennellii* and 62 % with *Solanum lycopersicum*. *CpORFX* shows homology to the cysteine-rich PLAC8 domain of unknown function that is found in animals and plants (Marchler-Bauer et al. 2005). No papaya ESTs were detected for *CpORFX*. The function of *CpORFX* in papaya is unknown though the wide range of papaya fruit sizes found presents the possibility it may have a role similar to that in tomato.

Fruit Shape

The shape of papaya fruit varies with fruit from female flowers being spherical to ovoid (Table 14.1), while fruit from hermaphrodite flowers being cylindrical or pear

shaped (Nakasone and Paull 1998). A mutation in the quantitative trait locus (QTL) *OVATE* changes the shape of tomato from round to pear shaped (Liu et al. 2002). This regulatory gene apparently has its impact early in flower development (Ku et al. 2000a, b). A similar QTL in eggplant has a similar effect on shape (Ku et al. 1999). Ovate is a negative regulatory hydrophilic protein with a putative bipartite nuclear localization signal (Liu et al. 2002). Three ovate proteins are predicted in the papaya genome with homology to ovate sequences in *Arabidopsis* and *S. lycopersicum* (Blas et al. 2012). Papaya ESTs are only found for one of the gene (*CpOVATE1*) with 38 % identity to *S. lycopersicum OVATE* and 39 % identity with *Arabidopsis*. All three predicted papaya genes had the ~70 aa plant-specific motif (DUF623) of unknown function, also reported for the tomato *OVATE* gene (Liu et al. 2002). *CpOVATE 2* has 31 % identity with *Arabidopsis* Ovate family (*AtOFP8*) and 58 % identity with tomato. *CpOVATE3* which was on the same linkage group as the papaya sex-related gene has 67 % identity with tomato ovate protein and an ovate-like protein from tobacco with 65 % identity.

Another major gene controlling the elongated fruit shape of tomato is *SUN* that acts in a dosage-dependent manner (Xiao et al. 2008). The gene belongs to the IQD family with *AtIQD1* being the only member with a known function. A homolog to *SUN* is found in papaya (Blas et al. 2012). This papaya homolog has four introns, the same as *SUN* and is 446 aa long versus 405 aa for *SUN* (Xiao et al. 2008). The homology of the predicted papaya gene is moderate (37 %) for *SUN* and 50 % for *Arabidopsis IQD11* and 35 % for *IQD12*. Another predicted papaya homolog is two-thirds the length and has only two introns possibly due to a gap between the two sequenced contigs. This second predicted gene is 45 % homologous to *SUN* and 53 % to *Arabidopsis IQD12*. The variation in papaya shape from round to elongated, between female and hermaphrodite fruit, presents a unique model to ascertain the role of *SUN* homologs in determining fruit shape.

Expansins

The cell wall proteins termed expansin are involved in cell wall relaxation and growth (McQueen-Mason et al. 1992). A number of expansin genes are recognized (Cosgrove 2007). Expansins have been shown to be expressed during tomato fruit growth (Rose et al. 2000), and different isoforms are expressed during fruit ripening (Rose et al. 1997; Brummell et al. 1999, 2004). Papaya fruit has been shown to express an expansin (*CpEXPA1*), and four have been reported in banana fruit (Asha et al. 2007). Immunoblots have also detected expansins in ripening pear, persimmon, kiwi fruit, strawberry, and pineapple but not detected in pepper (Rose et al. 2000). The papaya genome contains at least 15 *CpEXPA* (Expansin A), three *CpEXPB* (Expansin B), and one *CpEXPLA* (Expansin Like A) (Table 14.2). Secretory sequences are not predicted on four of the *CpEXPA*s. All *CpEXP*s had similar intron positions and lengths to those described for *Arabidopsis* (Choi et al. 2006). Sampedro et al. (2005, 2006) proposed that the number in the last common

Table 14.2 Expansin families in different plant species, gene numbers, and the last common ancestor from Sampedro et al. (2006)

Species	EXPA	EXPB	EXLA	EXLB
Last common ancestor	12	2	1	2
Papaya	15	3	1	0
Arabidopsis	26	6	3	1
Poplar	27	2	2	4
Rice	34	19	4	1

The estimates for papaya do not include incomplete sequences [reproduced with kind permission from Springer Science+Business Media from Paull et al. (2008)]

ancestor for expansin *EXLB* is four, though none are predicted for papaya (Paull et al. 2008). The monocot/dicot ancestor has 15–17 expansin genes (Sampedro et al. 2005). Papaya is classified as a basal clade in the order Brassicales (Ronse de Craene and Haston 2006) with at least 19 expansin genes close to the number predicted for the monocot/dicot ancestor (15) (Sampedro et al. 2005).

The expansins are believed to have arisen and diversified early if not before colonization of plants on land (Li et al. 2002; Choi et al. 2006). Expansins are found in monocots, pines, ferns, and mosses. The closely related expansin-like sequences are similarly widely found, and the absence of any *EXLB* predicted in papaya is unexpected. One *EXLB* is found in *Arabidopsis* and rice, and four in poplar, and one predicted in pine (Sampedro et al. 2005, 2006). This suggests that *EXLB*s were present before the angiosperms/gymnosperms separation (Table 14.2). The *Arabidopsis* (*At*4g17030) and *Pinus EXLB* are distantly related (Sampedro et al. 2006), but no match was found to any predicted papaya gene sequences or in the papaya EST database. When predicted papaya protein models are queried using the *Arabidopsis* expansin-like B sequence, homology is found to *EXPLA* and *EXPB* sequences predicted in papaya. A number of other papaya sequences are predicted with an EG45 motif though these peptides had less than 100 amino acids and had no CBD motif.

Cell Wall Synthesis Genes

The major components of plant cell walls are structural proteins, cellulose, matrix polysaccharides, pectins, and in secondary cell walls, lignin (Carpita and Gibeaut 1993; Fry 2004). Matrix polysaccharides and pectins are two of the most important components of the cell wall, but little is known about their biosynthesis, assembly, and degradation (Cosgrove 1999). Matrix polysaccharides comprise the cross-linking glycan molecules that are bound to the cell wall through covalent and non-covalent bonds (Carpita and Gibeaut 1993; Fry 2004). Pectins are a family of complex polysaccharides that all contain the acidic 1,4-linked α-D-galacturonic acid and surround the cellulose microfibrils and cross-linked matrix polysaccharides (Carpita and Gibeaut 1993; Cosgrove 2001).

Cell wall synthesis genes potentially associated with growth and development involve numerous glycosyltransferase (GT) genes in different families

Table 14.3 Major carbohydrate transferases in the papaya genome to *Arabidopsis* and tomato

Activity and family	Papaya	*Arabidopsis*	Poplar	Grape	Tomato	Peach
β-Glycosyltransferase, β-xylosyltransferase, β-rhamnosyltransferase, β-glucuronic acid transferase, β-galactosyltransferase (GT 1)	55	121	93	72	162	159
Glycosyltransferase, cellulose synthase (GT 2)	21	42	6	10	9	11
Sucrose synthase (GT 4)	14	24	18	30	28	31
α-Glucosyl and α-galactosyltransferase, α-1,4-galacturonosyltransferase (GT 8)	25	41	30	37	42	32
Branching, protein glycosyltransferase (GT 14)	5	11	20	26	33	30
Trehalose 6-phosphate transferase (GT 20)	5	11	5	10	10	9
Mannosyltransferase (GT 22)	2	3	2	3	3	2
Sialyltransferase (GT 29)	3	3	3	3	3	3
β-Galactosyltransferase, β-GalNAc transferase (GT 31)	8	33	17	19	18	19
α-Galactosyltransferase (GT 34)	3	8	1	3	5	7
Xyloglucan α-1,2-fucosyltransferase (GT 37)	3	10	6	2	3	2
β-Glucuronosyltransferase (GT 43)	4	4	4	4	5	4
Exostosin, β-glucuronosyltransferase (GT 47)	28	39	35	35	46	52
β-1,3-D-Glucan synthase (GT 48)	3	13	4	9	9	12
Self-glucosylating β-glucosyltransferase, L-arabinopyranose mutase (GT 75)	3	5	6	5	8	7

The carbohydrate catalytic group and families are given in parenthesis (Coutinho and Henrissat 1999). CAZy web site for *Arabidopsis* http://www.cazy.org/e1.html (Accessed 2011 December 29) and http://cellwall.genomics.purdue.edu/families/2.html (Accessed 2011 December 30). Tomato, apple, peach, and the moss (*Physcomitrella patens*) from http://www.plantgdb.org/ (Accessed 2012 March 03). The predicted genes were submitted to Bioenergy Science Center CAZymes Analysis Toolkit (Park et al. 2010) (http://cricket.ornl.gov/cgi-bin/cat.cgi) for identification of carbohydrate active enzymes; the results were then manually curated

(Henrissat et al. 2001). Papaya has at least 55 putative β-glucosyl, xylosyl, and rhamnosyltransferases (GT1) genes compared to 121 in *Arabidopsis* and 162 in tomato and higher numbers in other species. The higher number in tomato probably represents the triplicate replication that occurred on two separate occasions in this genome (The Tomato Genome Consortium 2012). These GT1 members are involved in hemicellulose synthesis. Twenty-one glycosyltransferases that are cellulose synthase-related genes (GT2) are identified in papaya versus the 48 *Populus* and 42 *Arabidopsis* and other sequenced species with the exception of poplar where only 6 are predicted (Table 14.3). These results suggest that an increase in the number of cellulose-related structural genes (Ces, Csl) might be due to genome duplication in *Arabidopsis* and specifically required for the biosynthesis of wood in poplar.

Galacturonosyltransferases are in family GT8 and are associated with pectin synthesis; 41 genes have been reported in *Arabidopsis* though only 25 have been found in papaya based upon sequence comparison with all other species so far having higher numbers than papaya (Table 14.3). There is more similarity for GT47 the glucuronosyltransferases, with papaya having 28 predicted genes and *Arabidopsis* having 39 genes. Papaya has at least 3 genes and *Arabidopsis* 10 genes for xyloglucan α-1,2-fucosyltransferase (GT37). Three genes for β-1,3-glucan (callose) synthase (GT48) are found in papaya with 13 reported for *Arabidopsis* and 9 genes in tomato. For most GT families, papaya has fewer genes for cell wall synthesis and metabolism than *Arabidopsis* and often similar number or less than that reported for other species. The exceptions are that papaya has the same number of genes for sialyltransferase (GT29) and β-glucuronosyltransferase (GT43) as other species listed, suggesting selected gene loss. The fewer genes for cell wall synthesis possibly reflect, in part, the absence of genome duplication in papaya.

Cell Wall Expansion and Degradation

Plant cell walls are capable of both plastic and elastic extension and control the rate and direction of cell expansion (Fry 2004). Modification of cell wall components during plant growth and development can significantly alter the cellular mechanics and the control of growth and morphogenesis. Enzymes that modify the cell wall structure are thought to play a central role in turnover and modification. The enzymes have been classified into a number of groups based upon catalytic activity [hydrolases (GH), esterases (CE), lyases (PL)] and then placed into families based upon their amino acid sequence and structure. This classification is used to analyze the *Arabidopsis* genome (Coutinho and Henrissat 1999; Henrissat et al. 2001). In fruit ripening, endo- and exo-polygalacturonase (PG), pectin methyl esterase, glucanase, galactosidase, and xylanase have been detected in papaya fruit (Paull and Chen 1983; Chen and Paull 2003; Thumdee et al. 2010). In papaya fruit, α-galactosidase protein and activity are detected and may be involved in pectin-related textural change during ripening (Soh et al. 2006). As tomato fruit ripens, the PG and β-galactosidase II mRNA increases in parallel with an increase in enzyme activities (Smith and Gross 2000). There are three beta-galactosidase genes expressed during papaya fruit ripening (Othman et al. 2011). The authors suggested that pPBGII and pBG(a) cDNA clones characterized in this work may be involved in fruit softening during papaya ripening while the fruit-specific pBG(b) may be related to early ripening stage.

As with the GT families, papaya has fewer members in most catalytic families than *Arabidopsis* or other species (Table 14.4). The exceptions are the chitinases (GH18), α-mannosidase (GH47), and carboxylesterase (CE1). In the case of the chitinases (GH18, GH19), this may reflect increased need for fungal resistance (Zhu et al. 2003). Pectin methyl esterases that remove methyl groups from the carbonyl group are present in multiple copies with 58 potential genes (CE8) in papaya (Table 14.4), 67 genes in *Arabidopsis*, and at least 79 in tomato and a higher number in other species.

Table 14.4 Major carbohydrate hydrolases, esterases, and lyases in papaya compared to *Arabidopsis* and tomato

Catalytic group Activity and family	Papaya	*Arabidopsis*	Poplar	Grape	Tomato	Peach
Hydrolases						
β-Glucosidase, β-galactosidase, β-mannosidase, (GH 1)	21	48	16	40	25	77
β-glucosidase, 1,4-β-xylosidase, 1,3-β-glucosidase (GH 3)	11	16	8	29	18	22
Endo-β-1,4-mannosidase, endo-1,4-β-glucanase (GH 5)	4	13	4	9	9	NP
Endo-1,4-β-glucanase (GH 9)	14	26	17	21	23	NP
Endo-1,4-β-xylanase (GH 10)	5	12	4	4	11	6
α-Amylase (GH 13)	7	10	8	9	10	10
β-Amylase (GH 14)	7	9	13	9	7	10
Xyloglucan endo-transglycosylase/ hydrolase (GH 16)	16	33	30	36	38	29
Glucan endo-1,3-β-glucosidase (GH 17)	36	51	59	41	51	50
Chitinase, yieldins (GH 18)	15	10	13	20	12	23
Chitinase (GH 19)	8	14	24	13	21	14
Polygalacturonase (GH 28)	45	68	35	60	56	62
α-Glucosidase (GH 31)	4	5	2	8	8	6
Invertase (GH 32)	3	8	7	8	12	10
β-Galactosidase (GH 35)	10	18	15	25	17	20
α-Galactosidase (GH 36)	7	6	NP	NP	NP	NP
α-Mannosidase (GH 38)	2	4	2	6	3	5
α-Mannosidase (GH 47)	8	5	2	8	5	6
Alkaline/neutral invertase (GH 100)	3	9	11	11	11	16
Carbohydrate esterase families						
Acetyl xylan esterase, feruloyl esterase, carboxylesterase (CE 1)	2	1	2	1	1	1
Pectin methylesterase (CEF 8)	58	67	51	50	79	72
Carboxylesterase, arylesterase (CE 10)	11	24	38	36	38	41
Pectin acetylesterase (CE 13)	5	12	13	7	18	12
Pectate lyase						
Pectate/pectin lyases (PL 1)	21	26	11	17	23	20
Rhamnogalacturonan lyase (PL 4)	2	8	NP	NP	NP	NP
α-L-guluronate lyase (PL 7)	NP	NP	2	3	1	5

The carbohydrate catalytic group and families are given in parenthesis (Coutinho and Henrissat 1999) (*GH* glycohydrolase family; *CE* carbohydrate esterase family). CAZy web site for *Arabidopsis* http://www.cazy.org/e1.html (Accessed 2011 December 29) and tomato, apple, peach, and the moss (*Physcomitrella patens*) from http://www.plantgdb.org/ (Accessed 2012 March 03). The predicted genes were submitted to Bioenergy Science Center CAZymes Analysis Toolkit (http://cricket.ornl.gov/cgi-bin/cat.cgi) for identification of carbohydrate active enzymes; the results were then manually curated. *NP* not predicted

Forty-five genes for polygalacturonase (GH28) are found in papaya with 68 in *Arabidopsis* and at least 56 in tomato, some of which are exclusively expressed in fruit ripening (The Tomato Genome Consortium 2012) with similar numbers in other species.

Hemicellulose degradative enzymes include endo-1,4-β-glucanases (GH5, GH9), endo-1,4-β-xylanases (GH10), and GH16, xyloglucan endotransglycosylase/hydrolase (Table 14.4). Fourteen genes for β-1,4-endoglucanase (GH9) are found in papaya, 26 in *Arabidopsis*, and 23 in tomato. Similarly, *Arabidopsis* has 12 predicted endoxylanase-like genes (Simpson et al. 2003) in GH10, while papaya has 5 putative genes and tomato 10. Fewer genes for xyloglucan endotransglycosylase/hydrolase (GH16, XET, or XTH) are found in papaya with 16 than the 33 in *Arabidopsis* and 38 in tomato. Of the 38 in tomato, fifteen are predominately expressed in fruit development and ripening (The Tomato Genome Consortium 2012). In *Arabidopsis*, there are 51 β-1,3-glucanases (GH17) and only 36 genes so far found in papaya. Multiple copies of β-galactosidase (GH35) are found in papaya, *Arabidopsis*, and tomato, 10 copies in papaya, 18 in *Arabidopsis,* and 17 in tomato. Papaya has 21 pectin lyase (PL 1) genes, while *Arabidopsis* has 26 and tomato 23. The gene number in papaya for starch degradation is also fewer than in *Arabidopsis* (Table 14.4). Seven genes for α-amylase (GH13) are found in papaya versus 10 in *Arabidopsis* while there are 7 and 9, respectively, for β-amylase (GH14). In conclusion, the fewer number of cell wall degrading genes in papaya make it a simpler model system to study the network of genes involved in cell wall rearrangement that plays a crucial role in plant development.

Fruit Ripening

Skin Chlorophyll Degradation

The pathway of chlorophyll breakdown during fruit ripening is expected to be similar to that occurring during senescence and involves the removal of the phytol residue and the central Mg by chlorophyllase and a dechelatase (Barry 2009). The product of these two activities is pheophorbide a, which is the substrate for pheophorbide a oxygenase (PAO) and red chlorophyll catabolite reductase (RCCR) (Hortensteiner 2006). Proteomic analysis indicated that PAO located in plastid inner envelope membrane and RCCR was found in chloroplasts and mitochondria (Kleffmann et al. 2004; Yao and Greenbreg 2006). The breakdown product from PAO and RCCR is then exported from the degenerating chloroplast for further breakdown in the cytoplasm. The predicted Mg dechelatase gene has not been identified (Kunieda et al. 2005).

One chlorophyllase (*CpCLH*) gene is predicted in the papaya genome versus two genes in *Arabidopsis thaliana* and three in *Brassica oleracea* (Hortensteiner 2006). The papaya *Cp*CLH protein has a predicted chloroplast transit peptide of 21 amino acids and has 60 % homology to *B. oleracea* BoCLH3 protein and 61 % to *At*CLH2. The *Cp*CLH has a 2Fe–2S ferredoxin, iron-sulfur binding site and is expressed in papaya based on ESTs. A magnesium chelatase subunit gene with three ESTs is predicted in papaya. This gene has 91 % identity with soybean (Paull et al. 2008).

PAO genes have been identified by functional genomics. A single *PAO* gene is predicted for papaya (*CpPAO*). The predicted gene is expressed based upon EST data and has a 72 aa chloroplast transit peptide. *CpPAO* was similar to the *Arabidopsis AtPAO* gene, first described as accelerated cell death (*ACD1*) in *Arabidopsis*. *ACD1* has homology to tomato lethal leaf spot 1-like protein and a putative cell death suppressor protein in *Oryza sativa*. The *PAO* genes of all three species contain the predicted iron-sulfur binding domain (Paull et al. 2008). There are reports on stay-green gene, SGR, and its role in regulating the upstream process of PAO gene (Aubry et al. 2008) resulting in stay-green phenotype. Recently Hu et al. (2011) reported LeSGR gene identification and its expression during tomato fruit ripening. However, no similarity was found to any predicted papaya gene sequences to *Arabidopsis* stay-green or LeSGR.

As with a single gene in *Arabidopsis*, described as ACD2 (Wüthrich et al. 2000; Mach et al. 2001), a single *RCCR* gene is predicted in papaya (*CpRCCR*). The papaya *Cp*RCCR protein has a predicted chloroplast transit peptide of 61 amino acids, and ESTs indicate it is expressed. RCCR has slightly greater identity (74 %) to the tomato *RCCR* (Pruzinska et al. 2007) than to the *ArabidopsisACD2* (71 %).

Carotenoid Biosynthesis

Carotenoids serve several functions in plants (Cunningham and Gantt 1998). During fruit ripening as chlorophyll is degraded, the underlying carotenoids are unmasked, and *de novo* biosynthesis occurs in the papaya mesocarp chromoplasts. Carotenoids are derived from the 5-carbon compound isopentenyl pyrophosphate (IPP). The C_{40} backbone of all carotenoids is assembled from two C_{20} geranylgeranyl pyrophosphate (GGPP) molecules which in turn are each derived from four C_5 IPP molecules; see volatile production discussion that follows.

Synthesis of phytoene from two GGPP molecules is the first step in the carotenoid-specific biosynthesis pathway and is catalyzed by phytoene synthase (PSY). PSY is a key regulator in carotenoid biosynthesis and has been found to be the rate-limiting enzyme in ripening tomato fruits, canola seeds, and marigold flowers (Bramley et al. 1992; Fraser et al. 1994; Hirschberg 2001). PSY is expected to have close membrane association as phytoene is lipid soluble and is localized, along with its subsequent end products, inside the chloroplasts and chromoplasts (Cunningham and Gantt 1998). Additionally, two forms of PSY, a chromoplast- and chloroplast-specific form, have been found in tomato: PSY-1 and PSY-2, respectively (Cunningham and Gantt 1998), with different tissue expression and only PSY-1 controls fruit tissue pigmentation (The Tomato Genome Consortium 2012). Desaturation of phytoene into z-carotene and lycopene is mediated by phytoene desaturase (PDS) and ζ-carotene desaturase (ZDS), respectively. This desaturation converts the colorless phytoene into the pink-hued lycopene. Like PSY, the desaturases PDS and ZDS are closely associated with the plastid membrane but are not integral membrane proteins (Schledz et al. 1996). Plastid transit peptides were

detected for PSY, PDS, ZDS, and LCY-b from the papaya genome predicted genes. The published sequences for both PDS and ZDS appear to be lacking part of the leader sequences and do not start with a methionine required by the transit peptide prediction software. When the predicted peptide sequences for PSY and ZDS from the papaya genome are used, both of which start with methionine, transit peptides are predicted.

Cyclization of lycopene via lycopene e-cyclase (LCY-e) or lycopene b-cyclase (LCY-b) results in a α- or β-carotene, respectively. Yamamoto (1964) showed that 59.3 % of the total carotenoids in yellow-fleshed papaya are comprised of b-carotene or its derivatives, i.e., cryptoxanthin, indicating the b-ring cyclization pathway for synthesis of the yellow-pigmented carotenoids. Yellow fruits contain only trace amounts of lycopene (Schweiggert et al. 2011), while lycopene predominates in red papaya (51 % of total carotenoids). Tubular plastids are abundant in yellow papaya with larger crystalloid structures present in red papaya chromoplasts.

The family of lycopene cyclases in plants shows evidence of multiple gene duplication events and divergence of catalytic function (Blas et al. 2010). Plant lycopene cyclases share a common phylogenetic origin with bacterial crtY and crtL lycopene cyclases. Plant and bacterial lycopene cyclases are polypeptides of ~400 aa; however, the plant lycopene cyclases have an additional 100 aa N-terminal transit sequence. Five regions of conserved amino acid sequence have been identified: one putative dinucleotide-binding region and four motifs of unknown function (Armstrong and Hearst 1996; Cunningham and Gantt 1998).

Several papaya genes that encode enzymes in the carotenoid biosynthesis pathway have been previously identified: phytoene synthase (*PSY*), phytoene desaturase (*PDS*), ζ-carotene desaturase (*ZDS*), and a chloroplast-specific lycopene β-carotene (*LCY-b*) (Paull et al. 2008; Blas et al. 2010). BLASTn search of an EST database generated from a multiple-tissue type cDNA library supports the activity of these single genes in papaya. Putative papaya homologs were identified for lycopene ε-cyclase (*LCY-e*), β-ring carotene hydroxylase (*CRTR*), and zeaxanthin epoxidase (*ZE*) based on amino acid sequences obtained from GenBank (Paull et al. 2008) and cloning of a chromoplast-specific lycopene α-cyclase, *CpCYC-b* (Blas et al. 2010). Expression of these putative papaya carotenoid biosynthesis genes was supported by EST data in NCBI that include fruit EST (Devitt et al. 2006).

Sugar Accumulation

Photosynthates are transported into fruits mainly as sucrose with some fruits accumulating starch that is broken down to sugars during ripening while other fruit such as papaya depending upon a continued supply of sucrose for sweetness when ripe. During early fruit growth the sucrose is thought to be unloaded through the symplast pathway involving SS (sucrose synthase) and SPS (sucrose phosphate synthase) activities and metabolized in respiration and used for growth (Patrick 1997; Sturm and Tang 1999; Zhang et al. 2006). The activities of SS and SPS and their

mRNA levels parallel the increase in growth. After fruit growth stops, sucrose accumulation is thought to involve a cell wall invertase to cleave the sucrose arriving via the phloem to hexoses, thus maintaining the sucrose source to sink gradient. The hexoses are then taken up into the fruit cells and the vacuole via hexose transporters (Caspari et al. 1994; Patrick 1997; Wu et al. 2004; Zhang et al. 2004, 2006).

Sugars, mainly as sucrose, begin to accumulate in papaya fruit about 110 days after anthesis during the last 28–42 days of fruit development (Chan 1979; Zhou and Paull 2001). Flesh total soluble solids can be as low as 5 % and up to 19 % (Table 14.1). The fact that invertase gene expression (Zhou et al. 2003; Zhu et al. 2003) and protein expression (Nogueira et al. 2012) and sugar (hexose) transporters (Sangwanangkul and Paull 2005) are upregulated and the protein detected just before and during sugar accumulation suggests the participation of both invertases and hexose transporter activities in fruit sugar accumulation.

Invertase (INV) is responsible for hydrolyzing sucrose into fructose and glucose and can be found in the cytoplasm, vacuole, and apoplast (Sturm 1999). The family is classified into acid, alkaline, and cell wall invertases (Tymowska-Lalanne and Kreis 1998; Ji et al. 2005; Bocock et al. 2008). *Arabidopsis* has one cell wall invertase and one acid invertase (Tsuchisaka et al. 2007). In tomato, 16 cell wall invertase sequences have been found (Fridman and Zamir 2003), while only 2 cell wall invertases are found in the papaya genome, *CpCWINV1* and *CpCWINV2* (Paull et al. 2008). These two papaya cell wall invertases have the GH32 domain, and the secretory peptide occupied the first 26 amino acids. *CpCWINV1* has 100 % homology in sequence with the published papaya invertase (Zhou et al. 2003; Zhu et al. 2003). The second cell wall invertase (*CpCWIN2*) had high homology to the *Arabidopsis* cell wall *INV1*. A possible third papaya acid invertase *CpCWINV3* has the required GH32 domain, but no signal peptide and no papaya ESTs are published. This acid invertase prediction showed moderate homology with coffee invertase and *Arabidopsis* cell wall invertase. Another predicted acid invertase had less than 20 % peptide sequence homology with Japan pear and *Arabidopsis* acid invertases.

Sucrose synthase (SUS) catalyzes the reversible conversion of sucrose and UDP to UDP-glucose and fructose and is only found in the cytoplasm (Martin et al. 1993; Baud et al. 2004). In *Arabidopsis*, five known sucrose synthases and another predicted sucrose synthase are found (Baud et al. 2004) with *AtSUS1* having 67–72 % homology to sucrose synthase genes from other species. Citrus is reported to have three *SUS* genes (Komatsu et al. 2002), and papaya had four predicted SUS genes (*CpSUS1* to *CpSUS4*) (Paull et al. 2008). All the predicted papaya SUS genes had sucrose synthase and glycosyltransferase (GT4) domains. The *CpSUS1* has 99 % homology with the reported partial sequence of *SUS* in papaya and 90 % homology with that of citrus and 85 and 87 % to *ArabidopsisAtSUS1* and *AtSUS3*, respectively (Paull et al. 2008). *CpSUS2* showed similarity with citrus *SUS2* 85 % homology to *AtSUS3*, while *CpSUS3* had high homology with onion (*E*-value 0.0) and tobacco (*E*-value 0.0). The fourth papaya sucrose gene (*CpSUS4*) had high homology with *AtSUS5*. Another SUS gene was predicted in papaya with a GT4 domain and sucrose synthase domains, but the sequence was not complete and only about half the expected peptide length.

Three sucrose phosphate synthase (SPS) genes were predicted in papaya (*CpSPS1*, *CpSPS2*, *CpSPS3*). All had GT4 and sucrose synthase domains (Paull et al. 2008). The *CpSPS1* and *CpSPS2* showed homology with the sucrose synthase in citrus and *Arabidopsis*. In higher plants, three families of sucrose phosphate synthase are described, and the functional differences between the families are not obvious (Langenkamper et al. 2002). Four putative *SPS* genes occur in *Arabidopsis*, two genes belong to Family A and one gene each belongs to Family B and Family C. The *CpSPS1* had 87 % homology to citrus *SPS* and 72 % to *Arabidopsis*'s *AtSPS1F*. The second papaya SPS, *CpSPS2*, was similar to *ArabidopsisAtSPS2F* with a homology of 77 and 80 % to the citrus *SPS*. The *AtSPS4F* gene in *Arabidopsis* had a homology of 73 % with the *CpSPS3* in papaya.

Papaya contains at least four hexose transporter genes (*CpHT1* to *CpHT4*). In contrast to papaya, tomato contains three named loci (*HT1*, *HT2*, and *HT3*) and seven putative loci (Gear et al. 2000; Dibley et al. 2005); *Arabidopsis* contains four named loci (*SGB1*, *TMT1*, *TMT2*, *TMT3*) and five putative loci (Büttner 2007). A fifth papaya hexose transporter was indicated, but the genomic sequence of the ORF is incomplete. Papaya ESTs are found for two of the predicted hexose transporters *CpHT2* and *CpHT3* (Paull et al. 2008). The number of introns varied widely; *CpHT1*, *CpHT2*, *CpHT3*, and *CpHT4* had 2, 12, 11, and 1, respectively. *CpHT3* had a 60 % chance of being localizing to the plastid. The remaining papaya hexose transporters did not show specific localization or export. Papaya contains at least three incomplete hexose transporter genes or pseudogenes that had only 7–11 of the 12 expected transmembrane domains while having domains for MFS and sugar transporter. One of the pseudogenes had four ESTs identified. Two of the three pseudogenes had secretory sequences for the endoplasmic reticulum.

Respiration

Plant respiration differs from most animal respiration in possessing a number of additional components including (a) the presence of an alternative oxidase to cytochrome c oxidase that is cyanide insensitive, (b) an internal rotenone-insensitive NAD(P)H non-proton-pumping dehydrogenase, and (c) an external NAD(P)H non-proton-pumping dehydrogenases (Vanlerberghe and McIntosh 1997; Mackenzie and McIntosh 1999; Elhafez et al. 2006). Alternate oxidase (AOX) is a terminal quinol oxidase that is non-proton pumping and transfers an electron to oxygen while dissipating the energy as heat (McDonald 2008). AOX is resistant to cyanide (Henry and Nyns 1975) and sensitive to salicylhydroxamic acid (SHAM) (Schonbaum et al. 1971). AOX activity is taxonomically widespread and found in all kingdoms (McDonald 2008). In dicots, two types of AOX, AOX1 and AOX2, are found. Monocots seem to possess only AOX1 (Considine et al. 2001) which has been associated with a number of physiological functions such as thermogenesis in sacred lotus during flowering (Watling et al. 2006). Other roles include balancing carbon metabolism and electron transport, as may happen during climacteric fruit

ripening (Theologis and Laties 1978; Considine et al. 2001), control of reactive oxygen species generation, O_2 scavenging, and resistance to toxins and pathogenicity (Moore et al. 2002; McDonald 2008). Though AOX1 and AOX2 are present in multigene families in most plants (McDonald 2008), only two genes for alternative oxidase AOX1 were predicted for papaya, *CpAOX1* and *CpAOX2*. The predicted *Cp*AOX1 protein had a predicted signal peptide of 22 amino acids, two transmembrane regions, and was supported by finding papaya ESTs. The two characteristic AOX conserved cysteine residues occurred at Cys 105 and Cys 155, and two introns similar to *Oryza sativa AOX* were present. The predicted *CpAOX1* gene has 79 and 71 % identity to the two *Arabidopsis* genes *Q9ZRT8* and *Q39219*, respectively. *CpAOX2* similarly has two transmembrane regions and no secretory sequence is predicted, though EST data suggested it is expressed. *CpAOX2* has 88 % identity to *Vigna unguiculata AOX* (Q93X12) and 82 % with *Arabidopsis* (O22049).

Plant mitochondria can oxidize NADH and NADPH without proton pumping (Melo et al. 2004; Geisler et al. 2007). These dehydrogenases are referred to as Type II NAD(P)H (Michalecka et al. 2003; Melo et al. 2004). Type II dehydrogenases operate in parallel to Type I proton-pumping multi-subunit complex I. Both types are found in the electron transfer chains of several bacteria and in fungal and plant mitochondria. Type II dehydrogenases are found on the external (NDB) and internal (NDA) faces of the inner mitochondrial membrane and transfer electrons from NAD(P)H to quinone. It has been proposed that there are four distinct types of NAD(P)H dehydrogenases: two on either side of the inner membrane. One was to oxidize NADH and the other for NADPH (Roberts et al. 1995; Melo et al. 2004). However, Rasmusson et al. (1999) reported two Type II dehydrogenases for potato located on inner and outer surfaces of the inner mitochondrial membrane. Only three open reading frames (ORF) with similarity to Type II dehydrogenase genes were found in the papaya genome. The ORF for an internal NAD(P)H dehydrogenase, *CpNDA1*, had 67 % identity to *ArabidopsisNDA2*. A papaya EST was found for *CpNDA1*. One *CpNDB* gene was also predicted, *CpNDB1*, that had 79 % identity to *ArabidopsisNDB4*. A mitochondrial transit peptide was predicted for the *CpNDB1* protein but not for *CpNDA1*. All had predicted ORFs motifs for the pyridine-redox superfamily.

Ethylene Synthesis

Arabidopsis and papaya show evolutionary similarity in the number of genes associated with ethylene synthesis, receptors, and response pathways. Papaya had four S-adenosyl-L-methionine synthases (SAMS), while 12 are reported for *Arabidopsis* (Yamagami et al. 2003) and 8 in tomato (Nakatsuka et al. 1998). Two of the papaya sequences from different regions of the papaya genome had the closest BLASTp homology to *Catharanthus roseus* SAM-2 (Paull et al. 2008). Unlike other SAMs, which do not have introns, one of these sequences appeared to have three. This may be a problem with the protein prediction, since corresponding gaps also appeared in

the amino acid alignment with *C. roseus* SAM-2. A tBLASTn search using the assembled nucleotide sequence showed no match between amino acids 97 and 102, plus one large region of low homology between amino acids 254–328; otherwise, there was very good homology with many SAMs genes.

The 11 tomato 1-aminocyclopropane-1-carboxylic acid synthase (ACS) (LeACS) genes differ in tissue-specific expression, developmental control, and ethylene induction (Nakatsuka et al. 1998; Alba et al. 2005; The Tomato Genome Consortium 2012). The papaya genome was predicted to contain seven ACS genes, each with aminotransferase I and II domains which were predicted to be an ACS domain (Paull et al. 2008). Two other sequences translatable to the aminotransferases I and II domains were found; however, the closest BLASTp matches to these two sequences was not annotated as ACS, and their homology to non-ACS aminotransferases was high and was not included in the papaya ACS list. Three sequences matched papaya sequences already deposited in GenBank, including *CpACS1* and *CpACS2* (98 and 98 % identity, respectively). All seven putative ACS sequences contained the conserved dodecapeptide binding site (Yip et al. 1990), although, one had an Asn substitution for the key Lys that bound either PLP or the 2-aminobutyrate of adenosylmethionine. The possible existence of nonfunctional sequences with ACS homology in the papaya genome would not be novel; *Arabidopsis* has at least one nonfunctional gene (*AtACS1*) shown to encode a protein with no enzyme activity (Tsuchisaka et al. 2007) that is nevertheless included in the gene family. In addition, *AtACS3* is not even transcribed, and *AtACS10* and *AtACS12* encode aminotransferase domains but have no ACS activity; none of these are included in the ACS family (Yamagami et al. 2003).

At a minimum, three ACO genes were predicted in papaya (Paull et al. 2008), with six additional sequences that showed partial homology to ACO-like genes or genes with the 2-oxoglutarate FeII oxygenase domain. This domain encompasses more than ACO's, but it is the motif that describes ACO's. As with the ACS, papaya has two well-studied *CpACO* genes, with multiple GenBank representatives. The *Cp*ACO genes identified (Paull et al. 2008) reflect BLASTp matches to *CpACO1* and *CpACO2* from *Carica papaya*, plus a translated supercontig with 79 % identity to a poplar protein shown to have ACO activity (Andersson-Gunneras et al. 2003). The sequence that matched *CpACO1* lacks the first 83 amino acids from *ACO1*, due to sequencing ambiguities, but otherwise matched the remaining 235 amino acids. All three sequences retained the three conserved subdomains of ACOs and 2-oxoglutarate FeII oxygenases (Trentmann and Kende 1995), as well as the HxD site for enzymatic activity. *CpACO*s share the motif KxxR within the conserved C-terminal site that was identified as essential for *Petunia hybrida ACO1* enzyme activity (Yoo et al. 2006). The protein encoded on *CpACO3* differs at this site by three amino acids, which is more than any of the ACOs from 24 different plants, including papaya (**LP**KEPRF**R** versus the consensus QAKEPRFE). Five other predicted models had between 42 and 71 % homology to *Arabidopsis* proteins with 2-oxoglutarate FeII oxygenase domains, but these *Arabidopsis* proteins are not clearly involved in ACO activity.

Table 14.5 Predicted papaya ethylene receptors, structure, and phylogenetic relationship with known *Arabidopsis* ethylene receptors

	CDS	Intron	aa	Homology	ESTs
CpETR1	25,248	4	629	*Carica papaya* AAG41977	9
CpETR2	56,390	5	738	*Prunus persica* ethylene receptor Q9M7M1 ETR1	4
CpETR3	5,560	9	767	*Lycopersicon esculentum* ethylene receptor homolog AF118844_1	8

The whole genome shotgun sequence accession number in NCBI (WGS Accession), coding sequence for amino acids (CDS), the number of introns in the nucleotide sequence (introns), amino acids (aa), presence of discrete portion of protein possessing its own function and the superfamily (domain), similarity in peptide sequence with another species (homology), extent the sequences were invariant (identity), *E*-value was the expectation value for homology, and the number of papaya expressed sequence tags (EST) detected of at least 500 bases and 99 % identity [reproduced with kind permission from Springer Science+Business Media from Paull et al. (2008)]

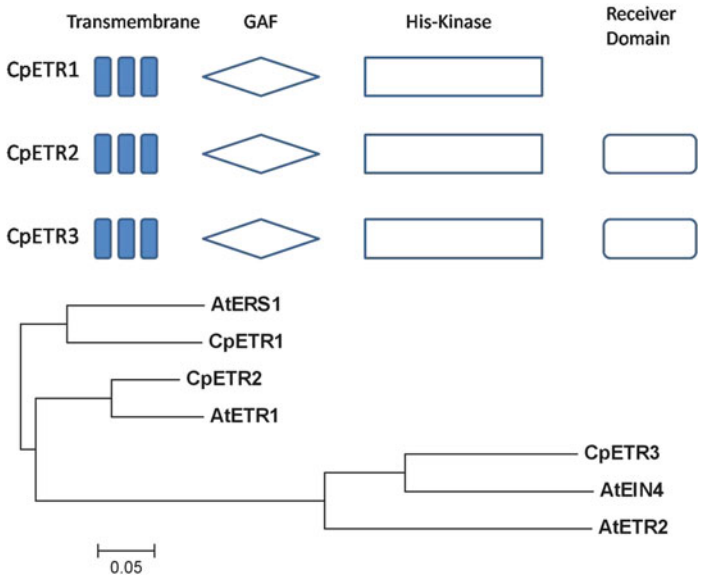

Fig. 14.1 Predicted papaya ethylene receptors, structure and phylogenetic relationship with known *Arabidopsis* ethylene receptors. The whole genome shotgun sequence accession number in NCBI (WGS Accession), coding sequence for amino acids (CDS), the number of introns in the nucleotide sequence (introns), amino acids (aa), presence of discrete portion of protein possessing its own function and the superfamily (domain), similarity in peptide sequence with another species (homology), extent the sequences were invariant (identity), *E*-value was the expectation value for homology, and the number of papaya expressed sequence tags (EST) detected of at least 500 bases and 99 % identity [reproduced with kind permission from Springer Science+Business Media from Paull et al. (2008)]

Ethylene Response

Papaya ripening has both ethylene insensitive and ethylene sensitive components (Chen and Paull 2003; Manenoi and Paull 2007) as has been found for tomato (Picton et al. 1993). At least three ethylene receptor genes (*CpETR1* to *CpETR3*) were predicted in the papaya genome (Table 14.5 and Fig. 14.1). All three have the expected

protein domains, though *CpETR1* had no receiver domain (Paull et al. 2008). *Cp*ETR1 was homologous (100 %) to the papaya receptor cDNA sequence deposited by the late Dr. Hamid Lazan and his group (Che Husin et al. 2000) from Malaysia. This sequence has 78 % homology with the *Arabidopsis ERS1* (Paull et al. 2008). *Arabidopsis ETR1* has the greatest homology with *Cp*ETR2 (86 %). Based upon gene and protein structure, both *CpETR1* and *CpETR2* would fall into the ethylene receptor subfamily 1 (Hua et al. 1998) with *CpETR3* being in subfamily 2 (Table 14.5 and Fig. 14.1). Subfamily 1 receptors maybe required for most ethylene responses (Wang et al. 2003). Originally, we reported four two-component ethylene receptors (Ming et al. 2008); however, a more thorough analysis indicated that one of the previously reported *Cp*ETRs lacked the expected GAF domain and a full histidine domain (Wang et al. 2003; Hall et al. 2007). This partial sequence is a 139 aa-long peptide with 75 % homology to the apple ethylene receptor. A fifth potential gene was short (224 aa) and only had a DUF623 domain of unknown function and low homology to the nearest ethylene receptor of *Litchi chinensis*. Papaya's three predicted ethylene receptors were fewer than the five ETRs found in *Arabidopsis* (Schaller and Kieber 2002) and the six in tomato (Klee and Tieman 2002). A seventh pseudogene has been reported in tomato with no expression data (The Tomato Genome Consortium 2012). Since ethylene receptors seem to be a negative regulator of action (Hua and Meyerowitz 1998), degradation plays a significant role in the control of ripening (Kevany et al. 2007). The relatively small complement of papaya ethylene receptors means that fewer interactions may occur between receptors so that their role in ripening should be more easily discerned.

Ethylene receptors are disulfide-linked dimers and ethylene binding involves a copper cofactor (Rodríguez et al. 1999). The gene RAN (response to antagonist) plays a role in making the ethylene receptor apoprotein functional by transporting a copper ion to the ethylene receptor site in the membrane-spanning regions. As with *Arabidopsis*, only one RAN homolog was found in papaya (Paull et al. 2008), and *CpRAN* had 76 % homology to *Arabidopsis RAN1*.

The ethylene response pathway after ethylene binding involves a RAF-related kinase *CTR1* that has a role in the negative response by forming a complex with the receptor (Hua and Meyerowitz 1998). Ethylene binding inhibits CTR1 kinase activity (Ouaked et al. 2003) and thus relieves the repression of the ethylene response pathway (Alonso et al. 2003; Alonso and Stepanova 2004). A signal is then transmitted from the positive regulator EIN2 to EIN3/EILs and induces transcription of ethylene response factors (ERF). *At*CTR1 is one of the six *Arabidopsis* MAPK kinases and three *Le*CTR1-like genes in tomato (Frye et al. 2001; Adams-Phillips et al. 2004a; The Tomato Genome Consortium 2012) with no evidence that more than one (*At*CTR1, *Le*CTR1) being involved in the ethylene signal transduction pathway. At least one *CpCTR1* gene was found in the papaya genome that had 56 % homology with *LeCTR1*. A second, incomplete CTR sequence was also found (Paull et al. 2008). One possible *CpEIN2* gene was found; although the nucleotide and amino acid sequence were short (Paull et al. 2008), The alignments with CTR1-like kinase and the protein kinase family had fewer gaps, and about one hundred amino acid matches. Four possible EIN3/EIL1 genes were also found, with two (*CpEIL1*, *CpEIN3*) having high EIN2/EIL homologies to published sequences, and these were expressed as papaya ESTs. *Arabidopsis* has nine EIN3 and EIN3-like genes (Binder et al. 2007), and tomato has at least five (Stepanova and Alonso 2005).

Eight ERF genes were found with high homology, and another 12 predicted ERF-like sequences that had low homology matches or had no significant protein domains. One hundred and twenty-two ethylene-responsive binding factors in the AP2/ERF superfamily have been found in *Arabidopsis* (Nakano et al. 2006). *AtERF1* falls into group IX with 17 members and 8 in group VIII. Using the conserved 60 amino acid sequences for AP2/ERF (Nakano et al. 2006), an additional 92 potential genes with this sequence were found in papaya which suggests more ERF may be present in the papaya genome.

Laticifers and Protease

The papaya fruit, as well as other aerial parts of the plant, has a dense network of laticifers (Roth and Clausnitzer 1972). The articulated laticifers begin as a column of cells that form vessel-like structures that retain the usual organelles. The milky latex is about 85 % water with the remaining 15 % being composed of 25 % insoluble matter of unknown composition. The soluble fraction contains carbohydrates (~10 %), salts (~10 %), lipids (~5 %), and biomolecules mainly proteins (~40 %) (El Moussaoui et al. 2001). The biomolecules include cysteine proteinases (papain, chymopapain), cystatin, β-1,3-glucanase, chitinase, lipases, and other proteins. The latex is thought to function as an induced defense mechanism (Salas et al. 2008; Shindo and Van Der Hoorn 2008). The latex is harvested from large fruited green varieties by scratching the skin, allowing the exuded latex to coagulate and dry, then collecting the dried latex by scraping. The latex is purified and its proteolytic activity is used for meat tenderization and chillproofing of beer (El Moussaoui et al. 2001). During fruit ripening some of the laticifers break and release latex under the cuticle disrupting this barrier and increasing fruit water loss rate (Paull and Chen 1989).

Four cysteine proteinases account for 80 % of the enzyme fraction (El Moussaoui et al. 2001). The proteinases are papain, chymopapain, caricain (proteinase omega), and glycyl endopeptidase (proteinase IV). Papaya proteinases are synthesized as proenzymes with a signal sequence. The prosequence is cleaved and proteinase activated. All four proteinases are members of the peptidase C1A subfamily of cysteine proteinases. Papain has been most intensively studied (El Moussaoui et al. 2001) although it is a minor component (5–8 %) of the endopeptidases in papaya latex (Baines and Brocklehurst 1982; Azarkan et al. 2003). One gene for a propapain (*CpPAPA*) precursor was predicted. The encoded protein had 26 amino acids ER secretory sequence and 95 % identity to the published papaya cDNA sequence (Paull et al. 2008). A papaya EST was found for *CpPAPA*.

A single papaya chymopapain gene (*CpCHYP*) was predicted that had 99 % identity to chymopapain isoforms I, III, and V and 100 % to chymopapain isoforms II and IV (Paull et al. 2008). Chymopapain is distinguished from papain by the proteolytic activity remaining after papain is removed (Jansen and Balls 1941). The five prochymopapain isoforms (I–V) were identified from sequenced leaf cDNAs (Taylor et al. 1999); all the isoforms have a free cysteine at position 251. The translated isoform cDNAs differ in one or two amino acid substitutions. At position 222, a cysteine is

replaced by a tyrosine in isoforms III and V, while at position 266, valine is replaced by phenylalanine in isoforms II, III, and V. These single amino acid substitutions require only a single base change. The predicted *CpCHYP* protein had the nine amino acids reported for isoforms III–V. It is possible that the chymopapain isoforms differ because of errors in PCR, sequencing, or translation. The *Cp*CHY protein, possibly isoform III, had 66 ESTs. A chymopapain isoform III EST was also reported by Devitt et al. (2006).

A number of other proteinases have been reported in papaya latex (El Moussaoui et al. 2001; Azarkan et al. 2006; Shindo and Van Der Hoorn 2008). Two caricain genes were predicted (proteinase omega) and one glycyl endopeptidase gene (Paull et al. 2008). The predicted gene numbers are in agreement with those reported (El Moussaoui et al. 2001), though one of the caricain gene cDNAs has not been deposited in GenBank. The *CpCARP1* was 100 % identical to the described sequence, although it is only 130 amino acids long due to incomplete sequencing of the WGS. A glutamine cyclotransferase (glutamine cyclase) gene was predicted in papaya (*CpGTR*) as previously reported from papaya latex (Oberg et al. 1998; El Moussaoui et al. 2001). The predicted papaya gene (Paull et al. 2008) was longer than the predicted 288 amino acid peptide from the cDNA in GenBank though analysis suggested intron boundaries may have not been accurately predicted. One cystatin (*CpCYST*) and two Kunitz-type trypsin inhibitors (*CpTINH1*, *CpTINH2*) were predicted in the papaya genome. The cystatin had the I25A domain and was classified as member of the CY superfamily. Papaya ESTs were found for both the cystatin and trypsin inhibitors (Paull et al. 2008).

At least 27 β-1,3-glucanase proteins (GH17) were predicted in papaya, and activity has been detected in papaya latex (El Moussaoui et al. 2001). In addition to β-1,3-glucanase, chitinase II (GH19) has been reported for papaya latex. A gene with 61 % identity to the partial cDNA (P81241) was found in the papaya genome (Paull et al. 2008).

Other than the genes encoding proteinases and those laticifer proteins described previously, the genes encoding other proteins potentially found in papaya latex have not been well described. A search of the papaya genome for latex-associated protein genes found six genes including one for caspase reported as a rubber latex-abundant protein and lysophospholipase. In addition, a number of potential latex genes were found for which predicted protein had allergen domains. Othman and Nuraziyan (2010) reported the specific expression of subtilase in papaya mesocarp that reached the highest level at ripening stage. Recently, there was a report on a new phospholipase D (CpPLD1) in papaya latex (Abdelkafi et al. 2012). The sequence alignment showed four conserved regions (I–IV) defined by most of the PLD superfamily.

Volatile Production

All plant parts emit volatiles, which have multiple functions. Of the many roles in plants, ecological interactions such as defense against herbivores, pathogens, and as attractants for animals to disperse pollen and seeds are presumed to be principal

functions (Pichersky and Gershenzon 2002; Phillips et al. 2006). Additionally, volatiles as components of flavor and aroma perception contribute substantially to the utility of plant parts as human foodstuffs. Odorants are volatile and interact with human olfactory reception (Buck and Axel 1991), but they must be somewhat lipophilic as well as water soluble, possess adequate vapor pressure, and occur in sufficient concentration for olfactory receptor interaction. At least 166 compounds have been identified in papaya fruit volatiles (Flath and Forrey 1977; Flath et al. 1990; MacLeod and Pieris 1983; Pino et al. 2003). The most commonly identified papaya volatiles have been methyl butanoate, ethyl butanoate, 3-methyl-1-butanol, and 1-butanol. The esters of lower fatty acids are considered to contribute much to the typical papaya flavor (Pino et al. 2003). The amounts and relative content of volatiles have been shown to vary with the stage of ripeness (Katague and Kirch 1965; Flath et al. 1990), for example, linalool production increases nearly 400-fold with only a sevenfold increase in benzyl isothiocyanate during ripening (Flath et al. 1990). 2-Ethyl-1-hexanol is found specifically in green fruit, while ethyl octanoate is found only in fully ripe fruit (Fuggate et al. 2010). When the fruit has reached the edible ripe stage, butanol, 3-methylbutanol, benzyl alcohol, and α-terpineol are at their maximum concentrations (Almora et al. 2004). One of the two volatiles that most closely resembled those of papayas is linalool, a product of the plastid synthesis pathway. The other papaya volatile is methyl benzoate, described as having papaya qualities on odor assessment (MacLeod and Pieris 1983). The sweaty odor quality of some papaya cultivars is due probably to the production of methyl butanoate (MacLeod and Pieris 1983). The Hawaii variety had very little methyl butanoate (0.06 %) (Flath and Forrey 1977), while the Sri Lankan variety had 48.3 % (MacLeod and Pieris 1983). In addition, benzyl isothiocyanate contributes a pungent off-odor. The amount of each volatile component varies both with cultivar and locality (MacLeod and Pieris 1983; Franco and Rodriguez 1993). No difference in volatile profiles is found in papaya fruit during on-tree and postharvest ripening (Fuggate et al. 2010).

Most volatile esters are described as fruity (Burdock 2002), and the most likely precursors are amino acids and lipids. The precursors of volatile esters, with branched alkyl chain, are valine, isoleucine, and other amino acids from threonine (Yabumoto et al. 1977; Newcomb et al. 2006). Aliphatic esters and alcohols are produced from free fatty acids such as linoleic and linolenic (Baldwin et al. 2000). The pathways come together in the formation of aldehydes that are reduced to alcohols by alcohol dehydrogenase (Speirs et al. 1998) and ester formation by alcohol acyltransferase (Fellman et al. 2000). The mesocarp alcohol dehydrogenase activity mesocarp increases dramatically during the early ripening stages, while alcohol acetyltransferase is active throughout ripening (Fuggate et al. 2010).

Single genes were found in the papaya genome for all the enzymes in the biosynthesis from threonine to 3-keto-3-methylvalerate (Paull et al. 2008). At least three branched chain amino transferase genes were found that could convert 3-keto-3-methylvalerate via isoleucine to 3-oxo-methylpentanoic acid. The next step to 2-methylbutanal involves pyruvate decarboxylase, and three genes encoding this enzyme were predicted from the genome data; all were predicted to have

the characteristic domains. Four branched chain alpha-ketoacid dehydrogenases (alcohol dehydrogenases) were predicted with three that had papaya EST. The interconversion between 3-methyl butanol and 2-methylbutyl acetate possibly involves a single predicted alcohol acyltransferase gene and four predicted carboxylesterases (Paull et al. 2008).

Straight chain ester biosynthesis commonly occurs from fatty acids with linoleic acid being a common precursor. The first enzyme is lipoxygenase of which nine were predicted in papaya with another possible partial sequence found (Paull et al. 2008). Except for two of the predicted lipoxygenase, all had papaya ESTs. The predicted gene for the previously described papaya lipoxygenase (*CpLOX1*) had eight ESTs and a chloroplast transit peptide. The other *CpLOX*s did not have predicted chloroplast transit peptides, suggesting they may be located in the cytoplasm (Paull et al. 2008).

One papaya hydroperoxide lyase gene (P450 superfamily member) was found and no isomerase was predicted (Paull et al. 2008). Hex-3-enal to hex-2-enal isomerase has been reported to be unstable and no sequences are in the gene databases. The dehydrogenation of hex-3-enal to hex-3-enoic acid by aldehyde dehydrogenase could involve upwards of five genes. Two of the aldehydrogenases have predicted mitochondrial transit peptides and may be involved in volatile production. A number of alcohol dehydrogenase genes were predicted that could convert hex-3-enal to hex-3-enol (Paull et al. 2008).

Two volatiles with fruity/floral properties are 2-phenylacetaldehyde and 2-phenylethanol derived from phenylalanine. The first enzyme in the pathway is a bifunctional phenylacetaldehyde synthase possibly combining decarboxylation-amine oxidation leading to phenylacetaldehyde (Kaminaga et al. 2007). The last step is the reduction of 2-phenylacetaldehyde to 2-phenylethanol. The reduction involves NADPH and 2-phenylacetaldehyde reductase (Tieman et al. 2007). Genes were predicted in papaya for both steps in biosynthesis (Paull et al. 2008). Two predicted phenylacetaldehyde synthase (*CpPALDS1*, *CpPALDS2*) genes were found having high homology to the deposited gene sequences for rose and petunia (Paull et al. 2008). A single gene with moderate homology (48 %) to the tomato gene was predicted for the reduction to 2-phenylethanol (Tieman et al. 2007).

Terpenoids, widely distributed in lichens, algae, and higher plants, represent the most diverse family of natural products, as there are over 40,000 different structures identified thus far. Many terpenoids are nonvolatile and constitute important elements of plant functional processes, i.e., plant growth regulation, redox reactions, membrane structure, and photosynthesis (Bohlmann et al. 1998; Croteau et al. 2000; Aubourg et al. 2002). Volatile and flavor compounds from essential oils of various terpene classes are produced in leaves and fruit, while carotenoids give the fruit skin and flesh its unique color. Biosynthesis occurs through two different pathways: the acetate-mevalonate pathway in the cytoplasm (Bohlmann et al. 1998) and the non-mevalonate pathway in the plastids (Lichtenthaler 1999; 2000). The mevalonate pathway leads to sesquiterpenes and triterpenes (sterols), while monoterpenes, diterpenes, tetraterpenes (carotenoids), and polyterpenes are formed in the plastids.

Genes were predicted for most of the steps in terpene biosynthesis occurring both in the plastids and in the cytoplasm (Paull et al. 2008). Two 1-deoxy-D-xylulose-5-phosphate (DXP) synthases were predicted with the shorter predicted peptide lacking a chloroplast transit peptide. Another partial DXP synthase, one-third the length of the expected peptide, was found with 78 % homology to the rubber TPP enzyme domain. A full-length linalool synthase gene was predicted with two partial sequences. No geranylgeranyl pyrophosphate (GGPP) synthase was found that had a chloroplast transit peptide. A short sequence was found that has low homology (61 %, 105 aa) to GGPP synthase that catalyzes the farnesyl diphosphate to GGPP conversion.

Cytoplasmic terpene synthesis from mevalonate to dimethylally diphosphate has a single gene for each step (Paull et al. 2008). Two geranyl diphosphate genes were predicted with significant homology to published sequences. Two full-length squalene synthases were predicted along with two partial sequences that had high homology to a tomato homolog and *Panax quinquefolius* sequence.

The cleavage of carotenoids derived from the terpene synthesis pathway can lead to the production of volatiles. Carotenoid cleavage can occur at any conjugated double bonds by various enzymes such as CCD and NCEDs (discussed previously in the Fruit Growth section) to form an aldehyde or ketone in each product. The products include abscisic acid, strigolactone, and various aroma compounds (Vogel et al. 2008). While *Arabidopsis* contains nine CCD, five of which are directly involved in abscisic acid synthesis, the other CCDs lead to volatiles and nonvolatile apocarotenoids (Simkin et al. 2004), with one CCD able to cleave multiple substrates at different positions (Vogel et al. 2008). Only one *Cp*CCD was predicted for papaya and another partial CCD ORF was found. Tomato contains two closely related CCDs, *LeCCD1A* and *LeCCD1B*, that generate the aldehydes and ketones including geranyl acetone, pseudoionone, and β-ionone (Simkin et al. 2004). *CpCCD* may serve a similar role in papaya ripening and generate ketone volatiles such as 6-methyl-5-hepten-2-one in tomatoes (Vogel et al. 2008) and as detected in papaya fruit volatiles (Flath et al. 1990).

Plant Growth Regulator Related Genes and Fruit Ripening

A list of papaya homologs to functionally characterized *Arabidopsis* plant growth regulator (PGR) synthesis, reception, response, and degradation genes (Quecini et al. 2007; Michael et al. 2008; Peng et al. 2009) was generated using BLASTp. Papaya has fewer genes in most PGR categories, and changes in ~150 microarray probes have been evaluated during ripening (Ming et al. 2012). Two of three ethylene receptors are upregulated, while the receptor previously reported to be expressed in papaya mesocarp (Che Husin et al. 2000) was not expressed in either of the ripening stages tested. One ACC synthase gene's expression declined and another increased threefold. Most auxin-related genes declined with upregulation of auxin response factors (ARF8), three SAURs, and two auxin-binding proteins. GA-related genes were mostly downregulated. Two cytokinin oxidases and a zeatin 4-glucosyltransferase were upregulated.

Most brassinosteroid-related genes were downregulated. Genes associated with ABA showed little change. These results indicate that the changes in PGR during ripening show varying response and potential positive and negative interactions (Wu and Paull, 2012, unpublished data).

Fruit Ripening and Evolution

Fleshy fruit ripening is a complex gene-controlled process, involving molecular, physiological, and biochemical events that require developmental and temporal coordination and regulation (Fig. 14.2). Fruit ripening changes in gene clusters such as for cell wall separation and degradation, carotenoid synthesis, chlorophyll breakdown, and altered organic acid metabolism affect texture, color, taste, and aroma (Giovannoni 2001, 2004; Alexander and Grierson 2002; Kevany et al. 2007). Although our understanding of ripening processes has progressed significantly over the last 30 years, the initiation and coordination of ripening are only partially understood. Variation between species and cultivars does occur in coordination, timing, and regulation of ripening, but similar networks of regulatory genes are expected to be expressed.

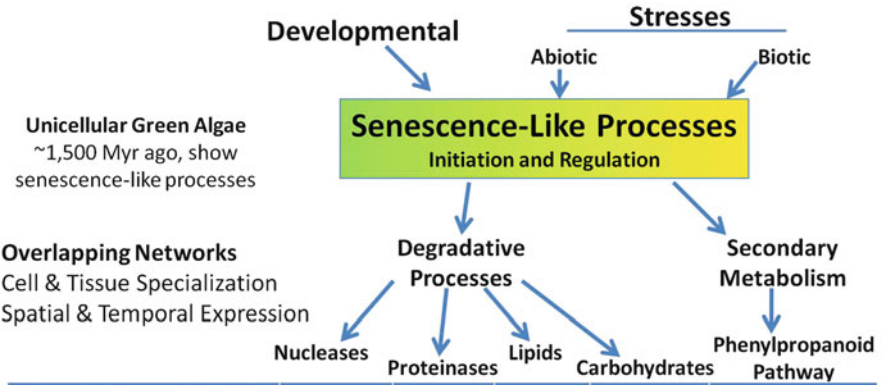

Fig. 14.2 Similarities between senescence, abscission, and fruit ripening in the networks of genes involved in molecular and biological processes (modules) in ultimate developmental stages that lead to final death of the organ. Papaya is a unique model system in which all processes can be sampled and gene expression determined

The premise that genes expressed in different species during fruit ripening may be conserved and regulated in similar ways (Adams-Phillips et al. 2004b; Brummell et al. 2004; Barry and Giovannoni 2007; Seymour et al. 2008) is strengthened by the similarities in the ripening processes in fleshy fruit originating from different tissues. The edible fruit tissues vary widely, from the floral receptacle (strawberry) to fused clustered flowers (pineapple), the entire pericarp (tomato), the exocarp and mesocarp (apple, peach), the mesocarp only (papaya), the endocarp (citrus), and aril tissue arising from the funicle (litchi, durian) or the seed coat (rambutan). Dehiscence does not occur in most fleshy fruit but is common in arillate fruits (akee, durian). In durian, similar cell wall changes and biochemical activities occur in the thick woody pericarp as occurs in ripening and dry fruit dehiscence but not in the non-climacteric edible aril. The genes associated with softening of the *Arabidopsis* valve margin during dehiscence occur in abscission and are similar to the softening that occurs in fleshy fruit ripening (Roberts et al. 2000; Seymour and Manning 2002; Rose et al. 2004; Brummell 2006). Fleshy fruit expansion is probably not directly associated with ripening since many fruits are not fleshy and undergo ripening.

A key unanswered question in plant biology is how fleshy fruit ripening and associated softening evolved seemingly independently throughout the higher plant families. Fleshy fruits, with their mutually beneficial interaction of providing nutrition to animals and improved seed dispersal, have arisen independently in different families, have disappeared and reappeared, are not evolutionarily conserved, and show no clear association with phylogeny (Knapp 2002; Givnish et al. 2005; Lorts et al. 2008). This suggests few constraints on fruit evolution from the ancestral dry follicle (Roth 1977; Knapp 2002) with only limited changes needed in the interconnected networks of clustered genes. Genome-wide changes that would have led to fleshy fruit in connected branches of the angiosperm phylogenetic tree did not occur. Either developmental processes or modules are recruited and adapted to new functions with the original transcription factors and hormones remaining in place. Alternatively, transcription factors and hormones have been recruited by preexisting developmental processes. A similar process has apparently occurred in gymnosperms; Lovisetto et al. (2012) showed that MADbox genes are involved in the formation of fruit-like structure and suggested that the same gene types have been recruited in phylogenetically distant species to make fleshy structures that also have different anatomical origins. The regulation and gene expression that occurs in fleshy fruit ripening is more likely associated with the recruiting of a more ancient developmental process (e.g., senescence) (Thomas et al. 2009). This recruitment could have occurred via adaptation to new functions with new interactions developed between the networks regulating the spatial-temporal expression of the various regulatory modules controlling senescence-like processes (Fig. 14.1).

Prospects

The genome of papaya (372 Mbp) is three times larger than that of *Arabidopsis* (125 Mbp) and two-fifths the size of the tomato genome (The Tomato Genome Consortium 2012). Papaya is predicted to code for 24,746 genes, which is 10–20 %

fewer than in *Arabidopsis* (Ming et al. 2008, 2012) and 35 % fewer than the estimated 31,760 genes in tomato (The Tomato Genome Consortium 2012). Unlike the tomato genome, the papaya genome is largely euchromatic. The lower gene number in papaya may reflect that, unlike *Arabidopsis*, the papaya genome has not undergone recent whole-genome duplication and has had fewer opportunities for fractionation of linkage arrangements by post-duplication gene loss. This suggests that gene structure, function, and arrangement in papaya may resemble the ancestral angiosperm more closely than do previously sequenced angiosperms. Comparison of the five sequenced clade members suggests a minimal angiosperm gene set of 13,311. Fewer introns have been found in papaya genes than in their orthologs in *Arabidopsis* (8,319 versus 11,718) and rice (6,874 versus 7,011). Approximately 2,000 transcription factors in over 60 families have been identified in *Arabidopsis*, with papaya having fewer members in the majority of families, but a significantly higher number of predicted MADS-box proteins (171 versus 141). The comparison of gene expression and activity during fruit ripening suggests that the same regulatory pathways and similar controls occur. The lack of whole genome duplication and reductions in most papaya gene families and biosynthetic pathways (Ming et al. 2012) makes papaya a valuable tool for the study of the complex regulatory networks active in fruit ripening. For example, fewer gene numbers for ethylene synthesis, receptors, and in the ethylene response pathways may provide insights as to how ripening is controlled.

Many major questions remain about the regulation and systems involved in leaf senescence and abscission and fruit ripening. What are the primary cues for the initiation of these processes? What regulators or gene clusters exert the earliest control over signal perception and response networks? Why do some organs senescence, abscise, and ripen in response to a given stimulus whereas others do not? What genes control ethylene-independent and ethylene-dependent pathways, and how is the process regulated and how do the network pathways interact? Global expression analysis and functional genomics can provide the foundation to addressing these system network questions.

References

Abdelkafi S, Abousalham A, Fendri I, Ogata H, Barouh N, Fouquet B, Scheirlinckx F, Villeneuve P, Carrière F (2012) Identification of a new phospholipase D in *Carica papaya* latex. Gene 499:243–249

Adams-Phillips L, Barry C, Kannan P, Lecercq J, Bouzayen M, Giovannoni JJ (2004a) Ctr1-mediated ethylene signal transduction in tomato is encoded by a multigene family whose members display distinct regulatory features. Plant Mol Biol 54:387–404

Adams-Phillips L, Barry C, Giovannoni J (2004b) Signal transduction systems regulating fruit ripening. Trends Plant Sci 9:331–338

Alba R, Payton P, Fei Z, McQuinn R, Debbie P, Martin GB, Tanksley SD, Giovannoni JJ (2005) Transcriptome and selected metabolite analysis reveal multiple points of ethylene control during tomato fruit development. Plant Cell 17:2954–2965

Alexander L, Grierson D (2002) Ethylene biosynthesis and action in tomato: a model for climacteric fruit ripening. J Exp Bot 53:2039–2055

Alferez F, Zacarias L (1999) Interaction between ethylene and abscisic acid in the regulation of citrus fruit maturation. In: Kanellis AK, Chang C, Klee H, Blecker AB, Pech JC, Grierson D (eds) Biology and biotechnology of the plant hormone ethylene. II. Kluwer Academic, Amsterdam, pp 183–184

Almora K, Pino JA, Hernandez M, Duarte C, Gonzalez J, Roncal E (2004) Evaluation of volatiles from ripening papaya (*Carica papaya* L., var. Maradol roja). Food Chem 86:127–130

Alonso JM, Stepanova AN (2004) The ethylene signaling pathway. Science 306:1513–1515

Alonso JM, Stepanova AN, Solano R, Wisman E, Ferrari S, Ausubel FM, Ecker JR (2003) Five components of the ethylene-response pathway identified in a screen for weak ethylene-insensitive mutants in *Arabidopsis*. Proc Natl Acad Sci USA 100:2992–2997

Andersson-Gunneras S, Hellgren JM, Bjorklund S, Regan S, Moritz T, Sundberg B (2003) Asymmetric expression of a poplar ACC oxidase controls ethylene production during gravitational induction of tension wood. Plant J 34:339–349

Armstrong GA, Hearst JE (1996) Carotenoids 2: genetics and molecular biology of carotenoid pigment biosynthesis. FASEB J 10:228–237

Asha VA, Sane AP, Sane PN (2007) Multiple forms of α-expansin genes are expressed during banana fruit ripening and development. Postharvest Biol Technol 45:184–192

Aubourg S, Lecharny A, Bohlmann J (2002) Genomic analysis of the terpenoid synthase (AtTPS) gene family of *Arabidopsis thaliana*. Mol Genet Genomics 267:730–745

Aubry S, Mani J, Hörtensteiner S (2008) Stay-green protein, defective in Mendel's green cotyledon mutant, acts independent and upstream of pheophorbide a oxygenase in the chlorophyll catabolic pathway. Plant Mol Biol 67:2443–2456

Auldridge ME, Block A, Vogel JT, Dabney-Smith C, Mila I, Bouzayen M, Magallanes-Lundback M, DellaPenna D, McCarty DR, Klee HJ (2006a) Characterization of three members of the *Arabidopsis* carotenoid cleavage dioxygenase family demonstrates the divergent roles of this multifunctional enzyme family. Plant J 45:982–993

Auldridge M, McCarty DR, Klee HJ (2006b) Plant carotenoid cleavage oxygenases and their apocarotenoid products. Curr Opin Plant Biol 9:315–321

Azarkan M, El Moussaoui A, van Wuytswinkel D, Dehon G, Looze Y (2003) Fractionation and purification of the enzymes stored in the latex of *Carica papaya*. J Chromatogr B 790: 229–238

Azarkan M, Dibiani R, Goormaghtigh E, Raussens V, Baeyens-Volant D (2006) The papaya Kunitz-type trypsin inhibitor is a highly stable β-sheet glycoprotein. Biochim Biophys Acta Proteins Proteomics 1764:1063–1072

Baines BS, Brocklehurst K (1982) Characterization of papaya peptidase A as a cysteine proteinase of *Carica papaya* L. with active center properties that differ from those of papain by using 2,2′-dipyridyl disulfide and 4-chloro-7-nitrobenzofurazan as reactivity probes. Biochem J 205:205–211

Baldwin EA, Scott JW, Shewmaker CK, Schuch W (2000) Flavor trivia and tomato aroma: biochemistry and possible mechanisms for control of important aroma components. Hortscience 35:1013–1022

Barry CS (2009) The stay-green revolution: recent progress in deciphering the mechanisms of chlorophyll degradation in higher plants. Plant Sci 176:325–333

Barry C, Giovannoni JJ (2007) Ethylene and fruit ripening. J Plant Growth Regul 26:143–159

Barry CS, Llop-Tous MI, Grierson D (2000) The regulation of 1-aminocyclopropane-1-carboxylic acid synthase gene expression during the transition from system-1 to system-2 ethylene synthesis in tomato. Plant Physiol 123:979–986

Bartley GE, Ishida BK (2003) Developmental gene regulation during tomato fruit ripening and in-vitro sepal morphogenesis. BMC Plant Biol 3, 4. http://www.biomedcentral.com/1471-2229/3/4

Baud S, Vaultier MN, Rochat C (2004) Structure and expression profile of the sucrose synthase multigene family in *Arabidopsis*. J Exp Bot 55:397–409

Biale JB (1964) Growth, maturation, and senescence in fruits. Science 146:880–888

Binder BM, Walker JM, Gagne JM, Emborg TJ, Hemmann G, Bleecker AB, Vierstra RD (2007) The *Arabidopsis* EIN3 binding F-Box proteins EBF1 and EBF2 have distinct but overlapping roles in ethylene signaling. Plant Cell 19:509–523

Blas AL, Ming R, Liu Z, Veatch OJ, Paull RE, Moore PH, Yu Q (2010) Cloning of papaya chromoplast specific lycopene â-cyclase, CpCYC-b, controlling fruit flesh color reveals conserved microsynteny and a recombination hotspot. Plant Physiol 152:2013–2022

Blas AL, Yu Q, Veatch OJ, Paull RE, Moore PH, Ming R (2012) Genetic mapping of quantitative trait loci controlling fruit size and shape in papaya. Mol Biol 29:457–466

Bocock PN, Morse AM, Dervinis C, Davis JM (2008) Evolution and diversity of invertase genes in *Populus trichocarpa*. Planta 227:565–576

Bohlmann J, Meyer-Gauen G, Croteau R (1998) Plant terpenoid synthases: molecular biology and phylogenetic analysis. Proc Natl Acad Sci USA 95:4126–4133

Booker J, Auldridge ME, Wills S, McCarty DR, Klee HJ, Leyser O (2004) MAX3/CCD7 is a carotenoid cleavage dioxygenase required for the synthesis of a novel plant signaling molecule. Curr Biol 14:1–20

Bramley P, Teulieres C, Blain I, Bird C, Schuch W (1992) Biochemical characterization of transgenic tomato plants in which carotenoid synthesis has been inhibited through the expression of antisense RNA to pTOM5. Plant J 2:343–349

Bruinsma J, Paull RE (1984) Respiration during postharvest development of soursop fruit. *Annona muricata* L. Plant Physiol 76:131–138

Brummell DA (2006) Cell wall disassembly in ripening fruit. Funct Plant Biol 33:103–119

Brummell DA, Harpster MH, Civello PM, Palys JM, Bennett AB, Dunsmuir P (1999) Modification of expansin protein abundance in tomato fruit alters softening and cell wall polymer metabolism during ripening. Plant Cell 11:2203–2216

Brummell DA, Dal Cin V, Crisosto CH, Labavitch JM (2004) Cell wall metabolism during maturation, ripening and senescence of peach fruit. J Exp Bot 55:2029–2039

Buck L, Axel R (1991) A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. Cell 65:175–187

Burdock GA (2002) Fenaroli's handbook of flavor ingredients, 4th edn. CRC, Boca Raton

Burg SP, Burg EA (1965) Relationship between ethylene production and ripening in bananas. Bot Gaz 126:200–204

Büttner M (2007) The monosaccharide transporter(-like) gene family in *Arabidopsis*. FEBS Lett 581:2318–2324

Carpita N, Gibeaut DM (1993) Structural models of primary cell walls in flowering plants: consistency of molecular structure with physical properties of the walls during growth. Plant J 3:1–30

Caspari T, Will A, Opekarova M, Sauer N, Tanner W (1994) Hexose/H^+ symporters in lower and higher plants. J Exp Biol 196:483–491

Chan HT (1979) Sugar composition of papaya during development. Hortscience 14:140–141

Chan YK, Paull RE (2007) Papaya *Carica papaya* L., Caricaceae. In: Janick J, Paull RE (eds) Encyclopedia of fruit and nuts. CABI, Wallingford, pp 237–247

Che Husin NM, Ali ZM, Lazan H (2000) Isolation and characterization of carERS1, an ethylene receptor cDNA clone from ripening papaya fruit. Trans Malaysian Soc Plant Physiol 9:136–138

Chen NJ, Paull RE (2003) Endoxylanase expressed during papaya fruit ripening: purification, cloning and characterization. Funct Plant Biol 30:433–441

Choi D, Cho DT, Lee Y (2006) Expansins: expanding importance in plant growth and development. Physiol Plant 126:511–518

Cong B, Tanksley SD (2006) FW2.2 and cell cycle control in developing tomato fruit: a possible example of gene co-option in the evolution of a novel organ. Plant Mol Biol 62:867–880

Considine MJ, Daley DO, Whelan J (2001) The expression of alternative oxidase and uncoupling protein during fruit ripening in Mango. Plant Physiol 126:1619–1629

Cosgrove DJ (1999) Enzymes and other agents that enhance cell wall extensibility. Annu Rev Plant Physiol Plant Mol Biol 50:391–417

Cosgrove DJ (2001) Wall structure and wall loosening. A look backwards and forwards. Plant Physiol 125:131–134
Cosgrove DJ (2007) Expansin home page. https://homes.bio.psu.edu/expansins/. Accessed 30 Dec 2011
Coutinho PM, Henrissat B (1999) Carbohydrate-active enzymes: an integrated database approach. In: Gilbert HJ, Davies G, Henrissat B, Svensson B (eds) Recent advances in carbohydrate bioengineering. The Royal Society of Chemistry, Cambridge, pp 3–12
Croteau R, Lutchan TM, Lewis NG (2000) Natural products (secondary metabolites). In: Buchanan B, Gruissem W, Jones R (eds) Biochemistry and molecular biology of plants. Amer Soc Plant Physiol, Rockville, pp 1250–1318
Cunningham FX Jr, Gantt E (1998) Genes and enzymes of carotenoid biosynthesis in plants. Annu Rev Plant Physiol Plant Mol Biol 49:557–583
Devitt LC, Sawbridge T, Holton TA, Mitchelson K, Dietzgen RG (2006) Discovery of genes associated with fruit ripening in *Carica papaya* using expressed sequence tags. Plant Sci 170:356–363
Dibley SJ, Gear ML, Xiao Y, Rosche EG, Offler CE, McCurdy DW, Patrick JW (2005) Temporal and spatial expression of hexose transporters in developing tomato (*Lycopersicon esculentum*) fruit. Funct Plant Biol 32:777–785
El Moussaoui A, Nijs M, Paul C, Wintjens R, Vincentelli J, Azarkan M, Looze Y (2001) Revisiting the enzymes stored in the laticifers of *Carica papaya* in the context of their possible participation in the plant defence mechanism. Cell Mol Life Sci 58:556–570
Elhafez D, Murcha MW, Clifton R, Soole KL, Day DA, Whelan J (2006) Characterization of mitochondrial alternative NAD(P)H dehydrogenases in *Arabidopsis*: intraorganelle location and expression. Plant Cell Physiol 47:43–54
Fellman JK, Miller TW, Mattison DS, Mattheis JP (2000) Factors that influence biosynthesis of volatile flavor compounds in apple fruits. Hortscience 35:1026–1033
Flath RA, Forrey RR (1977) Volatile components of papaya (*Carica papaya* L., Solo variety). J Agric Food Chem 25:103–109
Flath RA, Light DM, Jang EB, Mon TR, John JO (1990) Headspace examination of volatile emissions form ripening papaya (*Carica papaya* L., Solo variety). J Agric Food Chem 38:1060–1063
Franco MRB, Rodriguez ADB (1993) Volatile components of two pawpaw cultivars. Arquivos Biol Tecnol 36:613–632
Frary A, Nesbitt C, Frary A, Grandillo S, van der Knaap E, Cong B, Liu J, Meller J, Elber R, Alpert KB, Tanksley SD (2000) fw2.2: a quantitative trait locus key to the evolution of tomato fruit size. Science 289:85–88
Fraser PD, Truesdale M, Bird C, Schuch W, Bramley PM (1994) Carotenoid biosynthesis during tomato fruit development. Plant Physiol 105:405–413
Fridman M, Zamir D (2003) Functional divergence of a syntenic invertase gene family in tomato, potato, and *Arabidopsis*. Plant Physiol 131:603–609
Fry SC (2004) Primary cell wall metabolism: tracking the careers of wall polymers in living plant cells. New Phytol 161:641–675
Frye CA, Tang D, Innes RW (2001) Negative regulation of defense responses in plant by conserved MAPKK kinase. Proc Natl Acad Sci USA 98:373–378
Fuggate P, Wongs-Aree C, Noichinda S, Kanlayanarat S (2010) Quality and volatile attributes of attached and detached 'Pluk Mai Lie' papaya during fruit ripening. Sci Hortic 126:120–129
Gear ML, McPhillips ML, Patrick JW, McCurdy DW (2000) Hexose transporters of tomato: molecular cloning, expression analysis and functional characterization. Plant Mol Biol 44:687–697
Geisler DA, Broselid C, Hederstedt L, Rasmusson AG (2007) Ca2+-binding and Ca2+-independent respiratory NADH and NADPH dehydrogenases of *Arabidopsis thaliana*. J Biol Chem 282:28455–28464
Giovannoni J (2001) Molecular biology of fruit maturation and ripening. Annu Rev Plant Physiol Plant Mol Biol 52:725–749

Giovannoni JJ (2004) Genetic regulation of fruit development and ripening. Plant Cell 16: S170–S180

Givnish TJ, Pires JC, Graham SW, McPherson MA, Prince LM, Patterson TB, Rai HS, Roalson EH, Evans TM, Hahn WJ, Millam KC, Meerow AW, Molvray M, Kores PJ, O'Brien HE, Hall JC, Kress WJ, Sytsma KJ (2005) Repeated evolution of net venation and fleshy fruits among monocots in shaded habitats confirms a priori predictions: evidence from an ndhF phylogeny. Proc R Soc B Biol Sci 272:1481–1490

Gomez-Roldan V, Fermas S, Brewer PB, Puech-Pages V, Dun EA, Pillot JP, Letisse F, Matusova R, Danoun S, Portais JC, Bouwmeester H, Becard G, Beveridge CA, Rameau C, Rochange SF (2008) Strigolactone inhibition of shoot branching. Nature 455:189–194

Gussman CD, Goffredz JC, Gianfagna TJ (1993) Ethylene production and fruit-softening rates in several apple fruit ripening variants. Hortscience 28:135–137

Hall BP, Shakeel SN, Schaller GE (2007) Ethylene receptors: ethylene perception and signal transduction. J Plant Growth Regul 26:118–130

Henrissat B, Coutinho PM, Davies GJ (2001) A census of carbohydrate-active enzymes in the genome of *Arabidopsis thaliana*. Plant Mol Biol 47:55–72

Henry MF, Nyns EJ (1975) Cyanide-insensitive respiration. An alternative mitochondrial pathway. Sub-Cell Biochem 4:1–65

Hirschberg J (2001) Carotenoid biosynthesis in flowering plants. Curr Opin Plant Biol 4:210–218

Hortensteiner S (2006) Chlorophyll degradation during senescence. Annu Rev Plant Biol 57:55–77

Hu Z, Deng L, Yan B, Pan Y, Luo M, Chen X, Hu T, Chen G (2011) Silencing of the LeSGR1 gene in tomato inhibits chlorophyll degradation and exhibits a stay-green phenotype. Biol Plant 55:27–34

Hua J, Meyerowitz EM (1998) Ethylene responses are negatively regulated by a receptor gene family in *Arabidopsis thaliana*. Cell 94:261–271

Hua J, Sakai H, Nourizadeh S, Chen QG, Bleecker AB, Ecker JR, Meyerowitz EM (1998) EIN4 and ERS2 are members of the putative ethylene receptor gene family in *Arabidopsis*. Plant Cell 10:1321–1332

Ito PJ, Zee F, Hamilton RA, Paull R (1990) Genetic resources in tropical fruit and nut crops. Acta Hortic 269:125–130

Jansen EF, Balls AK (1941) Chymopapain: a new crystalline proteinase from papaya latex. J Biol Chem 137:459–460

Ji X, Van den Ende W, Van Laere A, Cheng S, Bennett J (2005) Structure, evolution, and expression of the two invertase gene families of rice. J Mol Evol 60:615–634

Kaminaga Y, Schnepp J, Peel G, Kish CM, Ben-Nissan G, Weiss D, Orlova I, Lavie O, Rhodes D, Wood K, Porterfield DM, Cooper AJL, Schloss JV, Pichersky E, Vainstein A, Dudareva N (2007) Plant phenylacetaldehyde synthase is a bifunctional homotetrameric enzyme that catalyzes phenylalanine decarboxylation and oxidation. J Biol Chem 281:23357–23366

Katague DB, Kirch ER (1965) Chromatographic analysis of volatile components of papaya fruits. J Pharm Sci 54:891–894

Kevany B, Taylor M, Dal Cin V, Klee HJ (2007) Ethylene receptor degradation controls the timing of ripening in tomato fruit. Plant J 51:458–467

Klee H, Tieman D (2002) The tomato ethylene receptor gene family: form and function. Physiol Plant 115:336–341

Kleffmann T, Russenberger D, Zychlinski A, Christopher W, Sjolander K, Gruissem W, Baginsky S (2004) The Arabidopsis thaliana chloroplast proteome reveals pathway abundance and novel protein functions. Curr Biol 14:354–362

Knapp S (2002) Tobacco to tomatoes: a phylogenetic perspective on fruit diversity in the Solanaceae. J Exp Bot 53:2001–2022

Komatsu A, Moriguchi T, Koyama K, Omura M, Akihama T (2002) Analysis of sucrose synthase genes in citrus suggests different roles and phylogenetic relationships. J Exp Bot 53:61–71

Ku HM, Doganlar S, Chen KY, Tanksley S (1999) The genetic basis of pear-shaped tomato fruit. TAG 9:844–850

Ku HM, Grandillo G, Tanksley SD (2000a) fs8.1, a major QTL, sets the pattern of tomato carpel shape well before anthesis. TAG 101:873–878
Ku HM, Vision T, Liu JP, Tanksley SD (2000b) Comparing sequenced segments of the tomato and *Arabidopsis* genomes: Large-scale duplication followed by selective gene loss creates a network of synteny. Proc Natl Acad Sci USA 97:9121–9126
Kunieda T, Amano T, Shioi Y (2005) Search for chlorophyll degradation enzyme, Mg-dechelatase, from extracts of *Chenopodium album* with native and artificial substrates. Plant Sci 169:177–183
Langenkamper G, Fung RWM, Newcomb RD, Atkinson RG, Gardner RC, MacRae EA (2002) Sucrose phosphate synthase genes in plants belong to three different families. J Mol Evol 54:322–332
Lelievre JM, Latche A, Jones B, Bouzayen M, Pech JC (1997) Ethylene and fruit ripening. Physiol Plant 101:727–739
Li Y, Darley CP, Ongaro V, Fleming A, Schipper O, Baldauf SL, McQueen-Mason SJ (2002) Plant expansins are a complex multigene family with an ancient evolutionary origin. Plant Physiol 128:854–864
Lichtenthaler HK (1999) The 1-deoxy-D-xylulose-5-phosphate pathway of isoprenoid biosynthesis in plants. Annu Rev Plant Physiol Plant Mol Biol 50:47–65
Lichtenthaler HK (2000) Non-mevalonate isoprenoid biosynthesis: Enzymes, genes and inhibitors. Biochem Soc Trans 28:785–789
Lippman Z, Tanksley SD (2001) Dissecting the genetic pathway to extreme fruit size in tomato using a cross between the small-fruited wild species *Lycopersicon pimpinellifolium* and *L. esculentum* var. Giant Heirloom. Genetics 158:413–422
Liu J, Eck JV, Cong B, Tanksley SD (2002) A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. Proc Natl Acad Sci USA 99:13302–13306
Lorts CM, Briggeman T, Sang T (2008) Evolution of fruit types and seed dispersal. A phylogenetic and ecological snapshot. J Syst Evol 46:396–404
Lovisetto A, Guzzo F, Tadiello A, Toffali K, Favretto A, Casadoro G (2012) Molecular analyses of MADS-Box genes trace back to gymnosperms the invention of fleshy fruits. Mol Biol Evol 29:409–419
Mach JM, Castillo AR, Hoogstraten R, Greenberg JT (2001) The *Arabidopsis* accelerated cell death gene ACD2 encodes red chlorophyll catabolite reductase and suppresses the spread of disease symptoms. Proc Natl Acad Sci USA 98:771–776
Mackenzie S, McIntosh L (1999) Higher plant mitochondria. Plant Cell 11:571–586
MacLeod AJ, Pieris NM (1983) Volatile components of papaya (*Carica papaya* L.) with particular reference to glucosinolate products. J Agric Food Chem 31:1005–1008
Manenoi A, Paull RE (2007) Papaya fruit softening, endoxylanase gene expression, protein and activity. Physiol Plant 131:470–480
Manenoi A, Bayogan ERV, Thumdee S, Paull RE (2007) Utility of 1-methylcyclopropene as a papaya postharvest treatment. Postharvest Biol Technol 44:55–62
Marchler-Bauer A, Anderson JB, Cherukuri PF, DeWeese-Scott C, Geer LY, Gwadz M, He S, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Liebert CA, Liu C, Lu F, Marchler GH, Mullokandov M, Shoemaker BA, Simonyan V, Song JS, Thiessen PA, Yamashita RA, Yin JJ, Zhang D, Bryant SH (2005) CDD: a conserved domain database for protein classification. Nucleic Acids Res 33:192–196
Martin T, Frommer WB, Salanoubat M, Willmitzer L (1993) Expression of an *Arabidopsis* sucrose synthase gene indicates a role in metabolization of sucrose both during phloem loading and in sink organs. Plant J 4:367–377
McDonald AE (2008) Alternative oxidase: an inter-kingdom perspective on the function and regulation of this broadly distributed 'cyanide-resistant' terminal oxidase. Funct Plant Biol 35:535–552
McQueen-Mason S, Durachko DM, Cosgrove DJ (1992) Two endogenous proteins that induce cell wall expansion in plants. Plant Cell 4:1425–1433
Melo AMP, Bandeiras TM, Teixeira M (2004) New insights into type II NAD(P)H:quinone oxidoreductases. Microbiol Mol Biol Rev 68:603–616

Michael TP, Breton G, Hazen SP, Priest H, Mockler TC, Kay SA, Chory J (2008) A morning-specific phytohormone gene expression program underlying rhythmic plant growth. PLoS Biol 6:e225

Michalecka AM, Svensson AS, Johansson FI, Agius SC, Johanson U, Brennicke A, Binder S, Rasmusson AG (2003) *Arabidopsis* genes encoding mitochondrial type II NAD(P)H dehydrogenases have different evolutionary origin and show distinct responses to light. Plant Physiol 133:642–652

Ming R, Hou S, Feng Y, Yu Q, Dionne-Laporte A, Saw JH, Senin P, Wang W, Ly BV, Lewis KLT, Salzberg SL, Feng L, Jones MR, Skelton RL, Murray JE, Chen C, Qian W, Shen J, Du P, Eustice M, Tong E, Tang H, Lyons E, Paull RE, Michael TP, Wall K, Rice D, Albert H, Li Wang M, Zhu YJ, Schatz M, Nagarajan N, Agbayani R, Guan P, Blas A, Man Wai C, Ackerman CM, Ren Y, Liu C, Wang J, Wang J, Na JK, Shakirov EV, Haas B, Thimmapuram J, Nelson D, Wang X, Bowers JE, Gschwend AR, Delcher AL, Singh R, Suzuki JY, Tripathi S, Neupane K, Wei H, Irikura B, Paidi M, Jiang N, Zhang W, Presting G, Windsor A, Navajas-Pérez R, Torres MJ, Feltus FA, Porter B, Li Y, Burroughs AM, Luo MC, Liu L, Christopher DA, Mount SM, Moore PH, Sugimura T, Jiang J, Schuler MA, Friedman V, Mitchell-Olds T, Shippen DE, dePamphilis CW, Palmer JD, Freeling M, Paterson AH, Gonsalves D, Wang L, Alam M (2008) Genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). Nature 452:991–996

Ming R, Yu Q, Moore PH, Paull RE, Zhu YJ, Schuler MA, Jiang J, Paterson AH (2012) Genome of papaya, a fast growing tropical fruit tree. Tree Genet Genomics 8:445–462

Moore AL, Albury MS, Crichton PG, Affourtit C (2002) Function of the alternative oxidase: is it still a scavenger? Trends Plant Sci 7:478–481

Nakano T, Suzuki K, Fujimura T, Shinshi H (2006) Genome-wide analysis of the ERF gene family in *Arabidopsis* and rice. Plant Physiol 140:411–432

Nakasone HY, Paull RE (1998) Papaya. In: Tropical fruit. CAB International, Wallingford, pp 239–269

Nakatsuka A, Murachi S, Okunishi H, Shiomi S, Nakano R, Kubo Y, Inaba A (1998) Differential expression and internal feedback regulation of 1-aminocyclopropane-1-carboxylate synthase, of 1-aminocyclopropane-1-carboxylase oxidase and ethylene receptor genes in tomato fruit during development and ripening. Plant Physiol 118:1295–1305

Newcomb RD, Crowhurst RN, Gleave AP, Rikkerink EHA, Allan AC, Beuning LL, Bowen JH, Kim E, Jamieson R, Janssen BJ, Laing WA, McArtney S, Nain B, Ross GS, Snowden KC, Souleyre EJF, Walton EF, Yauk YK (2006) Analyses of expressed sequence tags from apple. Plant Physiol 141:147–166

Nogueira SB, Labate CA, Gozzo FC, Pilau EJ, Nascimento FM, Lajolo JRO (2012) Proteomic analysis of papaya fruit ripening using 2DE-DIGE. J Proteomics 75:1428–1439

Oberg KA, Ruysschaert JM, Azarkan M, Smolders N, Zerhouni S, Wintjens R, Amrani A, Looze Y (1998) Papaya glutamine cyclase, a plant enzyme highly resistant to proteolysis, adopts an all-b conformation. Eur J Biochem 258:214–222

Othman R, Nuraziyan A (2010) Fruit-specific expression of papaya subtilase gene J. Plant Physiol 167:131–137

Othman R, Chong HL, Choo TS, Mohd Ali Z (2011) Three b-galactosidase cDNA clones related to fruit ripening in papaya (Carica papaya). Acta Physiol Plant 33:2301–2310

Ouaked F, Rozhon W, Lecourieux D, Hirt H (2003) A MAPK pathway mediates ethylene signaling in plants. EMBO J 22:1282–1288

Park BH, Karpinets TV, Syed MH, Leuze MR, Uberbacher EC (2010) CAZymes Analysis Toolkit (CAT): web-service for searching and analyzing carbohydrate-active enzymes in a newly sequenced organism using CAZy database. Glycobiology 20:1574–1584. http://cricket.ornl.gov/cgi-bin/cat.cgi

Patrick JW (1997) Phloem unloading: sieve element unloading and post-sieve element transport. Annu Rev Plant Physiol Plant Mol Biol 48:191–222

Paull RE, Chen NJ (1983) Postharvest variation in cell wall-degrading enzymes of papaya (*Carica papaya* L.) during fruit ripening. Plant Physiol 72:382–385

Paull RE, Chen NJ (1989) Waxing and plastic wraps influence water loss from papaya fruit during storage and ripening. J Am Soc Hortic Sci 114:937–942

Paull R, Irikura B, Wu PF, Turano H, Chen N, Blas A, Fellman J, Gschwend A, Wai CM, Yu Q, Presting G, Alam M, Ming R (2008) Fruit development, ripening and quality related genes in the papaya genome. Trop Plant Biol 1:246–277

Peng Z-Y, Zhou X, Li L, Yu X, Li H, Jiang Z, Cao G, Bai M, Wang X, Jiang C, Lu H, Hou X, Qu L, Wang Z, Zuo J, Fu X, Su Z, Li S, Guo H (2009) Arabidopsis hormone database: a comprehensive genetic and phenotypic information database for plant hormone research in Arabidopsis. Nucleic Acids Res 37:D975–D982

Phillips DR, Rasbery JM, Bartel B, Matsuda SPT (2006) Biosynthetic diversity in plant triterpene cyclization. Curr Opin Plant Biol 9:305–314

Pichersky E, Gershenzon J (2002) The formation and function of plant volatiles: perfumes for pollinator attraction and defense. Curr Opin Plant Biol 5:237–243

Picton S, Gray J, Barton S, AbuBakar U, Lowe A, Grierson D (1993) cDNA cloning and characterisation of novel ripening-related mRNAs with altered patterns of accumulation in the ripening inhibitor (*rin*) tomato ripening mutant. Plant Mol Biol 23:193–207

Pino JA, Almora K, Marbot R (2003) Volatile components of papaya (*Carica papaya* L., Maradol variety) fruit. Flavour Fragrance J 18:492–496

Pruzinska A, Anders I, Aubry S, Schenk A, Tapernoux-Luthe E, Muller T, Krautler B, Hortensteiner S (2007) In vivo participation of red chlorophyll catabolite reductase in chlorophyll breakdown. Plant Cell 19:369–387

Qiu YX, Nishina MS, Paull RE (1995) Papaya fruit growth, calcium uptake and fruit ripening. J Am Soc Hortic Sci 120:246–253

Quecini V, Torres GAM, Rosa Jr VEd, Gimenes MA, Machado JBdM, Figueira AvdO, Benedito V, Targon MLPN, Cristofani-Yaly M (2007) In silico analysis of phytohormone metabolism and communication pathways in citrus transcriptome. Genet Mol Biol 30:713–733

Rasmusson AG, Svensson AS, Knoop V, Grohmann L, Brennicke A (1999) Homologues of yeast and bacterial rotenone-insensitive NADH dehydrogenases in higher eukaryotes: two enzymes are present in potato mitochondria. Plant J 20:79–87

Roberts TH, Fredlund KM, Moller IM (1995) Direct evidence for the presence of two external NAD(P)H dehydrogenases coupled to the electron transport chain in plant mitochondria. FEBS Lett 373:307–309

Roberts JA, Whitelaw CA, Gonzalez-Carranza ZH, McManus MT (2000) Cell separation processes in plants - models, mechanisms and manipulation. Ann Bot 86:223–235

Rodrigo MJ, Marcos JF, Alférez F, Mallent MD, Zacarías L (2003) Characterization of Pinalate, a novel *Citrus sinensis* mutant with a fruit-specific alteration that results in yellow pigmentation and decreased ABA content. J Exp Bot 54:727–738

Rodrigo MJ, Alquezar B, Zacarias L (2006) Cloning and characterization of two 9-cis-epoxycarotenoid dioxygenase genes, differentially regulated during fruit maturation and under stress conditions, from orange (*Citrus sinensis* L. Osbeck). J. Exp Bot 57:633–643

Rodríguez FI, Esch JJ, Hall AE, Binder BM, Schaller GE, Bleecker AB (1999) A copper cofactor for the ethylene receptor ETR1 from *Arabidopsis*. Science 283:996–998

Ronse de Craene LP, Haston E (2006) The systematic relationships of glucosinolate-producing plants and related families: a cladistic investigation based on morphological and molecular characters. Bot J Linnean Soc 151:453–494

Rose JKC, Lee HH, Bennett AB (1997) Expression of a divergent expansin gene is fruit-specific and ripening regulated. Proc Natl Acad Sci USA 94:5955–5960

Rose JKC, Cosgrove DJ, Albersheim P, Darvill AG, Bennett AB (2000) Detection of expansin proteins and activity during tomato fruit ontogeny. Plant Physiol 123:1583–1592

Rose JKC, Saladié M, Catalá C (2004) The plot thickens: new perspectives of primary cell wall modification. Curr Opin Plant Biol 7:296–301

Roth I (1977) Fruits of angiosperms. Gebrüder Borntraeger, Berlin

Roth I, Clausnitzer I (1972) Desarrollo y anatomía del fruto y la semilla de *Carica papaya* L. (Lechosa). Acta Bot Venezuela 7:187–206
Salas CE, Gomes MTR, Hernandez M, Lopes MTP (2008) Plant cysteine proteinases: evaluation of the pharmacological activity. Phytochemistry 69:2263–2269
Sampedro J, Lee Y, Carey RE, de Pamphilis C, Cosgrove DJ (2005) Use of genomic history to improve phylogeny and understanding of births and deaths in a gene family. Plant J 44:409–419
Sampedro J, Carey RE, Cosgrove DJ (2006) Genome histories clarify evolution of the expansin superfamily: new insights from the poplar genome and pine ESTs. J Plant Res 119:11–21
Sangwanangkul P, Paull RE (2005) The role of hexose transporter in sugar accumulation of papaya fruit during maturation and ripening. Acta Hortic 740:313–316
Schaller GE, Kieber JJ (2002) Ethylene. In: Somerville C, Meyerowitz E (eds) The Arabidopsis book. American Society of Plant Biologists, Rockville
Schledz M, al-Babili S, von Lintig J, Haubruck H, Rabbani S, Kleinig H, Beyer P (1996) Phytoene synthase from Narcissus pseudonarcissus: functional expression, galactolipid requirement, topological distribution in chromoplasts and induction during flowering. Plant J 10:781–792
Schonbaum GR, Bonner WD Jr, Storey BT, Bahr JT (1971) Specific inhibition of the cyanide-insensitive respiratory pathway in plant mitochondria by hydroxamic acids. Plant Physiol 47:124–128
Schwab W, Davidovich-Rikanati R, Lewinsohn E (2008) Biosynthesis of plant-derived flavor compounds. Plant J 54:712–732
Schweiggert RM, Steingass CB, Heller A, Esquivel P, Reinhold C (2011) Characterization of chromoplasts and carotenoids of red- and yellow-fleshed papaya (*Carica papaya* L.). Planta 234:1031–1044
Seymour GB, Manning K (2002) Genetic control of fruit ripening. In: Knee MM (ed) Fruit quality and its biological basis. Academic, Sheffield, pp 253–274
Seymour G, Poole M, Manning K, King GJ (2008) Genetics and epigenetics of fruit development and ripening. Curr Opin Plant Biol 11:58–63
Shindo T, Van Der Hoorn RAL (2008) Papain-like cysteine proteases: key players at molecular battlefields employed by both plants and their invaders. Mol Plant Pathol 9:119–125
Simkin A, Schwartz S, Auldridge M, Taylor M, Klee HJ (2004) The tomato carotenoid cleavage dioxygenase 1 genes contribute to the formation of the flavor volatiles B-ionone, pseudoionone and geranylacetone. Plant J 40:882–892
Simpson DJ, Fincher GB, Huang AH, Cameron-Mills V (2003) Structure and function of cereal and related higher plant (1-4)-β-xylan endohydrolases. J Cereal Sci 37:111–127
Smith DL, Gross KC (2000) A family of at least seven ß-galactosidase genes is expressed during tomato fruit development. Plant Physiol 123:1173–1183
Soh CP, Ali ZM, Lazan H (2006) Characterization of an alpha-galactosidase with potential relevance to ripening related texture changes. Phytochemistry 67:242–254
Speirs J, Lee E, Holt K, Yong-Duk K, Scott NS, Loveys B, Schuch W (1998) Genetic manipulation of alcohol dehydrogenase levels in ripening tomato fruit affects the balance of some flavor aldehydes and alcohols. Plant Physiol 117:1047–1058
Stepanova AN, Alonso JM (2005) Ethylene signaling pathway. Sci STKE 276:cm3. doi:10.1126/stke.2762005cm3
Sturm A (1999) Invertases. Primary structures, functions, and roles in plant development and sucrose partitioning. Plant Physiol 121:1–8
Sturm A, Tang G-Q (1999) The sucrose-cleaving enzymes of plants are crucial for development, growth and carbon partitioning. Trends Plant Sci 4:401–407
Taylor MAJ, Al-Sheikh M, Revell DF, Sumner IG, Connerton IF (1999) cDNA cloning and expression of *Carica papaya* prochymopapain isoforms in *Escherichia coli*. Plant Sci 145:41–47
Theologis A, Laties GG (1978) Respiratory contribution of the alternate path during various stages of ripening in avocado and banana fruits. Plant Physiol 62:249–255

Thomas H, Huang L, Young M, Ougham H (2009) Evolution of plant senescence. BMC Evol Biol 9:163

Thumdee S, Manenoi A, Chen NJ, Paull RE (2010) Papaya fruit softening: role of hydrolases. Trop Plant Biol 3:98–109

Tieman DM, Loucas HM, Kim JY, Clark DG, Klee HJ (2007) Tomato phenylacetaldehyde reductases catalyze the last step in the synthesis of the aroma volatile 2-phenylethanol. Phytochemistry 68:2660–2669

Tomato Consortium (2012) The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485:635–641

Trentmann SM, Kende H (1995) Analysis of *Arabidopsis* cDNA that shows homology to the tomato E8 cDNA. Plant Mol Biol 29:161–166

Tsuchisaka A, Yamagami T, Alonso JM, Ecker JR, Theologis A (2007) The *Arabidopsis* 1-aminocyclopropane-1-carboxylic acid synthase (ACS) gene family. In: Proceedings of the 18th international conference on Arabidopsis research, 2007. TAIR Publication 501722104

Tymowska-Lalanne Z, Kreis M (1998) Expression of the *Arabidopsis thaliana* invertase gene family. Planta 207:259–265

Vanlerberghe GC, McIntosh L (1997) Alternative oxidase: from gene to function. Annu Rev Plant Physiol Plant Mol Biol 48:703–734

Vogel JT, Tan B-C, McCarty DR, Klee HJ (2008) The carotenoid cleavage dioxygenase 1 enzyme has broad substrate specificity, cleaving multiple carotenoids at two different bond positions. J Biol Chem 283:11364–11373

Wang W, Hall AE, O'Malley R, Bleecker AB (2003) Canonical histidine kinase activity of the transmitter domain of the ETR1 ethylene receptor from *Arabidopsis* is not required for signal transmission. Proc Natl Acad Sci USA 100:352–357

Watling JR, Robinson SA, Seymour RS (2006) Contribution of the alternative pathway to respiration during thermogenesis in flowers of the sacred lotus. Plant Physiol 140:1367–1373

Wu G-L, Zhang X-Y, Zhang L-Y, Pan Q-H, Shen Y-Y, Zhang D-P (2004) Phloem unloading in developing walnut fruit is symplasmic in the seed pericarp and apoplasmic in the fleshy pericarp. Plant Cell Physiol 45:1461–1470

Wüthrich KL, Bovet L, Hunziker PE, Donnison IS, Hörtensteiner S (2000) Molecular cloning, functional expression and characterisation of RCC reductase involved in chlorophyll catabolism. Plant J 21:189–198

Xiao H, Jiang N, Schaffner E, Stockinger EJ, van der Knaap E (2008) A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. Science 319:1527–1530

Yabumoto K, Yamaguchi M, Jennings WG (1977) Biosynthesis of some volatiles constituents of `muskmelon *Cucumis melo*. Chem Mikrobiol Technol Lebensraum 5:53–56

Yamagami T, Tsuchisaka A, Yamada K, Haddon WF, Harden LA, Theologis A (2003) Biochemical diversity among the 1-amino-cyclopropane-1-carboxylate synthase isozymes encoded by the *Arabidopsis* gene family. J Biol Chem 278:49102–49112

Yamamoto HY (1964) Comparison of the carotenoids in yellow- and red-fleshed *Carica papaya*. Nature 201:1049–1050

Yao N, Greenbreg JT (2006) Arabidopsis ACCELERATED CELL DEATH2 modulates programmed cell death. Plant Cell 18:397–411

Yip WK, Dong JG, Kenny JW, Thompson GA, Yang SF (1990) Characterization and sequencing of the active site of 1-aminocyclopropane-1-carboxylate synthase. Proc Natl Acad Sci USA 87:7930–7934

Yoo A, Seo YS, Jung JW, Sung SK, Kim WT, Lee W, Yang DR (2006) Lys296 and Arg299 residues in the C-terminus of MD-ACO1 are essential for a 1-aminocyclopropane-1-carboxylate oxidase enzyme activity. J Struct Biol 156:407–420

Zeevart JAD, Creelman RA (1988) Metabolism and physiology of abscisic acid. Annu Rev Plant Physiol Plant Mol Biol 39:439–473

Zhang L-Y, Peng Y-B, Pelleschi-Travier S, Fan Y, Lu Y-F, Lu Y-M, Gao X-P, Shen Y-Y, Delrot S, Zhang D-P (2004) Evidence for apoplasmic phloem unloading in developing apple fruit. Plant Physiol 135:574–586

Zhang X-Y, Wang X-L, Wang X-F, Xia G-H, Pan Q-H, Fan R-C, Wu F-Q, Yu X-C, Zhang D-P (2006) A shift of phloem unloading from symplasmic to apoplasmic pathway is involved in developmental onset of ripening in grape berry. Plant Physiol 142:220–232

Zhou LL, Paull RE (2001) Sucrose metabolism during papaya (*Carica papaya*) fruit growth and ripening. J Am Soc Hortic Sci 126:351–357

Zhou L, Chen CC, Ming R, Christopher DA, Paull RE (2003a) Apoplastic invertase and its enhanced expression and post-translational control during papaya fruit maturation and ripening. J Am Soc Hortic Sci 128:628–635

Zhu YJ, Qiu XH, Moore PH, Borth W, Hu J, Ferreira S, Albert HH (2003) Systemic acquired resistance induced by BTH in papaya. Physiol Mol Plant Pathol 63:237–248

Chapter 15
Genomics of Papaya Disease Resistance

Brad W. Porter, David A. Christopher, and Yun J. Zhu

Introduction

Papaya is a popular fruit in the tropics and subtropics that significantly contributes to the dietary intake of vitamins A and C (FAO 2009; Miller 1926). Although abundant year-round production of papaya is possible, its susceptibility to numerous diseases affects fruit quality and yield (Persley and Ploetz 2003) (Table 15.1). This may be partially attributable to its narrow genetic base. Presumably, papaya diverged from other species of the Caricaceae family as a result of being evolutionarily isolated in Central America (Aradhya et al. 1999). Archeological and paleoethnobotanical evidence indicates the presence of papaya in the region dating back to the Maya Classic Period (300–900 AD). The identification of maximal species richness of its close relative, the genus *Vasconcellea*, which occurs farther south in Colombia, Ecuador, and Peru, supports this hypothesis (Miksicek 1983; Lentz 1999; Scheldeman et al. 2007). As a consequence, *C. papaya* is the only member of the genus *Carica* (Aradhya et al. 1999; Kim et al. 2002) and is more vulnerable to disease than genera with greater genetic diversity.

Today, from a global production standpoint, papaya is no longer isolated anywhere in the world, and the severity and geographical distribution of some papaya diseases is highly variable. For example, papaya meleira virus (PMeV), which causes "sticky disease," is considered among the most severe diseases in Brazil (Ventura et al. 2004), but it is less prevalent elsewhere. Similarly, the acidic soils of Hawaii are thought to promote *Phytophthora* rot (Manshardt and Zee 1994), while in Malaysia where soil pH is also low, the disease is considered insignificant

B.W. Porter • D.A. Christopher
Department of Molecular Biosciences and Bioengineering,
University of Hawaii at Manoa, Honolulu, HI 96822, USA

Y.J. Zhu (✉)
Hawaii Agriculture Research Center, PO Box 100, Kunia, HI 96759, USA
e-mail: jzhu@harc-hspa.com

R. Ming and P.H. Moore (eds.), *Genetics and Genomics of Papaya*, Plant Genetics and Genomics: Crops and Models 10, DOI 10.1007/978-1-4614-8087-7_15,

Table 15.1 Major diseases of papaya

Pathogen	Species	Prominent distribution
Bacteria		
Bacterial canker and decline	*Erwinia*	Caribbean
Bacterial leaf spot	*Pseudomonas caricae-papayae*	Brazil
Internal yellowing	*Enterobacter cloacae*	Hawaii
Mushy canker	Erwinia	Northern Mariana Islands
Papaya bunchy top	*Rickettsia*	Puerto Rico, Caribbean, and Central and South America
Purple stain fruit rot	*Erwinia herbicola*	Brazil
Fungi and oomycetes		
Alternaria fruit spot	*Alternaria alternate*	Israel and Hawaii
Anthracnose	*Colletotrichum gloeosporioides*	Most production areas
Asperisporium black spot	*Asperisporium caricae*	Australia, Africa, Central America, South America, India, and the USA
Black/dry rot	*Mycosphaerella caricae*	Most production areas
Brown spot/Corynespora leaf spot	*Corynespora cassiicola*	Most production areas
Cercospora black spot	*Cercospora papayae*	Most production areas
Collar rot	*Calonectria ilicicola*	Hawaii
Fusarium fruit rot	*Fusarium solani*	Hawaii, India, Israel, and the Philippines
Lasiodiplodia fruit and stem rot	*Diplodia theobromae*	Hawaii and India
Leaf spot, dry rot, end rot of fruits, wet fruit rot	*Phomopsis caricae-papayae*, *Phomopsis* sp.	Most production areas
Phytophthora fruit, root, and stem rot	*Phytophthora palmivora*	Most production areas
Powdery mildew	*Oidium caricae*	Most production areas
Soft rot	*Rhizopus stolonifer*	Most production areas
Stemphylium fruit rot	*Stemphylium lycopersici*	Most production areas
Nematodes		
Reniform nematodes	*Rotylenchulus reniformis*	Most production areas
Root-knot nematodes	*Meloidogyne incognita, Meloidogyne javanica*	Most production areas
Phytoplasmas		
Papaya dieback	*Candidatus* Phytoplasma australiense	Australia
Yellow crinkle and mosaic	*Candidatus* Phytoplasma australasia	Australia
Viruses		
Leaf curl disease	*Papaya leaf curl virus*	India
Meleira or sticky disease	Virus	Brazil
Papaya droopy necrosis and papaya apical necrosis	Rhabdovirus	Florida and Venezuela

(continued)

Table 15.1 (continued)

Pathogen	Species	Prominent distribution
Papaya leaf distortion mosaic	*Papaya leaf distortion mosaic virus*	Japan, Saipan, and Taiwan
Papaya lethal yellowing disease	*Papaya lethal yellowing virus*	Brazil
Papaya mild yellowing disease	*Papaya mild yellowing virus*	Venezuela
Papaya mosaic	*Papaya mosaic virus*	USA, Mexico, and South America
Papaya ringspot	*Papaya ringspot virus*-type P	Most production areas
Tomato spotted wilt	*Tomato spotted wilt virus*	Hawaii

Data from Persley and Ploetz (2003) and Ventura et al. (2004)

(Personal communication: Dr. Chan Ying Kwok, Malaysian Agrifood Corporation). Such disparity in disease prevalence between environments might be explained by factors such as pathogen diversity, the concentration of disease vectors, the abundance of alternate hosts, or presence of natural barriers that affect pathogen movement and regional outbreaks, as was the case for papaya ringspot virus (PRSV-P) (Gonsalves 1998). In the case of pathogen diversity, the evolution of pathogenicity factors, including effector proteins, can intensify disease (Birch et al. 2006; Walton et al. 2009). Regardless of the mechanisms involved in a particular disease, local cultivars must be developed that can withstand pathogen pressure. To achieve this goal, sources of resistance can be obtained within *Carica* or related Caricaceae genera or by using bioengineering approaches.

Although *Carica* is monotypic, crosses of papaya cultivars have, in some cases, demonstrated that disease resistance is additive and selectable (Mosqueda-Vázquez et al. 1981; Mosqueda-Vázquez and Nakasone 1982). Markers linked to resistance loci are beginning to be developed (Noorda-Nguyen et al. 2010); with the papaya genome sequence now available (Ming et al. 2008), genetic resistance within the species may be more fully determined. Separately, 5 related genera, *Cylicomorpha*, *Horovitzia*, *Jacaratia*, *Jarilla*, and *Vasconcellea*, consist of 34 additional species (Scheldeman et al. 2007) that can be screened for resistance to papaya diseases (Tables 15.1 and 15.2). Transferring resistance genes from these species to papaya is difficult since hybrids often produce nonviable seed or parthenocarpic fruit caused by postzygotic barriers, such as abnormal endosperm development or ovule and embryo abortion (Mekako and Nakasone 1975; Manshardt and Wenslaff 1989a). Nevertheless, hybrids have been recovered (Manshardt and Wenslaff 1989b), and recently the introgression of PRSV resistance from a wild relative was successfully achieved (Siar et al. 2009). This advance provides encouragement that additional wild relative traits may be introgressed into papaya in the future.

Bioengineering has played a major role in securing the production of papaya, most notably through coat protein-mediated resistance to ringspot virus (Fitch et al. 1992; reviewed by Gonsalves in Chap. 7 in this text). New transgenic strategies for controlling carmine spider mite and *Phytophthora palmivora* of papaya have

Table 15.2 First reports and possible emerging diseases of papaya

Disease	Location	Reference
Powdery mildew	Taiwan	Tsay et al. (2011)
Bacterial crown rot	Tonga	Fullerton et al. (2011)
Postharvest anthracnose	South Florida	Tarnowski and Ploetz (2010)
Fruit rot (*Colletotrichum magna*)	Brazil	Nascimento et al. (2010)
Scab	Taiwan	Chen et al. (2009)
Erwinia papayae/papaya dieback	Malaysia	Maktar et al. (2008)
Moroccan watermelon mosaic virus	Democratic Republic of Congo	Arocha et al. (2008)
Papaya leaf distortion mosaic virus infecting transgenic papaya resistant to papaya ringspot virus	Taiwan	Bau et al. (2008)
Atypical internal yellowing	Hawaii	Keith et al. (2008)
Ringspot virus	Côte d'Ivoire	Diallo et al. (2007)
16SrII group phytoplasma	Ethiopia	Arocha et al. (2007)
16SrII group phytoplasma	Cuba	Arocha et al. (2006)
Nivun Haamir dieback disease	Israel	Gera et al. (2005)
Ringspot virus	Bangladesh	Jain et al. (2004)
Leaf curl virus	Taiwan	Chang et al. (2003)
Phytoplasmas	Cuba	Arocha et al. (2003)
Ringspot virus	Iran	Pourrahim et al. (2003)
Black spot	Hawaii	Ogata and Heu (2001)
Papaya mosaic virus (Mexican isolate)	Mexico	Noa-Carrazana and Silva-Rosales (2001)
Collar rot	Baja California Sur, Mexico	Rodriguez-Alvarado et al. (2001)
Leaf blight, fruit rot, root rot	American Samoa	Roberts and Trujillo (1998)
Leaf curl disease	Pakistan	Nadeem et al. (1997)

also been successful (Zhu et al. 2004; McCafferty et al. 2006; Zhu et al. 2007). Characterization of papaya's resistance genes, and those of its wild relatives, will likely provide additional sources of resistance (Porter et al. 2009a). At the same time, ecological control strategies, such as defensive mutualism, wherein a symbiont provides protection against pathogens, must be utilized to promote integrated disease management strategies to preserve resistance (Jaizme-Vega et al. 2006; Newcombe et al. 2010).

The genomics of papaya disease resistance will be discussed in this chapter in the context of the major diseases of papaya and the resources available to mitigate them. Genetic variation of resistance available within *C. papaya* and its wild relatives will be reviewed in addition to past and forthcoming transgenic approaches. Finally, pathogen diversity, emerging diseases, and strategies for promoting durable resistance will be addressed.

Genetic Variation for Disease Resistance in *Carica papaya*

Papaya Ringspot Virus

Papaya ringspot virus exists as multiple strains occurring worldwide and is among the most destructive diseases of papaya (see Chap. 7 in this text; Ventura et al. 2004). *C. papaya* lacks complete resistance to PRSV-P, but conventional breeding has developed partially resistant cultivars. In Florida, Conover et al. (1986) derived "Cariflora" from partially resistant dioecious lines (K2 and K3). Another partially resistant cultivar, Sinta, is an F_1 semidwarf hybrid developed by the Institute of Plant Breeding (College of Agriculture, University of the Philippines, Los Baños) (Siar et al. 2009). The level of resistance provided by "Sinta" is proposed to be sufficient for viable commercial production in areas where PRSV-P infection occurs (Siar et al. 2009). Although the resistance of lines developed from germplasm available within the *C. papaya* species is only partial, these genetic resources are valuable. The PRSV-P resistance of "Cariflora" (Conover and Litz 1978) and, likely, "Sinta" is multigenic and now more useful by using genomic tools (Ming et al. 2008). Markers for quantitative trait loci (QTLs) controlling PRSV resistance may be developed for breeding. New sources of resistance may be used to enhance protection against diverse virus isolates and contribute to the durability of deployed transgenic resistance (Fitch et al. 1990, 1992; Fitch 1993; Fitch and Manshardt 1990).

Phytophthora Fruit, Root, and Stem Rot

P. palmivora is the causal organism of *Phytophthora* fruit, root, and stem rot of papaya and is thought to have originated in Asia (Persley and Ploetz 2003; Mchau and Coffey 1994). *P. palmivora* is classified as an oomycete which is distinct from fungi. Oomycetes are distinguished by being diploid and having nonseptate hyphae and cell walls that contain cellulose but little or no chitin (Latijnhouwers et al. 2003). Many *Phytophthora* species are devastating pathogens, and *P. palmivora*, with over 160 documented hosts (Erwin and Ribeiro 1996), is no exception. The pathogen produces infectious, biflagellate zoospores that are motile in water, making the disease particularly infective during wet conditions (Erwin and Ribeiro 1996). *P. palmivora* is particularly destructive in the southeast part of the island of Hawaii, which can receive >120 in. of rainfall per year (NOAA Climate Data 1971–2000).

Partial resistance to *P. palmivora* has been identified within the *C. papaya* species. After inoculating 1-month-old papaya seedlings with sporangia, Mosqueda-Vázquez et al. (1981) identified four partially resistant lines (Line 8, Waimanalo-23, Waimanalo-24, and Line 40) and two moderately resistant lines (Line 45-T_{22} and Kapoho). Subsequently, "Waimanalo"-23, "Waimanalo"-24, "Line 40," "Line 45-T_{22}," and the susceptible cultivar Higgins were crossed in diallel (crosses in all

possible combinations) to determine the combining ability of *P. palmivora* resistance (Mosqueda-Vázquez and Nakasone 1982). F_1 progeny and parents were screened, and it was determined that there was significant general combining ability, suggesting that resistance is additive and selectable (Mosqueda-Vázquez and Nakasone 1982). In a separate study, field and greenhouse screenings identified the cultivars Tailandia Roxao and Cross Paris, which are larger "Formosa" types, as partially resistant and a separate group of "Solo" papaya as susceptible (Dianese et al. 2007, 2010). Interestingly, in Hawaii, the emergence of *P. palmivora* occurred when "Solo" (accession no. 2853) replaced the traditionally grown, more-resistant large-fruited cultivars (Parris 1941; Takeguchi et al. 1999). This suggests that marketing/educational strategies used to promote the production of both large- and small-fruited papaya might help overall crop resistance. Finally, molecular resources are being developed for marker-assisted selection. A segregating F_2 population derived from a cross of "Kamiya" (partially resistant) and "SunUp" (susceptible) was screened using amplified fragment length polymorphism (AFLP) analysis (Noorda-Nguyen et al. 2010). Several polymorphic DNA fragments linked to resistance were identified (Noorda-Nguyen et al. 2010) and may be converted to cleaved amplified polymorphic sequences (CAPS) to be used as markers to breed *Phytophthora* resistance.

Other Diseases and Pests

Genetic resistance has been reported for many other diseases and pests of papaya. Collar rot of papaya, caused by the fungus *Calonectria ilicicola*, is of notable concern in wet regions on the island of Hawaii (Persley and Ploetz 2003). Greenhouse inoculations identified the cultivar Kapoho Solo as partially resistant compared to the susceptible cultivars, Sunrise Solo and Waimanalo (Nishijima and Aragaki 1973). The fungal pathogen, *Colletotrichum gloeosporioides*, causes the postharvest disease of papaya known as anthracnose. While symptoms occur on the fruit after harvest, infection first occurs during fruit development (Alvarez and Nishijima 1987). "Sunrise Solo" displayed some resistance to *C. gloeosporioides* (Nakasone and Aragaki 1982). The fungus, *Asperisporium caricae*, causes black spot. Leaf infection decreases plant development, whereas the blemishes on infected fruit lessen marketability (Ventura et al. 2004). Dianese et al. (2007) found genotype Sekati to have the lowest severity of *A. caricae* foliage infection and "Sekati," "Tailandia Roxao," and "Tailandia Verde" to have the lowest levels of fruit infection.

In Hawaii, papaya to be exported to California must receive hot water or forced hot air disinfestation treatment to control fruit flies (*Toxotrypana curvicauda*) (Manshardt and Zee 1994), which are considered the most damaging insect pests of papaya (Pantoja et al. 2002). While fruit fly resistance has not been identified, Aluja et al. (1994) found more field infestation of a variety designated "Hawaiian" than two other cultivars, Cera Amarilla and Cera Roja. For all the diseases and pests mentioned previously, a genetic basis of resistance is worth exploring, especially as resources become available to associate molecular markers with these traits (Ming et al. 2008).

Genetic Variation for Disease Resistance in Papaya's Wild Relatives

Wild relatives of papaya offer a source of genetic variation for traits such as fruit quality and disease resistance. The most diverse of the five genera related to *Carica* is *Vasconcellea*. It includes 21 species (Badillo 2000). Grown at higher elevations, they are commonly referred to as "highland papaya" (National Research Council 1989). *Vasconcellea* is thought to have originated in the region of Ecuador, Colombia, and Peru where maximum species diversity occurs (Scheldeman et al. 2007). Some *Vasconcellea* spp. are used in local cuisine for flavoring or cooked with sugar to make jams (National Research Council 1989). It may be possible to use these species for enhancing or altering papaya's flavor. Currently, the only highland papaya grown extensively outside of its region of origin is "Babaco," a sterile hybrid (*Vasconcellea* × *heilbornii*) (Kyndt and Gheysen 2007) that produces large parthenocarpic fruit that tastes like "strawberry with a hint of pineapple" (National Research Council 1989). "Babaco" has been evaluated for commercial production in a number of countries with some success, including New Zealand, Australia, Spain, France, the United Kingdom, Switzerland, Italy, the Netherlands, South Africa, and Canada (Scheldeman et al. 2007; Kempler and Kabaluk 1996). Before "Babaco" can be fully commercialized, a reduction in production cost and consumer education must be addressed (Kempler and Kabaluk 1996). In addition, greenhouse production of "Babaco" is limited by fusarium wilt (Ochoa et al. 2000). Possible sources of resistance for this pathogen and those affecting papaya are other members of *Vasconcellea*.

In the mid-1960s, a number of *Vasconcellea* species were screened for PRSV-P resistance. *V. cundinamarcensis* and *V. quercifolia* were found to be resistant (Conove 1964). A separate study found *V. cauliflora* and its F_1 hybrids from a cross with a susceptible species (*C. monoica*) resistant to PRSV-P (Horovitz and Jiménez 1967). Attempts to introgress this resistance into papaya through crosses with *V. cauliflora* have been mostly unsuccessful due to postzygotic barriers, including embryo abortion, abnormal endosperm development, and polyembryony (Manshardt and Wenslaff 1989a). In contrast, crosses of *C. papaya* to *V. quercifolia* results in fewer postzygotic disruptions and can be grown in the field (Manshardt and Wenslaff 1989b). In an attempt to improve the success rate of *C. papaya* × *C. cauliflora* hybridization, Magdalita et al. (1998) developed an efficient hybridization protocol, including the use of a more compatible *C. papaya* cultivar, higher quality pollen, and embryo isolation time at 90–120 days postfertilization. Combined with an improved embryo-rescue technique, this protocol resulted in a 94 % embryo germination rate, providing 485 hybrid plants with normal morphology (Magdalita et al. 1996). Unfortunately, although these hybrids were resistant to PRSV-P (Magdalita et al. 1997), none were fertile (Drew et al. 2005a). As a result, the focus for a source of resistance returned to *V. quercifolia*, which is more closely related to *C. papaya* (Jobin-Décor et al. 1997) so that there are fewer postzygotic barriers (Manshardt and Wenslaff 1989b).

Because the Philippine papaya industry was experiencing significant losses due to PRSV-P, the Institute for Plant Breeding (College of Agriculture, University of the Philippines at Los Baños) established a collaborative project with Griffith University (Nathan, Australia) in 2002 to prioritize the introgression of resistance from *V. quercifolia* into elite Philippine inbred lines (Siar et al. 2009). To initiate this process, a resistant male BC_1 plant (line 54) from a *C. papaya* × *V. quercifolia* cross was developed (Drew et al. 2005b). Unfortunately, although micropropagated clones of this line were resistant in Australia, the clones inoculated with a local PRSV-P strain in the Philippines succumbed to disease (Drew et al. 2005b). In a second attempt, a number of inbred lines and F_1 hybrids were crossed, and a BC_1 line found to be resistant after 12 months in a field in Los Baños was selected (Siar et al. 2009). Advanced backcrossing was conducted with this line, and the resulting plants were found to have only mild or delayed virus symptoms with little or no disease progression (Siar et al. 2009). This major accomplishment represents the first successful transfer of disease resistance from a wild relative to papaya and establishes a precedent for developing resistance to other diseases.

Significant levels of partial resistance to PRSV-P (Siar et al. 2009) and other pathogens may be improved by combining sources of resistance from the multiple *Vasconcellea* spp. One way to circumvent compatibility barriers and achieve this goal is by using a bridge species. *V. parviflora* is closely related to *C. papaya* and may be used for this purpose (Jobin-Décor et al. 1997). Resistance genes from more distantly related incompatible *Vasconcellea* spp. might be introgressed into *V. parviflora* and then into *C. papaya* (O'Brien and Drew 2010). In addition, to reducing the cost and variability associated with manual disease screening, molecular markers can be used to track the movement of *Vasconcellea* spp. resistance genes through breeding schemes; see Chap. 19 in this text for details of this process. Using an F_2 *V. cundinamarcensis* × *V. parviflora* mapping population, the PRSV-P resistance of *V. cundinamarcensis* was identified as being regulated by a single, dominant gene (*prsv-1*) (Dillon et al. 2005a). Using this population, a codominant marker (*PsiIk4*) linked to *prsv-1* was developed that can now be used to move resistance from *V. cundinamarcensis* and *V. pubescens* to papaya (Dillon et al. 2005b, 2006a; Drew et al. 2007; O'Brien and Drew 2010). Interestingly, the *PsiIk4* marker is not linked to *V. quercifolia* resistance, suggesting that separate gene(s) regulates this trait (Dillon et al. 2006b). This is encouraging because, if multiple sources of PRSV-P resistance exist, opportunities will exist for achieving more durable resistance by gene pyramiding. Although achieving resistance through interspecific and intergeneric hybridization requires years of work, it avoids the regulatory obstacles associated with transgenic approaches.

In addition to PRSV-P resistance, the diverse *Vasconcellea* species offer resistance for other pathogens as well. Black rot spore inoculation of green and ripened fruit in the field demonstrated that *V. goudotiana* has some resistance to *Mycosphaerella caricae* (Sanchez et al. 1991). *V. monoica*, *V. goudotiana*, and *V. cauliflora* are cited as being resistant to *Cercospora papayae* (black spot), while *V. quercifolia* is noted as being resistant to *Ascochyta caricae-papayae* (Ascochyta leaf spot). After conducting pathogenicity tests, Nishijima and Aragaki (1973) found a low incidence of collar rot

(*Calonectria ilicicola*) on *V. goudotiana*. As for "Babaco's" previously mentioned susceptibility to *Fusarium oxysporum*, *V. weberbaueri* and *V. monoica* offer a source of potential resistance (Scheldeman et al. 2003). *V. goudotiana* may be a possible source of *Phytophthora* resistance (Drew et al. 1998). In preliminary experiments, *V. goudotiana* exhibited rate-limiting resistance, characterized by mild symptoms associated with *P. palmivora* infection that was later outgrown (Zhu and Porter, unpublished data). This reaction is similar to that afforded by the nucleotide binding site-leucine-rich repeat (NBS-LRR) resistance gene, *RB*, isolated from wild potato (*Solanum bulbocastanum*) (Song et al. 2003). Additional studies must be conducted to further characterize this response. Finally, *V. cauliflora* may have another source of *Phytophthora* resistance (Erwin and Ribeiro 1996; Zentmyer and Mitchell 1985/1986).

High-throughput next-generation sequencing offers a means to survey transcriptomes for genes regulating this resistance, while microarray technology can monitor expression changes. We emphasize that because *Vasconcellea* spp. are an invaluable sources of diversity for papaya, it is of utmost concern that 5 of the 21 species are considered threatened (Scheldeman et al. 2007). Others suggest the number of threatened *Vasconcellea* spp. might be even higher due to the rate of deforestation, especially in the species-rich "hybrid zones" that exhibit high morphological variability (Kyndt and Gheysen 2007).

Transgenic Resistance in *C. papaya*

Coat Protein-Mediated Resistance to Virus

In papaya, coat protein-mediated resistance (CP-MR) has been remarkably effective; reviewed by Gonsalves 1998; see also Chap. 7 in this text. Preceding the development of this technology, one strategy for virus control involved exposing plants to a mild or "weaker" virus strain to achieve "cross protection" (Yeh et al. 1988). The exact mechanisms of cross protection are still being revealed (reviewed by Ziebell and Carr 2010), but the added labor costs, risk of mild symptom development, and risk of virus reversion to a more virulent strain led to low adoption rates (Gonsalves 1998). At the time, Sanford and Johnston (1985) proposed an alternative strategy, that if host cells themselves were engineered to produce key pathogen gene products, either in excess or in a dysfunctional form, pathogenicity could be disrupted. The laboratory of Dr. Roger Beachy validated this hypothesis in plants by expressing a tobacco mosaic virus coat protein gene in tobacco, resulting in delayed disease development and resistance (Abel et al. 1986). Like cross protection, the exact mechanism of CP-MR was unknown at the time, but the results were encouraging enough to justify evaluation of the strategy in papaya for controlling PRSV-P. Ultimately CP-MR was shown to be highly successful in papaya (discussed in Chap. 7 in this text), and, conceivably, such an analogous strategy could be applied

to control other diseases. However, before expanding the application of this approach, there are several lessons that can be learned from papaya CP-MR that must first be considered.

Of the two cultivars developed, SunUp, which is homozygous for the transgene, was found to be more resistant than Rainbow, an F_1 hybrid from SunUp which is hemizygous for the PRSV-P coat protein gene (Tennant et al. 2001; see Chap. 6 in this text for a historical discussion). The mechanism of resistance was discovered to be RNA-mediated homology-dependent posttranscriptional gene silencing (PTGS), which targets the virus in a dose-dependent manner (Tennant et al. 2001; Baulcombe 1996). While CP-MR has worked well for the virus strain in Hawaii, sequence divergence of the CP gene among p-type viruses was found to be as high as 12 % (Gonsalves 1998) so the same construct may not provide the same level of protection against other PRSV strains. When challenged with a virus isolate from Thailand having only 89.5 % homology to the Hawaii strain, "SunUp" resistance broke down (Tennant et al. 2001).

In addition to the problems associated with CP sequence divergence, potyvirus helper component-proteinase (HC-Pro) also contributes to the suppression of PTGS, providing another mechanism for resistance breakdown (Mangrauthia et al. 2010). Fortunately, PTGS is only one of many mechanisms of CP-MR. Expression of tobacco mosaic virus CP prevents the virus from uncoating and regulates viral movement protein production (Register and Beachy 1988; Ling et al. 1991; Bendahmane et al. 2002; Asurmendi et al. 2004). The CP-MR of potato virus X is not significantly dependent on PTGS (Bazzini et al. 2006). Therefore, a better understanding of the many control mechanisms will allow the design of multimodal virus protection constructs in the future. Meanwhile, constructs that target local strains and multiple virus types are providing resistance. Using the sequence of local PRSV isolates, CP-MR has been deployed in a number of countries including Jamaica, Venezuela, and Brazil (Tennant et al. 2005; Fermin et al. 2004; Júnior et al. 2005). In Taiwan, papaya lines have been developed with double resistance to PRSV and papaya leaf distortion mosaic virus (Kung et al. 2009; Kung et al. 2010).

In Hawaii, CP-MR currently targets a relatively homogeneous PRSV population (Tripathi et al. 2006, 2008). If a viral strain emerges that breaks down this resistance, additional transformation, perhaps combined with what has already been used, may be needed. As coevolution between transgenic systems and viruses occurs, resources such as selectable markers will need to be managed, especially when combining multiple constructs into the same plant line.

Transgene insertions occurred in three locations in the "SunUp" genome (Ming et al. 2008). If all three insertions contribute to resistance, this may allow for loss of function of some copies over time. In addition, gene divergence may occur, including alterations in promoter regions that could result in changes in gene regulation. These scenarios are interesting to consider from a plant–pathogen evolutionary standpoint and may be more plausible than expected considering the worldwide distribution of the technology.

Stilbene Synthase

For the control of root rot, transgenic expression of the grapevine stilbene synthase (a resveratrol synthase) gene (*Vst1*) was evaluated in papaya (Zhu et al. 2004). Using the native grapevine pathogen-inducible promoter, transgenic lines produced the phytoalexin resveratrol and displayed increased resistance (Zhu et al. 2004). However, these plants failed to set fruit (unpublished data). Similar deleterious effects associated with excessive stilbene production have been previously reported, including abnormal pollen development, parthenocarpy, and male sterility (Ingrosso et al. 2011; Fischer et al. 1997). These effects are the result of resveratrol synthase competing for the same substrates as chalcone synthase, 4-coumaroyl CoA, and malonyl CoA (Fischer et al. 1997). Chalcone synthase requires these precursors to synthesize the scaffold required for the production of all flavonoids (Ferrer et al. 2008). In addition, these substrates are required in other pathways for the production of structural compounds, including lignin and sporopollenin (Ingrosso et al. 2011). Because of overproduction or mislocalization of resveratrol, synthase has the potential to impact other pathways. Genes encoding these enzymes are frequently pathogen/stress inducible and regulated in specific tissues. The stilbene synthase gene of sorghum, *SbSTS1,* is induced by host and nonhost pathogens (Yu et al. 2005). In grapevine, stilbene synthase is found in infected cells and in the exocarp of the berry where infection is likely to occur (Schnee et al. 2008; Fornara et al. 2008). Successful heterologous production of resveratrol synthase in papaya, therefore, will in part require tissue-specific, pathogen-inducible promoters. Expression characterizations of a number of papaya genes with promoters fitting these criteria have been identified (Porter et al. 2008, 2009b).

Heterologous expression of grapevine stilbene synthase for the control of *Botrytis cinerea* infection of tobacco was first demonstrated more than 18 years ago (Hain et al. 1993). Since then, the strategy has been evaluated in a number of crops, including tomato, for the control of *Phytophthora infestans* (Thomzik et al. 1997), rice for the control of *Pyricularia oryzae* (Stark-Lorenzen et al. 1997), wheat and barley for a number of fungal pathogens (Leckband and Lörz 1998; Serazetdinova et al. 2005), alfalfa for the control of *Phoma medicaginis* (Hipskind and Paiva 2000), and other plant species (reviewed by Delaunois et al. 2009). However, to date, no crops transformed with stilbene synthase have received regulatory approval (CERA 2010). This may be due to the fact that the current state of the technology has only achieved partial resistance and failed to prove effective in the field. Transformation of tomato with grapevine stilbene synthase resulted in a range of disease reduction for *P. infestans* (between 38 and 68 %) but provided no significant control of *B. cinerea* and *Alternaria solani* (Thomzik et al. 1997). Control of *P. medicaginis* in alfalfa transformed with a cDNA encoding resveratrol synthase was demonstrated using leaf inoculations (Hipskind and Paiva 2000) but will require larger trials to determine production-scale disease control. Disease symptoms of wheat transformed with stilbene synthase following inoculation with *Puccinia recondita* f. sp. *tritici* were reduced by 19 ± 9 % to 27 ± 8 % (Serazetdinova et al. 2005). Finally, in papaya

expressing *Vst1*, 50 % of transgenic plants remained healthy following inoculation with *P. palmivora*, while 25 % of the untransformed controls remained healthy (Zhu et al. 2004).

Looking to the future, the use of stilbene synthases to control fungal and oomycete pathogens holds promise. Resveratrol synthase generates the backbone molecule, resveratrol, from which its derivatives, piceid, viniferins, and pterostilbene, are derived. Pterostilbene, a dimethylated derivative of resveratrol, was found to have threefold the activity of resveratrol and rapidly destroys the plasma membrane of *B. cinerea* (Adrian et al. 1997; Pezet and Pont 1990). Recently, a gene encoding a pathogen-regulated resveratrol O-methyltransferase (ROMT) for pterostilbene biosynthesis was isolated from grapevine (Schmidlin et al. 2008). Therefore, the use of *ROMT* in combination with resveratrol synthase is suggested to be a more effective strategy (as described next).

Genes encoding stilbene synthases are thought to have evolved independently from chalcone synthases in a diverse but relatively small number of plant species (Tropf et al. 1994; Austin and Noel 2003). Examples include peanut (Schöppner and Kindl 1984), pine (Schanz et al. 1992), grapevine (Sparvoli et al. 1994), whisk fern (Yamazaki et al. 2001), *Rheum tataricum* (Samappito et al. 2003), sorghum (Yu et al. 2005), *Polygonum cuspidatum* (Liu et al. 2011), and spruce (Hammerbacher et al. 2011). In the majority of cases, these genes are pathogen-inducible (Preisig-Müller et al. 1999; Yu et al. 2005; Hammerbacher et al. 2011). Regulation of this pathway, however, does not end at the production of resveratrol (or pinosylvin) backbone molecules. In *V. vinifera* and *Arachis hypogaea*, differential accumulation of resveratrol derivatives between genotypes demonstrates that regulation of enzymatic modifications, such as glycosylation, oxidation, and methylation (in the case of ROMT), is critical for effective defense responses (Pezet et al. 2004; Sobolev et al. 2007; Schmidlin et al. 2008). A transgenic approach involving multiple genes will likely be required to maximize disease resistance from stilbenes. This approach has begun to be evaluated in tobacco and *Arabidopsis* through the co-expression of genes for O-methyltransferase and stilbene synthase (Rimando et al. 2012). For the control of *P. palmivora* of papaya, an attractive model for evaluating early-stage multigene regulation of stilbenes might be *Arabidopsis* and *Hyaloperonospora arabidopsidis*, an oomycete pathogen (Chou et al. 2011).

Dahlia merckii *Antimicrobial Peptide 1 (Dm-AMP1)*

First recognized in mammalian granulocytes, defensins are small, cysteine-rich, amphipathic peptides that permeabilize pathogen membranes, particularly those of fungi (Zeya and Spitznagel 1963; reviewed by Ganz 2003). Similar peptides have been identified in invertebrates, plants, and fungi, suggesting these ubiquitous components of innate immunity likely evolved from a common, ancient progenitor (reviewed by Wilmes et al. 2011; Zhu 2007). The defensin, *D. merckii* antimicrobial peptide 1 (Dm-AMP1), was first isolated from *D. merckii* (bedding dahlia) seed

(Osborn et al. 1995). Bioassays conducted using this defensin inhibited germ tube elongation rate, reduced hyphal thickness, and destroyed the cytoplasm of some fungi and inhibited the growth of *Bacillus subtilis* (Osborn et al. 1995). Interestingly, the binding of radioactively labeled Dm-AMP1 to *Neurospora crassa* and *Saccharomyces cerevisiae* cells can be blocked by preincubation with "cold" Dm-AMP1, but not by unrelated defensins (Thevissen et al. 2000a). This suggested that Dm-AMP1 binds a specific site on the target plasma membrane, a hypothesis supported by the identification of mutant *S. cerevisiae* that is resistant to Dm-AMP1 and demonstrates ten-fold less binding efficiency relative to wild type (Thevissen et al. 2000a). To determine the genetic basis of this loss of binding, a genomic library was constructed from susceptible, wild-type yeast and used to transform resistant mutants (Thevissen et al. 2000b). A clone encoding an enzyme that catalyzes the formation of sphingolipids (terminal sphingolipid mannosyldiinositolphosphophoceramide) was able to restore susceptibility in the mutants, suggesting this plasma membrane component is the Dm-AMP1 binding site (Thevissen et al. 2000b). This was confirmed using an enzyme-linked immunosorbent assay (ELISA), which demonstrated that Dm-AMP1 directly interacts with sphingolipids (Thevissen et al. 2003).

Constitutive expression of *Dm-AMP1* in papaya provided resistance to *P. palmivora* (Zhu et al. 2007). Leaf protein extract containing Dm-AMP1 inhibited hyphae growth by 35–50 %, and inoculated leaf discs from transformed plants had 40–50 % less infected area than controls (Zhu et al. 2007). The disease ratings of papaya plants expressing *Dm-AMP1* were significantly less than that of controls following root-drench inoculation (Zhu et al. 2007). Similarly, *Dm-AMP1* expressed in rice significantly suppressed the growth of *Magnaporthe oryzae* and *Rhizoctonia solani* (Jha et al. 2009). In *Solanum melongena*, Dm-AMP1 inhibited *Botrytis cinerea* in leaves, and root exudates containing the protein reduced the growth of *Verticillium albo-atrum* (Turrini et al. 2004a). Field trials will need to be conducted to evaluate the efficacy of *Dm-AMP1* in larger-scale production, with particular attention paid to gene durability.

Previously, plant defensins were evaluated in other crop–pathogen systems (Terras et al. 1995). Expression of a pea defensin (*DRR230*) in Canola targeted *Leptosphaeria maculans* (Wang et al. 1999). Monsanto Company successfully demonstrated the use of an alfalfa defensin (*alfAFP*) in potato for the control of *Verticillium dahliae* (Gao et al. 2000). Although *alfAFP* proved particularly effective in controlling *V. dahliae*, Monsanto's potato biotechnology program was halted in 2001 due to lack of market support (Gao et al. 2000; Kilman 2001). Nevertheless, as Dm-AMP1 and other defensins progress toward production-scale applications, strategies to promote durability should be prioritized. These peptides play a key role in innate immunity. *S. cerevisiae* mutants were resistant to Dm-AMP1 (Thevissen et al. 2000a), so defensin vulnerability to pathogen mutation under high selection pressure could undermine endogenous resistance in papaya and other species. For long-term durability, simply expressing defensins constitutively at high levels may be found to be too simplistic an approach.

Natural expression is more complex. The radish defensin genes *Rs-AFP3* and *Rs-AFP4*, for example, are pathogen-inducible in leaves, while Rs-AFP1 and

Rs-AFP2 accumulate in specific cell layers of the seed and are released during germination (Terras et al. 1995). Similarly, *PDF1.2*, an *Arabidopsis* pathogen-inducible defensin gene, is regulated by a jasmonate-dependent/salicyclic acid-independent pathway (Penninckx et al. 1996; Thomma et al. 1998). This suggests that, in nature, defensins are highly regulated and that avoiding resistance breakdown may require regulated expression and/or more complex multigene strategies.

Manduca sexta *Chitinase*

Chitin is an abundant biological polymer found in many organisms including fungi, arthropods, and crustaceans. Modification and destruction of this structural polysaccharide occurs in part by hydrolysis of its glycosidic bonds, catalyzed by chitinases. Chitinases are near ubiquitous in nature, occurring in organisms with and without endogenous chitin, including mammals, amphibians, arthropods, nematodes, fungi, bacteria, and baculoviruses. In organisms with chitin, chitinases are generally used for developmental purposes, whereas other organisms have evolved chitinases for defense or pathogenicity. Chitotriosidase, for example, is a human chitinase secreted from phagocytes as part of the immune system for the degradation of chitin-containing pathogens (Boot et al. 2001). Examples of chitinases contributing to pathogenicity come from the malaria parasite and a baculovirus. *PfCHT1*, a gene from the human malaria parasite (*Plasmodium falciparum*), encodes a chitinase that contributes to disease transmission by allowing the pathogen to escape the midgut of mosquitoes (Vinet et al. 1999). Cathepsin (a cysteine protease) and chitinase A from the baculovirus AcMNPV act together in the liquefaction of insect hosts (Hawtin et al. 1997). Finally, in insects, chitinase activity is highly regulated in precise fashion for elaborate developmental processes such as molting. Recently, in the red flour beetle (*Tribolium castaneum*), it was shown that Knickkopf protein protects new cuticle formation from chitinase found in molting fluid (Chaudhari et al. 2011). Disruption of such processes can be deleterious. Downregulation of the gene encoding Knickkopf protein is lethal, making it a potential target for biocontrol (Chaudhari et al. 2011). Similarly, ectopic expression of chitinase in plants can be used as a control strategy as demonstrated by overexpression of *M. sexta* (tobacco hornworm) chitinase in tobacco for the control of tobacco budworm and hornworm (Ding et al. 1998).

C. papaya was transformed with *M. sexta* chitinase (MSCH) under the control of the constitutive (CaMV 35S) promoter (McCafferty et al. 2006). Ten weeks post-inoculation in the laboratory with carmine spider mites (*Tetranychus cinnabarinus* Boisd.), all transgenic lines had a significantly higher number of leaves relative to the susceptible donor cultivar "Kapoho" (McCafferty et al. 2006). However, only one transgenic line (T-24) had significantly fewer mites per leaf than the control. This most likely occurred as a result of the control having fewer leaves, forcing the mites to migrate to the transgenic plants (McCafferty et al. 2006). Conversely, in the field, all transgenic lines expressing *MSCH* had fewer mites than the control, which

suggests that when the mites have a choice, they prefer to avoid chitinase-expressing lines (McCafferty et al. 2006). These results are particularly encouraging. Rather than functioning as an insecticide, MSCH appears to deter feeding and encourage migration. Because *T. cinnabarinus* has a large host range, including many weed species (Goff 1986), movement of mites from transgenic plants to alternate hosts could, in theory, occur with minimal selection pressure, effectively promoting *MSCH* durability.

Recently, corn plants engineered to express the insecticidal *Bacillus thuringiensis* (Bt) toxin Cry3Bb1 for the control of western corn rootworm (*Diabrotica virgifera virgifera*) were found to be susceptible in some fields in Iowa, illustrating the consequences of high selection pressure (Gassmann et al. 2011). Plants expressing Cry34/35Ab1 were found to be resistant to the problem rootworm, but pathogen resistance could emerge for this line as well (Gassmann et al. 2011). A combination of resistance sources combining Cry3Bb1 and Cry34/35Ab1 (SmartStax) may delay the evolution of pathogen resistance (Gassmann et al. 2011; EPA 2009). Strategies such as combining genes for chitinase and scorpion toxin, which have been determined to cause high larvae mortality, should be evaluated to determine if this selective combination is durable (Wang et al. 2005). Finally, the environmental impact of transformations using chitinase genes should be considered. In papaya, confirmation is needed to ensure that pollinating insects are unaffected by *MSCH*. In addition, papaya expressing *MSCH* should be evaluated for resistance to fruit flies and mites other than *T. cinnabarinus*. Although aphids do not colonize papaya, they transmit PRSV-P to papaya in a nonpersistent manner by conducting exploratory probes (Pantoja et al. 2002; Kalleshwaraswamy and Kumar 2008). The possible influence that MSCH may have on this behavior should be explored as well.

Papaya Mutualistic and Protective Endophytes

To ensure that beneficial microbes are not affected by transgenic modifications for disease resistance, it is sometimes necessary to survey and select for lines that maintain compatibility with mutualistic endophytes. Up to 90 % of terrestrial plants form mycorrhizal-root associations (Fitter and Moyerson 1996), but some, including papaya, are considered highly dependent upon arbuscular mycorrhizal fungi (AMF) for inorganic phosphorus (P_i) uptake (Miyasaka and Habte 2001). In addition, some endophytes also provide protection against insects, nematodes, and other pathogens (Vega et al. 2008; Jaizme-Vega et al. 2006; Stein et al. 2008).

There are instances of transgenes affecting AMF. Tobacco constitutively expressing a pathogenesis-related protein (PR-2) delayed *Glomus mosseae* colonization, whereas *G. mosseae* was resistant to constitutive chitinase expression in tobacco and *Nicotiana sylvestris* (Vierheilig et al. 1993, 1995). Because defensins can inhibit a range of fungi (Osborn et al. 1995), transformations using genes such as *Dm-AMP1* might inhibit endophytes. Fortunately, Dm-AMP1's inhibition of pathogen growth has been shown to spare some beneficial mycorrhizae. *Solanum melongena*

transformed with *Dm-AMP1* inhibited the pathogenic fungi, *Botrytis cinerea* and *Verticillium albo-atrum,* while the arbuscular mycorrhizal fungus *G. mosseae* was able to established host recognition, initiate symbiosis, and promote host plant growth (Turrini et al. 2004a, b). Examination of the possible effects of Dm-AMP1 on other beneficial, nontarget microorganisms in other host systems, such as papaya and rice (Zhu et al. 2007; Jha et al. 2009), will determine if the observed AMF resistance is an exception or trend. In papaya, *G. mosseae* not only contributes significantly to plant phosphorus uptake but, along with *G. manihotis*, significantly reduces the reproduction of the parasitic nematode *Meloidogyne incognita* (Jaizme-Vega et al. 2006; Rodriguez-Romero et al. 2011). A comprehensive study of these AMF in papaya expressing *Dm-AMP1* will require phosphorus evaluations and nematode bioassays. In maize, one line (*Bt* 176) with high expression of *CryIAb* toxin negatively affected *G. mosseae* pre-symbiotic hyphal growth and appressoria development, but another line, *Bt* 11, was indistinguishable from the non-transgenic control (Turrini et al. 2004b).

This suggests that selection for AMF-compatible lines is possible. The next step is the development of more rapid high-throughput monitoring. Arnold et al. (2000) used plating techniques to isolate endophytes representing 347 genetically distinct taxa from the leaves of two tropical tree species, *Heisteria concinna* (Olacaceae) and *Ouratea lucens* (Ochnaceae). Screening techniques have been developed for evaluating the impact of transgenes on AMF (Turrini et al. 2004b), but a comprehensive DNA-based screen that captures difficult-to-culture microorganisms may be needed (Mlot 2004). While the elimination of endophytes is one concern, another possible consequence of transgene selection pressure is conversion of endophytes from mutualists to pathogens. Mutation of a single NADPH oxidase gene was shown to disrupt reactive oxygen species (ROS) production in the endophyte *Epichloë festucae*, causing the death of its host *Lolium perenne* (perennial ryegrass) (Tanaka et al. 2006). Conversely, Freeman and Rodriguez (1993) used UV mutagenesis to demonstrate conversion of the pathogen *Colletotrichum magna* into a protective endophyte (Freeman and Rodriguez 1992; Freeman and Rodriguez 1993; Redman et al. 1999). In the future, transgenic strategies designed to promote mutualistic and protective endophytes in papaya may enhance yield and pathogen resistance. One particularly attractive candidate for this application is *Piriformospora indica*. An AMF isolated from woody shrubs from Rajasthan's Thar Desert, *P. indica*, has been associated with disease resistance and higher yield (Verma et al. 1998; Verma and Sharma 1999; Waller et al. 2005; Shahollari et al. 2007; Stein et al. 2008).

The Nucleotide Binding Site-Leucine-Rich Repeat (NBS-LRR) Gene Family and *P. palmivora* Resistance

Solanum spp. and *P. infestans* provide an analogous host–pathogen system for guiding the development of *P. palmivora* resistance in papaya. Wild potatoes and *Vasconcellea* spp. both occur in the tropical highlands at average altitudes of

~1,500 m and ~2,800 m, respectively, with overlapping geographical regions of species richness (Hijmans and Spooner 2001; Scheldeman et al. 2007). The diversity of *Solanum* species is highest in Mexico, Peru, Bolivia, and Argentina, while the maximum diversity of *Vasconcellea* spp. is found in Ecuador, Colombia, and Peru (Hijmans and Spooner 2001; Scheldeman et al. 2007). To date, 21 *P. infestans* resistance genes have been cloned from *Solanum* spp., reflecting extensive coevolution with a pathogen that shares a center of origin in the central highlands of Mexico (Vleeshouwers et al. 2011; Grünwald and Flier 2005). Eighteen of these genes originate from species found in Mexico, and four originate from species from Argentina (Vleeshouwers et al. 2011).

The story of *P. palmivora* is somewhat more complex in that the duration of its coevolution with papaya's wild relatives (i.e., *Vasconcellea* spp.) is uncertain. It has been suggested that *P. palmivora* originated from Central or South America (Zentmyer 1988), but the diversity of isolates identified from coconut (*Cocos nucifera*), durian (*Durio zibethinus*), and other Southeast Asia hosts points instead to a Southeast Asia center of origin (Mchau and Coffey 1994). If this is true, *P. palmivora* may have only recently spread from Asia, and the evolution of *Vasconcellea* spp. resistance gene(s) specificity may be the result of more modern selection pressure. Nevertheless, what appears to be a rate-reducing form of resistance similar to that of *Rpi-blb1* (*RB*) (Song et al. 2003; van der Vossen et al. 2003) has been identified in *V. goudotiana* (Zhu and Porter, unpublished data). *Rpi-blb1* is generally considered a broad-spectrum, durable source of resistance, and it would be encouraging to find similar resistance for papaya. The only exceptions are two *P. infestans* isolates from Mexico (PIC99189 and PIC99177) lacking an effector variant (class I *ipiO*) that were recently determined to be virulent in the presence of *Rpi-blb1* (Champouret et al. 2009). Although resistance can break down and *P. infestans* has the reputation of being an "R gene destroyer" (Fry 2008), the *P. palmivora* resistance observed in *V. goudotiana* is worth exploring and if isolated, perhaps combined with known sources of partial resistance (Noorda-Nguyen et al. 2010; Dianese et al. 2007, 2010).

All *P. infestans* resistance genes cloned to date belong to the nucleotide binding site-leucine-rich repeat (NBS-LRR) gene family (Vleeshouwers et al. 2011). Possibly, the *P. palmivora* resistance genes observed in *V. goudotiana* and some papaya genotypes (Zhu and Porter, unpublished data; Noorda-Nguyen et al. 2010; Dianese et al. 2007, 2010) are also members of this family. From the draft genome of *C. papaya*, 54 NBS class resistance genes have been identified (Ming et al. 2008; Porter et al. 2009a). This is substantially fewer than the number found in other plant genomes (Table 15.3), including *Arabidopsis*, which has 174 NBS genes (*Arabidopsis* Genome Initiative 2000).

While few in number, papaya's NBS genes represent both Toll/interleukin-1 receptor (TIR) and non-TIR subclasses found as clusters and single genes throughout the genome (Fig. 15.1) (Porter et al. 2009a). Unlike *Arabidopsis*, whole genome duplication has not occurred in the papaya lineage since its divergence from *Arabidopsis* (Ming et al. 2008; Sémon and Wolfe 2007; see also discussion in Chap. 11 of this text). The lack of genome duplication may partially explain the

Table 15.3 The total number of predicted NBS-encoding genes identified in five sequenced angiosperm genomes

Species	Total number of predicted protein-encoding genes	Total number of predicted NBS-encoding genes	Genome size (Mb)	Source
Carica papaya	24,746	54	372	Ming et al. (2008)
Arabidopsis thaliana	25,498	174	125	Arabidopsis Genome Initiative (2000)
Vitis vinifera	30,434	535	487	Jaillon et al. (2007)
Oryza sativa	37,544	519	389	International Rice Genome Sequencing Project (2005)
Populus trichocarpa	45,555	416	485	Tuskan et al. (2006)

Values for *Arabidopsis thaliana*, *Vitis vinifera*, *Oryza sativa*, and *Populus trichocarpa* were previously summarized by Yang et al. (2008). The total number of predicted protein-encoding genes and the genome size of each species are also provided (see source column for references)

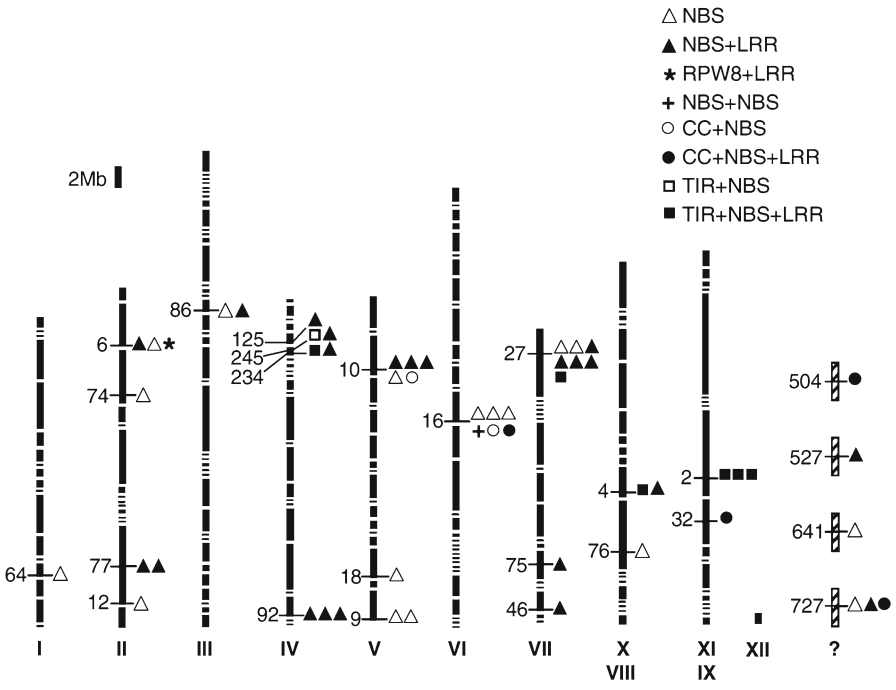

Fig. 15.1 Distribution of predicted *Carica papaya* NBS-encoding genes across linkage groups. The papaya genome sequence was anchored to the 12 papaya linkage groups as described by Ming et al. (2008) [reproduced with kind permission of Springer Science+Business Media from Porter et al. (2009a)]

scarcity of NBS-LRR genes, but it is also possible that papaya preferentially relies upon surveillance, or "guarding," of common effector targets to detect large numbers of pathogens with relatively few NBS genes (van der Biezen and Jones 1998; Dangl and Jones 2001; DeYoung and Innes 2006; McDowell and Simon 2006). Papaya also has a lower total number of genes than other sequenced angiosperms (Ming et al. 2008; *Arabidopsis* Genome Initiative 2000; Jaillon et al. 2007; International Rice Genome Sequencing Project 2005; Tuskan et al. 2006), so it is possible that fewer NBS genes are required for surveillance (Porter et al. 2009a).

Finally, some NBS genes may reside in the limited portion of the genome lacking sequence coverage, but because 90 % of the euchromatic chromosomal regions have been sequenced, finding a significant number of additional NBS genes is considered unlikely (Ming et al. 2008; Porter et al. 2009a). Now that papaya's NBS genes have been mapped (Fig. 15.1), susceptible and partially resistant cultivars (Noorda-Nguyen et al. 2010; Dianese et al. 2007, 2010) may be compared using targeted sequencing techniques (reviewed by Mamanova et al. 2010).

In the future, native R-genes may be ineffective for the control of *P. palmivora*, but a better understanding of the molecular basis of *Phytophthora* pathogenicity may provide opportunities to modify NBS genes or effector targets to achieve resistance. A first step in this process is determining pathogen host recognition and translocation of effectors from *Phytophthora* haustoria into the plant cell. Two N-terminal-conserved motifs identified in *P. infestans* effectors are RXLR and EER, which serve as a host cell uptake (penetration) signal. In *P. infestans*, 425 genes of this protein class have been identified (Birch et al. 2006; Whisson et al. 2007). The exact mechanism of effector entry is uncertain, and recent studies have reached contradictory conclusions (Ellis and Dodds 2011). Kale et al. (2010) suggest that phospholipid, phosphatidylinositol-3-phosphate (PI3P), found on the surface of plant cell plasma membranes mediates effector entry. Yaeno et al. (2011) suggest that PIP binding contributes intracellularly, promoting effector stabilization, accumulation, and virulence function. Resolving the exact mechanism of this process is important as it may lead to upstream resistance strategies to block effector entry. The virulence functions of *P. infestans* effectors are beginning to be revealed. AVR3a, for example, has been shown to act upstream at the plasma membrane by inhibiting the host ubiquitin E3-ligase, CMPG1, required for plant immunity (Bos et al. 2010; Gilroy et al. 2011). Interestingly, in papaya, *P. palmivora* infection is associated with reduction of a transcript encoding a putative aquaporin (Porter et al. 2009a, b). Similar aquaporin repression has been reported in other plant systems, including cotton following *Fusarium oxysporum* f. sp. *vasinfectum* inoculation (Dowd et al. 2004). Aquaporins play a role in hydraulic permeability and have been shown to be targets of bacterial effector regulation in animal disease (Guttman et al. 2007). Further investigation will be required to determine if *P. palmivora* effectors regulate papaya aquaporins, either directly or indirectly. Regardless, once effector targets are determined, they may be modified for resistance.

Emerging Diseases of Papaya

Adaptation and evolution increase pathogen diversity, a process that often begins with the spread of disease into new environments. The first reports of disease in papaya (Table 15.2) suggest that this phenomenon is active. In 2001, for example, black spot disease of papaya [*A. caricae* (Speg.) Maulbl.] was discovered on the island of Maui and subsequently on other Hawaiian Islands (Ogata and Heu 2001). Outbreaks of black spot now require the application of costly fungicides. Early detection of emerging diseases can provide an opportunity to implement cultural practices to help delay the spread of disease until tolerant cultivars are obtained for production. Maintaining genetic diversity in the field will hedge against losses and slow disease spread. Expanding niche markets, such as those that utilize larger-fruited papaya, is one example of how diversification may be achieved. The management of alternate hosts, including weeds, provides another means to mitigate and monitor pathogen movement (Chin et al. 2007). Genetic characterization of pathogen diversity can also be used to predict the likelihood of disease outbreaks (Gibb et al. 1998; Maoka and Hataya 2005). Ultimately, however, an understanding of the molecular basis of host–pathogen interaction will be needed to allow for resistance to be engineered or selected for.

Recently, a proteinase (NIaPro) of the virus nuclear inclusion body was shown to regulate PRSV host specificity (Chen et al. 2008), which offers insight for the possible disruption of host recognition. Separately, PRSV helper component-proteinase (HC-Pro) was found to interact with papaya calreticulin, suggesting the involvement of calcium signaling in infection or defense (Shen et al. 2010). This and other host–pathogen interactions may be regulated for creating resistance.

References

Abel PP, Nelson RS, De B, Hoffmann N, Rogers G, Fraley RT, Beachy RN (1986) Delay of disease development in transgenic plants that express the tobacco mosaic virus coat protein gene. Science 232:738–743

Adrian M, Jeandet P, Veneau J, Weston LA, Bessis R (1997) Biological activity of resveratrol, a stilbenic compound from grapevines, against *Botrytis cinerea*, the causal agent for gray mold. J Chem Ecol 23:1689–1702

Aluja M, Jimene A, Camino M, Aldana L, Castrejon V, Valdes YME (1994) Determinacion de la susceptibilidad de tres variedades de papaya (*Carica papaya*) al ataque de *Toxotrypana curvicauda* (Diptera: Tephritidae). Folia Entomol Mex 90:33–42

Alvarez AM, Nishijima WT (1987) Postharvest diseases of papaya. Plant Disease 71:681–686

Aradhya MK, Manshardt RM, Zee F, Morden CW (1999) A phylogenetic analysis of the genus *Carica* L. (Caricaceae) based on restriction fragment length variation in a cpDNA intergenic spacer region. Genet Resour Crop Evol 46:579–586

Arnold AE, Maynard Z, Gilbert GS, Coley PD, Kursar TA (2000) Are tropical fungal endophytes hyperdiverse? Ecol Lett 3:267–274

Arocha Y, Horta D, Peralta E, Jones P (2003) First report on molecular detection of phytoplasmas in papaya in Cuba. Disease notes. Plant Disease 87:1148

Arocha Y, Piñol B, Picornell B, Almeida R, Jones P (2006) First report of a 16SrII ("Candidatus Phytoplasma aurantifolia") group phytoplasma associated with a bunchy-top disease of papaya in Cuba. New disease report. Plant Pathol 55:821

Arocha Y, Bekele B, Tadesse D, Jones P (2007) First report of a 16SrII group phytoplasma associated with die-back diseases of papaya and citrus in Ethiopia. New disease report. Plant Pathol 56:1039

Arocha Y, Vigheri N, Nkoy-Florent B, Bakwanamaha K, Bolomphety B, Kasongo M, Betts P, Monger WA, Harju V, Mumford RA, Jones P (2008) First report of the identification of Moroccan watermelon mosaic virus in papaya in Democratic Republic of Congo. New disease report. Plant Pathol 57:387

Asurmendi S, Berg RH, Koo JC, Beachy RN (2004) Coat protein regulates formation of replication complexes during tobacco mosaic virus infection. Proc Natl Acad Sci USA 101:1415–1420

Austin MB, Noel JP (2003) The chalcone synthase superfamily of type III polyketide synthases. Nat Prod Rep 20:79–110

Badillo VM (2000) *Carica* L. vs. *Vasconcellea* St. Hil. (Caricaceae) con la rehabilitacion de este ultimo. Ernstia 10:74–79

Bau HJ, Kung YJ, Raja JA, Chan SJ, Chen KC, Chen YK, Wu HW, Yeh SD (2008) Potential threat of a new pathotype of *Papaya leaf distortion mosaic virus* infecting transgenic papaya resistant to *Papaya ringspot virus*. Phytopathology 98:848–856

Baulcombe DC (1996) Mechanisms of pathogen-derived resistance to viruses in transgenic plants. Plant Cell 8:1833–1844

Bazzini AA, Hopp HE, Beachy RN, Asurmendi S (2006) Posttranscriptional gene silencing does not play a significant role in *Potato virus X* coat protein-mediated resistance. Phytopathology 96:1175–1178

Bendahmane M, Szecsi J, Chen I, Berg RH, Beachy RN (2002) Characterization of mutant tobacco mosaic virus coat protein that interferes with virus cell-to-cell movement. Proc Natl Acad Sci USA 99:3645–3650

Birch PR, Rehmany AP, Pritchard L, Kamoun S, Beynon JL (2006) Trafficking arms: oomycete effectors enter host plant cells. Trends Microbiol 14:8–11

Boot RG, Blommaart EF, Swart E, Ghauharali-van der Vlugt K, Bijl N, Moe C, Place A, Aerts JM (2001) Identification of a novel acidic mammalian chitinase distinct from chitotriosidase. J Biol Chem 276:6770–6778

Bos JI, Armstrong MR, Gilroy EM, Boevink PC, Hein I, Taylor RM, Zhendong T, Engelhardt S, Vetukuri RR, Harrower B, Dixelius C, Bryan G, Sadanandom A, Whisson SC, Kamoun S, Birch PR (2010) *Phytophthora infestans* effector AVR3a is essential for virulence and manipulates plant immunity by stabilizing host E3 ligase CMPG1. Proc Natl Acad Sci USA 107:9909–9914

CERA (2010) GM crop database. Center for Environmental Risk Assessment (CERA), ILSI Research Foundation, Washington, DC. http://cera-gmc.org/index.php?action=gm_crop_database.

Champouret N, Bouwmeester K, Rietman H, van der Lee T, Maliepaard C, Heupink A, van de Vondervoort PJ, Jacobsen E, Visser RG, van der Vossen EA, Govers F, Vleeshouwers VG (2009) *Phytophthora infestans* isolates lacking class I *ipiO* variants are virulent on *Rpi-blb1* potato. Mol Plant Microbe Interact 22:1535–1545

Chang LS, Lee YS, Su HJ, Hung TH (2003) First report of Papaya leaf curl virus infecting papaya plants in Taiwan. Disease notes. Plant Disease 87:204

Chaudhari SS, Arakane Y, Specht CA, Moussian B, Boyle DL, Park Y, Kramer KJ, Beeman RW, Muthukrishnan S (2011) Knickkopf protein protects and organizes chitin in the newly synthesized insect exoskeleton. Proc Natl Acad Sci USA 108:17028–17033

Chen KC, Chiang CH, Raja JA, Liu FL, Tai CH, Yeh SD (2008) A single amino acid of NIaPro of *Papaya ringspot virus* determines host specificity for infection of papaya. Mol Plant Microbe Interact 21:1046–1057

Chen RS, Wang WL, Li JC, Wang YY, Tsay JG (2009) First report of papaya scab caused by *Cladosporium cladosporioides* in Taiwan. Disease notes. Plant Disease 93:426

Chin M, Ahmad MH, Tennant P (2007) *Momordica charantia* is a weed host reservoir for Papaya ringspot virus Type P in Jamaica. Disease notes. Plant Disease 91:1518

Chou S, Krasileva KV, Holton JM, Steinbrenner AD, Alber T, Staskawicz BJ (2011) *Hyaloperonospora arabidopsidis* ATR1 effector is a repeat protein with distributed recognition surfaces. Proc Natl Acad Sci USA 108:13323–13328

Conove RA (1964) Distortion ringspot, a severe virus disease of papaya in Florida. Fla Agric Exp Stations J Ser 2001:440–444

Conover RA, Litz RE (1978) Progress in breeding papayas with tolerance to papaya ringspot virus. Proc Fla State Hortic Soc 91:182–184

Conover RA, Litz RE, Malo SE (1986) "Cariflora" a papaya ringspot virus tolerant papaya for South Florida and the Caribbean. Hortscience 21:1072

Dangl JL, Jones JD (2001) Plant pathogens and integrated defence responses to infection. Nature 411:826–833

Delaunois B, Cordelier S, Conreux A, Clément C, Jeandet P (2009) Molecular engineering of resveratrol in plants. Plant Biotechnol J 7:2–12

DeYoung BJ, Innes RW (2006) Plant NBS-LRR proteins in pathogen sensing and host defense. Nat Immunol 7:1243–1249

Diallo HA, Monger W, Kouassi N, Yoro DT, Jones P (2007) First report of Papaya ringspot virus infecting papaya in Côte d'Ivoire. New disease report. Plant Pathol 56:718

Dianese AC, Blum LEB, Dutra JB, Lopes LF, Sena MC, Freitas LF, Yamanishi OK (2007) Reaction of papaya genotypes to black-spot and foot-rot. Fitopatol Bras 32:419–423

Dianese AC, Blum LEB, Dutra JB, Freitas LF, Lopes LF, Sena MC, Lima L, Yamanishi OK, Martins DMS (2010) Reaction of papaya (*Carica papaya*) genotypes to foot rot caused by *Phytophthora palmivora*. Acta Hortic 864:249–255

Dillon SK, Drew RA, Ramage C (2005a) Development of a co-dominant SCAR marker linked to a putative PRSV-P resistance locus in "wild papaya". Acta Hortic 694:101–104

Dillon S, Ramage C, Drew R, Ashmore S (2005b) Genetic mapping of a PRSV-P resistance gene in "highland papaya" based on inheritance of RAF markers. Euphytica 145:11–23

Dillon S, O'Brien CM, Drew RA, Ramage C (2006a) Application of SCAR markers linked to a putative PRSV-P resistance locus for assisted breeding of resistant *C. papaya* cultivars. Acta Hortic 725:627–633

Dillon S, Ramage C, Ashmore S, Drew RA (2006b) Development of a codominant CAPS marker linked to PRSV-P resistance in highland papaya. Theor Appl Genet 113:1159–1169

Ding X, Gopalakrishnan B, Johnson LB, White FF, Wang X, Morgan TD, Kramer KJ, Muthukrishnan S (1998) Insect resistance of transgenic tobacco expressing an insect chitinase gene. Transgenic Res 7:77–84

Dowd C, Wilson IW, McFadden H (2004) Gene expression profile changes in cotton root and hypocotyl tissues in response to infection with *Fusarium oxysporum* f. sp. *vasinfectum*. Mol Plant Microbe Interact 17:654–667

Drew RA, Magdalita PM, O'Brien CM (1998) Development of *Carica* interspecific hybrids. Acta Hortic 461:285–291

Drew R, Persley D, O'Brien C, Bateson M (2005a) Papaya ringspot virus in Australia and the development of virus resistant plants. Acta Hortic 692:101–106

Drew RA, Siar SV, Villegas VN, O'Brien CM, Sajise AGC (2005b) Development of PRSV-P resistant *Carica papaya* genotypes by introgression of genes from wild *Vasconcella* species. Acta Hortic 694:73–77

Drew R, Siar SV, Dillon S, Ramage C, O'Brien C, Sajise AGC (2007) Intergeneric hybridisation between *Carica papaya* and wild *Vasconcellea* species and identification of a PRSV-P resistance gene. Acta Hortic 738:165–169

Ellis JG, Dodds PN (2011) Showdown at the RXLR motif: Serious differences of opinion in how effector proteins from filamentous eukaryotic pathogens enter plant cells. Proc Natl Acad Sci USA 108:14381–14382

Environmental Protection Agency (2009) Pesticide Fact Sheet. Available: http://www.epa.gov/oppbppd1/biopesticides/pips/smartstax-factsheet.pdf

Erwin DC, Ribeiro OK (1996) *Phytophthora* diseases worldwide. APS, St. Paul
Fermin G, Inglessis V, Garboza C, Rangel S, Dagert M, Gonsalves D (2004) Engineered resistance against papaya ringspot virus in Venezuelan transgenic papayas. Plant Disease 88:516–522
Ferrer JL, Austin MB, Stewart C Jr, Noel JP (2008) Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. Plant Physiol Biochem 46:356–370
Fischer R, Budde I, Hain R (1997) Stilbene synthase gene expression causes changes in flower colour and male sterility in tobacco. Plant J 11:489–498
Fitch MMM (1993) High frequency somatic embryogenesis and plant regeneration from papaya hypocotyl callus. Plant Cell Tiss Org Cult 32:205–212
Fitch MMM, Manshardt RM (1990) Somatic embryogenesis and plant regeneration from immature zygotic embryos of papaya (*Carica papaya* L.). Plant Cell Rep 9:320–324
Fitch MMM, Manshardt RM, Gonsalves D, Slightom JL, Sanford JC (1990) Stable transformation of papaya via microprojectile bombardment. Plant Cell Rep 9:189–194
Fitch MMM, Manshardt RM, Gonsalves D, Slightom JL, Sanford JC (1992) Virus resistant papaya plants derived from tissues bombarded with the coat protein gene of papaya ringspot virus. Nat Biotechnol 10:1466–1472
Fitter AH, Moyerson B (1996) Evolutionary trends in root-microbe symbioses. Phil Trans R Soc Land B 351:1367–1375
Food and Agriculture Organization of the United Nations (2009) Statistics Division, FAOSTAT, Crops Production Data
Fornara V, Onelli E, Sparvoli F, Rossoni M, Aina R, Marino G, Citterio S (2008) Localization of stilbene synthase in *Vitis vinifera* L. during berry development. Protoplasma 233:83–93
Freeman S, Rodriguez RJ (1992) A rapid, reliable bioassay for pathogenicity of *Colletotrichum magna* on cucurbits and its use in screening for nonpathogenic mutants. Plant Disease 76:901–905
Freeman S, Rodriguez RJ (1993) Genetic conversion of a fungal plant pathogen to a nonpathogenic, endophytic mutualist. Science 260:75–78
Fry W (2008) *Phytophthora infestans*: the plant (and R gene) destroyer. Mol Plant Pathol 9:385–402
Fullerton RA, Taufa L, Vanneste JL, Yu J, Cornish DA, Park D (2011) First record of bacterial crown rot of papaya (*Carica papaya*) caused by an *Erwinia papayae*-like bacterium in the Kingdom of Tonga. Disease notes. Plant Disease 95:70
Ganz T (2003) Defensins: antimicrobial peptides of innate immunity. Nat Rev Immunol 3:710–720
Gao AG, Hakimi SM, Mittanck CA, Wu Y, Woerner BM, Stark DM, Shah DM, Liang J, Rommens CM (2000) Fungal pathogen protection in potato by expression of a plant defensin peptide. Nat Biotechnol 18:1307–1310
Gassmann AJ, Petzold-Maxwell JL, Keweshan RS, Dunbar MW (2011) Field-evolved resistance to Bt maize by western corn rootworm. PLoS One 6(7):e22629
Gera A, Mawassi M, Zeidan M, Spiegel S, Bar-Joseph M (2005) An isolate of "Candidatus Phytoplasma australiense" group associated with Nivun-Haamir-Die Back disease of papaya in Israel. New disease report. Plant Pathol 54:560
Gibb KS, Schneider B, Padovan AC (1998) Differential detection and genetic relatedness of phytoplasmas in papaya. Plant Pathol 47:325–332
Gilroy EM, Taylor RM, Hein I, Boevink P, Sadanandom A, Birch PR (2011) CMPG1-dependent cell death follows perception of diverse pathogen elicitors at the host plasma membrane and is suppressed by *Phytophthora infestans* RXLR effector AVR3a. New Phytol 190:653–666
Goff L (1986) Spider mites (Acari: Tetranychidae) in the Hawaiian Islands. Int J Acarol 12:43–49
Gonsalves D (1998) Control of papaya ringspot virus in papaya: a case study. Annu Rev Phytopathol 36:415–437
Grünwald NJ, Flier WG (2005) The biology of *Phytophthora infestans* at its center of origin. Annu Rev Phytopathol 43:171–190

Guttman JA, Samji FN, Li Y, Deng W, Lin A, Finlay BB (2007) Aquaporins contribute to diarrhoea caused by attaching and effacing bacterial pathogens. Cell Microbiol 9:131–141

Hain R, Reif HJ, Krause E, Langebartels R, Kindl H, Vornam B, Wiese W, Schmelzer E, Schreier PH, Stöcker RH, Stenzel K (1993) Disease resistance results from foreign phytoalexin expression in a novel plant. Nature 361:153–156

Hammerbacher A, Ralph SG, Bohlmann J, Fenning TM, Gershenzon J, Schmidt A (2011) Biosynthesis of the major tetrahydroxystilbenes in spruce, astringin and isorhapontin, proceeds via resveratrol and is enhanced by fungal infection. Plant Physiol 157:876–890, Epub 2011 Aug 24

Hawtin RE, Zarkowska T, Arnold K, Thomas CJ, Gooday GW, King LA, Kuzio JA, Possee RD (1997) Liquefaction of *Autographa californica* nucleopolyhedrovirus-infected insects is dependent on the integrity of virus-encoded chitinase and cathepsin genes. Virology 238:243–253

Hijmans RJ, Spooner DM (2001) Geographic distribution of wild potato species. Am J Bot 88:2101–2112

Hipskind JD, Paiva NL (2000) Constitutive accumulation of a resveratrol-glucoside in transgenic alfalfa increases resistance to *Phoma medicaginis*. Mol Plant Microbe Interact 13:551–562

Horovitz S, Jiménez H (1967) Cruzamientos interespecíficos e intergenéricos en caricaceae y sus implicaciones fitotécnicas. Agron Trop 3:323–343

Ingrosso I, Bonsegna S, Dedomenico S, Laddomada B, Blando F, Santino A, Giovinazzo G (2011) Over-expression of a grape stilbene synthase gene in tomato induces parthenocarpy and causes abnormal pollen development. Plant Physiol Biochem 49:1092–1099

Initiative AG (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796–815

International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. Nature 436:793–800

Jaillon O, Aury JM, Noel B, Policriti A, Clepet C et al (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449:463–467

Jain RK, Nasiruddin KM, Sharma J, Pant RP, Varma A (2004) First report of occurrence of papaya ringspot virus infecting papaya in Bangladesh. Disease notes. Plant Disease 88:221

Jaizme-Vega MC, Rodríguez-Romero AS, Núñez LAB (2006) Effect of the combined inoculation of arbuscular mycorrhizal fungi and plant growth-promoting rhizobacteria on papaya (*Carica papaya* L.) infected with the root-knot nematode *Meloidogyne incognita*. Fruits 61:151–162

Jha S, Tank HG, Prasad BD, Chattoo BB (2009) Expression of *Dm-AMP1* in rice confers resistance to *Magnaporthe oryzae* and *Rhizoctonia solani*. Transgenic Res 18:59–69

Jobin-Décor MP, Graham GC, Henry RJ, Drew RA (1997) RAPD and isozyme analysis of genetic relationships between *Carica papaya* and wild relatives. Genet Resour Crop Evol 44:471–477

Júnior MTS, Nickel O, Gonsalves D (2005) Development of virus resistant transgenic papayas expressing the coat protein gene from a Brazilian isolate of *Papaya ringspot virus*. Fitopatol Bras 30:357–365

Kale SD, Gu B, Capelluto DG, Dou D, Feldman E, Rumore A, Arredondo FD, Hanlon R, Fudal I, Rouxel T, Lawrence CB, Shan W, Tyler BM (2010) External lipid PI3P mediates entry of eukaryotic pathogen effectors into plant and animal host cells. Cell 142:284–295

Kalleshwaraswamy CM, Kumar NK (2008) Transmission efficiency of papaya ringspot virus by three aphid species. Phytopathology 98:541–546

Keith RC, Nishijima KA, Keith LM, Fitch MM, Nishijima WT, Wall MM (2008) Atypical internal yellowing of papaya fruit in Hawaii caused by *Enterobacter sakazakii*. Disease notes. Plant Disease 92:487

Kempler C, Kabaluk T (1996) Babaco (*Carica pentagona* Heilb.): a possible crop for the greenhouse. Hortscience 31:785–788

Kilman S (2001) Monsanto Co. shelves seed that turned out to be a dud of a spud. Wall Street J B21 (March 21)

Kim MS, Moore PH, Zee F, Fitch MMM, Steiger DL, Manshardt RM, Paull RE, Drew RA, Sekioka T, Ming R (2002) Genetic diversity of *Carica papaya* as revealed by AFLP markers. Genome 45:503–512

Kung YJ, Bau HJ, Wu YL, Huang CH, Chen TM, Yeh SD (2009) Generation of transgenic papaya with double resistance to *Papaya ringspot virus* and *Papaya leaf-distortion mosaic virus*. Phytopathology 99:1312–1320

Kung YJ, Yu TA, Huang CH, Wang HC, Wang SL, Yeh SD (2010) Generation of hermaphrodite transgenic papaya lines with virus resistance via transformation of somatic embryos derived from adventitious roots of in vitro shoots. Transgenic Res 19:621–635

Kyndt T, Gheysen G (2007) Evolutionary relationships between and within the highland papayas (Genus *Vasconcellea*) and the common papaya (*Carica papaya*). Acta Hortic 740:61–72

Latijnhouwers M, de Wit PJ, Govers F (2003) Oomycetes and fungi: similar weaponry to attack plants. Trends Microbiol 11:462–469

Leckband G, Lörz H (1998) Transformation and expression of a stilbene synthase gene of *Vitis vinifera* L. in barley and wheat for increased fungal resistance. Theor Appl Genet 96: 1004–1012

Lentz DL (1999) Plant resources of the ancient Maya, The paleoethnobotanical evidence. In: White CD (ed) Reconstructing ancient Maya diet. The University of Utah Press, Salt Lake City, pp 3–18 (Chapter 1)

Ling K, Namba S, Gonsalves C, Slightom JL, Gonsalves G (1991) Protection against detrimental effects of potyvirus infection in transgenic tobacco plants expressing the papaya ringspot virus coat protein gene. Biotechnology 9:752–758

Liu Z, Zhuang C, Sheng S, Shao L, Zhao W, Zhao S (2011) Overexpression of a resveratrol synthase gene (*PcRS*) from *Polygonum cuspidatum* in transgenic *Arabidopsis* causes the accumulation of *trans*-piceid with antifungal activity. Plant Cell Rep 30:2027–2036

Magdalita PM, Adkins SW, Godwin ID, Drew RA (1996) An improved embryo-rescue protocol for a *Carica* interspecific hybrid. Aust J Bot 44:343–353

Magdalita PM, Persley DM, Godwin ID, Drew RA (1997) Screening *Carica papaya*×*C. cauliflora* hybrids for resistance to papaya ringspot virus-type P. Plant Pathol 46:837–841

Magdalita PM, Drew RA, Godwin ID, Adkins SW (1998) An efficient interspecific hybridization protocol for *Carica papaya* L.×*C. cauliflora* Jacq. Aust J Exp Agric 38:523–530

Maktar NH, Kamis S, Mohd Yusof FZ, Hussain NH (2008) *Erwinia papayae* causing papaya dieback in Malaysia. New disease report. Plant Pathol 57:774

Mamanova L, Coffey AJ, Scott CE, Kozarewa I, Turner EH, Kumar A, Howard E, Shendure J, Turner DJ (2010) Target-enrichment strategies for next-generation sequencing. Nat Methods 7:111–118

Mangrauthia SK, Singh P, Praveen S (2010) Genomics of helper component proteinase reveals effective strategy for *papaya ringspot virus* resistance. Mol Biotechnol 44:22–29

Manshardt RM, Wenslaff TF (1989a) Zygotic polyembryony in interspecific hybrids of *Carica papaya* and *C. cauliflora*. J Am Soc Hortic Sci 114:684–689

Manshardt RM, Wenslaff TF (1989b) Interspecific hybridization of papaya with other *Carica* species. J Am Soc Hortic Sci 114:689–694

Manshardt RM, Zee FTP (1994) Papaya germplasm and breeding in Hawaii. Fruit Varieties J 48:146–152

Maoka T, Hataya T (2005) The complete nucleotide sequence and biotype variability of papaya leaf distortion mosaic virus. Phytopathology 95:128–135

McCafferty HR, Moore PH, Zhu YJ (2006) Improved *Carica papaya* tolerance to carmine spider mite by the expression of *Manduca sexta chitinase* transgene. Transgenic Res 15:337–347

McDowell JM, Simon SA (2006) Recent insights into *R* gene evolution. Mol Plant Pathol 7:437–448

Mchau GRA, Coffey MD (1994) Isozyme diversity in *Phytophthora palmivora*: evidence for a southeast Asian centre of origin. Mycol Res 98:1035–1043

Mekako HU, Nakasone HY (1975) Interspecific hybridization among 6 *Carica* species. J Am Soc Hortic Sci 100:237–242

Miksicek CH (1983) Macrofloral remains of the Pulltrouser area: settlements and fields. In: Turner II BL, Harrison PD (eds) Pulltrouser swamp: ancient Maya habitat, agriculture, and settlement in northern Belize. University of Texas Press, Austin, pp 94–104

Miller CD (1926) The vitamins (A, B, and C) of papaya. Biochem J 20:515–518

Ming R, Hou S, Feng Y, Yu Q, Dionne-Laporte A, Saw JH et al (2008) The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). Nature 452:991–996

Miyasaka SC, Habte M (2001) Plant mechanisms and mycorrhizal symbioses to increase phosphorus uptake efficiency. Commun Soil Sci Plant Anal 32(7/8):1101–1147

Mlot C (2004) Microbial diversity unbound. What DNA-based techniques are revealing about the planet's hidden biodiversity. Bioscience 54:1064–1068

Mosqueda-Vázquez R, Nakasone HY (1982) Diallel analysis of root rot resistance in papaya. Hortscience 17:384–385

Mosqueda-Vázquez R, Aragaki M, Nakasone HY (1981) Screening of *Carica papaya* L. seedlings for resistance to root rot caused by *Phytophthora palmivora* Butl. J Am Soc Hortic Sci 106:484–487

Nadeem A, Mehmood T, Tahir M, Khalid S, Xiong Z (1997) First report of Papaya leaf curl disease in Pakistan. Disease notes. Plant Disease 81:1333

Nakasone HY, Aragaki M (1982) Current status of papaya improvement program. Hawaii Inst Trop Agric Hum Resour Res Ext Ser 033:51–55

Nascimento RJ, Mizubuti ESG, Câmara MPS, Ferreira MF, Maymon M, Freeman S, Michereff SJ (2010) First report of papaya fruit rot caused by *Colletotrichum magna* in Brazil. Disease notes. Plant Disease 94:1506

National Oceanic and Atmospheric Administration (NOAA), climate data, climatological normals, normal precipitation inches, 30 year normals 1971–2000. http://www.ncdc.noaa.gov/oa/climate/online/ccd/nrmpcp.txt

National Research Council (1989) Lost crops of the Incas: little-known plants of the Andes with promise for worldwide cultivation. National Academy Press, Washington, DC

Newcombe G, Martin F, Kohler A (2010) Defense and nutrient mutualisms in *Populus*. In: Jansson S, Bhalerao RP, Groove, AT (eds) Genetics and genomics of *Populus*. Part III. Genetics and genomics of key *Populus* traits. Springer, New York, pp 247–277

Nishijima WT, Aragaki M (1973) Pathogenicity and further characterization of *Calonectria crotalariae* causing collar rot of papaya. Phytopathology 63:553–558

Noa-Carrazana JC, Silva-Rosales L (2001) First report of a Mexican isolate of Papaya mosaic virus in papaya (*Carica papaya*) and pumpkin (*Cucurbita pepo*). Disease notes. Plant Disease 85:558

Noorda-Nguyen K, Jia R, Aoki A, Yu Q, Nishijima W, Zhu YJ (2010) Identification of disease tolerance loci to *Phytophthora palmivora* in *Carica papaya* using molecular marker approach. Acta Hortic 851:189–196

O'Brien CM, Drew RA (2010) Parker-assisted hybridization and backcrossing between *Vasconcellea* species and *Carica papaya* for PRSV-P resistance. Acta Hortic 859:361–368

Ochoa J, Fonseca G, Ellis MA (2000) First report of *Fusarium* wilt of Babaco (*Carica*×*heilbornii* var. *pentagona*) in Ecuador. Plant Disease 84:199

Ogata DY, Heu RA (2001) Black spot of papaya disease, *Asperisporium caricae* (Speg.) Maulbl. State of Hawaii, Department of Agriculture. New pest advisory no. 2001-01

Osborn RW, De Samblanx GW, Thevissen K, Goderis I, Torrekens S, Van Leuven F, Attenborough S, Rees SB, Broekaert WF (1995) Isolation and characterization of plant defensins from seeds of Asteraceae, Fabaceae, Hippocastanaceae and Saxifragaceae. Fed Eur Biochem Soc Lett 368:257–262

Pantoja A, Follett PA, Villanueva-Jiménez JA (2002) Pests of papaya. In: Peña JE, Sharp JL, Wysoki M (eds) Tropical fruit pests and pollinators: biology, economic importance, natural enemies and control. CABI, New York, pp 131–156 (Chapter 5)

Parris GK (1941) Diseases of papaya in Hawaii and their control. In: Papaya production in the Hawaiian Islands. Hawaii Experiment Station Bulletin 87, pp 32–44

Penninckx IA, Eggermont K, Terras FR, Thomma BP, De Samblanx GW, Buchala A, Métraux JP, Manners JM, Broekaert WF (1996) Pathogen-induced systemic activation of a plant defensin gene in Arabidopsis follows a salicylic acid-independent pathway. Plant Cell 8:2309–2323

Persley DM, Ploetz RC (2003) Diseases of papaya. In: Ploetz, RC. (ed) Diseases of tropical fruit crops. CABI, Cambridge, pp 373–412 (Chapter 17)

Pezet R, Pont V (1990) Ultrastructural observations of pterostilbene fungitoxicity in dormant conidia of *Botrytis cinerea* Pers. J Phytopathol 129:19–30

Pezet R, Gindro K, Viret O, Spring JL (2004) Glycosylation and oxidative dimerization of resveratrol are respectively associated to sensitivity and resistance of grapevine cultivars to downy mildew. Physiol Mol Plant Pathol 65:297–303

Porter BW, Aizawa KS, Zhu YJ, Christopher DA (2008) Differentially expressed and new non-protein-coding genes from a *Carica papaya* root transcriptome survey. Plant Sci 174:38–50

Porter BW, Paidi M, Ming R, Alam M, Nishijima WT, Zhu YJ (2009a) Genome-wide analysis of *Carica papaya* reveals a small NBS resistance gene family. Mol Genet Genomics 281:609–626

Porter BW, Zhu YJ, Christopher DA (2009b) *Carica papaya* genes regulated by *Phytophthora palmivora*: a model system for genomic studies of compatible *Phytophthora*-plant interactions. Trop Plant Biol 2:84–97

Pourrahim R, Farzadfar S, Golnaraghi AR, Shahraeen N (2003) First report of Papaya ringspot virus on papaya in Iran. Disease notes. Plant Disease 87:1148

Preisig-Müller R, Schwekendiek A, Brehm I, Reif HJ, Kindl H (1999) Characterization of a pine multigene family containing elicitor-responsive stilbene synthase genes. Plant Mol Biol 39:221–229

Redman RS, Ranson JC, Rodriguez RJ (1999) Conversion of the pathogenic fungus *Colletotrichum magna* to a nonpathogenic, endophytic mutualist by gene disruption. Mol Plant Microbe Interact 12:969–975

Register JC 3rd, Beachy RN (1988) Resistance to TMV in transgenic plants results from interference with an early event in infection. Virology 166:524–532

Rimando AM, Pan Z, Polashock JJ, Dayan FE, Mizuno CS, Snook ME, Liu CJ, Baerson SR (2012) In planta production of the highly potent resveratrol analogue pterostilbene via stilbene synthase and O-methyltransferase co-expression. Plant Biotechnol J 10(3):269–283

Roberts PD, Trujillo E (1998) First report of *Phytophthora nicotianae* causing leaf blight, fruit rot, and root rot of papaya in American Samoa. Disease notes. Plant Disease 82:712

Rodriguez-Alvarado G, Fernandez-Pavìa SP, Geraldo-Verdugo JA, Landa-Hernandez L (2001) *Pythium aphanidermatum* causing collar rot on papaya in Baja California Sur, Mexico. Disease notes. Plant Disease 85:444

Rodriguez-Romero AS, Azcón R, Jaizme-Vega MDC (2011) Early mycorrhization of two tropical crops, papaya (*Carica papaya* L.) and pineapple [*Ananas comosus* (L.) Merr.], reduces the necessity of P fertilization during the nursery stage. Fruits 66:3–10

Samappito S, Page JE, Schmidt J, De-Eknamkul W, Kutchan TM (2003) Aromatic and pyrone polyketides synthesized by a stilbene synthase from *Rheum tataricum*. Phytochemistry 62:313–323

Sanchez M, Dianese JC, Costa CL (1991) Factors affecting the damage by *Phoma caricae-papayae* on papaya fruits, and detection of resistance to the fungus in *Carica gaudotiana*. Fitopatol Bras 16:121–129

Sanford JC, Johnston SA (1985) The concept of parasite-derived resistance-Deriving resistance genes from the parasite's own genome. J Theor Biol 113:395–405

Schanz S, Schröder G, Schröder J (1992) Stilbene synthase from Scots pine (*Pinus sylvestris*). Fed Eur Biochem Soc Lett 313:71–74

Scheldeman X, Motoche JPR, Van Damme V, Heyens V, Van Damme P (2003) Potential of highland papayas (*Vasconcella* spp.) in southern Ecuador. Lyonia 5:73–80

Scheldeman X, Willemen L, Coppens d'Eeckenbrugge G, Romeijn-Peeters E, Restrepo MT, Motoche JR, Jiménez D, Lobo M, Medina CI, Reyes C, Rodríguez D, Ocampo JA, Van Damme P, Goetgebeur P (2007) Distribution, diversity and environmental adaptation of highland papayas (Vasconcellea spp.) in tropical and subtropical America. Biodivers Conserv 16:1867–1884

Schmidlin L, Poutaraud A, Claudel P, Mestre P, Prado E, Santos-Rosa M, Wiedemann-Merdinoglu S, Karst F, Merdinoglu D, Hugueney P (2008) A stress-inducible resveratrol O-methyltransferase involved in the biosynthesis of pterostilbene in grapevine. Plant Physiol 148:1630–1639

Schnee S, Viret O, Gindro K (2008) Role of stilbenes in the resistance of grapevine to powdery mildew. Physiol Mol Plant Pathol 72:128–133

Schöppner A, Kindl H (1984) Purification and properties of a stilbene synthase from induced cell suspension cultures of peanut. J Biol Chem 259:6806–6811

Sémon M, Wolfe KH (2007) Consequences of genome duplication. Curr Opin Genet Dev 17:505–512

Serazetdinova L, Oldach KH, Lörz H (2005) Expression of transgenic stilbene synthases in wheat causes the accumulation of unknown stilbene derivatives with antifungal activity. J Plant Physiol 162:985–1002

Shahollari B, Vadassery J, Varma A, Oelmüller R (2007) A leucine-rich repeat protein is required for growth promotion and enhanced seed production mediated by the endophytic fungus *Piriformospora indica* in *Arabidopsis thaliana*. Plant J 50:1–13

Shen W, Yan P, Gao L, Pan X, Wu J, Zhou P (2010) Helper component-proteinase (HC-Pro) protein of *Papaya ringspot virus* interacts with papaya calreticulin. Mol Plant Pathol 11:335–346

Siar SV, Beligan GA, Drew RA, O'Brien C (2009) Final report: development of PRSV-P resistant papaya genotypes by introgression of genes from wild *Carica* species. Australian Centre for International Agricultural Research, Canberra

Sobolev VS, Guo BZ, Holbrook CC, Lynch RE (2007) Interrelationship of phytoalexin production and disease resistance in selected peanut genotypes. J Agric Food Chem 55:2195–2200

Song J, Bradeen JM, Naess SK, Raasch JA, Wielgus SM, Haberlach GT, Liu J, Kuang H, Austin-Phillips S, Buell CR, Helgeson JP, Jiang J (2003) Gene *RB* cloned from *Solanum bulbocastanum* confers broad spectrum resistance to potato late blight. Proc Natl Acad Sci USA 100:9128–9133

Sparvoli F, Martin C, Scienza A, Gavazzi G, Tonelli C (1994) Cloning and molecular analysis of structural genes involved in flavonoid and stilbene biosynthesis in grape (*Vitis vinifera* L.). Plant Mol Biol 24:743–755

Stark-Lorenzen P, Nelke B, Hänßler G, Mühlbach HP, Thomzik JE (1997) Transfer of a grapevine stilbene synthase gene to rice (*Oryza sativa* L.). Plant Cell Rep 16:668–673

Stein E, Molitor A, Kogel KH, Waller F (2008) Systemic resistance in *Arabidopsis* conferred by the mycorrhizal fungus *Piriformospora indica* requires jasmonic acid signaling and the cytoplasmic function of NPR1. Plant Cell Physiol 49:1747–1751

Takeguchi A, Hollyer J, Koga W, Hakoda M, Rohrbach K, Bittenbender HC, Buckley B, Friday J B, Bowen R, Manshardt R, Leary J, Teves G, Herring E, Zaleski H, Leonhardt K, Eger B (1999) History of agriculture in Hawaii. State of Hawaii Department of Agriculture. http://hawaii.gov/hdoa/ag-resources/history

Tanaka A, Christensen MJ, Takemoto D, Park P, Scott B (2006) Reactive oxygen species play a role in regulating a fungus-perennial ryegrass mutualistic interaction. Plant Cell 18:1052–1066

Tarnowski TLB, Ploetz RC (2010) First report of *Colletotrichum capsici* causing postharvest anthracnose on papaya in South Florida. Disease notes. Plant Disease 94:1065

Tennant P, Fermin G, Fitch MM, Manshardt RM, Slightom JL, Gonsalves D (2001) Papaya ringspot virus resistance of transgenic Rainbow and SunUp is affected by gene dosage, plant development, and coat protein homology. Eur J Plant Pathol 107:645–653

Tennant P, Ahmad MH, Gonsalves D (2005) Field resistance of coat protein transgenic papaya to *Papaya ringspot virus* in Jamaica. Plant Disease 89:841–847

Terras FR, Eggermont K, Kovaleva V, Raikhel NV, Osborn RW, Kester A, Rees SB, Torrekens S, Van Leuven F, Vanderleyden J, Cammue BPA, Broekaert WF (1995) Small cysteine-rich antifungal proteins from radish: their role in host defense. Plant Cell 7:573–588

Thevissen K, Osborn RW, Acland DP, Broekaert WF (2000a) Specific binding sites for an antifungal plant defensin from Dahlia (*Dahlia merckii*) on fungal cells are required for antifungal activity. Mol Plant Microbe Interact 13:54–61

Thevissen K, Cammue BP, Lemaire K, Winderickx J, Dickson C, Lester RL, Ferket KK, Van Even F, Parret AH, Broekaert WF (2000b) A gene encoding a sphingolipid biosynthesis enzyme

determines the sensitivity of *Saccharomyces cerevisiae* to an antifungal plant defensin from dahlia (*Dahlia merckii*). Proc Natl Acad Sci USA 97:9531–9536

Thevissen K, François IE, Takemoto JY, Ferket KK, Meert EM, Cammue BP (2003) DmAMP1, an antifungal plant defensin from dahlia (*Dahlia merckii*), interacts with sphingolipids from *Saccharomyces cerevisiae*. FEMS Microbiol Lett 226:169–173

Thomma BP, Eggermont K, Penninckx IA, Mauch-Mani B, Vogelsang R, Cammue BP, Broekaert WF (1998) Separate jasmonate-dependent and salicylate-dependent defense-response pathways in *Arabidopsis* are essential for resistance to distinct microbial pathogens. Proc Natl Acad Sci USA 95:15107–15111

Thomzik JE, Stenzel K, Stöcker R, Schreier PH, Hain R, Stahl DJ (1997) Synthesis of a grapevine phytoalexin in transgenic tomatoes (*Lycopersicon esculentum* Mill.) conditions resistance against *Phytophthora infestans*. Physiol Mol Plant Pathol 51:265–278

Tripathi S, Suzuki J, Gonsalves D (2006) Development of genetically engineered resistant papaya for *Papaya ringspot virus* in a timely manner-A comprehensive and successful approach. In: Ronald P (ed) Plant-pathogen interactions: methods and protocols, vol 354. The Humana Press, Totowa, pp 197–240

Tripathi S, Suzuki JY, Ferreira SA, Gonsalves D (2008) *Papaya ringspot virus*-P: characteristics, pathogenicity, sequence variability and control. Mol Plant Pathol 9:269–280

Tropf S, Lanz T, Rensing SA, Schröder J, Schröder G (1994) Evidence that stilbene synthases have developed from chalcone synthases several times in the course of evolution. J Mol Evol 38:610–618

Tsay JG, Chen RS, Wang HL, Wang WL, Weng BC (2011) First report of powdery mildew caused by *Erysiphe diffusa*, *Oidium neolycopersici*, and *Podosphaera xanthii* on papaya in Taiwan. Disease notes. Plant Disease 95:1188

Turrini A, Sbrana C, Pitto L, Castiglione MR, Giorgetti L, Briganti R, Bracci T, Evangelista M, Nuti MP, Giovannetti M (2004a) The antifungal DM-AMP1 protein from *Dahlia merckii* expressed in *Solanum melongena* is released in root exudates and differentially affects pathogenic fungi and mycorrhizal symbiosis. New Phytol 163:393–403

Turrini A, Sbrana C, Nuti MP, Pietrangeli BM, Giovannetti M (2004b) Development of a model system to assess the impact of genetically modified corn and aubergine plants on arbuscular mycorrhizal fungi. Plant Soil 266:69–75

Tuskan A, Difazio S, Jansson S, Bohlmann J, Grigoriev I et al (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science 313:1596–1604

Van der Biezen EA, Jones JD (1998) Plant disease-resistance proteins and the gene-for-gene concept. Trends Biochem Sci 23:454–456

van der Vossen E, Sikkema A, Hekkert BL, Gros J, Stevens P, Muskens M, Wouters D, Pereira A, Stiekema W, Allefs S (2003) An ancient R gene from the wild potato species *Solanum bulbocastanum* confers broad-spectrum resistance to *Phytophthora infestans* in cultivated potato and tomato. Plant J 36:867–882

Vega FE, Posada F, Aime MC, Pava-Ripoll M, Infante F, Rehner SA (2008) Entomopathogenic fungal endophytes. Biol Control 46:72–82

Ventura JA, Costa H, da Silva Tatagiba J (2004) Papaya diseases and integrated control. In: Naqvi SAMH (ed) Diseases of fruits and vegetables, vol II. Kluwer Academic, Dordrecht, pp 201–268

Verma LR, Sharma RC (1999) Fungal diseases of papaya and their management. In: Verma LR, Sharma RC (eds) Diseases of horticultural crops, fruits. Indus Publishing Company, New Delhi, pp 479–492 (Chapter 21)

Verma S, Varma A, Rexer KH, Hassel A, Kost G, Sarbhoy A, Bisen P, Bütehorn B, Franken P (1998) *Piriformospora indica*, gen. et sp. Nov., a new root-colonizing fungus. Mycologia 90:896–903

Vierheilig H, Alt M, Neuhaus JM, Boller T, Wiemken A (1993) Colonization of transgenic *Nicotiana sylvestris* plants, expressing different forms of *Nicotiana tabacum* chitinase, by the root pathogen *Rhizoctonia solani* and by mycorrhizal symbiont *Glomus mosseae*. Mol Plant Microbe Interact 6:261–264

Vierheilig H, Alt M, Lange J, Gut-Rella M, Wiemken A, Boller T (1995) Colonization of transgenic tobacco constitutively expressing pathogenesis-related proteins by the vesicular-arbuscular mycorrhizal fungus *Glomus mosseae*. Appl Environ Microbiol 61:3031–3034

Vinet M, Dave SK, Specht CA, Brameld KA, Xu B, Hayward R, Fidock DA (1999) The chitinase PfCHT1 from the human malaria parasite *Plasmodium falciparum* lacks proenzyme and chitin-binding domains and displays unique substrate preferences. Proc Natl Acad Sci USA 96:14061–14066

Vleeshouwers VG, Raffaele S, Vossen JH, Champouret N, Oliva R, Segretin ME, Rietman H, Cano LM, Lokossou A, Kessel G, Pel MA, Kamoun S (2011) Understanding and exploiting late blight resistance in the age of effectors. Annu Rev Phytopathol 49:507–531

Waller F, Achatz B, Baltruschat H, Fodor J, Becker K, Fischer M, Heier T, Hückelhoven R, Neumann C, von Wettstein D, Franken P, Kogel KH (2005) The endophytic fungus *Piriformospora indica* reprograms barley to salt-stress tolerance, disease resistance, and higher yield. Proc Natl Acad Sci USA 102:13386–13391

Walton JD, Avis TJ, Alfano JR, Gijzen M, Spanu P, Hammond-Kosack K, Sánchez F (2009) Effectors, effectors et encore des effectors: the XIV international congress on molecular-plant microbe interactions, Quebec. Mol Plant Microbe Interact 22:1479–1483

Wang Y, Nowak G, Culley D, Hadwiger LA, Fristensky B (1999) Constitutive expression of pea defense gene *DRR206* confers resistance to blackleg (Leptosphaeria maculans) disease in transgenic canola (*Brassica napus*). Mol Plant Microbe Interact 12:410–418

Wang J, Chen Z, Du J, Sun Y, Liang A (2005) Novel insect resistance in *Brassica napus* developed by transformation of chitinase and scorpion toxin genes. Plant Cell Rep 24:549–555

Whisson SC, Boevink PC, Moleleki L, Avrova AO, Morales JG, Gilroy EM, Armstrong MR, Grouffaud S, van West P, Chapman S, Hein I, Toth IK, Pritchard L, Birch PR (2007) A translocation signal for delivery of oomycete effector proteins into host plant cells. Nature 450:115–118

Wilmes M, Cammue BP, Sahl HG, Thevissen K (2011) Antibiotic activities of host defense peptides: more to it than lipid bilayer perturbation. Nat Prod Rep 28:1350–1358

Yaeno T, Li H, Chaparro-Garcia A, Schornack S, Koshiba S, Watanabe S, Kigawa T, Kamoun S, Shirasu K (2011) Phosphatidylinositol monophosphate-binding interface in the oomycete RXLR effector AVR3a is required for its stability in host cells to modulate plant immunity. Proc Natl Acad Sci USA 108:14682–14687

Yamazaki Y, Suh DY, Sitthithaworn W, Ishiguro K, Kobayashi Y, Shibuya M, Ebizuka Y, Sankawa U (2001) Diverse chalcone synthase superfamily enzymes from the most primitive vascular plant, *Psilotum nudum*. Planta 214:75–84

Yang S, Zhang X, Yue JX, Tian D, Chen JQ (2008) Recent duplications dominate NBS-encoding gene expansion in two woody species. Mol Genet Genomics 280:187–198

Yeh S-D, Gonsalves D, Wang H-L, Namba R, Chiu R-J (1988) Control of papaya ringspot virus by cross protection. Plant Disease 72:375–380

Yu CKY, Springob K, Schmidt J, Nicholson RL, Chu IK, Yip WK, Lo C (2005) A stilbene synthase gene (*SbSTS1*) is involved in host and nonhost defense responses in sorghum. Plant Physiol 138:393–401

Zentmyer GA (1988) Origin and distribution of four species of *Phytophthora*. Trans Br Mycol Soc 91:367–378

Zentmyer GA, Mitchell DJ (1985/1986) *Phytophthora* diseases of fruit trees in the tropics. Rev Trop Plant Pathol 2:287–309

Zeya HI, Spitznagel JK (1963) Antibacterial and enzymatic basic proteins from leukocyte lysosomes: Separation and identification. Science 3595:1085–1087

Zhu S (2007) Evidence for myxobacterial origin of eukaryotic defensins. Immunogenetics 59:949–954

Zhu YJ, Agbayani R, Jackson MC, Tang CS, Moore PH (2004) Expression of the grapevine stilbene synthase gene *VST1* in papaya provides increased resistance against diseases caused by *Phytophthora palmivora*. Planta 220:241–250

Zhu YJ, Agbayani R, Moore PH (2007) Ectopic expression of *Dahlia merckii* defensin *DmAMP1* improves papaya resistance to *Phytophthora palmivora* by reducing pathogen vigor. Planta 226:87–97

Ziebell H, Carr JP (2010) Cross-protection: a century of mystery. Adv Virus Res 76:211–264 (Chapter 6)

Chapter 16
Genomics of Papaya Sex Chromosomes

Robert VanBuren and Ray Ming

Introduction: The Evolution of Sex Chromosomes

Systems for sex determination in plants and animals are remarkably diverse. Sex determination can be controlled by simple mechanisms such as a single gene on homologous chromosomes with two distinct loci or more complex mechanisms in the case of differentiated sex chromosomes. Sex chromosomes have evolved numerous times in diverse lineages of plants and animals and are in the process of evolving in several plant species (Graves and Shetty 2001; Tanurdzic and Banks 2004; Ming et al. 2011). Most sex chromosomes in mammals are ancient, but sex chromosomes in plants, insects, and some fish are much younger, having evolved more recently. Sex chromosomes originate from autosomes and arise from the suppression of recombination at sex determination loci (Westergaard 1958; Bergero and Charlesworth 2009). Suppression of recombination fixes the sex determination loci, allowing for the formation of two distinct chromosome types. The lack of recombination relaxes purifying selection in the Y (male heterogametic) or W (female heterogametic) chromosome, typically causing the slow accumulation of deleterious mutations (Bachtrog 2003; Bachtrog and Charlesworth 2002; Bergero and Charlesworth 2009). Ultimately, the accumulated mutations results in chromosome degeneration and eventual loss of the heterogametic chromosome.

A number of sex chromosomes in diverse stages of evolution have been identified (Fig. 16.1). In humans, the ancient Y chromosome has diverged and degenerated a great deal from the X. The Y chromosome is three times smaller than the X and has lost 1,393 of its original 1,438 genes (Skaletsky et al. 2003; Ross et al. 2005). At the other extreme, the young Y chromosome of white campion, *Silene latifolia*, has expanded dramatically and is currently 570 Mb in size, 150 Mb larger than its X

R. VanBuren • R. Ming (✉)
Department of Plant Biology, University of Illinois at Urbana-Champaign, 1201 W. Gregory Drive, 148 ERML, MC-051, Urbana, IL 61801, USA
e-mail: rming@life.uiuc.edu

R. Ming and P.H. Moore (eds.), *Genetics and Genomics of Papaya*, Plant Genetics and Genomics: Crops and Models 10, DOI 10.1007/978-1-4614-8087-7_16,

Fig. 16.1 Organization of sex chromosomes in select species. The size and degeneracy of sex chromosomes changes rapidly at different stages of sex chromosome evolution. Papaya, in the early stages of sex chromosome evolution, has relatively small sex determination regions and divergence from the X (Y = 50 Mb, X = 46 Mb). *Silene latifolia* shows dramatic expansion in the heterogametic Y chromosome (Y = 570 Mb, X = 420 Mb). Human sex chromosomes are the most ancient, with significant deletion and degeneracy in the Y chromosome (Y = 66 Mb, X = 155 Mb)

counterpart (Delph et al. 2010; Bergero et al. 2007). Research on the evolution of sex chromosomes is progressing rapidly, yielding tremendous insight into the mechanisms that govern the formation, divergence, and maintenance of sex chromosomes. Papaya is an excellent model for studying sex chromosome evolution, as its young sex chromosomes evolved 3–4 million years ago (MYA) (Yu et al. 2008). The young age and sequence similarity of the papaya sex chromosomes allow for the identification of the chromosomal rearrangement events leading to their inception. An XY system determines papaya sex, with two slightly different Y chromosomes: Y^h in hermaphrodites and Y in males (Liu et al. 2004). Sequencing of the papaya sex chromosomes was recently completed and has uncovered a number of exciting findings into the origin and rapid evolution of sex chromosomes.

The vast majority of land plants are hermaphroditic, producing male and female gametes in the same plant. Roughly 6 % of flowering plants, however, are dioecious with separate sexes (Renner and Ricklefs 1995). Dioecious plants have unisexual flowers that result from the abortion of stamens in female plants and carpels in male plants. The phylogenetic distribution of dioecious plants is widespread yet patchy, indicating that dioecy has evolved independently many times (reviewed in Ming et al. 2011). The self-incompatibility of dioecious plants is evolutionarily advantageous, as it increases genetic diversity and prevents self-crossing (Charlesworth 1985).

Papaya is unusual among crop plants as it is trioecious, with three sex types. Varieties of papaya are either dioecious with male and female plants or gynodioecious with hermaphrodite and female plants. The flowers of male, female, and hermaphrodite papaya are morphologically distinct, as are hermaphrodite and female fruits (Ming et al. 2007). Male papaya flowers develop in the leaf axis and form large, branching clusters with long peduncles. Male flowers are small and slender, with a whorl of five stamens surrounding an aborted carpel (Fig. 16.2a–c). Female and hermaphrodite flowers also arise at the leaf axis; however, they develop singly and have short peduncles. Carpeloid flowers are larger than male flowers and have a well-developed carpel surrounded by stamens (in the case of hermaphrodites) or lack stamens completely (in the case of females). Upon fertilization, hermaphrodite and female flowers develop fruit, which are usually harvestable in 4–5 months.

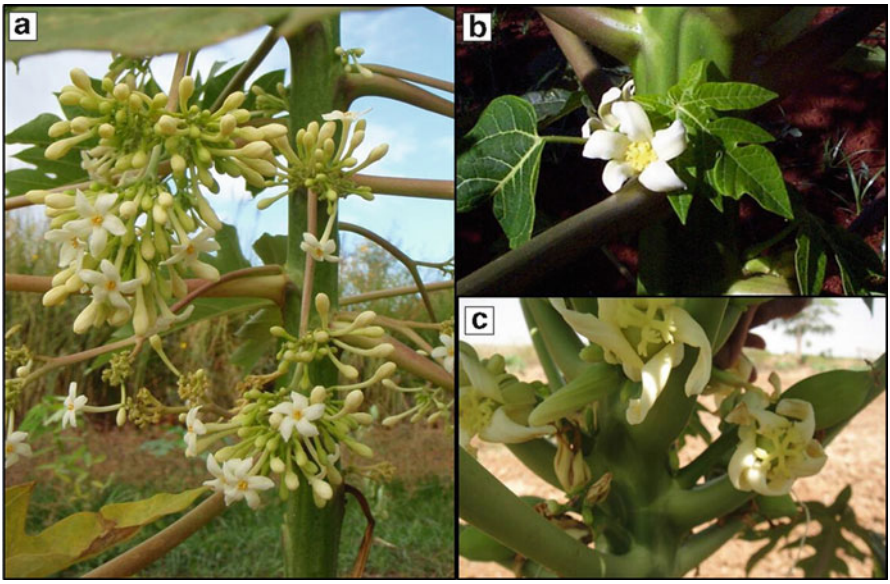

Fig. 16.2 (**a**–**c**) Flower morphology in trioecious papaya. The flowers of male, female, and hermaphrodite papaya are morphologically distinct. (**a**) Male flowers are borne in large, branching clusters, with five well-developed stamens surrounding an aborted carpel. (**b**) Hermaphrodite flowers develop singly, with short peduncles and five stamens surrounding a fully developed carpel. (**c**) Female flowers are similar to hermaphrodite flowers except they have no traces of stamens

Hermaphrodite and female fruits are morphologically distinct, as hermaphrodite fruits tend to be elongated, and female fruits are more rounded.

The initial female and male sterile mutations that trigger sex chromosome evolution can occur in a number of developmental genes and on any part of any autosomal chromosome (Wellmer et al. 2004). The development of unisexual flowers leads to diocey, but the continued recombination between the sex determination genes leads to a mixture of sex types in the resulting offspring. In the case of wild strawberry, *Fragaria virginiana*, two genes control sex (Spigler et al. 2008). One gene leads to male sterility and another leads to female sterility. Recombination between these genes results in a mixture of male, female, hermaphrodite, and neuter individuals. These sex determination genes have a 5.6 cM distance between them and are arguably in the early stages of sex chromosome evolution (Spigler et al. 2008, 2010). Recombination between the sex determination genes maintains selective pressures, purifying selection, and genetic diversity, preventing the accumulation of deleterious mutations. However, if recombination between these loci is suppressed, they will co-segregate, producing only male and female offspring, facilitating the inception of sex chromosomes. Suppression of recombination can be triggered by large-scale inversions, deletions, and chromosomal rearrangements.

The genomics of sex chromosomes are of enormous evolutionary interest, and until recently, the human X and Y chromosomes represented the only known

complete pair of sequenced sex chromosomes. The chimpanzee Y chromosome has also been sequenced, allowing for comparative analysis between the diverged human and chimp Y chromosomes (Hughes et al. 2010). The X chromosome in chimpanzees, however, remains unsequenced, preventing within species comparison between the diverged X and Y chromosomes.

Genome sequencing projects typically use the homogametic sex for shotgun sequencing, to alleviate the problems associated with assembling the heterogenic X and Y chromosomes. Although a number of complex plant and animal genomes have been sequenced, the heterogametic sex determination regions remain largely unsequenced. Furthermore, constructing physical maps of Y or Z chromosome is extremely difficult, as the lack of recombination makes it impossible to order BACs via an integrated genetic and physical map. Thus, the most accurate and efficient method is a slow and reiterative BAC by BAC approach. These constraints have restricted the sequencing of most Y and W chromosomes.

Papaya Is a Model for Sex Chromosome Evolution

Many theories regarding the molecular forces that facilitate the initiation of sex chromosomes and the mechanisms driving their divergence and evolution have been proposed (Charlesworth and Charlesworth 1978; Charlesworth et al. 1993, 2005). These theories are difficult to study, however, without complete sequence information. The nascent sex chromosomes of papaya evolved only 3–4 MYA and offer a unique opportunity to study the characteristics of and mechanisms that govern the early stages of sex chromosome evolution. Mammalian, reptilian, and avian sex chromosomes are ancient and highly degenerated, making it impossible to trace their complex evolutionary history (Skaletsky et al. 2003; Tsuda et al. 2007; Hughes et al. 2010). The papaya sex chromosomes are young and still share a high degree of synteny and gene content. Conserved regions allow for the rearrangements, inversions, and deletions to be easily identified. Furthermore, the papaya sex chromosomes are relatively small (8.1 and 3.5 Mb for HSY and X, respectively), and the papaya genome has been sequenced with an integrated, high-density genetic map (Ming et al. 2008).

Sequencing the papaya sex chromosomes is also important agriculturally. The dominant sex determination gene that promotes maleness is likely confined to the HSY (reviewed in Ming et al. 2007). Identification of the stamen-promoting gene would allow for the creation of a true-breeding hermaphrodite papaya. Growing hermaphrodite papaya can be problematic, as the seeds segregate in a 2:1 ratio of hermaphrodite to female. To encourage maximum production efficiency, papaya farmers often plant 10 or more papaya seeds per mound to ensure that mound will have a mature, fruit-producing hermaphrodite tree. When the trees flower, all females and all but one hermaphrodite tree are removed by hand. The initial plant crowding and competition wastes water and fertilizer, stunts development, delays flowering, and makes the plants more susceptible to disease and abiotic stresses.

Creating a true-breeding hermaphrodite papaya would allow farmers to plant one or two seeds per mound, without having to segregate flowering trees for sex type. True-breeding hermaphrodites would revolutionize the modern papaya industry. However, the lack of recombination in the sex determination region makes a map-based cloning approach to finding the sex determination genes impossible, so complete sequencing of the HSY is the most efficient way to identify candidate genes.

A BAC by BAC approach was used to physically map and sequence the papaya sex-specific X and HSY regions. The HSY and X represent the second set of completely sequenced sex chromosomes in any plant or animal. These sequences have uncovered a number of novel insights into the nature of sex chromosome evolution, as well as putative sex determination genes. In this chapter, an overview of the physical mapping and sequencing of the HSY and X is briefly reiterated, and the major findings, implications, and future prospects of the papaya sex chromosomes are explored.

Evidence of Sex Chromosomes in Papaya

The first indication of sex chromosomes in papaya arose from the unexpected ratios of offspring inheritance. When hermaphrodite flowers are self-pollinated, the resulting seeds segregate in a 2:1 ratio of hermaphrodite to female. If female flowers are fertilized with hermaphrodite pollen, the seeds segregate in a 1:1 hermaphrodite to female ratio and when fertilized with male pollen segregate in a 1:1 ratio of male to female. Furthermore, when the occasional male fruit is selfed, the seeds segregate in a 2:1 ratio of males to females (Storey 1938). These intriguing ratios led to the hypothesis that papaya has three distinct sex chromosomes, with a lethal factor in the male- and hermaphrodite-specific regions (Storey 1941). Males and hermaphrodites are heteromorphic with sex chromosomes denoted as Y for males and Y^h for hermaphrodites, and females are homomorphic with two X chromosomes. Any combination of the YY, Y^hY, or Y^hY^h genotypes is lethal, explaining the 2:1 ratio in selfed hermaphrodite and male trees (Fig. 16.3a, b).

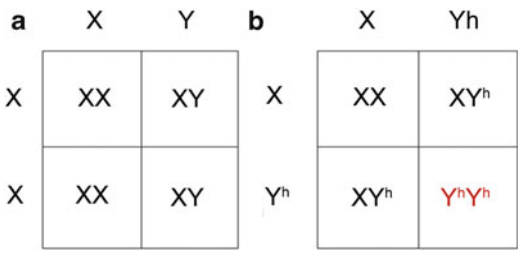

Fig. 16.3 (**a**, **b**) Sex chromosome segregation ratios. (**a**) A cross between male and female flowers produces offspring with a ratio of 50 % male (XY) and 50 % female (XX). (**b**) Self-crossed hermaphrodites produce 25 % females (XX), 50 % hermaphrodites (XY^h), and 25 % aborted seeds (Y^hY^h)

Cytologically, it is impossible to distinguish the heteromorphic sex chromosomes in papaya. Papaya chromosomes are all similar in size, and the sex-specific regions are too small to accurately resolve at the chromosomal level. Early karyotype analyses were useful for determining the number of chromosomes in papaya, but the similarities between chromosomes made it difficult to assign individual chromosome numbers, let alone differentiate sex chromosomes (Ming et al. 2008). Recently, however, DAPI staining of pachytene chromosomes revealed consistent heterochromatin patterns in the sex-specific regions (Zhang et al. 2008). The HSY has five distinct knob structures that lie within the centromeric region and span 13 % of the chromosome length. The homologous X chromosome has one only one knob structure, which is shared with the HSY (Zhang et al. 2008). The shared knob structures pair accurately during recombination, but the area around knob 4 protrudes away from the X region in a curved shape. This curved shape is similar to the loop structure seen when large duplications or deletions occur in one of the two chromosomes. The lack of homologous pairing in this region indicates that there is a suppression of recombination. These regions were later verified as the X and HSY using fluorescence in situ hybridization with known X and HSY BACs (Zhang et al. 2008; Wang et al. 2012).

Before DAPI chromosome staining, the strongest evidence supporting the theory of nascent sex chromosomes in papaya came from high-density linkage mapping. The high-density genetic map was initially created for the papaya genome sequencing project but also served to identify co-segregating markers. 1,501 markers were screened in 54 F2 offspring derived from crossing Kapoho and SunUp cultivars (Blas et al. 2009). The markers were mapped to 12 linkage groups, corresponding to the 9 papaya chromosomes. A total of 225 of the markers on linkage group 1 co-segregated with sex type. This suggested severe suppression of recombination around the sex determination loci, a hallmark of sex chromosomes. The genetic map provided the foundation for anchoring the draft papaya genome sequences, but the co-segregating markers also served as a starting point for physical mapping of the sex chromosomes.

Physical Mapping and Sequencing of the Papaya Sex Chromosomes

The papaya HSY and X are on chromosome 1 and are relatively small, representing around 13 % of the total chromosome length (Zhang et al. 2010). The sequences flanking the HSY and X represent the pseudo-autosomal regions. The pseudo-autosomal regions are homologous in both sexes and undergo normal recombination and selective pressures. The border sequences of the HSY represent the area of transition between the sex-specific region and the pseudo-autosome. Genetically, the borders are where recombination picks up and genetic marker co-segregation ends. On the molecular level, the borders are where the polymorphism between the sexes drops to 0, resulting in identical DNA sequences. There is little genetic

diversity in papaya haplotypes, and within genome polymorphisms are estimated to be 0.06 % based on the papaya draft genome sequence (Ming et al. 2008). Thus, the pseudo-autosomal regions should be nearly identical, making it easy to accurately identify the border boundaries.

The positions of the border regions between sex determination and pseudo-autosome were first defined genetically, by mapping the location of the non-recombining region. An F2 population with 1,460 plants was established, and 32 simple sequence repeat (SSR) markers at border A and 43 SSR markers at border B were designed to fine map recombination in the X and HSY regions. Five SSR markers near border A and two markers near B were polymorphic between the parental cultivars, allowing them to be useful in mapping. In border A, one of the markers (spctg177-12) had recombinants, while the other four SSR markers co-segregated with sex (Wang et al. 2012). This accurately positioned the non-recombining region in border A between spctg177-12 and next distal most co-segregating SSR marker (84M10ctg-34a/b). Both of the SSR markers at border B had recombinant offspring, placing the non-recombining region between a previously identified co-segregating marker, and 58C24-25b, the recombining SSR marker closest to the sex-specific region. Fine mapping of the non-recombining region established the initial borders of the HSY and X regions.

The molecular position of border A has not been determined in the HSY, but it has been mapped and sequenced in the X. Border A is adjacent to knob 1 in the HSY and X and represents a highly repetitive, heterochromatic region. Although border A was successfully mapped in the X, repeated attempts have failed to fill the gaps in the HSY. The male-specific region (MSY) shares 98.8 % sequence identity with the HSY, and physical mapping of the MSY has filled this gap in border A. Sequencing BACs up to the genetic border B revealed continued heterogeneity between the X and HSY, so the physical map was extended beyond the genetic border. The similar sequence composition of these regions made it difficult to pick and distinguish X- and HSY-derived BACs. However, after additional BAC sequencing, a region was identified where the SNP and Indel rate gradually fell to 0 between the HSY and X BACs, indicating that the molecular border B had been reached and the transition between sex chromosome and the pseudo-autosomal region had been found. Interestingly, the molecular border extends 277 kb beyond the genetic border. This suggests a gradual expansion of the HSY into the pseudo-autosomal region, and as expected, the autosomal regions in X and Y^h chromosomes are homologous. The border sequences at the molecular and genetic level accurately defined the sex-specific regions, forming the breaking points for the physical maps.

Constructing physical maps of the papaya sex chromosomes was a tedious and laborious process, presenting a number of challenges. The sex chromosome regions are pericentromeric and contain a high percentage of repeats (Liu et al. 2004; Zhang et al. 2008). The HSY contains five heterochromatic knob structures that are highly repetitive, making it difficult to design probes for picking BACs and gap filling (Na et al. 2012). The HSY and X sequences have increasing sequence similarity near the border regions where the sex chromosomes transition back into autosomes. Identifying BACs in these regions was difficult, as probes often hybridized to BACs

from both sexes, and BAC ends were nearly identical. Recombination is suppressed in the HSY and it was impossible to generate an accurate genetic map of the region to anchor and order BACs. Several co-segregating SCAR markers provided the starting points for generating the physical map, and additional BACs were added to the contigs using BAC end sequences as probes. BACs were confirmed to by X or HSY specific using fluorescence in situ hybridization (FISH). FISH was also used to estimate the size of gaps between BACS. After years of constructing the physical maps in a BAC by BAC approach, a minimum tiling path was generated consisting of 68 BACs in the HSY totaling 8.1 Mb and 44 BACs in the X totaling 5.3 Mb (Wang et al. 2012).

Surprisingly, five knob structures were mapped without a gap in the HSY, and one shared knob was mapped in the X. One large gap remains in the HSY near border A and knob 1, where the autosome transitions into the sex chromosome. The border A region is heterochromatic and highly repetitive. Despite exhaustive efforts to fill these gaps, no BACs were found. Several small gaps remain in the X, but the corresponding regions in the HSY are gapless. HSY and X BACs were sequenced using the Sanger method. Shotgun libraries were constructed from each BAC, and inserts were sequenced and assembled. Each BAC was sequenced to an average of 8–20× coverage, depending on sequence complexity. Repeat-rich regions proved difficult to sequence and assemble even using a BAC by BAC approach. Despite this challenge, BAC sequences were assembled into pseudomolecules with 15 contigs in the HSY and 12 contigs in the X.

Structure of the HSY and X

The papaya HSY has undergone a significant number of small-scale inversions, duplications, translocations, deletions, and repeat accumulations but, in terms of gene content, can be divided into three distinct regions (Fig. 16.4). The first two regions are large-scale inversions in the HSY with respect to the X, and the third is a collinear region where gene order is shared between the X and HSY. The two inversions form two evolutionary strata indicating that they occurred independently, with an estimated divergence from the X of 6.6 MYA for inversion one and 1.9 MYA for inversion two. Inversions in the HSY contribute to the suppression of recombination needed to maintain sex chromosomes and co-segregation of the sex determination loci. The evolutionary strata are based on 70 alignable coding sequences and, to a lesser extent, syntenic intergenic regions. The first large-scale inversion is the oldest and most diverged and likely established the initial suppression of recombination and the inception of sex chromosomes. Inversion 1 is large and encompasses almost half of the HSY (3.6 Mb). It contains 27 paired genes that have an inverted orientation in the HSY compared to the X.

The critical sex determination genes that promote stamen development and female sterility are likely confined to first inversion in the HSY. According to Ming et al. (2007), the ancestral papaya progenitor was most likely monecious and

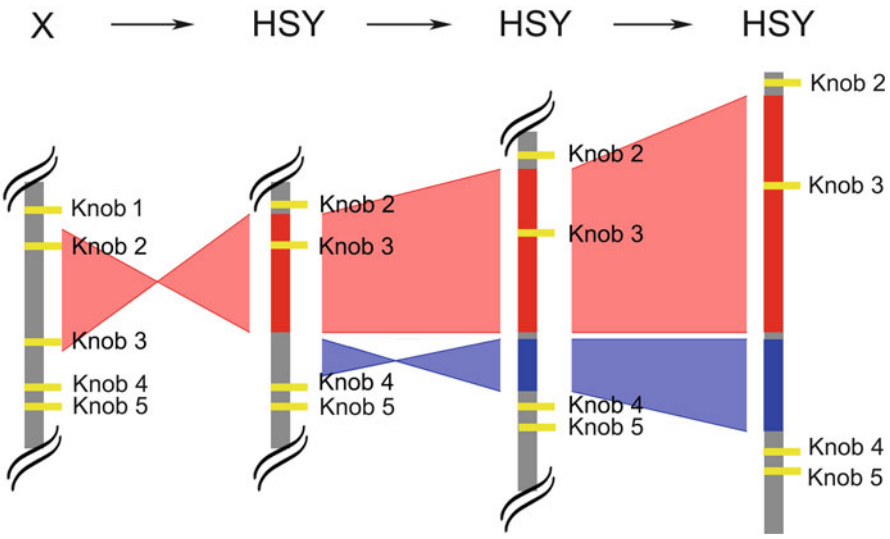

Fig. 16.4 Structure and expansion of the HSY. The inception of sex chromosomes in papaya was triggered by a large-scale inversion (*red*). The inverted region went through subsequent rearrangements and drastic expansion in the HSY. A second inversion extended the HSY (*blue*). This inversion occurred later but also caused significant expansion in the HSY

acquired mutations in two closely linked genes controlling sex. One gene induced carpel abortion and the other induced stamen abortion. Although the sex organ genes were closely linked, they still segregated in offspring, resulting in a mixture of sex types. A large-scale inversion encompassing both of these genes occurred in the HSY, initiating sex chromosome evolution. The Y chromosome contained two dominant genes that led to carpel abortion and promoted stamen development. The X chromosome contained two recessive genes, one that suppressed stamens and a second that promoted carpels. The Y^h chromosome differs from the Y chromosome in that it has a recessive gene promoting carpels, which explains the hermaphrodite sex.

The second inversion is slightly smaller than the first, at 3.1 Mb, and is much younger than the first occurring approximately 4.7 million years later. This second inversion further expanded the area of recombinational suppression and the sex-specific regions. It is unlikely that the second inversion contains the sex determination genes, as it occurred over one million years after the first inversion. The second inversion may, however, contain genes that have other sex-specific functions. We recognize at least two additional genes that are sex specific: one controlling long peduncles and a second that results in embryo lethality.

Male papaya flowers are born in clusters with long peduncles extending away from the leaf axis. Female and hermaphrodite flowers have much shorter peduncles and develop close to the leaf axis. The long peduncles of male flowers are potentially desirable for papaya cultivation. If female or hermaphrodite flowers had long peduncles, the fruits would be closer to the ground and easier to harvest.

The functional gene for long peduncles is present on the Y chromosome, but not the Y^h, as hermaphrodites have short peduncles. Cloning the gene responsible for long peduncles would allow for the creation of a hermaphrodite tree with low-hanging fruit. The HSY also contains a gene that results in embryo abortion. Papaya seeds with a combination of two Y or Y^h chromosomes abort 25–50 days after pollination, indicating that the Y chromosome contains a mutant gene essential for early embryo development. These two genes likely reside in the first or second inversions.

The collinear region is the youngest part of the HSY, with the X and HSY diverging some time after the two large inversions. As the term collinear implies, genes within this region are in the same order in the HSY and the X and have a higher sequence homology than genes in other regions. The collinear region is syntenous in the HSY and X, with few small-scale duplications, deletions, or inversions. The collinear region marks the end of the sex chromosome as it extends past border A and into the pseudo-autosome. Although infrequent, recombination likely occurs in this region with gene loci exchanged between the X and HSY. It is unlikely that the collinear region contains the sex determination genes, the gene controlling peduncle length, or the embryo lethal gene. If the sex determination genes were in this region, they could recombine, allowing for females with the Y^h chromosome and males or hermaphrodites with XX. Environmental factors, especially cool temperature, can influence sex expression, but sex reversals due to recombination have not been documented.

All of the sex-coding genes lie within the inversions and collinear regions; yet these regions only account for about half of the HSY (Fig. 16.5). The expansion blocks in the HSY are barren with an extremely low gene density and extensive repeat accumulation, rearrangements, and duplications. These regions are not syntenic and represent expanded areas in the HSY not seen in the X. The two largest expansion blocks span 1.7 Mb and 2.4 Mb and are nestled between gene islands in the inversions and collinear region. The expansion blocks contain heterochromatic knob structures 3, 4, and 5, explaining the high repeat content and low gene density. Repeat percentages in two of these blocks are 84.6 and 87.2 %, much higher than the 77 % average seen in the entire the HSY. Most of the sex-specific repeats are also confined to this region.

The most prominent characteristic of newly evolved sex chromosomes is the rapid expansion of heterogametic sex. Accumulation of repetitive sequences and degeneracy of the W or Z chromosome is explained by a phenomenon called Meullers ratchet (Meuller 1964). Under Meullers ratchet, slightly deleterious mutations can accumulate in the non-recombining regions due to transposable element insertion, a high rate of mutation, and chromosomal rearrangements (Felsenstein 1974). These factors are usually removed by purifying gene selection and chromosomal recombination. However, in sex chromosomes, there is no recombination, and purifying selection is too weak to repair or replace the deleterious mutations. Thus, mutations accumulate, unhindered, until the Y or W chromosome degenerates completely and is lost (Westergaard 1958; Steinemann and Steinemann 2005). Meullers ratchet is often cited as the driving force for evolving recombination mechanisms and also as the primary cause for sex chromosome degeneration. In the initial phases of sex

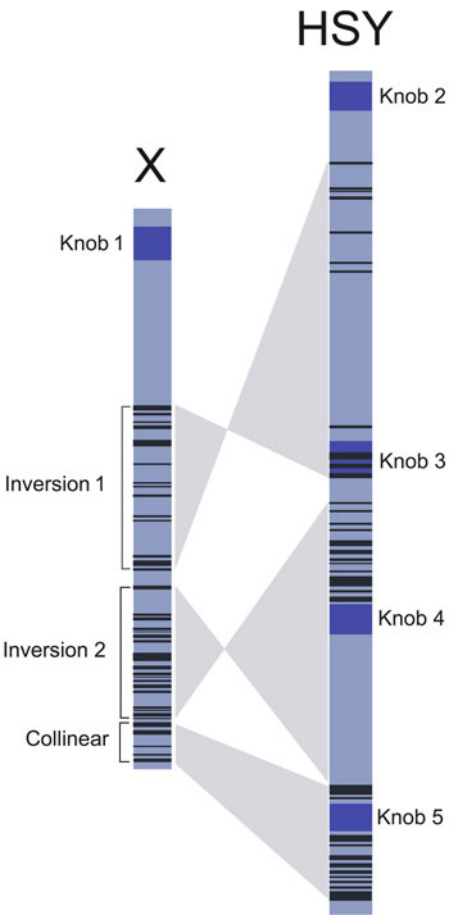

Fig. 16.5 Gene content, expansion, and rearrangements in the X and HSY. The 70 alignable transcripts are positioned in *black*, and knob structures are shown in *gray*. Knob 1 is shared between X and HSY (although it has not been sequenced in the HSY), and knobs 2–5 are HSY specific. Transcripts are contained in the two inversion regions and the collinear region. Subsequent rearrangements have changed the gene order significantly in X and HSY, and the expansion blocks in the HSY are gene poor

chromosome evolution, repetitive elements rapidly accumulate. The nascent papaya sex chromosomes are young, and the HSY is expanding rapidly. Repeat accumulation leads to the degeneration of X and Y gene pairs, until functional copies of Y paired genes are no longer present. Then, the process of expansion reverses, and the Y chromosome slowly shrinks (Steinemann and Steinemann 2005). The human Y chromosome has been subject to the pressures of Meullers ratchet for 166 million years. As a result, it has lost 1,393 of its original 1,438 genes and is three times smaller than its X counterpart (Skaletsky et al. 2003).

Repetitive Element Accumulation and Expansion

The papaya HSY is highly repetitive, and these repeats are the cause of its rapid and dramatic expansion. Overall, 77.9 % of the HSY and 59.4 % of the X is composed

Table 16.1 Repetitive elements in the papaya sex chromosomes

Sequence source	HSY 8.1 Mb			X 3.5 Mb		
Repeat class/family	No. of elements	Length occupied (bp)	Percentage of sequences (%)	No. of elements	Length occupied (bp)	Percentage of sequences (%)
Retroelements	3,962	3,491,027	43.3	1,470	1,544,912	44.7
Ty1/copia	349	358,804	4.5	284	273,411	7.9
Ty3/gypsy	2,541	2,404,485	29.8	835	950,748	27.5
DNA transposons	9	6,732	0.1	10	1,357	0
Unclassified	5,504	2,902,233	36	1,601	776,581	22.5
Total interspersed repeats	9,475	6,399,992	77.9	3,081	2,322,850	59.4

of repetitive elements (Table 16.1). These percentages are higher than the papaya whole genome average of 51.9 %, indicating that expansion and repeat accumulation has occurred extensively in both the HSY and X (Ming et al. 2008; Wang et al. 2012). Most repeats are long terminal repeats (LTR) which are characteristic of retrotransposon insertions. Most of the retrotransposons are *Ty3-gypsy* elements with an increased abundance of 20 % in the HSY and 10 % in the X compared to the genome-wide average. Although repetitive elements are particularly abundant in most plant genomes, purifying selection with homologous chromosomes has helped to slow their accumulation. Suppression of recombination in the HSY eliminates purifying selection. Consequently, the absence of purifying selection resulted in the accumulation of an enormous amount of repeats in a relatively short evolutionary time frame of 7 million years. DNA transposons are also found in the HSY and X, but are much less abundant than retrotransposons. Despite the fact that inversion 1 is considerably older and more diverged than other regions in the HSY, it contains relatively the same percentage of repeats. This could, however, be due to degradation of early retroelement insertions.

In addition to the abundance of Ty3-gypsy elements in the HSY, sex-specific repeats have also been found. The presence of sex-specific repeats is intriguing and unexpected, as during the 3 million years of evolution and divergence from the autosome, new repeats, not seen anywhere else in the genome, emerged. One novel sex-specific repeat was found in the X where it represents 3.5 % of the total sequence, and 20 novel repeats were found in the HSY where they represent 10.7 % of the total sequence. It is difficult to speculate where these repeats originated, but further annotation and classification may shed light on their origin. It is also possible that these sex-specific repeats are in reality from unassembled regions in the papaya draft genome. The draft genome is far from complete, and the majority of sequence gaps represent highly repetitive regions.

Although the majority of repeats in the sex chromosomes are LTRs and retrotransposons, there are other types of repeats, duplications, and organelle genome insertions. Inverted repeats are especially abundant in the HSY, with 5,972 repeats ranging in size from 500 bp to 28 kb. The inverted repeats are distributed unevenly in the HSY, with most localized at the knob structures and near border A.

Nine of the ten largest inverted repeats are located near border A, and they represent a significant portion of the finished sequence. The X chromosome is also rich in inverted repeats with 2,642 repeats ranging from 500 bp to 17 kb. The largest X inverted repeats are also near knob one and border A. Inverted repeats mark the boundaries of transposons, and the high proportion of inverted repeats correlates to the high percentage of transposons. In the human Y, enormous, highly homologous inverted repeats encompass 15 genes that are predominantly testes specific (Rozen et al. 2003). The palindromic inverted repeats in the human Y have almost identical genes in each of their two arms, which are maintained by gene conversion despite the lack of meiotic recombination. This indicates that the human Y has evolved mechanisms to prevent further degeneration of essential genes. Although the HSY in papaya has a large number of inverted repeats, palindrome sequences have not evolved to combat Meullers ratchet and prevent chromosome degeneration.

The HSY has a disproportionally high percentage of organelle-derived sequences. Organelle genome integration is a common feature of plant nuclear genomes. Organelle genomes were once comparable in size to cyanobacteria but, in the roughly 2 billion years since their inception, have dramatically shrunk to the size of plasmids. Nuclear genomes are the recipients of these lost sequences and have subsequently expanded in both size and metabolic function (Timmis et al. 2004). Organelle to nucleus transfers are still occurring, with both beneficial and deleterious effects. The Arabidopsis genome for instance, contains a large 620 kb mitochondrial DNA insertion, and 17 chloroplast insertions totaling 11 kb (The Arabidopsis Genome Initiative 2000). The papaya genome contains 786 kb and 858 kb of interspersed chloroplast and mitochondria sequences, respectively. Overall, the chloroplast sequences represent 0.28 % of the genome, and mitochondria represent 0.3 % of the papaya genome. Purifying selection reduces the fixation of organelle sequences in the nuclear genome. Without this selection, however, the rate of organelle integration is much higher. Chloroplast DNA insertions in the HSY are numerous, three times the genome average, encompassing 94 kb (1.15 %) of the overall sequence. Most of the integrated sequences have greater than 95 % similarity to the chloroplast genome, indicating that insertion occurred recently. A majority of the chloroplast-derived sequences are located in the expanded regions of the HSY and in the knob structures. The X region, however, contains fewer chloroplast sequences than the genome-wide average, with 10 kb chloroplast sequences or 0.18 % of the total X region. Mitochondria insertions in the X and HSY are less than the genome-wide average.

Despite the youthful nature of papaya sex chromosomes, extensive Y chromosome degradation has already occurred. The earliest and strongest evidence of this is the lethality of two Y or Y^h chromosomes in zygotes. Hermaphrodite papaya fruits can contain upwards of 500 seeds, but roughly 25 % of them are small, white, and aborted. When hermaphrodite flowers are selfed, 50 % of the seeds are hermaphrodite (XY genotype), 25 % are female (XX genotype), and 25 % are aborted (Y^hY^h) and fail to germinate. Embryo abortion of the Y^hY^h genotype indicates that a metabolic housekeeping gene or a gene involved in early embryo development has mutated and is nonfunctional in the Y^h chromosome. Without an X chromosome to

balance this recessive mutation and facilitate normal development, these seeds abort 25–50 days after fertilization. The suppression of recombination in the Y^h inhibits the purifying selection needed to eliminate the recessive loss of function mutation from the population. The lethal factor on the Y and Y^h chromosomes is an indication of the degenerative powers of Mueller's ratchet.

Gene Content and Trafficking in the HSY and X

Evidence of Muller's ratchet is prominent in the papaya sex chromosomes as the HSY has an extremely low gene density relative to the autosome, and it has a higher percentage of pseudogenes compared to the X. A total of 96 transcripts have been annotated in the HSY, and 98 have been annotated in the X. Transcribed gene products were identified using a combination of papaya-expressed sequence tags (EST), gene models, and two gene prediction programs. Coding sequences were verified using RT-PCR, and the resulting products were sequenced to alleviate the high false discovery rate. Sixty four of the gene products are paired between the X and HSY, 29 are X specific and 22 are HSY specific. Most of the paired genes code for full-length, functional proteins, but several are classified as pseudogenes. Pseudogenes represent transcripts with premature stop codons, truncated protein products, or frame shift mutations that result in a nonfunctional protein. Pseudogenes can result from transposon insertion into genic regions, non-synonymous substitutions, and deletions that cause frameshift mutations. The predicted number of pseudogenes in the HSY may be an underestimation, as proteins can lose function due to non-synonymous mutations that result in a change in amino acids, but not necessarily a truncated protein with a premature stop codon.

Of the transcribed units, nearly all of the paired genes (98 %) have homology to known proteins in the NCBI database. The paired genes have a range of functions, from metabolic housekeeping genes such as formate dehydrogenase, to specific developmental genes like flowering locus T (Wang et al. 2012). When sex chromosomes initially evolve, they cause the co-segregation of not only the sex determination genes, but also any flanking autosomal genes (Rice 1987). Suppression of recombination subjects the paired genes to the forces of Meullers ratchet, causing them to degenerate. In the human Y chromosome, most of the genes from the ancestral autosome have either been deleted, truncated, or acquired missense mutations. Fifteen of the genes are protected from degeneration via gene conversion between the palindromic inverted repeats (Rozen et al. 2003). The palindromic sequences protect essential genes, but the remainder will likely degrade. Gene content has been extensively eroded in mammalian heterogametic chromosomes because of their ancestral states. Recently emerged sex chromosomes show varying levels of gene erosion. The average expression ratio of paired genes on the Y chromosome of Selene is less than one, indicating that degeneration is already underway. Degeneration of paired genes in papaya is apparently less, as expression ratios between the sexes are around one (VanBuren, unpublished data).

Twenty two Y-specific genes reside in the HSY. Most of the Y-specific genes are the result of retrotransposon-mediated gene trafficking from the autosome. Newly acquired retrotransposons are abundant in the HSY, and when they transpose, they often bring flanking genome sequences with them. Transcription of active retrotransposons can extend beyond the LTR and into the surrounding genome sequence. Reverse transcriptase converts this RNA intermediate into extrachromosomal DNA which is then inserted randomly back into the genome (Havecker et al. 2004). The newly integrated retrotransposons can bring along full-length genes. Deleterious retrotransposons that integrate into genic regions are removed from the population by purifying selection. The lack of recombination in the HSY prevents the removal of deleterious insertions, contributing to its degeneration. Several protein-coding genes are truncated in the HSY, likely the result of transposon integration. Retrotransposon insertion can be beneficial, however, as it can foster gene duplication. Six of the HSY-specific genes have a high homology to the autosome, suggesting that they could be the result of retrotransposon-mediated trafficking. Genes from the autosome can undergo duplication and translocation, allowing them to take on new functions.

Much of our understanding into the nature of sex chromosome evolution has come from studying gene loss, rearrangements, and repeat content in the heterogametic sex chromosome (Y or W). In mammals, studying the homogametic sex chromosome is more difficult, as the autosomal regions corresponding to the sex chromosomes have long since disappeared. Genome-wide comparison of the Z chromosome in chicken and X chromosome in human revealed significantly lower gene content and higher repeat composition in the sex chromosomes (Bellott et al. 2010). Studying X chromosome expansion in papaya is straightforward, as a closely related monecious species serves as a valuable reference for comparison. *Vasconcellea monoica* is the only monecious members in the *Caricaceae* family, bearing complete flowers and an absence of sex chromosomes. Furthermore, *V. monoica* diverged from papaya roughly 28 MYA, making it a valuable outgroup (Wu et al. 2010; Gschwend et al. 2012). Since *V. monoica* has an orthologous autosomal region corresponding to the X in papaya, changes in gene content, organization, and repeat accumulation can be accurately deciphered.

The X chromosome in papaya has changed significantly in comparison to the orthologous region in *V. monoica.* This difference dismisses the notion that X chromosomes remain relatively static throughout evolution (Gschwend et al. 2012). Significant repeat accumulations, gene losses and gains, and rearrangements are found in the papaya X chromosome. Eleven BACs (totaling 1.1 Mb) were sequenced from *V. monoica*, corresponding to 2.56 Mb of the papaya X. Eighteen transcription units were annotated in the X and 19 were annotated in monoica. Ten of the genes are monoica specific and 9 are papaya specific, suggesting significant divergence and changes in gene content. Five syntenic regions are shared between monoica and papaya and show marked expansion in papaya. The *V. monoica* genome is estimated to be 630 Mb (Ming, unpublished data), which is almost 40 % larger than the papaya genome estimate (Arumuganathan and Earle 1991). The *V. monoica* genome contains a higher percentage of repeats and should show expansion in comparison to the X,

as seen in syntenic papaya genome regions. Instead, the X regions show expansion, indicating that the X has diverged significantly from the autosomal precursor. The structural changes displayed by the X are likely due to a lack of recombination in the heterogametic male or hermaphrodite sex, as the X and Y or Y^h cannot recombine. Orthologous sequence comparisons provide direct evidence that the X chromosome is evolving faster than the autosome.

Concluding Remarks and Future Prospects

The papaya X and HSY represent only the second complete pair of sequenced sex-specific regions. Physical mapping of the sex chromosomes was a difficult and reiterative process that began with the identification of several co-segregating sex-specific markers. Three large contigs consisting of BACs joined through fingerprint mapping were anchored to the papaya HSY using sex-specific SCAR markers. A BAC by BAC approach was applied to extend the contigs until nearly complete physical maps were generated. BACs were sequenced and assembled to produce an 8.1 Mb pseudomolecule in the HSY and a 3.5 Mb pseudomolecule in the X. Repeat content is significantly higher in the HSY, and retrotransposons are the driving force of the expansion. Gene loss, truncation, and rearrangements are a prominent feature of the HSY. Syntenic analysis between alignable transcripts revealed three distinct regions. The first two regions are large-scale inversions that form separate evolutionary strata. The third is a collinear region where gene content and order are relatively the same. The first large-scale inversion likely spurred the initiation of sex chromosomes, and the second inversion expanded the size of the HSY. Expansion and gene loss was also detected in the X. An orthologous, autosomal region corresponding to the X was sequenced and annotated from *V. monoica*, revealing repeat expansion and gene trafficking. The fine resolution of the sequenced sex-specific regions clearly shows two large-scale inversions that fostered the suppression of recombination necessary for emerging sex chromosomes. Suppression of recombination subjected the HSY to the powerful forces of Meuller's ratchet, causing rapid expansion and degeneration. The young age of the papaya sex chromosomes makes them an important tool for studying the early events of sex chromosome evolution.

Discovering the sex determination gene controlling stamen development could revolutionize the papaya industry. Successful transformation of the stamen-promoting gene will complement the loss of function mutation in the female X chromosomes, resulting in hermaphrodite papaya. Agriculturally, the segregation of hermaphrodite plants is laborious and costly. Female trees are undesirable, as they cannot self-fertilize, so they must be manually removed by hand. A map-based approach for identifying the sex determination genes is not possible in papaya because of the lack of recombination in the sex determination regions. Thus, the most feasibly way to identify candidate genes is through physical mapping, sequencing, and complete annotation of the sex chromosomes. Annotation has revealed a number of strong candidate genes

for sex determination. Analyzing expression of these genes in sex reversal mutants will help facilitate the identification of sex determination genes, ultimately resulting in the generation of a true-breeding hermaphrodite papaya.

References

Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. Plant Mol Biol Rep 9:208–218

Bachtrog D (2003) Adaptation shapes patterns of genome evolution on sexual and asexual chromosomes in *Drosophila*. Nat Genet 34:215–219

Bachtrog D, Charlesworth B (2002) Reduced adaptation of a non-recombining neo-Y chromosome. Nature 416:323–326

Bellott DW, Skaletsky H, Pynitkova T, Mardis ER et al (2010) Convergent evolution of chicken Z and human X chromosomes by expansion and gene acquisition. Nature 466:612–616

Bergero R, Charlesworth D (2009) The evolution of restricted recombination in sex chromosomes. Trends Ecol Evol 24:94–102

Bergero R, Forrest A, Kamau E, Charlesworth D (2007) Evolutionary strata on the X chromosomes of the dioecious plant *Silene latifolia*: evidence from new sex-linked genes. Genetics 175:1945–1954

Blas AL, Yu Q, Chen C, Veatch O, Moore PH, Paull RE, Ming R (2009) Enrichment of a papaya high-density genetic map with AFLP markers. Genome 52:716–725

Charlesworth D (1985) Distribution of dioecy and self-incompatibility in angiosperms. In: Greenwood PJ, Slatkin M (eds) Evolution—essays in honour of John Maynard Smith. Cambridge University Press, Cambridge, pp 237–268

Charlesworth B, Charlesworth D (1978) A model for the evolution of dioecy and gynodioecy. Am Nat 112:975–997

Charlesworth B, Morgan MT, Charlesworth D (1993) The effect of deleterious mutations on neutral molecular variation. Genetics 134:1289–1303

Charlesworth D, Charlesworth B, Marais G (2005) Steps in the evolution of heteromorphic sex chromosomes. Heredity 95:118–128

Delph LF, Arntz AM, Scotti-Saintagne C, Scotti I (2010) The genomic architecture of sexual dimorphism in the dioecious plant *Silene latifolia*. Evolution 64:2873–2886

Felsenstein J (1974) The evolutionary advantage of recombination. Genetics 78:737–756

Graves JAM, Shetty S (2001) Sex from W to Z: evolution of vertebrate sex chromosomes and sex determining genes. J Exp Zool 290:449–462

Gschwend AR, Yu Q, Tong EJ, Zeng F, Han J, VanBuren R et al (2012) Rapid divergence and expansion of the X chromosome in papaya. Proc Natl Acad Sci USA 109:13716–13721

Havecker ER, Gao X, Voytas DF (2004) The diversity of LTR retrotransposons. Genome Biol 5:225

Hughes JF, Skaletsky H, Pyntikova T, Graves TA et al (2010) Chimpanzee and human Y chromosomes are remarkably divergent in structure and gene content. Nature 463:536–539

Liu Z, Moore PH, Ma H, Ackerman CM, Makandar R, Yu Q, Pearl HM, Kim MS, Charlton JW, Stiles JI, Zee FT, Paterson AH, Ming R (2004) A primitive Y chromosome in papaya marks incipient sex chromosome evolution. Nature 427:348–352

Meuller HJ (1964) The relation of recombination to mutational advance. Mutat Res 106:2–9

Ming R, Yu Q, Moore PH (2007) Sex determination in papaya. Semin Cell Dev Biol 18:401–408

Ming R, Hou S, Feng Y, Yu Q et al (2008) The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya Linnaeus*). Nature 452:991–996

Ming R, Bendahmane A, Renner SS (2011) Sex chromosomes in land plants. Annu Rev Plant Biol 62:485–514

Na JK, Wang J, Murray JE, Gschwend AR, Zhang W, Yu Q et al (2012) Construction of physical maps for the sex-specific regions of papaya sex chromosomes. BMC Genomics 13:176

Renner SS, Ricklefs RE (1995) Dioecy and its correlates in the flowering plants. Am J Bot 82:596–606

Rice WR (1987) Genetic hitchhiking and the evolution of reduced genetic activity of the Y sex chromosome. Genetics 116:161–167

Ross MT, Grafham DV, Coffey AJ, Scherer S et al (2005) The DNA sequence of the human X chromosome. Nature 434:325–337

Rozen S, Skaletsky H, Marszalek JD, Minx PJ et al (2003) Abundant gene conversion between arms of massive palindromes in human and ape Y chromosomes. Nature 423:873–876

Skaletsky H, Kuroda-kawaguchi T, Minx PJ, Cordum HS et al (2003) The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature 423:825–837

Spigler RB, Lewers KS, Main DS, Ashman TL (2008) Genetic mapping of sex determination in a wild strawberry, *Fragaria virginiana*, reveals earliest form of sex chromosome. Heredity 101:507–517

Spigler RB, Lewers KS, Johnson AL, Ashman TL (2010) Comparative mapping reveals autosomal origin of sex chromosome in octoploid *Fragaria virginiana*. J Hered 101:107–117

Steinemann S, Steinemann M (2005) Y chromosomes: born to be destroyed. Bioessays 27: 1076–1083

Storey WB (1938) Segregations of sex types in Solo papaya and their application to the selection of seed. Proc Am Soc Hortic Sci 35:83–85

Storey WB (1941) The botany and sex relations of the papaya. Hawaii Agric Exp Sta Bul 87:5–22

Tanurdzic M, Banks JA (2004) Sex-determining mechanisms in land plants. Plant Cell 16:61–71

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796–815

Timmis JN, Ayliffe MA, Huang CY, Martin W (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat Genet Rev 5:123–135

Tsuda Y, Nishida-Umehara C, Ishijima J, Yamada K, Matsuda Y (2007) Comparison of the Z and W sex chromosomal architectures in elegant crested tinamou (*Eudromia elegans*) and ostrich (*Struthio camelus*) and the process of sex chromosome differentiation in palaeognathous birds. Chromosoma 116:159–173

Wang J, Na JK, Yu Q, Gschwend AR, Han J et al (2012) Sequencing papaya X and Yh chromosomes reveals molecular basis of incipient sex chromosome evolution. Proc Natl Acad Sci USA 109:13710–13715

Wellmer F, Riechmann JL, Alves-Ferreira M, Meyerowitz EM (2004) Genome-wide analysis of spatial gene expression in *Arabidopsis* flowers. Plant Cell 15:1314–1326

Westergaard M (1958) The mechanism of sex determination in dioecious flowering plants. Adv Genet 9:217–281

Wu X, Wang J, Na JK, Yu Q, Moore RC, Zee F, Huber SC, Ming R (2010) The origin of the non-recombining region of sex chromosomes in *Carica* and *Vasconcellea*. Plant J 63:801–810

Yu Q, Hou S, Feltus FA, Jones MR et al (2008) Low X/Y divergence in four pairs of papaya sex-linked genes. Plant J 53:124–132

Zhang W, Wang X, Yu Q, Ming R, Jiang J (2008) DNA methylation and heterochromatinization in the male-specific region of the primitive Y chromosome of papaya. Genome Res 18:1938–1943

Zhang W, Wai CM, Ming R, Yu Q, Jiang J (2010) Integration of genetic and cytological maps and development of a pachytene chromosome-based karyotype in papaya. Trop Plant Biol 3:166–170

Part IV
Early Applications and Future Prospect

Chapter 17
Physical Mapping of Papaya Sex Chromosomes

Jianping Wang, Jong-Kuk Na, and Ray Ming

Introduction

The vast majority of the flowering plants (angiosperms) are hermaphrodite; that is, they produce flowers that contain both female and male reproductive structures: stamen and pistil. Each individual perfect flower can be self-or cross-pollinated to produce seeds. However, about 6 % (14,620 of 240,000) of the flowering plant species are dioecious (Renner and Ricklefs 1995). In dioecious species, such as asparagus, separate unisexual (male or female) flowers occur on different individual plants. As obligate outcrossers, dioecious female plants can produce seeds only after fertilization by pollen from a male plant. Although the dioecious condition with different sex types of individuals is not predominant in flowering plants, it is distributed widely across 75 % of the angiosperm families, including all six dicotyledonous and all five monocotyledonous subclasses (Yampolsky and Yampolsky 1922). In monoecious species, such as maize, the unisexual female and male flowers occur on the same individual plant at different parts. In gynomonoecious and andromonoecious species, both bisexual and unisexual flowers are produced on the same individual plants.

These various plant sexual reproductive modes or systems are the result of diverse sex determination mechanisms. Developmental examination of unisexual flowers revealed that a complete set of floral organs including sepals, petals,

J. Wang (✉)
Department of Agronomy, Genetics Institute, Plant Molecular and Cellular Biology Program, University of Florida, 2223 Mowry Road, Gainesville, FL 32610, USA
e-mail: wangjp@ufl.edu

J.-K. Na
Department of Agricultural Bio-Resources, National Academy of Agricultural Science, RDA, Suwon, Kyunggi-Do 441-707, Republic of Korea

R. Ming
Department of Plant Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

R. Ming and P.H. Moore (eds.), *Genetics and Genomics of Papaya*, Plant Genetics and Genomics: Crops and Models 10, DOI 10.1007/978-1-4614-8087-7_17,

stamens, and pistils are initiated, as hermaphroditic floral primordia, in both sex types of flower. The further selective development and abortion of stamens and pistils respectively lead to the physical separation of male and female unisexual flower (Irish and Nelson 1989). Therefore, sex determination in most angiosperm species occurs via selective abortion of flower organs (Tanurdzic and Banks 2004). In monoecious maize and cucumber, for example, cloning of the sex-determining genes leads to the finding that the gibberellin and ethylene play a pivotal role in sex determination process through regulating and controlling stamen abortion and feminization of flowers (Spray et al. 1996; Perl-Treves 1999). Sex expression in cucumber and maize is also influenced by environmental factors such as light intensity, day length, temperature, or mineral nutrition (Frankel and Galun 1977; Chailakhyan 1979), which may affect hormone status and then trigger abortion of male or female organs to determine the sex (Zhang et al. 2005; Adam et al. 2011).

In some dioecious species, specialized sex chromosomes have evolved relatively rapidly and independently in many taxonomic groups (Bull 1983; Nicolas et al. 2005). Chromosomal sex determination systems may be an evolutionary consequence of natural selection in favor of dioecy. The evolving process of sex chromosomes involves two critical events: linkage of two sex determination genes (one aborting stamens and the other aborting carpels) and suppression of recombination between the two genes caused by chromosome rearrangements or other mechanisms such as DNA methylation. Once the two sex determination genes are completely suppressed from recombination, a pair of incipient chromosomes arises (Ming et al. 2011). For example, in wild strawberry (*Fragaria virginiana*), the two sex determination genes, male sterility and female fertility, are linked on the same chromosome with a genetic distance of 5.6 cM (Spigler et al. 2008, 2010). Without suppression of recombination between the two genes, the male sterile or female sterile mutations could revert to hermaphrodite or neuter individuals, which represent the earliest stage of the sex chromosome evolution. While in dioecious *Silene* species, the sex chromosomes are morphologically distinct and exhibit a progressive suppression of recombination that gradually diminishes toward the pseudoautosomal region between the diverged sex chromosomes (Nicolas et al. 2005), thus representing a prototypical later stage of sex chromosome evolution.

The unisexual flower development reassembles the bisexual perfect flower development at the early stages. This developmental similarity suggests that the plant species with unisexual flowers have evolved from species having perfect flowers and that the sex chromosomes have evolved from a pair of autosomes, regardless of taxonomic classification, due to the common selective pressures of gender maintenance and sexual selection (Charlesworth 2002). Examining young sex chromosomes at various early evolutionary stages allows comprehensive analyses of the evolutionary processes as nucleotide signatures are still present in the DNA sequence. The remnants of nucleotide signatures in young sex chromosomes is in contrast to the situation of more ancient sex chromosomes in which historic processes must be inferred from fewer nucleotide signatures. Six stages of sex chromosomes evolution were proposed based on current findings of various early stages of sex chromosomes (Ming et al. 2011). At stage 1, the male sterile and female fertility loci and their

complementary loci occur in close proximity on a chromosome, as in wild strawberry. At stage 2, the recombination between the two sex-determining loci is suppressed and a small male-specific region on the nascent Y chromosome (MSY) is formed, and the YY (or WW) genotype is viable, as in asparagus. At stage 3, the suppression recombination extends to the neighboring regions of MSY and the YY genotype is lethal, as in papaya. At stage 4, the MSY spreads to the majority of the Y chromosome and further degeneration occurs, as in white campion. At stage 5, severe degeneration of the Y chromosome causes the loss of function for most genes and there is a shrinking of the Y chromosome, as in *Cycas revoluta*. At stage 6, the suppression of recombination spreads to the entire Y chromosome and sex determination is controlled by X to autosome ratio, as in sorrel. Functional analysis of the plant sex chromosomes at various evolutionary stages could uncover the molecular mechanisms of chromosomal sex determination and document the events and driving forces at these various evolutionary stages of sex chromosomes.

Papaya Sex Determination

The tropical fruit crop papaya (*Carica papaya* L.) is a trioecious species with three sex types of individual plants: male, female, and hermaphrodite. Though natural papaya populations are mostly dioecious, the hermaphrodite plants are preferred for papaya fruit production due to their desirable fruit size and shape and higher yield since every tree produces fruit from self-pollination.

The papaya inflorescences of the three sex types are morphologically distinct. Female flowers are large and born singularly on the stem; male flowers are smaller and cluster in a large inflorescence characterized by having long peduncles from the branches, and hermaphroditic flower is born individually from the stem with an intermediate shape between the female and male flower (Fig. 17.1a–c).

Genetic analysis of sex segregation ratio led to the hypothesis that papaya sex is determined by a single locus with three different alleles: M, M^h, and m that segregate in Mendelian inheritance (Hofmeyr 1938; Storey 1938). Females (mm) were thought to be homozygous recessive at the sex-determining locus, whereas males (Mm) and hermaphrodites (M^hm) were heterozygous. The genotypes with

Fig. 17.1 (**a–c**) The papaya flowers. (**a**) Female flower on gynoecious plant. (**b**) Male flower on androecious plant. (**c**) Perfect flower on hermaphrodite plant

homozygous dominant alleles, MM, M^hM^h, and MM^h, are lethal. Self-pollinated hermaphrodites result in a 2:1 segregation of hermaphrodite to female, and cross-pollination between male and hermaphrodite results in a 1:1:1 segregation of male, hermaphrodite, and female.

Among the nine pairs of chromosomes in papaya, there is no morphologically distinctive chromosome pair as seen in humans or the plant white campion. Based on the observations of co-segregation of the long peduncles with male flowers and the lethal factor in hermaphrodites and males, Storey (1953) proposed that papaya sex is determined by a group of closely linked genes within a small region on the sex chromosome where recombination within this region is suppressed. Hofmeyr (1939, 1967) hypothesized that the M1 (M) and M2 (M^h) lack vital genes, which cause lethality of the homozygous dominant genotypes. Meanwhile, Horovitz and Jiménez (1967) proposed that sex determination in papaya is an XY sex chromosome system based on the interspecific hybridization patterns among *Vasconcellea* species where homozygous dominant genotypes were not viable because embryos aborted several weeks after pollination (Chiu 2000; Chiu et al. 2003).

A high-density genetic map of papaya revealed that recombination was suppressed with 225 sex-co-segregated markers accounting for 66 % of the markers on the linkage group in the region containing the sex determination locus (Ma et al. 2004). A physical map of this recombination suppressed region, using a bacterial artificial chromosome (BAC) library, resulted in an estimated size of 4–5 Mb or a 10 % of the chromosome at this region (Liu et al. 2004). Selective sequencing of the BAC clones identified by male-specific markers revealed a decreased gene density and an increased transposable element density in this non-recombining male-specific region. These findings led to the conclusion that a pair of incipient sex chromosomes has evolved in papaya (Liu et al. 2004). Additional cytogenetic evidence supporting the existence of sex chromosomes in papaya is that a pair of chromosomes was observed separating precociously in 60–70 % of pollen mother cells (Kumar et al. 1945; Storey 1953), which is likely due to the lack of homology over the non-recombination region (Liu et al. 2004). The papaya females are thus concluded to be homogametic XX, while males are heterogametic XY. The hermaphroditic plant carries a slightly different Y, designated as Y^h; therefore, it is designated as the heterogametic genotype of XY^h. The non-recombining male- or hermaphrodite-specific region of the Y or Y^h chromosome is designated as MSY or HSY, respectively. The HSY region spans approximately 13 % of the papaya Y^h chromosome with the centromere embedded in this region (Zhang et al. 2008).

Some characteristics of papaya sex chromosomes, especially the HSY, have been obtained by genomic and cytogenetic evidence and small-scale sequence analysis. A physical map of the hermaphrodite papaya genome was constructed by fingerprinting 38,522 BAC clones (Yu et al. 2009). Integration of the physical and genetic maps revealed that the suppressed recombination rates gradually recover to the genomic average in the borders of the HSY region. Proceeding outward from the centromere, the recombination rates dramatically exceed the genome-wide average by sevenfold at both sides of the HSY, demonstrating the dynamics of recombination surrounding the HSY.

Physical Mapping of Papaya Sex Chromosomes

The papaya genome was sequenced using genomic DNA from a female plant of the transgenic variety SunUp to avoid the complication for assembling the heterozygous HSY and its corresponding region (Ming et al. 2008). The HSY sequence is not included in the draft genome, and the X chromosome sequence is incomplete because of sequence gaps. Obtaining the complete genomic sequences of HSY and its X counterpart is an essential step for identification of sex determination genes in papaya, because there is no recombination within the HSY. Having physical maps with large insert genomic clones covering the sex determination region in both the HSY and its X counterpart is a necessary resource for sequencing and correctly assembling this pericentromeric and heterochromatic region. The draft genome sequence of the papaya female SunUp yielded scaffolds of the X chromosome (Ming et al. 2008), but many gaps remain. Physical mapping of the HSY and the X corresponding region will provide minimum tiling paths for complete sequencing of the sex determination region and to finish the genome of papaya when sequencing technologies of ultra long reads become available to sequence across heterochromatic regions.

Physical Mapping of the HSY

The first effort toward HSY physical mapping was initiated by using male-specific or sex-co-segregating markers to screen the hermaphrodite BAC library (Liu et al. 2004). Specifically, the sex-co-segregating markers, T12 and W11 (Deputy et al. 2002), were used as probes for library screening; four positive BACs were derived from marker W11. The BAC end sequences (BES) were used for PCR-based configuration to determine the two most extended BAC ends which were then used as probes for the next round of BAC library screening to identify overlapping BACs and to extend the existing physical map. This stepwise chromosome walking process extended the HSY physical map to about 900 kb but was slowed down by frequently encountering repetitive sequences while chromosome walking. Therefore, 85 of the 225 sex-co-segregating AFLP markers ranging from 200 to 500 bp (Ma et al. 2004) were selected as additional markers. These markers were developed into *Carica papaya* sex-linked markers (CPSM) for the BAC library screening to identify additional BACs that could be used as new starting points to extend the physical map. The extensive library screening using CPSMs resulted in five contigs having a combined size of 2.5 Mb (Liu et al. 2004).

Seven HSY BACs were then sequenced to further assist physical mapping (Yu et al. 2007, 2008). Sequences of these seven BACs, along with BES and the genome-wide physical map based on fingerprints of the entire hermaphrodite BAC library, were then available for accelerated construction of the HSY physical map (Yu et al. 2009). BACs on the contigs from chromosome walking were searched in the fingerprint map to identify BACs that could extend the existing contigs or fill the remaining gaps. The BACs presumably extending from the contigs were selected and validated by multiple PCR confirmations. The BACs at the extreme ends of the

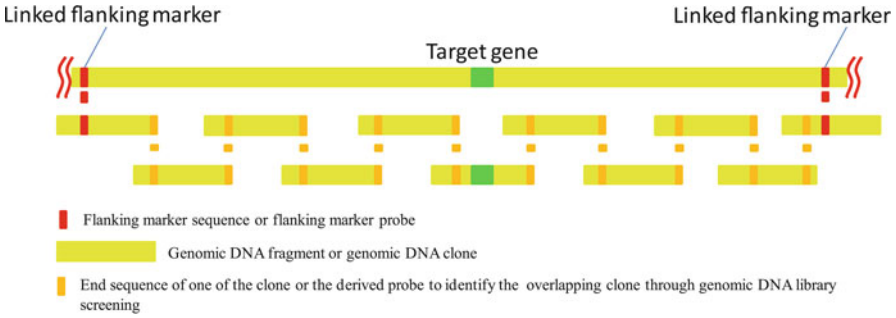

Fig. 17.2 Schematic illustration of the map-based cloning through chromosome walking

existing contigs were then used as seed BACs for the next round of BAC library screening or superpool screening (Fig. 17.2). Superpool screening consisted of using pooled BAC DNA from twelve 384-well plates to screen row, column, and diagonal pools from each superpool to identify positive BACs.

HSY physical mapping, using the previously described chromosome walking strategies, then reached 8.5 Mb, consisting of four contigs with three gaps. No further progress was made from the existing library screening indicating that the available genomic resources were exhausted. Subsequently, two additional BAC libraries, one from female SunUp and the other from male AU9, were constructed for extending the contigs and filling the gaps (Gschwend et al. 2011). Two BACs from the male AU9 BAC library were identified which filled two gaps, leaving one gap remaining. An additional male BAC clone further extended the extended contig and narrowed the remaining gap (Fig. 17.3a). The final HSY physical map spanned about 8.5 Mb, including shared border regions with the X chromosome with one gap near the border unfilled (Na et al. 2012).

Overall, more than 300 probes were used for screening three BAC libraries resulting in thousands of positive BAC clones. A majority of positive clones from the library screening had been excluded due to less than satisfactory extension, redundancy, or false positives as a result of repetitive elements. The HSY physical map constructed was based on multiple lines of evidence including Southern hybridization, PCR validation, sequencing of PCR products, and ultimately mapping of positive BACs to the Y^h chromosome using florescent in situ hybridization (FISH). This map covered the four HSY-specific heterochromatic Knobs with no gaps in these regions (Zhang et al. 2008). A total of 68 BACs on the minimum tiling path of the HSY physical map were selected as candidate BACs for sequencing the HSY (Fig. 17.3a).

Physical Mapping of X Corresponding Region

Before the papaya female genome became available, five X-specific genes were used to identify X-specific BAC clones from the hermaphrodite library. Two clones: SH53E18 and SH61H12, having genic regions with high sequence identity

(>95 %) to two HSY BACs, SH 95B12 and SH85B24, respectively, were selected, sequenced, and used as seed BACs to start chromosome walking on the X physical map using the same procedure as the one used for the HSY physical mapping (Yu et al. 2008). The papaya female genome sequence and the integrated genetic map, constructed by sequence tagged SSR markers, then became available as the source sequences of sex-co-segregating SSR markers (Chen et al. 2007). These sequences were used for probe design and BAC library screening to identify new starting points for chromosome walking and to maximize the efficiency of extending the physical map. Meanwhile, the scaffolds of the draft female genome and the integrated physical map were also used to identify additional staring points of BAC contigs in the X region (Na et al. 2012). The papaya draft genome and physical map are the primary resources that accelerated physical mapping of the X corresponding regions. A number of contigs were selected from the physical map based on two scaffolds of the draft genome. Validation and gap filling were carried out to complete the physical map of the X counterpart of HSY.

The final X physical map consisted of 44 BAC clones and extended about 5.4 Mb with an unfilled gap in the middle, which was closed on the corresponding HSY (Na et al. 2012) (Fig. 17.3). The remaining gap is likely the centromere since repetitive sequences peaked at this region in both papaya and its close relative, the monoecious *Vasconcellea monoica* (Gschwend et al. 2012). All BACs on the minimum tiling path were confirmed by FISH mapping to the X chromosome before sequencing individual BACs using the Sanger method.

Physical Mapping of the HSY Borders

Completely mapping the borders of the non-recombining sex determination region would provide crucial information about how suppression of recombination progresses along young sex chromosomes and also ensure the completeness of the physical maps of papaya's sex determination region. Initially two overlapping BACs SH86B15 and SH85C03 were identified through chromosome walking for the HSY physical map. Soon after, the same BACs were identified during X chromosome walking suggesting that the mapped location of these shared BAC clones might be in the border region. The combination of the two sex-linked SSR markers and chromosome walking identified additional BAC clones located at the two borders of the sex determination region through library screening using probes from the SSR source sequences (Chen et al. 2007). Two clones, SH94G23 and SH68N07, were identified and designated as border A and border B, respectively. At border B, BAC SH68N07 and SH86B15 were present in the same fingerprint map and physically located about 1 Mb apart. Therefore, the minimal tiling path between these two BACs was established by the fingerprint map and confirmed by PCR configuration among the overlapped BACs. At border A, SH69A05 was used as a seed BAC for chromosome walking in both directions until one end reached the most extended BAC on X physical map. However, further chromosome walking could not reach to the most extended clone at the HSY physical map and a gap remained unfilled as

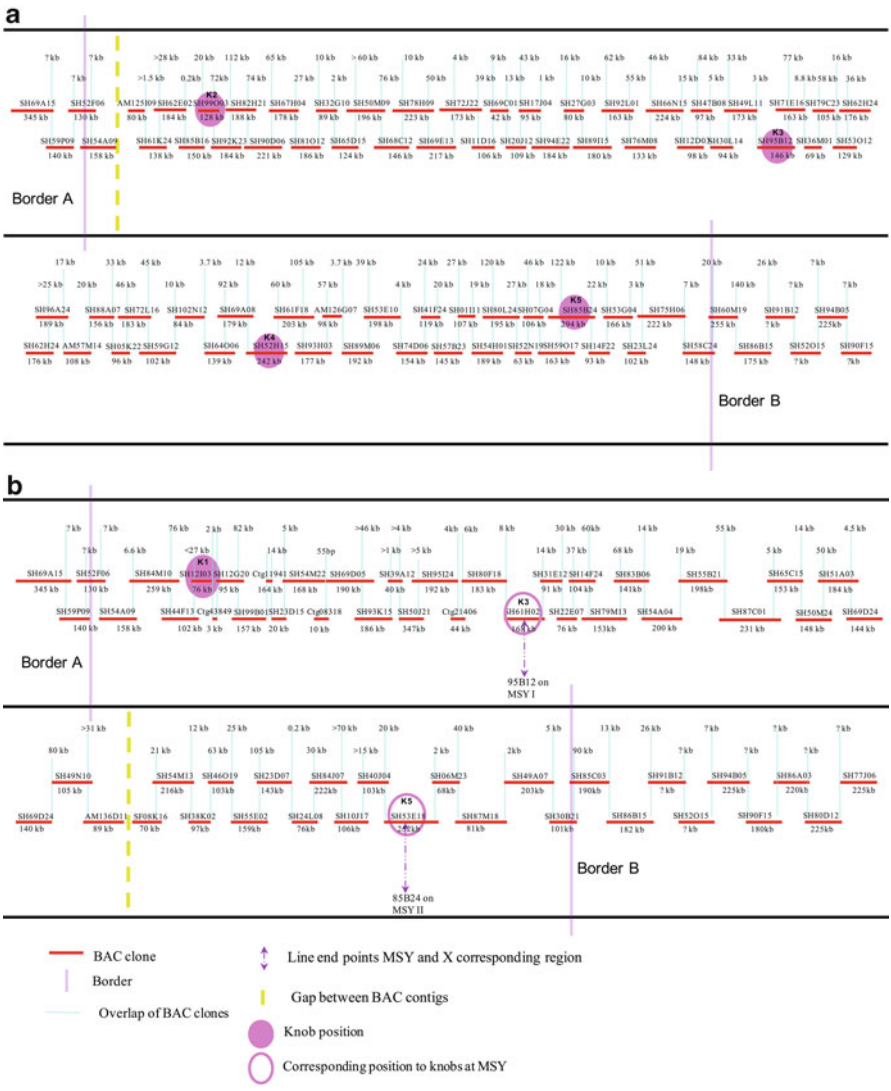

Fig. 17.3 (**a**, **b**) The physical maps of the HSY and its corresponding X region with BAC clones on the minimum tiling path. *Red lines* represent individual BAC clone with clone ID on the *top* and estimated insert size at the *bottom*. *Blue dotted lines* indicate the overlap of neighboring BAC clones. *K1–K5* indicate the position of heterochromatic knob-like structures (Zhang et al. 2008). *Purple lines* define the borders of the non-recombination region genetically by fine mapping. *SH* SunUp hermaphrodite BACs, *SF* SunUp female BACs, *AM* AU9 male BACs [reproduced with permission of BioMed Central from Na et al. (2012)]

shown in the HSY physical map (Fig. 17.3a, b; Na et al. 2012). Attempt to fill this gap identified clones that were all on the X chromosome without any evidence of being Y^h specific. This region contains knob 1, the only knob structure shared by both X and HSY, about A 900 kb (Wang et al. 2012).

Table 17.1 The physical location of SSR markers on the physical map and the number of recombinants at the given locus in the populations

	SSR markers	Marker location in BAC	Number of recombinants
Border A	CPM1055	SH69A15	4
Region	spctg177-12a	SH52F06	2
	84M10ctg6-34a/b	SH84M10	0
Non-recombining region on Y chromosome			
Border B	P3K4642	SH53E10	0
Region	58C24-25b	SH58C24	1
	P3K8303	SH85C03	3
	BorderB-2001a	SH86B15	3
	P3K2608	SH68N07	87

Fine Mapping of the HSY

The high-density genetic map exhibiting suppression of recombination in the HSY in linkage group 1 (LG1) was designated as the papaya sex chromosomes (Ma et al. 2004; Chen et al. 2007). Two SSR markers, CPM1055C0 and P3K2608C0, were linked with sex at a genetic distance of 2–3 cM to the sex determination region (Chen et al. 2007). A fine mapping approach was used to narrow down the genetic distance and define the borders of the sex determination region at a higher resolution by developing additional markers at the border region and screening for recombinants in a large F_2 mapping population consisting of 1,460 progeny (Na et al. 2012). Among the sex segregating population, 452 were phenotyped by their flower morphology as female and 1,008 as hermaphrodite.

The two flanking SSR markers, CPM1055C0 locus on border A and P3K2608C0 locus on border B, identified four recombinants from the F_2 populations at the border A and 87 recombinants at border B, respectively (Table 17.1). Seventy-five SSR markers (32 for border A and 43 for border B, respectively) were developed from available BAC end sequences, the draft genome sequences, and newly sequenced BAC sequences near the two borders of the physical maps to narrow down the interval between these two borders. Among the new SSR markers, five were polymorphic at border A region and two were polymorphic at border B region, respectively (Na et al. 2012). Genotyping the previously mentioned recombinants using these polymorphic markers revealed that the sex determination region is physically located between markers spctg177-12a and 58C24-25b, which identified two recombinants at border A and one at border B (Table 17.1). These two markers are located on BAC SH52F06 and SH58C24, respectively (Fig. 17.3a, b).

Conclusion

The sex determination region of sex chromosomes is known for being heterochromatic. The HSY of the recently evolved sex chromosomes in papaya, although to a lesser degree, is no exception. For this reason, physical mapping of the HSY and its X counterpart provided the foundation for sequencing these regions towards the ultimate goal of identifying the sex determination genes, which have implications for papaya improvement and understanding the origin of dioecy in *Caricaceae*. The current physical map through reiterative chromosome walking of HSY consists of 72 BAC clones on the minimum tiling path, while there are 44 BAC clones on the minimum tiling path of X. These maps were used for the complete sequencing of the papaya sex chromosomes, making this plant species the second example after the human X and Y chromosomes of a sequenced sex chromosome pair (both X and Y^h) (see Chap. 16 and this chapter).

Physical mapping by chromosome walking is labor intensive and costly. However, there is no alternative approach for accurately sequencing and assembling the sex determination regions that are so highly repetitive. Future breakthroughs in single molecule sequencing yielding ultra long reads may solve this problem and expedite further sex chromosome research.

References

Adam H, Collin M, Richaud F, Beule T, Cros D, Omore A, Nodichao L, Nouy B, Tregear JW (2011) Environmental regulation of sex determination in oil palm: current knowledge and insights from other species. Ann Bot 108:1529–1537

Bull J (1983) Evolution of sex determining mechanisms. Benjamin-Cummings, Menlo Park

Chailakhyan MKH (1979) Genetic and hormonal regulation of growth, flowering, and sex expression on plants. Am J Bot 66:717–736

Charlesworth D (2002) Plant sex determination and sex chromosomes. Heredity 88:94–101

Chen C, Yu Q, Hou S, Li Y, Eustice M, Skelton RL, Veatch O, Herdes RE, Diebold L, Saw J, Feng Y, Qian W, Bynum L, Wang L, Moore PH, Paull RE, Alam M, Ming R (2007) Construction of a sequence-tagged high-density genetic map of papaya for comparative structural and evolutionary genomics in brassicales. Genetics 177:2481–2491

Chiu CT (2000) Study on sex inheritance and horticultural characteristics of hermaphrodite papaya. Master thesis, National Pingtung University of Science and technology, Pingtung, p 61

Chiu CT, Yen CR, Chang LS, Hsiao CH, Ko TS (2003) All hermaphrodite progeny are derived by self-pollinating the sunrise papaya mutant. Plant Breeding 122:431–434

Deputy JC, Ming R, Ma H, Liu Z, Fitch MM, Wang M, Manshardt R, Stiles JI (2002) Molecular markers for sex determination in papaya (*Carica papaya* L.). Theor Appl Genet 106:107–111

Frankel R, Galun E (1977) Pollination mechanisms, reproduction and plant breeding. In: Frankel R, Gall GAE, Grossman M, Linskens HF, de Zeeuw D (eds) Mono-graphs on theoretical and applied genetics. Springer, New York. pp 141–157

Gschwend AR, Yu Q, Moore P, Saski C, Chen C, Wang J, Na JK, Ming R (2011) Construction of papaya male and female BAC libraries and application in physical mapping of the sex chromosomes. J Biomed Biotechnol 2011:929472

Gschwend AR et al (2012) Rapid divergence and expansion of the X chromosome in papaya. Proc Natl Acad Sci USA 109:13716–13721

Hofmeyr JDJ (1938) Genetical studies of *Carica papaya* L. S Afr Dept Agric Sci Bull 187:64
Hofmeyr JDJ (1939) Sex-linked inheritance in *Carica papaya* L. S Afr J Sci 36:283–285
Hofmeyr JDJ (1967) Some genetic and breeding aspects of *Carica papaya*. Agron Trop 17:345–351
Horovitz S, Jiménez H (1967) Cruzamientosinterespecificos e intergenericos en caricaceas y susimplicacionesfitotechicas. Agron Trop 17:323–343
Irish EE, Nelson T (1989) Sex determination in monoecious and dioecious plants. Plant Cell 1:737–744
Kumar LSS, Abraham A, Srinivasan VK (1945) The cytology of *Carica papaya* Linn. Indian J Agric Sci 15:242–253
Liu Z, Moore PH, Ma H, Ackerman CM, Ragiba M, Yu Q, Pearl HM, Kim MS, Charlton JW, Stiles JI, Zee FT, Paterson AH, Ming R (2004) A primitive Y chromosome in papaya marks incipient sex chromosome evolution. Nature 427:348–352
Ma H, Moore PH, Liu Z, Kim MS, Yu Q, Fitch MM, Sekioka T, Paterson AH, Ming R (2004) High-density linkage mapping revealed suppression of recombination at the sex determination locus in papaya. Genetics 166:419–436
Ming R, Hou S, Feng Y, Yu Q, Dionne-Laporte A, Saw JH, Senin P, Wang W, Ly BV, Lewis KL et al (2008) The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). Nature 452:991–996
Ming R, Bendahmane A, Renner SS (2011) Sex chromosomes in land plants. Annu Rev Plant Biol 62:485–514
Na JK, Wang J, Murray JE, Gschwend AR, Zhang W, Yu Q, Pérez RN, Feltus FA, Chen C, Kubat Z, Moore PH, Jiang J, Paterson AH, Ming R (2012) Construction of physical maps for the sex-specific regions of papaya sex chromosomes. BMC Genomics 13:176
Nicolas M, Marais G, Hykelova V, Janousek B, Laporte V, Vyskot B, Mouchiroud D, Negrutiu I, Charlesworth D, Moneger F (2005) A gradual process of recombination restriction in the evolutionary history of the sex chromosomes in dioecious plants. PLoS Biol 3:e4
Perl-Treves R (1999) Male to female conversion along the cucumber shoot: approaches to studying sex genes and floral development in *Cucumis sativus*. In: Ainsworth CC (ed) Sex determination in plants. Bios Scientific Publishers, Oxford, pp 189–216
Renner SS, Ricklefs RE (1995) Dioecy and its correlates in the flowering plants. Am J Bot 82:596–606
Spigler RB, Lewers KS, Main DS, Ashman TL (2008) Genetic mapping of sex determination in a wild strawberry, *Fragaria virginiana*, reveals earliest form of sex chromosome. Heredity 101(6):507–517
Spigler RB, Lewers KS, Johnson AL, Ashman TL (2010) Comparative mapping reveals autosomal origin of sex chromosome in octoploid *Fragaria virginiana*. J Hered 101(Suppl 1):S107–S117
Spray CR, Kobayashi M, Suzuki Y, Phinney BO, Gaskin P, MacMillan J (1996) The dwarf-1 (dt) Mutant of Zea mays blocks three steps in the gibberellin-biosynthetic pathway. Proc Natl Acad Sci USA 93:10515–10518
Storey WB (1953) Genetics of the papaya. J Hered 44:70–78
Tanurdzic M, Banks JA (2004) Sex-determining mechanisms in land plants. Plant Cell 16(Suppl): S61–S71
Wang J, Na J-K, Yu Q, Gschwend AR, Han J, Zeng F, Aryal R, VanBuren R, Murray JE, Zhang W, Pérez RN, Feltus FA, Lemke C, Tong EJ, Chen C, Wai CM, Singh R, Wang M-L, Min X, Alam M, Charlesworth D, Moore PH, Jiang J, Paterson AH, Ming R (2012) Sequencing papaya X and Yh chromosomes revealed molecular basis of incipient sex chromosome evolution. PNAS 109:13710–13715
Yampolsky C, Yampolsky H (1922) Distribution of the sex forms in the phanerogamic flora. Bibl Genet 3:1–62
Yu Q, Hou S, Hobza R, Feltus FA, Wang X, Jin W, Skelton RL, Blas A, Lemke C, Saw JH, Moore PH, Alam M, Jiang J, Paterson AH, Vyskot B, Ming R (2007) Chromosomal location and gene paucity of the male specific region on papaya Y chromosome. Mol Genet Genomics 278: 177–185

Yu Q, Hou S, Feltus FA, Jones MR, Murray JE, Veatch O, Lemke C, Saw JH, Moore RC, Thimmapuram J, Liu L, Moore PH, Alam M, Jiang J, Paterson AH, Ming R (2008) Low X/Y divergence in four pairs of papaya sex-linked genes. Plant J 53:124–132

Yu Q, Tong E, Skelton RL, Bowers JE, Jones MR, Murray JE, Hou S, Guan P, Acob RA, Luo MC, Moore PH, Alam M, Paterson AH, Ming R (2009) A physical map of the papaya genome with integrated genetic map and genome sequence. BMC Genomics 10:371

Zhang X, Feng B, Zhang Q, Zhang D, Altman N, Ma H (2005) Genome-wide expression profiling and identification of gene activities during early flower development in *Arabidopsis*. Plant Mol Biol 58:401–419

Zhang W, Wang X, Yu Q, Ming R, Jiang J (2008) DNA methylation and heterochromatinization in the male-specific region of the primitive Y chromosome of papaya. Genome Res 18:1938–1943

Chapter 18
Cloning Major Genes Controlling Fruit Flesh Color in Papaya

Rishi Aryal and Ray Ming

Introduction

Papaya produces a climacteric fleshy fruit that when mature ripens in response to ethylene. The fruit matures at about 5 months after fertilization. The mature fruit can ripen if left on the tree; however, picking at the color-break stage and allowing postharvest ripening is a common practice to reduce fruit diseases or predation. This practice has the added benefit of synchronizing harvest supply with the market demand. The unripe fruit has green skin (epicarp) and white flesh (mesocarp and endocarp). Upon ripening, the green skin slowly turns to golden yellow, while the flesh color changes from white to yellow or red, depending on the variety (Fig. 18.1). Along with color changes, additional physiological changes occur as symptoms of fruit ripening, including cell wall degradation producing soft texture fruits, increased sugar and acid content, and the production of volatiles having characteristic fragrances. Papaya fruit softening as it ripens starts near the endocarp and proceeds outward. This might be due to higher accumulation of cell wall degrading enzyme (beta-galactosidase) near the endocarp (Lazan et al. 1995). The softening is more rapid in red-fleshed papaya (referred to as red papaya hereafter) than in yellow-fleshed papaya (referred to as yellow papaya hereafter); hence, the yellow fruit is preferred for having longer postharvest duration whereas red fruit is preferred by some consumers.

Development of colored fruits is an evolutionary achievement of plants. Fruit color serves to attract animals which consume the fruit and subsequently disperse the seeds. Major pigments responsible for plant fruit and flower colors have been identified. The most common non-chlorophyll pigments in plant fruits and flowers are either vacuole-localized anthocyanins or chromoplast-localized carotenoids and

R. Aryal • R. Ming (✉)
Department of Plant Biology, University of Illinois at Urbana-Champaign, 1201 W. Gregory Drive, 148 ERML, MC-051, Urbana, IL 61801, USA
e-mail: rming@life.uiuc.edu

R. Ming and P.H. Moore (eds.), *Genetics and Genomics of Papaya*, Plant Genetics and Genomics: Crops and Models 10, DOI 10.1007/978-1-4614-8087-7_18,

Fig. 18.1 Ripe fruits of papaya, *yellow-fleshed* (Kapoho) and *red-fleshed* (SunUp). The characteristic *yellow* and *red color* in papaya flesh is determined by the accumulation of β-carotene and lycopene, respectively. The skin color of both varieties is *golden yellow*, irrespective of the underlying flesh color which indicates a different mechanism of carotene regulation in photosynthetic skin and the non-photosynthetic tissues of the fruit

xanthophylls. Anthocyanins are a principal constituent in most berry fruits (Hakkinen et al. 1999; Cooke et al. 2005; Juranic and Zizak 2005; Zafra-Stone et al. 2007). Lycopene is the major constituent of the ripe tomato fruit, giving it a brilliant red color (Rao et al. 1998; Arias et al. 2000). In peppers, a range of colors is produced by the accumulation of different pigments, β-carotene, capsanthin, and capsorubin (Camara and Monbger 1978; Minguez-Mosquera and Hornero-Mendez 1993; Rodriguez-Uribe et al. 2011). The yellow color of navel orange is due to carotene content, while the red color in grapefruit is due to accumulation of lycopene (Khan and Mackinney 1953; Xu et al. 2006; Alquezar et al. 2009). In red watermelon, the color of endocarp is determined by lycopene (Perkins-Veazie et al. 2001, 2006).

Carotenoid biosynthesis in chloroplasts is regulated by light; however, its biosynthesis in chromoplasts is regulated developmentally (Gillaspy et al. 1993; Bramley 2002). In papaya fruit, the skin always turns golden yellow on ripening, indicating that the coloration is due to the presence of beta-carotene. However, the flesh color varies between red and yellow, depending on the variety genotype. In fact, the chromoplasts in the fruit skin are derived from green chloroplasts, while the flesh chromoplasts are the derivative of colorless leucoplasts. The papaya gene that synthesizes β-carotene in chloroplasts (*CpLCY-B*) does not show increased expression during fruit ripening (Skelton et al. 2006). The loss-of-function mutation of the chromoplast-specific version of the *CpCYC-b* gene disrupts the β-carotene biosynthesis in fruit flesh (Blas et al. 2010; Devitt et al. 2010) but does not affect the skin color. This indicates different mechanisms for color development in skin and

flesh. The yellow color in ripe papaya fruit skin appears to be a regular senescence process triggered by ethylene production by mature fruit, rather than a regulated increase in carotene production. On the other hand, the flesh color is from the de novo synthesis of color pigments by the enzymes expressed specifically in the fruit-ripening stage.

Dual regulation of carotene biosynthesis in photosynthetic chloroplast and non-photosynthetic chromoplast tissues has also been reported in citrus fruits. The chloroplast lycopene β-cyclase, *Csβ-LCY1* of sweet orange (*Citrus sinensis*), and its grapefruit homolog show constant low-level expression in both oranges and grapefruit, which is similar to the expression pattern of *CpLCY-B* in papaya. The fact that *Csβ-LCY2* expression increases during fruit ripening shows its chromoplast-specific function during ripening. Furthermore, the nonfunctional allele of *Csβ-LCY2* homolog in grapefruit causes the accumulation of lycopene and bright-red pulp of some varieties of grapefruit, but the peel does not show parallel accumulation of red color, indicating differential regulation of gene expression in peel and pulp (Alquezar et al. 2009).

Flesh color of papaya is an important trait determining its commercial and nutritional value. The color difference is primarily due to differential accumulation of the two pigments, lycopene and β-carotene. Red papaya has higher lycopene content, while β-carotene is a major pigment constituent in yellow papaya (Yamamoto 1964; Chandrika et al. 2003). The specific coloration of fruits is not only appealing to the eye but also a source of different precursors in various metabolic pathways. Beta-carotene is a direct precursor of vitamin A and is thus also referred to as provitamin A. This compound is necessary as an animal food supplement because animals cannot synthesize carotene pigments in their body. For this reason the yellow papaya is recommended as dietary supplement for vitamin A in tropical countries (Chandrika et al. 2003) addition to having a higher provitamin A content, yellow papaya also has a longer shelf life over red papaya and is thus preferred for long-term storage (Blas et al. 2010). On the other hand, red fruit is a rich source of the antioxidant pigment lycopene. In recent years, people are paying considerable attention to dark pigments in their daily diet as these pigments act as antioxidants to sequester reactive oxygen species (ROS) from several metabolic processes. Lycopene is considered a more effective antioxidant than β-carotene (Levy et al. 1995) and has a higher quenching rate constant with singlet oxygen ($kq = 31 \times 10^9\ M^{-1}\ s^{-1}$) than that of β-carotene ($kq = 14 \times 10^9\ M^{-1}\ s^{-1}$) (Mascio et al. 1989). Antioxidants provide protection against cancers and cardiovascular diseases. In developed countries, where vitamin A deficiency is not a common dietary disorder, red papaya may be preferred as a source of antioxidants.

Given the indispensable roles of carotenes in plants and their important health benefits, there is a growing interest in studying the genes involved in carotenoid metabolism. Attempts to clone and characterize carotenoid metabolic genes have been made in many fruits, including tomato (Pecker et al. 1996; Ronen et al. 2000), kiwifruit (Ampomah-Dwamena et al. 2009), fig (Araya-Garay et al. 2011), and papaya (Skelton et al. 2006; Blas et al. 2010; Devitt et al. 2010). Several attempts have been made to manipulate pigment amounts in different crops to meet nutritional and commercial requirements. In this chapter, we will describe the research conducted to explore the genes and genetic pathway related to fruit flesh color determination in papaya.

Carotene Biosynthetic Pathway

Biosynthesis of carotenoid pigments occurs in the plastids. In chloroplasts, carotenoids function as an accessory pigment in photosynthetic light harvesting, so they are localized primarily in the photosynthetic membranes in association with light-harvesting and reaction center complexes. In the chromoplasts of ripening fruits and flower petals and in the chloroplasts of senescing leaves, carotenoids are found in the stroma which is associated with membranes, or in oil bodies, or other structures (Cunningham and Gantt 1998). Carotenoids are mainly 40-carbon isoprenoids with a polyene chain that contains up to 15 conjugated double bonds. The building blocks of carotenoids are the 5-carbon compound isopentenyl diphosphate (IPP) and its isomer, dimethylallyl pyrophosphate (DMAPP). A 20-carbon molecule, geranylgeranyl diphosphate (GGDP), is produced by GGDP synthase (GGPS) from IPP and DMAPP. Phytoene synthase (PSY) combines two GGDP molecules to produce the first 40C precursor in the carotenoid biosynthetic pathway. All carotenoids are produced from this 40C phytoene backbone by series of enzymatic reactions (Fig. 18.2). Phytoene is first converted to ζ-carotene

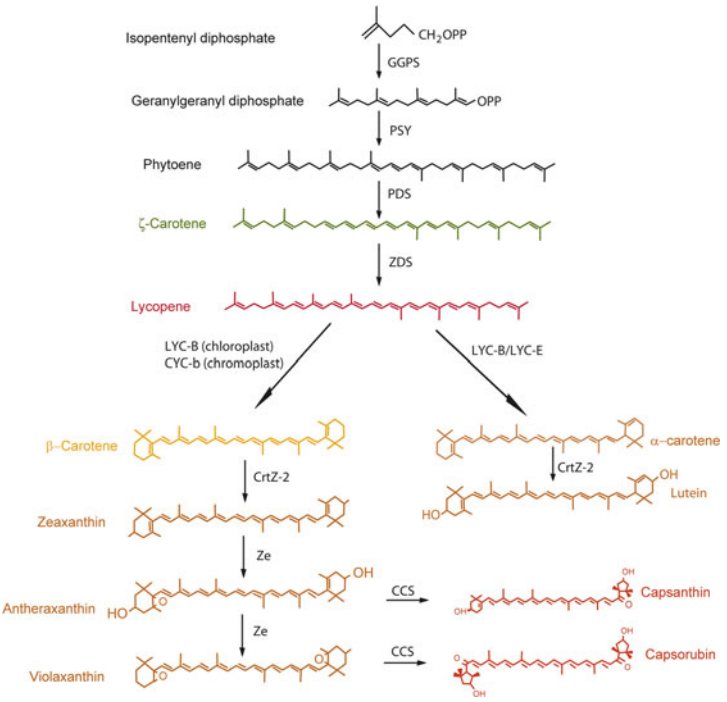

Fig. 18.2 The carotenoid biosynthesis pathway in plants. GGPS geranylgeranyl diphosphate, *PSY* phytoene synthase, PDS phytoene desaturase, ZDS zeta-carotene desaturase, LCY-B/CYC-b/ LCY-E lycopene beta-cyclase, CrtZ-2 beta-carotene hydrolase, Ze zeaxanthin epoxidase, CCS capsanthin-capsorubin synthase. Genes encoding the proteins PDS, ZDS, LYC-B, and CYC-b has been cloned from papaya

(zeta-carotene) by the enzyme phytoene desaturase (PDS), which is then converted to lycopene by ζ-carotene desaturase (ZDS). Lycopene is the precursor for two biochemical pathways, leading to lutein via alpha-carotene and/or abscisic acid via beta-carotene and a series of xanthophylls. In pepper, the enzyme capsanthin-capsorubin synthase (CCS) converts the xanthophylls antheraxanthin and violaxanthin to capsanthin and capsorubin, respectively, giving the characteristic red color of the pepper.

Mapping Genes Involved in Carotene Biosynthesis

A high-density genetic map with 1,501 markers mapped the fruit color locus near the end of linkage group 7 (LG7), at the distance of 3.4 and 3.7 cM from the nearest flanking amplified fragment length polymorphism (AFLP) markers (Ma et al. 2004). On a different genetic map, with 706 simple sequence repeat (SSR) markers, the fruit flesh color locus was mapped at the end of LG5 at the distance of 12.7 cM from the nearest marker P3K2152 (Chen et al. 2007). Enrichment of high-density SSR genetic map by the addition of 277 AFLP markers maintained the flesh color locus at LG5 (Blas et al. 2009).

Several genomic resources of papaya are available publicly, which facilitate the cloning and mapping of various candidate genes. A papaya draft genome covering more than 90 % of the euchromatic region has been published (Ming et al. 2008). A bacterial artificial chromosome (BAC)-based physical map from hermaphrodite SunUp was also constructed and integrated into the high-density SSR genetic map and genomic sequences (Yu et al. 2009). A chloroplast-specific lycopene β-cyclase (GenBank accession number DQ415894) was mapped to LG4. *CpPDS* and *CpZDS* genes were mapped to LG6 using the genomic sequence and integrated maps (Yan et al. 2011).

Cloning Genes in Carotene Biosynthesis

Several researches have been done to explore the mechanism and regulation of flesh color in papaya. Since lycopene is the major pigment determining the fruit flesh color, the major focus has been on the genes involved in the synthesis of lycopene (Yan et al. 2011) or the conversion of lycopene to β-carotene (Skelton et al. 2006; Blas et al. 2010; Devitt et al. 2010). As the cloning of these genes was primarily based on homology, the naming of these genes sometimes differs in different publications. For consistency and to avoid confusion, we will use the names used in Fig. 18.2 throughout this chapter. Each of these genes are prefixed “Cp” to indicate they are from *Carica papaya*.

The first attempt to clone papaya flesh color gene was carried out by Skelton et al. (2006). They used degenerate primers designed across the conserved region of

lycopene β-cyclase genes from several other angiosperm species. The degenerate primers were used to amplify the orthologous gene from papaya. The fragment amplified by those primers using SunUp hermaphrodite genomic DNA as a template was used to hybridize the papaya BAC library. Papaya lycopene β-cyclase (*CpLCY-B*) was mapped in BAC clone 68A06 by southern hybridization. Sequencing the mRNA from the *CpLCY-B* specific primer yielded a 1731 bp transcript, including 1281 bp ORF (open reading frame) and untranslated regions from both ends. The cloned cDNA sequence showed 85 and 81 % sequence identity with citrus and tomato lycopene β-cyclase genes, respectively. The complete genomic segment of the DNA containing *CpLCY-B* coding region and its promoter region was cloned. The total genomic segment of 4,118 bp, containing *CpLCY-B* and its putative regulatory elements, is available in GenBank (DQ415894).

Expression analysis ruled out the possibility of *CpLCY-B* being the gene candidate for regulation of fruit flesh color. SunUp (red) and Kapoho (yellow) fruits were analyzed at different stage of maturity (color-break, 30 % ripe, and 90 % ripe) by quantitative real-time PCR (qPCR). Since the *CpLCY-B* converts lycopene to β-carotene, higher expression of this gene was expected in yellow papaya. However, this gene did not show any difference in expression between red and yellow papaya. The same level of expression of *CpLCY-B* in both genotypes excluded the possibility of this gene being the determinate of flesh color. Also, a surge of expression of the flesh color gene was expected upon the start of fruit ripening, but the level of expression was relatively constant throughout the fruit developmental stages. This indicated the possibility of a different gene to carry out the rapid rise in carotene biosynthesis during fruit ripening. Further comparison of expression level in fruit, flower, and leaves showed that *CpLCY-B* was expressed sevenfold higher in leaves and threefold higher in flowers than in fruit. Higher expression of *CpLCY-B* gene in leaves than in flowers and fruit shows that it is functional in green tissue. Since carotenes are also the major component of light-harvesting system in photosynthesis, *CpLCY-B* appears to function in chloroplasts to maintain the carotene level necessary for photosynthesis. Although flowers of both cultivars of papaya have white petals, the flower is light green at the bud stage. The slightly higher expression level of *CpLCY-B* in flower may be due to the remaining chloroplasts in petals and sepals.

Since *CpLCY-B* was chloroplast specific, it led to the search for another gene being responsible for biogenesis of carotene in fruit flesh. Two different lycopene β-cyclase genes functioning distinctly in leaf and fruit have been reported from other fruit species. In tomato, fruit-specific lycopene β-cyclase "B" is responsible for the fruit color determination instead of the chloroplast-specific lycopene β-cyclase (*CrtL-b*). In citrus plants, *CsLCY2* functions in fruit-specific chromoplasts while *CsLCY1* is active in chloroplasts. Thus, the search for chromoplast-specific genes in carotenoid biosynthesis was the next approach of various labs to uncover the mechanism of fruit flesh color determination in papaya.

A detailed study on the genes regulating flesh color was carried out using the combination of map-based cloning and candidate gene approaches (Blas et al. 2010). Map-based cloning of the flesh color gene from red SunUp and yellow AU9 varieties began by screening a papaya hermaphrodite BAC library using tomato

CYC-b cDNA and a sequence-characterized amplified region (SCAR) marker that had mapped 3.4 cM away from the fruit flesh color gene (Ming et al. 2001; Ma et al. 2004). Two BACs were positive to the tomato CYC-b probe and five BACs hybridized to the SCAR marker. The seven selected BACs were mapped to a single fingerprinted contig, FPC-1648, of the papaya physical map (Yu et al. 2009). Co-localization of the genetic SCAR marker and tomato *CYC-b* provided strong support for the locus to contain the lycopene β-cyclase gene controlling the fruit flesh color. Genscan analysis of the fingerprinted contig FPC-1648 predicted a gene homologous to tomato *CYC-b* and to citrus *CCS*. The complete gene sequence was obtained from the yellow-colored cultivar Kapoho by chromosome walking. The sequence was verified later by sequencing BAC SH18O09 from red SunUp and BAC DM105M02 from yellow AU9. The cloned gene was named *CpCYC-B* for its structural and functional homology with tomato *CYC-b*.

The *CpCYC-b* sequence shows high similarity with the chromoplast-specific lycopene β-cycles genes from citrus (Blas et al. 2010) and tomato (Devitt et al. 2010). The conserved region of *CpCYC-b* shares 60 % amino acid identity with *CpLCY-B*. Both *CpCYC-b* and *CpLCY-b* contain a 5′ signal sequence targeted to plastids, thus confirming two different genes are involved in carotene biosynthesis in plastids. Phylogenetic analysis of these two genes in both studies placed these genes in two distinct clades, each clade representing chloroplast-specific and chromoplast-specific lycopene β-cyclase from different plants. Furthermore, *CpCYC-b* shows closer homology to the *CCS* gene from carrot and capsicum than to the *CpLCY-b* gene as additional confirmation of its role in the chromoplast.

The coding region of *CpCYC-b* showed polymorphisms between red and yellow color varieties. A thymine dinucleotide insertion at 834 bp away from the start codon in the preexisting sequence of five thymine nucleotides was observed in the red cultivar SunUp. Insertion of the thymine dinucleotide created a frameshift mutation, rendering a premature stop codon and causing nonfunctional protein translation (Fig. 18.3). The mutation in the fruit-specific lycopene β-cyclase gene disrupts conversion of lycopene to β-carotene, resulting in the accumulation of high amounts of lycopene and is thus responsible for the red flesh of this variety. The sequence comparison from 15 papaya varieties showed that all red varieties have the same inserted thymine dinucleotide. In addition to the dinucleotide insertion, three synonymous substitutions were found between the *CpCYC-b* genes of red and yellow varieties (Blas et al. 2010).

Five genes involved in carotene biosynthesis, *CpPDS* (phytoene desaturase), *CpZDS* (zeta-carotene desaturase), *CpLCY-e, CpCYC-b,* and *CpCHY-b*, were tested for their expression in leaf and fruit during different stages of fruit ripening in both red and yellow papaya (Blas et al. 2010). All genes, except *CpCYC-b,* showed relatively constant expression in leaf and immature fruit and a slight increase in expression in the fruit after complete maturation. However, the level of expression of *CpCYC-b* increased markedly higher in mature yellow fruit than in the red fruit. The relative expression level of *CpCYC-b* in yellow fruit compared to its respective leaf tissue was 11.5-fold and was 2.6-fold higher than in the red papaya at the green mature and 30 % ripe fruit stages. Similar elevated expression of

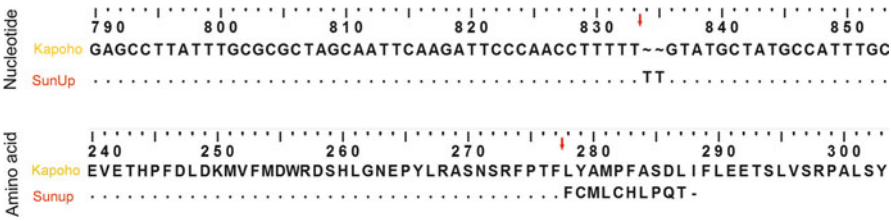

Fig. 18.3 Nucleotide and amino acid sequence alignment of chromoplast-specific lycopene β-cyclase (*CpCYC-b)* from *yellow-fleshed* Kapoho and *red-fleshed* SunUp papaya. The *red arrow* indicates the position of an indel mutation that causes the protein to terminate after 287 amino acids

chromoplast-specific lycopene β-cyclase in fruit was also recorded by Devitt et al. 2010 (see discussion that follows). The *CpCYC-b* gene in the red variety (SunUp) has a frameshift mutation and does not code for a functional protein. The mutation in the coding region does not necessarily account for the observed lower expression level of this gene in red papaya, since the frameshift should not affect transcription itself. The observed reduction of expression of the mutated gene might be due to a cellular reflex to reduce nonfunctional transcripts. Alternately, it may be the result of rapid degradation of the nonfunctional transcript, since it is not occupied by ribosomes.

The functionality of the cloned chromoplast-specific lycopene β-cyclase was also confirmed by color complementation in bacteria. Color complementation is a protocol developed to clone genes in the carotenoid biosynthetic pathway by producing a visual analysis of gene function (Lotan and Hirschberg 1995). In this method, the gene under study is expressed in a bacterial strain capable of producing all of the precursors for that gene function but lacking that particular gene. Function of the gene is analyzed based on the color produced at the end of the experiment. Full-length coding regions from both yellow and red papaya, as well as some truncated ORFs from the red variety, were cloned in an expression vector and transformed into *Escherichia coli*. Successful transformation of the proper construct was confirmed by restriction digestion with SacI and also by PCR amplification with *CpCYC-b* specific primers. Expression of full-length coding region from Kapoho (the yellow variety) catalyzed the conversion of lycopene to beta-carotene as indicated by a red to yellow color shift, while the full-length and shorter ORFs of *CpCYC-b* from SunUp (the red variety) was not able to catalyze such conversion. The conversion was further confirmed by measuring the lycopene to β-carotene content by absorption spectrum. The reaction mixture containing clones from red papaya showed lycopene absorption, while the one containing clones from yellow papaya showed accumulation of beta-carotene. These results show that *CpCYC-b* from yellow flesh papaya is functional and catalyzes the conversion of lycopene to β-carotene.

The flesh color-determining gene was also analyzed in context of neighboring genes and genomic context in two flesh color varieties. The SunUp papaya BAC (SH18O09) and AU9 BAC (DM105M02) were sequenced and subjected to gene

prediction. The two BACs share 99 % sequence similarity over a 64,996 bp overlapping region. Two insertion/deletions (indels) of 1,805 and 225 bp were observed in this region, which could be either an insertion in AU9 or a deletion from SunUp. Deletion from SunUp is more likely due to fragmentation and rearrangement caused by particle bombardment during the transformation process. Both indels are in nongenic regions and cannot account for flesh color difference (Blas et al. 2010). The aligned BAC contained ten genes, which was confirmed by the papaya EST sequences and RT-PCR. Comparison of the gene order in this region shows microsynteny with the homologous genes in *Arabidopsis*, tomato, and grape.

The commercial value of fruit flesh color is established by consumer's choice. Consumers in different locals prefer different colored fruit. It is important that papaya producers select and grow specific lines to produce the most marketable colored fruit. Since there is no morphological difference between plants producing two different color fruit, molecular markers to screen for fruit color could be used for papaya producers to maintain the crop production based on marked demand. Screening of two alleles based on two-nucleotide difference is not feasible so it is imperative to have reliable markers linked with the fruit color trait. Reliability of the screening result depends on the distance of the marker from the gene and recombination frequency in that area. Blas et al. (2010) studied the recombination frequency of the chromosomal region containing flesh color gene. Three SCAR markers, CPFC1, CPFC2, and CPFC3 located 580 bp, 9.0 kb, and 36.9 kb away from *CpCYC-b* genes, were used to determine the recombination frequency in this region. The genetic distance between *CpCYC-b* and CPFC1 was determined to be 0.9 cM, the distance between CPFC1 and CPFC2 is 2.4 cM, and between *CpCYC-b* and CPFC2 is 1.9 cM. From these data, the average rate of recombination in this region was determined as 3.7 kb per cM. The papaya genome-wide recombination rate is 400 kb/cM. Interestingly, fruit color coding region shows more than 100 times higher recombination rate than the genome average. The high recombination rate on this region is attributed to two features commonly associated with recombination hot spots (Mezard 2006): a high gene density (Blas et al. 2010) and a terminal position on the chromosome (Ma et al. 2004; Chen et al. 2007; Blas et al. 2009).

A similar approach to clone the lycopene β-cyclase genes was carried out in Tainung (red) and Hybrid 1B (yellow) papaya (Devitt et al. 2010). PCR amplification was performed from papaya genomic DNA using degenerate primers designed from edible tomato, wild tomato, and *Arabidopsis lycopene* β-*cyclase* as well as from the *CCS* gene from bell pepper. These two primers gave two distinct amplicones, *lcy-β1* and *lcy-β2*. The amplified *lcy-β1* was the same as the *CpLCY-B* obtained by Skelton et al. (2006) and showed homology with chloroplast-specific lycopene β-cyclase from tomato. The second gene (*lyc-β2*) showed 52 % amino acid sequence similarity with *CpLCY-B*. Phylogenetic analysis using amino acid sequences from *lycopene* β-*cyclase* puts the *lcy-β1* with the chloroplast-specific (*CrtL-B*) and *lcy-β2* with the chromoplast-specific *lycopene β-cyclase* from tomato.

As in SunUp, a dinucleotide insertion causing a frameshift mutation was also seen in Tainung (a red cultivar). In addition to the dinucleotide insertion, one single-nucleotide polymorphism (SNP) was recorded at position 607, with cytosine in yellow

Hybrid 1B, while adenine in red Tainung. Genotyping 11 other varieties, 7 red and 4 yellow, showed that the SNP does not account for the difference in flesh color.

The mutated *lcy-β2* gene in red Tainung was expressed much lower than the functional copy in yellow Hybrid 1B. The expression of *lcy-β2* was about 15-fold higher in color-break and 60-fold higher in ripe Hybrid 1B fruit compared to leaves. The *lcy-β1,* being coded by different gene showed similar level of expression in both red and yellow fruits. The expression of *lcy-β2* was 18-fold higher in ripe yellow fruit compared to *lcy-β1* while it was only fivefold higher in red Tainung. Similarly, the expression of *lcy-β2* gene increases as the ripening process progresses in yellow fruit; a greater than a fourfold increase was observed in expression of *lyc-β2* in Hybrid 1B between the color-break stage to ripe stage, but expression of *lyc-β2* remained constant during ripening in red Tainung.

Another possibility for higher accumulation of lycopene in red papaya is higher activity of the lycopene synthesis gene in this variety. *PDS* and *ZDS* are the major genes involved in lycopene biosynthesis. Homology search showed a single copy for each of these two genes in the papaya genome. To further explore the genes responsible for papaya fruit flesh color, Yan et al. (2011) cloned the genes upstream of lycopene synthesis. Expression of these genes in different tissues of red and yellow papaya was reported.

The coding regions of the *CpPDS* and *CpZDS* genes were amplified from papaya cDNA using primers designed in the conserved region of the respective homologous genes from higher plants and papaya EST sequences. The initially amplified sequence was used to design primers for 3′ and 5′ RACE (rapid amplification of cDNA ends) and the complete transcripts of these genes were obtained. The *CpPDS* full-length transcript was 2,164 bp with a 1,749 bp ORF, 150 bp 5′ UTR, and 262 bp 3′ UTR. A similar approach with *ZDS* yielded a 2,096 bp transcript with a 1,716 bp ORF, 168 bp 5′ UTR, and 209 bp 3′ UTR. The predicted *PDS* protein was 583 aa with 65.13 kDa; the *ZDS* protein was 572 aa with 63.12 kDa. The amino acid sequence of *CpPDS* showed 78 % similarity with *Arabidopsis* 81 % with tomato, 80 % with citrus, and 76 % sequence identity with carrot. The *CpZDS* shares 86 and 84 % similarity with tomato and citrus, 82 % with *Arabidopsis*, and 79 % similarity with carrot. Papaya is phylogenetically closer to *Arabidopsis* than to tomato and citrus, so the better alignment with *Arabidopsis* shows more conservation toward the fruit function. In addition, the lower similarity with carrot might be due to root-specific functions in carrot.

The expression levels of *CpPDS* and *CpZDS* were analyzed in leaf, flower, and several stages of fruit development in yellow (Dwarf Solo) and red (Sunrise Solo) papaya. The expression levels of these genes were similar in leaves, flowers, and unripe fruits in both yellow Dwarf Solo and red Sunrise Solo. The expression of both genes was low at early fruit developmental stages but started increasing at the color-break stage and increased up to tenfold in ripe fruit compared to remaining constant in the leaves as the fruit ripened. Both genes showed increased levels of expression in red papaya over yellow. Although the data suggest increased activity of these genes in the fruit-ripening stages, the difference in expression level between

red and yellow papaya was less than twofold, which might not account for the difference in the lycopene content in the two varieties. In fact, the authors noted that there must be other mechanisms responsible for the color difference. The data thus show that the higher accumulation of lycopene in red papaya is due to inefficient conversion to β-carotene due to ineffective lycopene *β*-cyclase, rather than increased production from its precursors.

Phytoene synthase (PSY) is the rate-limiting enzyme in the carotenoid biosynthesis pathway. Although the expression level of PSY remains to be checked, the elevated expression of PDS and ZDS in red fruit might be responsible for additional attributes of the fruit such as firmness. The red papaya fruit softens much faster than the yellow and the red has a relatively short shelf life. The carotenoid molecules, in addition to having light-harvesting and color development functions in fruit and flowers, are regulators of diverse metabolic pathways (Qin et al. 2007). It has been reported that leucoplast-derived carotenoids in *Arabidopsis* provide precursor for the synthesis of signal molecules (Van Norman and Sieburth 2007). The carotenoid biosynthetic pathway shares its early steps with gibberellin and cytokinin biosynthetic pathways (Kleinig 1989; Qin et al. 2007). Overexpression of the carotenoid pathway gene phytoene synthase significantly decreased gibberellin biosynthesis in tomato (Fray et al. 1995). It is too early to conclude a role of carotenoid biosynthesis in the softer flesh quality of red papaya, but it is worth further investigations to test the hypothesis that it does play a role.

Prospects

Papaya fruit flesh color trait is inherited in Mendelian fashion with yellow flesh being dominant over red flesh. Cloning the major genes in carotenoid biosynthesis in papaya revealed the mechanism for fruit flesh color determination. Differential expression of the genes involved in carotenoid biogenesis in red and yellow papaya is not sufficient to account for the flesh color difference. The major cause of the fruit color difference in papaya is a dinucleotide insertion mutation in the gene encoding enzyme lycopene β-cyclase (*CpCYC-b*). The *CpCYC-b* enzyme converts lycopene to β-carotene in the chromoplasts of fruit. Mutation of this gene disrupts this biochemical reaction, resulting in the accumulation of lycopene in red fruit. The normal functioning *CpCYC-b* gene encodes the competent enzyme for converting lycopene to beta-carotene giving yellow fruit.

Fruit color has been important in establishing the market value of many fruits based on their appearance, shelf life, and dietary content. Red papaya is preferred for better taste and higher antioxidant content, while yellow papaya is a better source of vitamin A with a longer shelf life of the fruit. Understanding the mechanism of fruit color determination and the production of markers to screen fruit color in early seedlings can provide a tool for farmers to manage their crop based on market demand.

References

Alquezar B, Zacarıas L, Rodrigo MJ (2009) Molecular and functional characterization of a novel chromoplast-specific lycopene b-cyclase from Citrus and its relation to lycopene accumulation. J Exp Bot 60(6):1783–1797

Ampomah-Dwamena C, McGhie T, Wibisono R, Montefiori M, Hellens RP, Allan AC (2009) The kiwifruit lycopene beta-cyclase plays a significant role in carotenoid accumulation in fruit. J Exp Bot 60(13):3765–3779

Araya-Garay JM, Feijoo-Siota L, Veiga-Crespo P, Villa TG (2011) cDNA cloning of a novel gene codifying for the enzyme lycopene β-cyclase from *Ficus carica* and its expression in *Escherichia coli*. Appl Microbiol Biotechnol 92:769–777

Arias R, Lee TC, Logendra L, Janes H (2000) Correlation of lycopene measured by HPLC with the L, a, b color readings of a hydroponic tomato and the relationship of maturity with color and lycopene content. J Agric Food Chem 48:1697–1702

Blas A, Qingyi Y, Cuixia C, Veatch O, Moore P, Paull R, Ming R (2009) Enrichment of a papaya high-density genetic map with AFLP markers. Genome 52(8):716–725

Blas AL, Ming R, Liu Z, Veatch OJ, Paull RE, Moore PH, Yu Q (2010) Cloning of the papaya chromoplast-specific lycopene β-cyclase, CpCYC-b, controlling fruit flesh color reveals conserved microsynteny and a recombination hot spot. Plant Physiol 152:2013–2022

Bramley PM (2002) Regulation of carotenoid formation during tomato fruit ripening and development. J Exp Bot 53:2107–2113

Camara B, Monbger R (1978) Free and esterified carotenoid in green and red fruits of *Capsicum annuum*. Phytochemistry 17:91–93

Chandrika UG, Jansz ER, Wickramasinghe SMDN, Warnasuriya ND (2003) Carotenoids in yellow- and red papaya (*Carica papaya* L.). J Sci Food Agric 83:1279–1282

Chen C, Yu Q, Hou S, Li Y, Eustice M, Skelton RL, Veatch O, Herdes RE, Diebold L, Saw J, Feng Y, Qian W, Bynum L, Wang L, Moore PH, Paull RE, Alam M, Ming R (2007) Construction of a sequence-tagged high-density genetic map of papaya for comparative structural and evolutionary genomics in brassicales. Genetics 177:2481–2491

Cooke D, Steward WP, Gescher AJ, Marczylo T (2005) Anthocyans from fruits and vegetables – does bright colour signal cancer chemopreventive activity? Eur J Cancer 41:1931–1940

Cunningham FX, Gantt E (1998) Genes and enzymes of carotenoid biosynthesis in plants. Annu Rev Plant Physiol Plant Mol Biol 49:557–583

Devitt LC, Fanning K, Dietzgen RG, Holton TA (2010) Isolation and functional characterization of a lycopene beta-cyclase gene that controls fruit colour of papaya (*Carica papaya* L.). J Exp Bot 61:33–39

Fray RG, Wallace A, Fraser PD, Valero D, Hedden P et al (1995) Constitutive expression of a fruit phytoene synthase gene in transgenic tomatoes causes dwarfism by redirecting metabolites from the gibberellin pathway. Plant J 8:693–701

Gillaspy G, Ben-David H, Gruissem W (1993) Fruits: a developmental perspective. Plant Cell 5:1439–1451

Hakkinen SH, Karenlampi SO, Heinonen M, Mykkanen HM, Torronen AR (1999) Content of the flavonols quercetin, myricetin, and kaempferol in 25 edible berries. J Agric Food Chem 47:2274–2279

Juranic Z, Zizak Z (2005) Biological activities of berries: from antioxidant capacity to anti-cancer effects. Biofactors 23:207–211

Khan M, Mackinney G (1953) Carotenoids in grapefruit, *Citrus paradisi*. Plant Physiol 28:550–552

Kleinig H (1989) The role of plastids in isoprenoid biosynthesis. Annu Rev Plant Physiol 40:39–59

Lazan H, Selamat MK, Ali ZM (1995) β-Galactosidase, polygalacturonase and pectinesterase in differential softening and cell wall modification during papaya fruit ripening. Physiol Plant 95(1):106–112

Levy J, Bosin E, Feldman B, Giat Y, Miinster A, Danilenko M, Sharoni Y (1995) Lycopene is a more potent inhibitor of human cancer cell proliferation than either α-carotene or β-carotene. Nutr Cancer 24(3):257–266

Lotan T, Hirschberg J (1995) Cloning and expression in *Escherichia coli* of the gene encoding β-C-4-oxygenase, that converts β-carotene to the ketocarotenoid canthaxanthin in *Haematococcus pluvialis*. FEBS Lett 364(2):125–128

Ma H, Moore PH, Liu Z, Kim MS, Yu Q, Fitch MMM, Sekioka T, Paterson AH, Ming R (2004) High-density linkage mapping revealed suppression of recombination at the sex determination locus in papaya. Genetics 166:419–436

Mascio DP, Kaiser S, Sies H (1989) Lycopene as the most efficient biological carotenoid singlet oxygen quencher. Arch Biochem Biophys 274(2):532–538

Mezard C (2006) Meiotic recombination hotspots in plants. Biochem Soc Trans 34:531–534

Ming R, Moore PH, Zee F, Abbey CA, Ma H, Paterson AH (2001) Construction and characterization of a papaya BAC library as a foundation for molecular dissection of a tree-fruit genome. Theor Appl Genet 102:892–899

Ming R, Hou S, Feng Y, Yu Q, Dionne-Laporte A et al (2008) The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). Nature 452:991–996

Minguez-Mosquera MI, Hornero-Mendez D (1993) Separation and quantification of the carotenoid pigments in red peppers (*Capsicum annuum* L.), paprika, and oleoresin by reversed- phase HPLC. J Agric Food Chem 41:1616–1620

Pecker I, Gabbay R, Cunningham FX Jr, Hirschberg J (1996) Cloning and characterization of the cDNA for lycopene beta-cyclase from tomato reveals decrease in its expression during fruit ripening. Plant Mol Biol 30:807–819

Perkins-Veazie P, Collins JK, Pair SD, Roberts W (2001) Lycopene content differs among red watermelon cultivars. J Sci Food Agric 81:983–987

Perkins-Veazie P, Collins JK, Davis AR, Roberts W (2006) Carotenoid content of 50 watermelon cultivars. J Agric Food Chem 54:2593–2597

Qin G, Gu H, Ma L, Peng Y, Deng XW, Chen Z, Li-Jia Qu L (2007) Disruption of phytoene desaturase gene results in albino and dwarf phenotypes in *Arabidopsis* by impairing chlorophyll, carotenoid, and gibberellin biosynthesis. Cell Res 17:471–482

Rao AV, Waseem Z, Agarwal S (1998) Lycopene contents of tomatoes and tomato products and their contribution to dietary lycopene. Food Res Int 31:737–741

Rodriguez-Uribe L, Guzman I, Rajapakse W, Richins RD, O'Connell MA (2011) Carotenoid accumulation in orange-pigmented Capsicum annum fruit, regulated at multiple levels. J Exp Bot 63(1):517

Ronen G, Carmel-Goren L, Zamir D, Hirschberg J (2000) An alternative pathway to β-carotene formation in plant chromoplasts discovered by map-based cloning of beta and old-gold color mutations in tomato. PNAS 97(20):11102–11107

Skelton RL, Yu Q, Srinivasan R, Manshardt R, Moore PH, Ming R (2006) Tissue differential expression of lycopene beta-cyclase gene in papaya. Cell Res 16:731–739

Van Norman JM, Sieburth LE (2007) Dissecting the biosynthetic pathway for the bypass1 root-derived signal. Plant J 49:619–628

Xu CJ, Fraser PD, Wang WJ, Bramley PM (2006) Differences in the carotenoid content of ordinary citrus and lycopene accumulating mutants. J Agric Food Chem 54:5474–5481

Yamamoto HY (1964) Comparison of the carotenoids in yellow- and red *Carica papaya*. Nature 201:1049–1050

Yan P, Gao XZ, Shen WT, Zhou P (2011) Cloning and expression analysis of phytoene desaturase and zeta-carotene desaturase genes in *Carica papaya*. Mol Biol Rep 38:785–791

Yu Q, Tong E, Skelton RL, Bowers JE, Jones MR, Murray JE, Hou S, Guan P, Acob RA, Luo MC, Moore PH, Alam M, Paterson AH, Ming R (2009) A physical map of the papaya genome with integrated genetic map and genome sequence. BMC Genomics 10:371

Zafra-Stone S, Yasmin T, Bagchi M, Chatterjee A, Vinson JA, Bagchi D (2007) Berry anthocyanins as novel antioxidants in human health and disease prevention. Mol Nutr Food Res 51:675–683

Chapter 19
Molecular Markers in Papayas

Chutchamas Kanchana-udomkan, Rebecca Ford, and Rod Drew

Background

Molecular markers are effective tools and have been used to facilitate genetic improvement in many crop species including *Carica papaya* (Eustice et al. 2008). The main purpose of using molecular markers within a breeding program is to either determine the relatedness among genotypes for germplasm resource management and parental choice, for true-to-type, and for hybrid identification or to identify and select for particular sequences that are associated with traits of interest. DNA markers (Table 19.1) are generally stable, unaffected by environment, and present at all stages of plant growth and in all tissue types. They have been adopted within papaya breeding programs as accurate selection tools for traits of interest (Eustice et al. 2008; Ma et al. 2004; Porter et al. 2009b; de Oliveira et al. 2010a; Deputy et al. 2002; Dillon et al. 2006).

Recent research has led to some important developments in this field. In the search for DNA markers linked to the genes that condition the traits of interest, a genetic and physical map of the papaya genome was developed (Yu et al. 2009). The papaya genome was sequenced and has been used to identify a library of SSR [simple sequence repeat (SSR); microsatellite] loci (Eustice et al. 2008; Santos et al. 2003; Wang et al. 2008). In addition, several gene sequences with associated functions have become available through the papaya genome project database (Ming et al. 2008).

This chapter will focus on the development and application of molecular markers that have been used to assess genetic diversity and to improve papaya breeding objectives.

C. Kanchana-udomkan • R. Drew (✉)
School of Biomolecular and Physical Sciences, Griffith University, 170 Kessels Road, Nathan, QLD 4111, Australia
e-mail: r.drew@griffith.edu.au

R. Ford
Department of Agriculture and Food Systems, Melbourne School of Land and Environment, Melbourne, VIC 3010, Australia

R. Ming and P.H. Moore (eds.), *Genetics and Genomics of Papaya*, Plant Genetics and Genomics: Crops and Models 10, DOI 10.1007/978-1-4614-8087-7_19,

Table 19.1 DNA markers and their primer sequences linked to sex type identification in *Carica papaya*

Marker name	Primer sequence (5′ ---> 3′)	Product size	Type of marker	Sex detection	Note	Reference
OPT12	OPT12: GGGTGTGTAG	Na	RAPD		*Sex1*: 7 cM	Sondur et al. (1996)
OPT1C	OPT1: GGGCCACTCA	Na	RAPD		*Sex1*: 7 cM	Sondur et al. (1996)
T1	T1F: TGCTCTTGATATGCTCTCTG T1R: TACCTTCGCTCACCTCTGCA	1.3 kb	SCAR	All sex	*Sex1*: n/a	Deputy et al. (2002)
T12	T12F: GGGTGTGTAGGCACTCTCCTT T12R: GGGTGTGTAGCATGCATGATA	800 bp	SCAR	Hermaphrodite and male	*Sex1*: 0.3 cM	Deputy et al. (2002)
W11	W11F: CTGATGCGTGTGTGGCTCTA W11R: CTGATGCGTGATCATCTACT	800 bp	SCAR	Hermaphrodite and male	*Sex1*: 0.3 cM	Deputy et al. (2002)
PSDM	IBRC-RP07: TTGGCACGGG	450 bp	RAPD	Hermaphrodite and male	*Sex1*: n/a	Urasaki et al. (2002b)
SCARps	SDP-1: GCACGATTTAGATTAGATGT SDP-2: GGATAGCTTGCCCAGGTCAC	225 bp	SCAR	Hermaphrodite and male	*Sex1-H* and *Sex1-M*: n/a	Urasaki et al. (2002b)
Papain	P5′: GGGCATTCTCAGCTGTTGTA P3′: CTCCCTTGAGCGACAATAAC	221 bp	EST		Papain	Urasaki et al. (2002a)
SCARpm	SDP-2: GGATAGCTTGCCCAGGTCAC SDP-3: GGTAAGAGTTTTTCCCAAGC	347 bp	SCAR	Hermaphrodite and male		Urasaki et al. (2002a)
$BC210_{438}$	BC210: GCACCGAGAG	438 bp	RAPD	Hermaphrodite		Lemos et al. (2002)
$OP\text{-}Y7_{900}$	OP-Y7: AGAGCCGTCA	900 bp	RAPD	Male		Chaves-Bedoya and Nuñez (2007)
SCAR SDSP	CFw: AAACTACCGTGCCATTATCA CRv: AGAGATGGGTTGTGT CACTG	369 bp	SCAR	Hermaphrodite and male		Chaves-Bedoya and Nuñez (2007)
$(GATA)_4$	Genomic DNA digested with HinfI	5 kb	RFLP	Hermaphrodite and male		Parasnis et al. (1999)

$(GATA)_4$	Genomic DNA digested with HaeIII	4 kb	RFLP	Hermaphrodite and male		Parasnis et al. (1999)
OPC09-1.7	OPC09: CTCACCGTCC	1.7 kb	RAPD			Niroshini et al. (2008)
OPE03-0.4	OPE03: CCAGATGCAC	0.4 kb	RAPD			Niroshini et al. (2008)
C09/20	C09/20FP: CTCACCGTCCATTTTAATTA C09/20RP: CTCACCGTCCGCGGCATCAATGTA	1.7 kb 978 bp	SCAR	Hermaphrodite and male		Niroshini et al. (2008)
$(GACA)_4$	(GACA)4: GACAGACAGACAGACA		ISSR			Gangopadhyay et al. (2007)
OPF2-0.8	OPF2: GAGGATCCCT	831 bp	RAPD	Male	Acc#AF148540	Parasnis et al. (2000)
	F-Napf-76: GAGGATCCCTATTAGTGTAAG R-Napf-77: GAGGATCCCTTTTGCACTCTG	831 bp	SCAR	Male		Parasnis et al. (2000)
	F-Napf-70: GGATCCCT ATTAG R-Napf-71: GAGGATCCCTTTTGC	0.83 kb	SCAR	Male	Mass screening	Parasnis et al. (2000)
GC	F-GN-C: CGAAATCGGTAGACGATACG R-GN-D: GGGGATAGAGGGACTTGAAC	0.6 kb		Sex neutral		Parasnis et al. (2000)
OPA06	OPA 06: 5′ GGT CCC TGAC 3′	365 bp 360 bp	DAF	Hermaphrodite and male		Somsri and Bussabakornkul (2008)

Development of Molecular Marker in Papaya

The first markers were morphological, which Hofmeyr (1938) mapped onto the initial papaya genetic map in 1938. Subsequently, from the 1970s to the 1990s, biochemical markers (isozymes) were used to study the development of papaya fruits and mature leaves (Tan and Weinheimer 1976), hybridity of *C. papaya* and *Carica cauliflora* (Moore and Litz 1984), and genetic relationships between *C. papaya* and wild *Vasconcellea* relatives (Jobin-Decor et al. 1997).

The first report of the use of DNA markers in a papaya genomic study was in 1992. Southern blot detection of restriction fragment length polymorphisms (RFLPs), produced by digesting total genomic DNA with restriction enzymes and detection by micro- and minisatellite probes, were used to detect polymorphisms between *C. papaya* genotypes and between *C. papaya* and its related species (Sharon et al. 1992). The report suggested that *C. papaya* genotypes and their related species could be identified and distinguished by their unique DNA fingerprints.

The invention of the polymerase chain reaction (PCR) in the early 1990s then led to rapid advances in the development and application of molecular markers. The first PCR-based markers for papaya were randomly amplified polymorphic DNA (RAPD) markers, which are dominant markers that are amplified by arbitrary short primer sequences. These were used to evaluate genetic relationships among papaya cultivars (Stiles et al. 1993) and between papaya and wild related *Vasconcellea* species. RAPD sourced markers were also applied to determine sex of papaya prior to flowering. However, RAPD markers can be difficult to reproduce on different equipment and by different researchers. In addition, although assumptions have been made in the past, RAPD markers cannot be used to determine allele differences at a particular locus without fragment sequencing. A variation of this technique, which employs increased annealing temperature and polyacrylamide gel detection, randomly amplified DNA fingerprinting (RAF), identified markers linked to papaya ringspot virus type P (PRSV-P) resistance in the related *Vasconcellea pubescens* (Dillon et al. 2005).

Subsequently, amplified fragment length polymorphism (AFLP) markers, integrating RFLP with PCR, were applied to assess genetic relationships within the Caricaceae (Kim et al. 2002; Van Droogenbroeck et al. 2002; Ocampo Pérez et al. 2007; Ratchadaporn et al. 2007). AFLP markers have also been placed on papaya genome maps (Ma et al. 2004; Blas et al. 2009). Approximately 1,500 AFLP markers were mapped onto 12 linkage groups (LGs) (Ma et al. 2004). However, using the same population, only 20 % of these markers were mapped along with other types of markers into nine major and five minor linkage groups (Blas et al. 2009). These markers, RAPD, RAF, and AFLP, which are nonspecific and dominant markers, could be useful for papaya crop improvements as they do not require much genetic information, i.e., DNA sequence analysis. However, the amplification patterns of these markers are complicated and they cannot identify single loci and alleles. Hence, more reliable, stable, and useful markers were needed.

SSRs, otherwise known as microsatellite markers, have many advantages over anonymous dominant markers. They are targeted to flanking sequences to amplify

the tandem short repeat units dispersed throughout the genome. Therefore, they are generally locus specific and may be codominant. Also, due to the phenomenon of conserved slippage, they may be highly polymorphic among individuals within a species, although they are generally not well transferred between species. The *C. papaya* SSR markers have been used for sex identification (Parasnis et al. 1999), in genetic diversity studies (Pérez et al. 2006; Ocampo Pérez et al. 2007; de Oliveira et al. 2008, 2010a, b; Eustice et al. 2008), for the construction of genetic maps (Blas et al. 2009; Chen et al. 2007) and integration to a physical map, and for comparing cytogenetic markers to merge linkage fragments (Yu et al. 2009; Wai et al. 2010).

Application

Molecular Markers and Papaya Genomic Studies

Papaya belongs to the order Brassicales which comprises 17 families including Caricaceae, which contains papaya, and *Arabidopsis*, the model plant in Brassicaceae. The family Caricaceae contains six genera and 35 species including *C. papaya*. Papaya is the only member in the genus *Carica*, has a relatively small genome size of 372 Mb (Arumuganathan and Earle 1991), and is diploid 2n = 18. Papaya has been identified as a model for biotechnology applications in tropical fruit species because it is an economically important fruit crop in tropical and subtropical regions, it has a small genome size, and it has a short generation time (9–15 months). The genome was successfully sequenced in 2008 for 'SunUp', which is a commercial virus-resistant transgenic genotype of papaya (Ming et al. 2008). The papaya genome was the fifth flowering plant to be sequenced. Compared to the other four plant genomes that were sequenced, papaya contains 24,746 genes, which is 20, 34, 46, and 19 % less than *Arabidopsis*, rice, poplar, and grape, respectively.

The first genetic map of papaya comprised three morphological markers: sex type, flower color, and stem color, covering 41 cM (Hofmeyr 1939) of the genome. The second papaya genetic map (F_2: Hawaiian cultivar 'Sunrise' × UH breeding line 356) was established 60 years later comprising 61 RAPD markers and one morphological locus (*SEX1*) within 11 linkage groups and comprising a total map distance of 999.3 cM (Sondur et al. 1996). In 2004, another map was produced using 54 F_2 plants derived from cultivars Kapoho × SunUp and containing 1,498 AFLP markers, a PRSV-P coat protein marker, and two morphological markers that determined sex type and fruit flesh color (Ma et al. 2004). These markers were mapped into 12 linkage groups with a total length of 3,294.2 cM and an average distance of 2.2 cM. This map was then integrated into a recent genetic map (Blas et al. 2009).

SSR markers have been used widely in papaya research and SSR libraries for papaya have been developed (Eustice et al. 2008; Santos et al. 2003; Wang et al. 2008). The patterns of SSR distribution are similar within genomic or genic regions and the most abundant motif repeats are dinucleotides. The AT/TA motif repeats are predominant across several studies (Eustice et al. 2008; Santos et al. 2003;

Wang et al. 2008; Nagarajan et al. 2008). The most abundantly detected trinucleotide motif differs among studies. AAG/TTC (44.1 %; Eustice et al. 2008), AAT/TTA (55.8 %; Wang et al. 2008), and TAC/ATG, AGA/TCT, and ATT/TAA were reported to be very common in the papaya genome sequence (Nagarajan et al. 2008). This is likely due to the different methods used to isolate the SSR sequences. SSRs were identified from the papaya genome sequence using Microsatellite Analysis Server (Eustice et al. 2008), Perl program, Simple Sequence Repeat Identification Tool (SSRIT; Wang et al. 2008), and the Tandem Repeats Finder software (Nagarajan et al. 2008).

SSR markers were included within a subsequent map constructed with 54 F_2 plants that were derived from AU9 × SunUp. This comprised 706 SSR markers and one morphological marker for fruit flesh color within 12 linkage groups over 1,068.9 cM and an average distance of 1.51 cM (Chen et al. 2007). The nine major linkage groups represent nine chromosomes in the papaya genome, covered 993.5 cM with 683 map loci with an average marker density of 1.45 cM. The three short linkage groups covered 75.4 cM with 24 mapped loci and an average distance of 3.1 cM between adjacent markers.

Codominant markers have greatly increased the resolution and accuracy of papaya genetic maps. The map containing SSR markers was far more compact (1,068.9 cM) than that produced by AFLP markers (3,294.2 cM). The threefold reduction in genetic distance was due to the ability of codominant markers to separate the three classes of genotypes within the F_2 population: homozygous dominant, homozygous recessive, and heterozygous.

A recent genetic map of papaya was saturated with 712 SSR, 277 AFLP, and one morphological marker spanning 945.2 cM (Blas et al. 2009). This was constructed with the same 54 F_2 AU9 × SunUp population and comprised 14 linkage groups, nine major and five minor. The 9 major linkage groups incorporated 939 marker loci over a total 849.4 cM at 1.6 cM or less intervals. The 5 minor linkage groups comprised 51 loci and spanned 95.8 cM with 3.0 cM or less intervals. Comparing the previous map (Chen et al. 2007) to this recent map, the addition of AFLP markers allowed six unmapped SSR markers from previous map to be linked, but did not join the gap between three previous minor linkage groups and the nine major linkage groups. However, the number of gaps that were greater than 5 cM between adjacent loci was reduced from 48 to 27, and the total map length was reduced by approximately 11.5 %. The addition of AFLP markers resulted in an order of locus rearrangement.

Prior to the availability of the whole genome sequence, bacterial artificial chromosome clones containing papaya genomic DNA were produced and assembled. The first papaya BAC library consisted of 39,168 clones from two separate ligation reactions (Ming et al. 2001). The average insert size of 18,700 clones from the first ligation was 86 kb, while 20,468 clones from the second ligation contained inserts twice as large, averaging 174 kb. The entire BAC library was estimated to provide a 13.7 × papaya genome coverage. In 2006, a total of 50,661 BAC end sequence (BES) chromatograms were generated from 26,017 BAC clones (Lai et al. 2006) from the BAC clone library by Ming et al. (2001). After eliminating all unused

sequences, 35,472 high-quality sequences from 20,842 BAC clones were generated. The total number of high-quality bases was 17,483,563 or 4.7 % of the papaya genome. Two years later, a three times draft genome of cultivar 'SunUp' was reported within which 1.6 million high-quality reads were assembled into contigs containing 271 Mb and scaffolds spanning 370 Mb (Ming et al. 2008). Subsequently, 652 BAC and whole genome sequence-derived SSRs were used to anchor 167 Mb of contigs and 235 Mb of scaffolds to the 12 linkage groups of papaya on the current genetic map (Ming et al. 2008).

A physical map of the papaya genome that integrated with the genetic map and genome sequence was published in 2009 (Yu et al. 2009). The BAC-based physical map of papaya covered 95.8 % of the genome, while 72.4 % was aligned to a sequence-tagged SSR genetic map (Chen et al. 2007). This BAC library initially included 39,168 BAC clones; however, after evaluation and reviews, 26,466 BAC clones were assembled into 963 contigs having an average number of fragments for each clone of 69.4. The average physical distance per centimorgan was approximately 348 kb. The integrated genetic and physical map when aligned with the genome sequence revealed recombination hot spots as well as regions suppressed for recombination across the genome, particularly on the sex chromosome, namely, LG1 (Yu et al. 2009). A total of 1,181 overgos representing conserved sequences of *Arabidopsis* and genetically mapped *Brassica* loci were anchored on the integrated genetic and physical map and the draft genome sequence of papaya. These overgos are direct links among papaya, *Arabidopsis,* and *Brassica* genomes for comparative genomic research among species within the order Brassicales. The combined information of physical and genetic maps will enhance the capacity for map-based cloning and identification of underlying genes controlling traits of interest in papaya. It will also expedite the mapping and cloning of target genes and promote marker-assisted selection for papaya breeding.

Recently, chromosome-specific cytogenetic markers were developed and merged with linkage groups of papaya using the integrated technique of fluorescence in situ hybridization (FISH) and BAC clones harboring mapped SSR markers as probes (Wai et al. 2010). Minor linkage groups 10, 11, and 12 from the previous map were assigned to major LGs 8, 9, and 7, respectively. Thus, the nine linkage groups in the genetic map corresponded to the haploid number of papaya chromosomes. This integrated map will facilitate genome assembly, quantitative trait locus mapping, and the study of cytological, physical, and genetic distance relationships between papaya chromosomes. It is an even more powerful and accurate tool for trait selection.

Molecular Markers and Genetic Diversity in *Caricaceae*

Variations in *C. papaya* in terms of phenotypic, morphological, and horticultural characteristics such as fruit size, fruit shape, flesh color, texture, flavor and sweetness, stamen carpellody and carpel abortion, sex type, length of juvenile period, plant stature, and plant canopy size can be detected at high levels between different

genotypes of papaya in the field. However, most of this morphological diversity in papayas has not been correlated to genetic diversity, specific genes, or molecular markers despite many studies on genetic diversity. Different techniques have been used to study genetic diversity in papaya and relationships between plants within *Caricaceae* including AFLPs, isozymes, RAPDs, microsatellites, and SSRs.

Kim et al. (2002) studied genetic relationships between papayas and related *Vasconcellea* species using samples which had a wide range of morphologies and climate adaptation variations, from tropical, subtropical, and temperate regions. But only 12 % of genetic variation was detected among this diverse group of material using AFLP markers, and it was not representative of the wide range of morphological characteristics that were observed in the field. The accessions that Kim et al. (2002) used in their study consisted of breeding lines, unimproved germplasm, and related species. Five cultivars of papayas were initially screened for polymorphisms by 64 sets of *Eco*RI-*Mse*I primers with three nucleotide extensions. The number of polymorphic markers ranged from 0 to 9 with an average of 3.2. Nine primers were selected to assess all samples and generated 186 polymorphic markers (42 %) from 445 readable fragments. The estimation of genetic similarity using pairwise comparison among 63 papaya accessions ranged between 0.74 and 0.98 (mean 0.88). Cluster analysis of 71 papaya accessions and related species showed the genetic relationship among individual genotypes which developed in different geographic regions. The first cluster included all 15 Solo-type cultivars and breeding lines. The second cluster included dioecious Australian cultivars and Indian cultivars that grow in subtropical or temperate regions. The third cluster was a group of cultivars originating from different countries. AFLP markers were used to study genetic relationships between papaya and wild relatives by Van Droogenbroeck et al. (2002) who analyzed 95 accessions of papayas and wild relatives from Ecuador. Five primer combinations were used and revealed 951 bands ranging in length from 50 to 500 base pairs. Only 512 bands were scored, of which 16 were monomorphic. All papaya genotypes were clustered separate to two other clusters comprising individuals from *Jacaratia* or *Vasconcellea*. Genetic similarity ranged from 0.39 to 0.81 among the *Vasconcellea* species that were assessed.

Genetic relationships between *C. papaya* and wild relatives were studied by Jobin-Decor et al. (1997) by comparing isozyme and RAPD techniques. A total of 47 bands were generated by nine enzymes and 188 bands were generated from 14 RAPD primers. Both techniques gave similar measures of genetic distance of 70 % dissimilarity between *C. papaya* and the other *Carica* species (later renamed *Vasconcellea* species) and approximately 50 % dissimilarities among *Vasconcellea* species. RFLP markers were used to study phylogenetic relationships using the chloroplast DNA (cpDNA) of 12 wild and cultivated species of *Carica* (Aradhya et al. 1999). Twenty-three accessions, representing 14 taxa, were analyzed in the cpDNA intergenic spacer region by amplification of the region via PCR technique and then the PCR product was digested by 14 restriction enzymes. A total of 138 fragments accounting for 137 restriction sites were examined and the results confirmed the close association among South American wild *Carica* (*Vasconcellea*) species.

From a microsatellite-enriched library developed using $(GA)_n$ and $(GT)_n$ probes (Pérez et al. 2006), 45 primer pairs giving the best resolution and allelic differentiation were used for evaluation of 29 accessions of *C. papaya* and 11 accession of *Vasconcellea*. Of these, 24 revealed polymorphisms between these two genera. A total of 99 alleles were observed in papaya with an average of 3.8 alleles per locus. In the *Vasconcellea* samples, 22 alleles were identified from four loci. These two genera had a clear allelic divergence for the loci that they shared. This strong differentiation gave further support to the hypothesis of the early divergence of *Vasconcellea* from *Carica* (Pérez et al. 2006). Many more SSR markers were identified from the genome sequence project (Eustice et al. 2008). These were tested for polymorphism in seven genotypes of papaya. Of the 938 SSR markers that were defined, 812 were from genomic sequences and 126 from genic sequences (Eustice et al. 2008). Overall, 52.9 % were polymorphic. SSR primers developed in 2008 (Oliveira et al. 2008) and were screened on 30 papaya accessions and 18 landraces (de Oliveira et al. 2010b). Of the 100 SSR primers, 81 successfully amplified PCR products of high quality and were selected for further studies. Of these, 59 produced easily scorable markers and detected a total of 237 alleles with 2–11 per locus. In a separate study using 27 of the same SSR loci developed in 2008 (Oliveira et al. 2008), the relationships among 83 papaya lines were assessed (de Oliveira et al. 2010b). Of the 27 primers, 20 were polymorphic and identified a total of 86 alleles, with an average of 3.18 alleles per primer. Since cultivated *C. papaya* is proposed to have a low or narrow genetic base (Ratchadaporn et al. 2007; Stiles et al. 1993), many important genes conditioning traits of interest (such as disease resistances and abiotic stress tolerances) may have been lost or excluded in the pursuit of other traits (such as fruit color and sweetness). Indeed, many disease resistances are found in wild relative *Vasconcellea* species and future breeding may require interspecific recombination to reintrogress these back into the elite cultivated genomes (Coppens et al. 2013; also Chap. 4 in this book).

A number of taxonomy studies have confirmed the diversity between *C. papaya* and *Vasconcellea* species and have supported the separation into two genera by Badillo (2000). Most studies supported the early divergence of *C. papaya* from the wild relatives, and this has been verified by different molecular marker techniques: isozyme and RAPD (Jobin-Decor et al. 1997), RFLP (Aradhya et al. 1999), and AFLP (Kim et al. 2002; Van Droogenbroeck et al. 2002). The most closely related species were reported as *V. stipulata* and *V. pubescens* (Jobin-Decor et al. 1997; Sharon et al. 1992) and *V. goudotiana* and *V. pubescens* (Kim et al. 2002). The species most distant from *C. papaya* were reported to be *V. cauliflora* (Jobin-Decor et al. 1997) and *V. goudotiana* (Sharon et al. 1992; Kim et al. 2002). In particular, Kim et al. (2002) reported that the average genetic similarity between papaya and other *Vasconcellea* species was 0.43 and the average similarity among six different species of *Vasconcellea* was 0.73, much closer to each other than to *C. papaya*. *C. goudotiana* was the most distantly related species to papaya (0.36 similarity), while it was closely related to *C. pubescens* (*V. pubescens*) with a similarity of 0.87.

Previously, Jobin-Decor et al. (1997) had reported *C. papaya* to be distinct from the other *Carica* species, *C. cauliflora*, *C. parviflora*, *C. pubescens*, *C. goudotiana*,

C. stipulata, and *C. quercifolia,* with a mean dissimilarity of 0.73 and 0.69 using isozyme and RAPD analysis, respectively. As this work was done before 2000, all these species were considered to be in *Carica* genus at that time. The other *Carica* (*Vasconcellea*) species were more closely related to each other with a mean dissimilarity of 0.46. The closest two species were *V. stipulata* and *V. pubescens* with the dissimilarity of 0.13 and 0.18 and 0.87 and 0.82 using isozyme and RAPD analysis, respectively. However, from RAPD analysis, the *Vasconcellea* species most distant from *C. papaya* was *cauliflora*. It was reported that *C. papaya* and other species had band sharing between 25 and 48 % (Sharon et al. 1992). *C. goudotiana* is more distant from *C. papaya* with band sharing of 25 %. *C. stipulata* and *C. pubescens* were very closely related species with band sharing of 71 %.

In summarizing their genetic diversity, it should be noted that the *Vasconcellea* species have recently been divided into three clades (Coppens et al. 2013; Chap. 4 in this book). They are (1) *V. weberbaueri, V. stipulata, V. x heilbornii,* and *V. parviflora*; (2) *V. chilensis, V. candicans, V. quercifolia,* and *V. glandulosa*; and (3) a clade holding all other species of the genus (Coppens et al. 2013; Chap. 4 in this book).

Molecular Marker and Sex Determination in *C. papaya*

C. papaya is a polygamous species with three basic sex types: female, male, and hermaphrodite. Although the male plants occasionally produce hermaphrodite flowers on the abaxial end of inflorescences, they do not produce commercial fruit. However, in dioecious plantings, which are common in subtropical regions, male plants are still needed for pollen. Usually the ratio of male and female plants is 1:10; thus multiple seedlings are planted at each site and then the plants are thinned to achieve the required ratio of female and male plants. In tropical regions hermaphrodite trees are preferred because every tree is capable of producing marketable fruits. In plantations comprising hermaphrodite plants, female plants are unwanted and removed after flowering. Commercial papaya growers have to plant 3–5 plants per site, then evaluate sex type after flowering and cull the undesired plants. The process is time consuming, laborious, and cost ineffective. In addition, competition between multiple plants at each planting site causes poor root systems, elongated plants, and increased height to the first flower and fruit. Therefore, the use of DNA markers to discriminate sex of papaya plants at the earliest plant growth stage has the potential to greatly increase efficiency within the papaya production system. Thus much research has been applied to this subject in recent years.

Genetic control of sex of papaya has been studied since 1938. Hofmeyr (1938) and Storey (1938) independently proposed the hypothesis that sex determination in papaya is controlled by a single dominant gene with three alleles, named M^1, M^2, and *m* by Hofmeyr and *M*, M^h, and *m* by Storey. They proposed the genotype of male, hermaphrodite, and female plants are *Mm* (M^1m), M^hm (M^2m), and *mm,* respectively, and explained that homozygous dominant alleles are lethal. Therefore, segregation of sex type from selfed hermaphrodite trees is 2 hermaphrodites: 1 female.

Whereas seeds from female trees segregate at the ratio of 1:1 female: hermaphrodite if the plant is crossed with a hermaphrodite tree, but that of 1:1 female: male when it is crossed with a male tree. Subsequently, other researchers have proposed other hypotheses for genetic control of sex in papaya and this was reviewed by Ming et al. (2007).

Most research on DNA marker-assisted sex selection in papaya has been done on Hawaiian papaya genotypes. The first report of a sex-linked marker in papaya was reported by Sondur et al. (1996). They created a genetic linkage map using RAPD markers and investigated the genetics of sex determination in papaya using an F_2 population of Hawaiian cultivar Sunrise × IH breeding line 356. Of 596 10-base primers screened, two; OPT1 and OPT12, produced two marker bands; OPT1C and OPT12, flanking the *SEX1* locus at 7 cM for both markers. RAPD and DNA amplification fingerprinting (DAF) were compared and showed that DAF reactions produced at least five times more fragments than equivalent RAPD reactions in terms of ability to detect variation. They also revealed that DAF reactions were more reliable (Somsri et al. 1998). Bulk segregant analysis was used to define a large number of DAF markers present in only male or hermaphrodite pooled DNAs. Preliminary analyses for linkage associations indicated these markers were closely linked to the sex-determining alleles. Ten years later, Somsri and Bussabakornkul (2008) employed DAF to study the relationships between 14 cultivars of papaya in Thailand. Using 11 primers, a total of 129 distinct fragments were amplified. Primer OPA06 could be used to identify the sex type of papaya plants. This primer produced two polymorphic bands: at 365 bp from the hermaphrodite bulk DNA and 360 bp from the male bulk DNA. Neither band was detected for females. Evaluation of the accuracy of OPA 06 analysis was verified using 254 plants of different generations and their original parents, and the analysis correctly identified sex type for 88.18 % of the plants. In the final experiment, 47 hermaphrodite plants of the Khaeg Dum cultivar, that were grown in tissue culture, were examined using OPA 06, and the sex type was identified correctly for 100 % of the plants.

From 2000 onwards, the sequence-characterized amplified region (SCAR) technique has been used to increase specificity of priming sites from RAPD primers to the target DNA for sex determination of papaya. Eighty RAPD primers from Operon kits were screened on 12 different papaya varieties and the marker OPF2-0.8 kb was identified as male specific (Parasnis et al. 2000). The marker was converted into a SCAR marker, by designing a 20-bp primer pair as a sex-specific primer. They also developed an internal control for the PCR reaction using primer GC, which is neutral for sex, meaning the marker presents in all sex types of papaya. For mass screening, they developed a single step DNA extraction and used a 15-bp SCAR primer at a lower annealing temperature for sex detection.

The information from Sondur et al. (1996) was applied by cloning three RAPD products and SCAR primers were designed on these sequences. Two RAPD markers, OPT1C and OPT12, flanked the *SEX1* gene (Sondur et al. 1996); however, Deputy et al. (2002) identified another marker, W11, that did not show recombination in the population that they used. So, W11 was included in this study along with T1 and T12. A SCAR T1 primer was designed on the interval sequences of the T1 marker;

however, SCAR T1 can amplify all sex types: female, male, and hermaphrodite at 1,300 bp. This could be because one or more point mutations exist in the original 10-base primer site that can distinguish females from hermaphrodites. Therefore, T1 was used as positive (or internal) control for PCR amplification. T12 and W11 were designed on the original 10-base plus a further 10–11 bases to make the primer more specific to papaya sequences. Both primers successfully differentiated sex type male/hermaphrodite from female plants generating the marker at 800 bp for both primers. The linkage analysis of SCAR markers W11 and T12 in 182 F_2 plants from the cross of SunUp × Kapoho indicated that these markers were within 0.3 cM of *SEX1*.

RAPD techniques were used to determine the sex of 11 Hawaiian cultivars of *C. papaya* for three sex types: male, female, and hermaphrodite (Urasaki et al. 2002b). Twenty-five arbitrary 10-mer primers were tested with papaya DNA. The IBRC-RP07 primer produced a fragment named PSDM at 450 bp in all male and hermaphrodite but not female plants. The fragment was analyzed and a SCAR marker named SCARps was designed and produced a PCR product of 225 bp in the male and hermaphrodite plants only. A multiplex PCR assay was developed for a sex-specific SCARpm marker and a marker for a papain gene as an internal control to minimize false negatives (Urasaki et al. 2002a). The marker was tested for amplification on a hermaphrodite plant of *C. papaya* 'Sunrise Solo'. This marker differentiated hermaphrodite and male plants from female plants.

Lemos et al. (2002) screened 152 RAPD primers on female and hermaphrodite plants of *C. papaya* cv. Baixinho de Santa Amália, cv. Sunrise Solo, and cv. Improved Sunrise Solo 72/12. Primer BC210 produced a marker band at 438 bp ($BC210_{438}$) that was present in hermaphrodite but not in female plants. The marker was tested with 195 different samples from three cultivars and was present only in hermaphrodite plants. Published sex-linked markers were validated on a selection of Brazilian commercial genotypes, two varieties of a Solo group and two hybrids of the Formosa group (de Oliveira et al. 2007). Four SCAR markers (Deputy et al. 2002; Parasnis et al. 2000; Urasaki et al. 2002b; Chaves-Bedoya and Nuñez 2007) revealed the presence of both false positives and negatives in some varieties, while the RAPD marker $BC210_{438}$ (Lemos et al. 2002) could predict papaya sex type correctly.

In another study, the sex-linked markers that were described previously were tested on three Columbian papayas, but could not distinguish between male/hermaphrodite and female plants (Chaves-Bedoya and Nuñez 2007). Therefore, 32 arbitrary 10-mer Operon primers were screened and the OP-Y7 primer that generated a PCR product that had 900 bp and was present only in male plants and absent in female and hermaphrodite plants. The marker was analyzed and converted into a SCAR marker that could differentiate female plants from male and hermaphrodite plants. The SCAR SDSP marker at 369 bp was present in male and hermaphrodite but not female plants.

In Sri Lanka, Niroshini et al. (2008) screened 100 arbitrary decamer primers in ten plants of each papaya sex type. The plants were selected from home gardens near Kadawatha, Sri Lanka. Of the 100 primers that were tested, two primers OPC09

and OPE03 produced two DNA maker bands specific to male/hermaphrodite plants at 1.7 and 0.4 kb, respectively. The markers were then sequenced and SCAR primers were designed by use of extension oligonucleotides at both ends of the sequence. SCAR primer C09/20 amplified two fragments of length 1.7 kb and 978 bp in both male and hermaphrodite plants. However, SCAR primer E03/20FP and E03/20RP that were designed from OPE03-0.4 did not detect polymorphisms among plants of different sex types.

Other markers based on microsatellites have been used for sex detection in papaya by Santos et al. (2003). It was possible to design primers from the library from sequences enriched with the probe $(TCA)_{10}$. Thirty-two pairs of SSR primers were designed; however, none of them could identify sex type in this study. Commercial cultivars of papaya and a wild species *V. cauliflora* were screened by digesting genomic DNA with various restriction enzymes and microsatellite sequences were used as probes by Parasnis et al. (1999). The microsatellite repeats $(GATA)_4$ and $(GAA)_6$ detected sex-specific differences in *Hin*fI or *Hae*III digested samples. However, only the repeat $(GATA)_4$ showed male/hermaphrodite specific bands at 5 and 4 kb after digesting the genomic DNA with *Hin*fI and *Hae*II, respectively, while the repeat $(GAA)_6$ could detect polymorphisms for sex in some cultivars only.

Inter simple sequence repeat (ISSR) and RAPD techniques were used to determine sex of 200 seedlings of a local variety of papaya (Gangopadhyay et al. 2007) in Kolkata, India. Ten RAPD primers (OPA01-OPA05 and OPB01-OPB05) failed to show polymorphisms among the three sex types. Of three ISSR primers used in this study, one ISSR primer $(GACA)_4$ could distinguish female or hermaphrodite from male plants.

In conclusion, much marker research has been applied to the identification of DNA markers for the differentiation and selection of male, female, and hermaphrodite plants. Although many of these markers can be used for marker-assisted selection in breeding programs, there is a need to adapt some of these markers into low-cost techniques that can identify sex of seedlings and thus facilitate commercial production of papayas. Over planting then thinning to achieve the desired sex ratios after plants reach flowering stage is still a major and potentially unnecessary cost for papaya producers in both tropical and subtropical regions.

Molecular Marker-Assisted Selection for Papaya Breeding

Disease Resistance

Lack of disease resistance genotypes in *Carica* is the major problem for crop improvement, while resistance to all diseases that attack papaya can be found in wild relative *Vasconcellea* species (Coppens et al. 2013; Chap. 4 in this book). Therefore the relationship between these two genera has been studied with the aim of transferring resistant genes from *Vasconcellea* spp. to papaya.

Papaya ringspot virus, caused by papaya ringspot virus type P (PRSV-P), is widely reported as the most devastating disease of papaya production worldwide. The sustainable method to control this disease in the papaya industry is to produce improved papaya varieties that are resistant to the pathogen. A transgenic papaya variety that was resistant to PRSV-P was successfully developed in 1990 (Fitch et al. 1990); however, GMOs plants are not accepted in many countries, and transgenic resistant varieties can be virus-strain specific in their resistance (Tennant et al. 1994). Therefore, conventional breeding for PRSV-P resistance in papaya is a viable option for long-term control of the disease, and until GMO food crops are more universally accepted worldwide.

Resistance to PRSV-P has been reported in *V. cauliflora, V. stipulata, V. pubescens, and V. quercifolia* (Manshardt and Drew 1998). The gene for PRSV-P resistance was successfully backcrossed from *V. pubescens* into *V. parviflora* from F_3 interspecific hybrids containing the homozygous dominant allele of the gene (O'Brien and Drew 2009). This suggested that a single dominant gene controlled PRSV-P resistance in *V. pubescens* and was consistent with earlier reports on the PRSV-P resistance in a generic hybrid between *V. pubescens* and *C. papaya* (Drew et al. 1998). This is consistent with reports of single gene dominance regulating the PRSV-P resistance in F_1 intergeneric hybrids of *C. papaya* and *V. cauliflora* (Magdalita et al. 1997). Similar results were reported in interspecific hybrids between *V. pubescens* × *V. parviflora* (Dillon et al. 2006). Genetic mapping studies of these hybrids further supported the concept that PRSV-P resistance in *V. pubescens* is controlled by a single dominant gene (Dillon et al. 2006). This was confirmed with molecular markers which were linked to a single locus resistance gene (*prsv-1*) that was identified in *V. pubescens* (Dillon et al. 2005, 2006).

A genetic map of PRSV-P resistance gene(s) based on RAF markers was constructed from100 F_2 plants of *V. cundinamarcensis* (*V. pubescens*) and *V. parviflora* (Dillon et al. 2005). Mapping of dominant markers in repulsion phase in F_2 populations can cause an incorrect estimation of genetic distance; therefore, RAF markers were mapped to separate parental maps. For *V. pubescens*, markers were mapped to ten linkage groups that covered 745.4 cM with an average distance of 9.68 cM between adjacent markers. The PRSV-P resistance locus (*prsv-1*) was mapped within 4 and 2.8 cM of adjacent markers Pbw15_40 and OPA15_8 on LG7. For *V. parviflora*, markers were mapped to ten linkage groups that covered 630.2 cM separated by an average distance of 7.95 cM between adjacent markers. The markers Pbw15_40 and OPA15_8, flanking *prsv-1* locus, were near but not on the resistance gene-coding region, as they did not colocate with the resistant phenotype.

Markers linked to *prsv-1* have been used in the marker-assisted breeding programs described previously because of their dominant inheritance and because resistance to the Australian strain of PRSV-P imparted by *prsv-1* has been shown to be robust (Drew et al. 1998; Magdalita et al. 1997). Five DNA markers, which were developed by use of RAF on bulked segregants of virus resistant and susceptible populations, were identified in the cross of *V. parviflora* × *V. pubescens* (Dillon et al. 2006). The markers were mapped to the same linkage group, LG7, flanking *prsv-1*

at the distance of 2.1, 5.4, 9.7, and 12.0 cM for the marker Opa_16r, Opk4_1r, Opk4_2r, and Opb8_1r, respectively, while another marker Opa11_5R colocated with *prsv-1*. The two candidate markers, Opa11_5R and Opk4_1r, were sequenced and converted to SCAR markers. A SCAR marker, Opk4_1r, was converted into a CAPS marker, *Psilk4*, by digesting the amplicon with *Psi*1 and was shown to be diagnostic for the three alleles of *prsv-1*. The SCAR marker Opk4_1r detected similar band sizes for *V. pubescens, V. cauliflora,* and *V. goudotiana* with the size of 360 bp, 360 bp, and 361 bp, respectively. However, the amplicons of *V. parviflora*, *V. quercifolia,* and *V. stipulata* were slightly larger with the size of 372 bp. The application of this SCAR and CAPS marker for marker-assisted breeding was confirmed in research on interspecific populations of *V. pubescens* × *V. parviflora*; F_2 and F_3 populations produced from the *V. pubescens* and *V. parviflora* F_1 and BC_1 and BC_2 generations when *V. pubescens* × *V. parviflora* F_3 RR plants were backcrossed to *V. parviflora* (O'Brien and Drew 2010). The Opk4_1r SCAR marker amplified in other *Vasconcellea* spp. *quercifolia, goudotiana,* and *cauliflora;* however, the CAPS marker was not consistent in determining the allele of prsv-1 in crosses involving *Vasconcellea* spp. other than *pubescens* and *parviflora* (O'Brien and Drew 2010).

Transgenic papayas resistant to PRSV-P have been developed in many countries, and DNA markers have been used to detect the transgenic plants. A detection protocol for characterization of PRSV coat protein transgenic papaya lines was demonstrated by use of PCR (Fan et al. 2009). PCR patterns using primers designed from the left or right flanking DNA sequence of the transgene insert in transgenic papaya lines were specific and reproducible.

In addition to PRSV-P, papayas are susceptible to many other pathogens. However, there are few reports on the use of DNA markers to identify other disease genes. However, tolerance to *Phytophthora palmivora* in papaya was identified, and molecular markers linked to this resistance were developed by use of AFLPs (Noorda-Nguyen et al. 2010). Several polymorphic bands linked with the tolerance trait in a F_2 population, derived from an F_1 of the most tolerant Hawaiian cultivar Kamiya, crossed with a highly susceptible cultivar SunUp, have been identified. These markers were further characterized to form SCAR markers. Nineteen genes were selected for gene expression analysis for resistance to *P. palmivora* (Porter et al. 2009a). Of these genes, a predicted peroxidase, β-1,3-glucanase, ferulate 5-hydroxylase, and hypersensitive-induced response protein were pathogen upregulated, while a second peroxidase (Cp9) and aquaporin (Cp15) were downregulated.

Hybrid Identification

To be able to identify hybrids at an early stage of plant growth is preferable in crop improvement studies, especially confirmation of F_1 hybrids in populations to facilitate further crossing. Seventeen RAPD primers were selected and screened to confirm hybridity of 120 putative interspecific cross of *C. papaya* × *C. cauliflora*

(whereas, nowadays, they are intergeneric cross of *C. papaya* × *V. cauliflora*; Magdalita et al. (1997)). A range of 1–5 primers consistently confirmed that all 120 plants were genetic hybrids. A single primer cannot guarantee accurate results; thus more than one marker is necessary to analyze for hybridity. This is because chromosome elimination can occur during meiosis; thus absence of a single marker may represent elimination of part of a chromosome.

Fruit Quality Traits

Even though papaya genomic research is exceptionally advanced in many aspects, there are many other characteristics of papaya that need to be studied. Very little information on other characteristics of papaya has been studied in the past few years. One of the characteristics that need to be identified for papaya marketing is fruit flesh color. The Australian papaya industry is clearly split between growers who grow either the yellow-fleshed (commonly described as Pawpaw) or the red-fleshed varieties (known as Papaya) or both types (Australian Papaya Industry Strategic Plan, 2008–2012). A single major gene for yellow flesh that is dominant over red flesh color has been found in a simple Mendelian segregation in flesh fruit color of papaya by recognized papaya breeders (Blas et al. 2010). The carotenoid composition profiles of red- and yellow-fleshed Hawaiian Solo papayas showed a strong accumulation of lycopene in red-fleshed fruit, while none were detected in yellow flesh (Yamamoto 1964). The flesh color locus was mapped near the end of LG7 and the two flanking markers were located at 3.4 and 3.7 cM, respectively (Ma et al. 2004). Recently, a high-density genetic map of papaya using SSR markers was established, fruit flesh color was mapped at the end of LG5, and the closet marker was located at 13 cM (Chen et al. 2007). Blas et al. (2010) reported the cloning and characterization of the papaya chromoplast-specific lycopene β-cyclase, *CpCYC-b*, and genomic analysis of the surrounding region included a recombination hot spot in papaya. They found tomato chromoplast-specific lycopene β-cyclase (*CYC-b*) and SunUp *CpCYC-b* shared 75 % sequence identity over a 682-bp genomic sequence length. Quantitative RT-PCR analysis and subsequent functional analysis in bacteria confirmed the role of *CpCYC-b* in controlling fruit flesh color in papaya. The elevated expression of *CpCYC-b* and papaya β-carotene hydroxylase (*CpCHY-b*) between yellow-fleshed Kapoho and red-fleshed SunUp validated activation of the carotenoid biosynthesis pathway in yellow-fleshed papaya. The disruption of this pathway in red flesh varieties is caused by a frame-shift mutation induced by a 2-bp insertion (Blas et al. 2010). A PCR-based marker was developed, CPFC1 marker (Table 19.2), which is 530 bp away from *CpCYC-b*. The marker showed approximately 98 % recombination frequency to flesh color in 219 F_2 (KD3 × 2H94). It should be noted that this tightly linked marker, only 580 bp away from the target gene, is still not 100 % accurate due to the extremely high recombination rate in this region of the genome.

Table 19.2 DNA markers for other genetic traits in *Carica papaya*

Characteristics	Markers name	Primer sequence	Reference
PRSV-P coat protein region detection	Cp	F: GAGAAGTGGTATGAGGGAGTG R: CCATACCTGCCGTCACAATCA	Dillon et al. (2005)
PRSV-P resistance; *prsv-1* locus	Opa5_11R	PBA115R F: CAATCGCCGTAGGAAAATTC PBA115R R: CAATCGCCGTAGAGGAGGAGG	Dillon et al. (2006)
PRSV-P resistance; *prsv-1* locus	Opk4_1r	PBK41R F: CCGCCCAAACTGCGGAACAC PBK41R R: CCGCCCAAACCCCCAACTAG	Dillon et al. (2006)
Flesh color; *CpCYC-b*	CPFC1	F: GACGTGTTAGTGTCCGACAA R: GACCAGGAAGCAAATTTTGTAA	Blas et al. (2012)

Quantitative Trait Loci in Papaya

Many ergonomically important traits, i.e., fruit size, fruit shape, flesh flavor, and skin quality, are quantitative traits that are influenced by multiple genes. Various quantitative trait loci (QTLs) have been identified in many crops; however, surprisingly not many QTLs have been identified in papaya breeding research. Recently, in 2012, a QTL analysis for papaya fruit size and shape has been reported (Blas et al. 2012). Fifty-four SSR markers, the morphological flesh color locus, and CPF1 and CPF2 SCAR markers were mapped to 11 LGs using a population of 219 F_2 plants (KD×2H94). Fourteen QTLs having phenotypic effects ranging from 5 to 23 % were identified across six linkage groups. These loci contain homologs to the tomato fruit QTL *ovate*, *sun,* and *fw2.2* regulating fruit size and shape.

Potential and Future

Papaya genomic research is exceptionally advanced in many aspects. There is a recent genetic map which correlates with the number of chromosomes. A physical map was analyzed, a draft genome of the plant was sequenced and anchored on the genetic map, the genome was thoroughly analyzed for SSRs and *NBS* gene families, and SSR libraries have been developed. Thus much basic information has been revealed in the past few years. However, there are many other characteristics of papaya that need to be studied, and the molecular markers that have been identified, particularly the SSR markers, need to be linked to the many important traits of interest for MAS in papaya breeding.

Future papaya genomics approaches to develop more precise tools for trait selection are likely to involve the identification and mapping of candidate genes from the full genome sequence. These must then be tested for functional validation to the trait

of interest, potentially through expression analysis. Meanwhile, with affordability becoming a reality, it is likely that expression of representative RNA sequences of the genome will be analyzed and functional genes isolated directly from the transcriptome in response to a particular target trait, such as disease resistance. We predict that in the very near future, that suits of expressed papaya genes and their predicted pathways will be commonly available for e-mapping and trait-associated marker-assisted selection.

References

Aradhya MK, Manshardt RM, Zee F, Morden CW (1999) A phylogenetic analysis of the genus *Carica* L. (*Caricaceae*) based on restriction fragment length variation in a cpDNA intergenic spacer region. Genet Resour Crop Evol 46(6):579–586

Arumuganathan K, Earle E (1991) Nuclear DNA content of some important plant species. Plant Mol Biol Rep 9(3):208–218

Badillo VM (2000) *Carica* L. vs. *Vasconcella* St. Hil. (*Caricaceae*): con la rehabilitacion de este ultimo. Ernstia 10:74–79

Blas AL, Yu Q, Chen C, Veatch O, Moore PH, Paull RE, Ming R (2009) Enrichment of a papaya high-density genetic map with AFLP markers. Genome Res 52(8):716–725

Blas AL, Ming R, Liu Z, Veatch OJ, Paull RE, Moore PH, Yu Q (2010) Cloning of the papaya chromoplast-specific lycopene β-cyclase, *CpCYC-b*, controlling fruit flesh color reveals conserved microsynteny and a recombination hot spot1[W][OA]. Plant Physiol 152(4):2013–2022

Blas A, Yu Q, Veatch O, Paull R, Moore P, Ming R (2012) Genetic mapping of quantitative trait loci controlling fruit size and shape in papaya. Mol Breeding 29(2):457–466. doi:10.1007/s11032-011-9562-1

Chaves-Bedoya G, Nuñez V (2007) A SCAR marker for the sex types determination in Colombian genotypes of *Carica papaya*. Euphytica 153(1):215–220

Chen C, Yu Q, Hou S, Li Y, Eustice M, Skelton RL, Veatch O, Herdes RE, Diebold L, Saw J, Feng Y, Qian W, Bynum L, Wang L, Moore PH, Paull RE, Alam M, Ming R (2007) Construction of a sequence-tagged high-density genetic map of papaya for comparative structural and evolutionary genomics in brassicales. Genetics 177(4):2481–2491

Coppens G, d'Eeckenbrugge C, Drew R, Kyndt T, Scheldeman X (2013) Vasconcellea for papaya improvement. In: Ming R, Moore P (eds) Genetics and genomics of papaya. Springer Science+Business Media, New York

de Oliveira EJ, Dantas JLL, Castellen MDS, de Lima DS, Barbosa HD, Motta TBN (2007) Molecular markers for sex identification in papaya. Pesquisa Agropecuaria Brasileira 42(12):1747–1754

de Oliveira EJ, Dantas JLL, Castellen MDS, Machado MD (2008) Identification of microsatellites for papaya by DNA data bank exploration. Revista Brasileira De Fruticultura 30(3):841–845

de Oliveira EJ, Amorim VBO, Matos ELS, Costa JL, Castellen MD, Padua JG, Dantas JLL (2010a) Polymorphism of microsatellite markers in papaya (*Carica papaya* L.). Plant Mol Biol Rep 28(3):519–530

de Oliveira EJ, dos Santos Silva A, de Carvalho FM, dos Santos LF, Costa JL, de Oliveira Amorim VB, Dantas JLL (2010b) Polymorphic microsatellite marker set for *Carica papaya* L. and its use in molecular-assisted selection. Euphytica 173(2):279–287

Deputy JC, Ming R, Ma H, Liu Z, Fitch MMM, Wang M, Manshardt R, Stiles JI (2002) Molecular markers for sex determination in papaya (*Carica papaya* L.). Theor Appl Genet 106(1):107–111

Dillon S, Ramage C, Drew R, Ashmore S (2005) Genetic mapping of a PRSV-P resistance gene in "highland papaya" based on inheritance of RAF markers. Euphytica 145(1):11–23

Dillon S, Ramage C, Ashmore S, Drew RA (2006) Development of a codominant CAPS marker linked to PRSV-P resistance in highland papaya. Theor Appl Genet 113(6):1159–1169

Drew RA, O'Brien C, Magdalita PM (1998) Development of *Carica* interspecific hybrids. Acta Hortic 461:285–292

Eustice M, Yu Q, Lai CW, Hou S, Thimmapuram J, Liu L, Alam M, Moore PH, Presting GG, Ming R (2008) Development and application of microsatellite markers for genomic analysis of papaya. Tree Genet Genome 4(2):333–341

Fan M-J, Chen S, Kung Y-J, Cheng Y-H, Bau H-J, Su T-T, Yeh S-D (2009) Transgene-specific and event-specific molecular markers for characterization of transgenic papaya lines resistant to papaya ringspot virus. Transgenic Res 18(6):971–986

Fitch MMM, Manshardt RM, Gonsalves D, Slightom JL, Sanford JC (1990) Stable transformation of papaya via microprojectile bombardment. Plant Cell Rep 9(4):189–194

Gangopadhyay G, Roy SK, Ghose K, Poddar R, Bandyopadhyay T, Basu D, Mukherjee KK (2007) Sex detection of *Carica papaya* and *Cycas circinalis* in pre-flowering stage by ISSR and RAPD. Curr Sci 92(4):524–526

Hofmeyr JDJ (1938) Genetical studies of *Carica papaya* L. I. The inheritance and relation of sex and certain plant characteristics. II. Sex reversal and sex forms. South African Department of Agriculture and Science Bulletin No 187

Hofmeyr JDJ (1939) Sex-linked inheritance in *Carica Papaya* L. S Afr J Sci 36:283–285

Jobin-Decor MP, Graham GC, Henry RJ, Drew RA (1997) RAPD and isozyme analysis of genetic relationships between *Carica papaya* and wild relatives. Genet Resour Crop Evol 44(5): 471–477

Kim MS, Moore PH, Zee F, Fitch MMM, Steiger DL, Manshardt RM, Paull RE, Drew RA, Sekioka T, Ming R (2002) Genetic diversity of *Carica papaya* as revealed by AFLP markers. Genome Res 45(3):503–512

Lai CWJ, Yu Q, Hou S, Skelton RL, Jones MR, Lewis KLT, Murray J, Eustice M, Guan P, Agbayani R, Moore PH, Ming R, Presting GG (2006) Analysis of papaya BAC end sequences reveals first insights into the organization of a fruit tree genome. Mol Genet Genomics 276(1):1–12

Lemos EGM, Silva CLSP, Zaidan HA (2002) Identification of sex in *Carica papaya* L. using RAPD markers. Euphytica 127(2):179–184

Ma H, Moore PH, Liu Z, Kim MS, Yu Q, Fitch MMM, Sekioka T, Paterson AH, Ming R (2004) High-density linkage mapping revealed suppression of recombination at the sex determination locus in papaya. Genetics 166(1):419–436

Magdalita PM, Drew RA, Adkins SW, Godwin ID (1997) Morphological, molecular and cytological analyses of *Carica papaya* X *C. cauliflora* interspecific hybrids. Theor Appl Genet 95(1):224–229

Manshardt RM, Drew RA (1998) Biotechnology of papaya. Acta Hortic 461:65–74

Ming R, Moore PH, Zee F, Abbey CA, Ma H, Paterson AH (2001) Construction and characterization of a papaya BAC library as a foundation for molecular dissection of a tree-fruit genome. Theor Appl Genet 102(6):892–899

Ming R, Yu Q, Moore PH (2007) Sex determination in papaya. Semin Cell Dev Biol 18(3): 401–408

Ming R, Hou S, Feng Y, Yu Q, Dionne-Laporte A, Saw JH, Senin P, Wang W, Ly BV, Lewis KLT, Salzberg SL, Feng L, Jones MR, Skelton RL, Murray JE, Chen C, Qian W, Shen J, Du P, Eustice M, Tong E, Tang H, Lyons E, Paull RE, Michael TP, Wall K, Rice DW, Albert H, Wang M-L, Zhu YJ, Schatz M, Nagarajan N, Acob RA, Guan P, Blas A, Wai CM, Ackerman CM, Ren Y, Liu C, Wang J, Wang J, Na J-K, Shakirov EV, Haas B, Thimmapuram J, Nelson D, Wang X, Bowers JE, Gschwend AR, Delcher AL, Singh R, Suzuki JY, Tripathi S, Neupane K, Wei H, Irikura B, Paidi M, Jiang N, Zhang W, Presting G, Windsor A, Navajas-Pérez R, Torres MJ, Feltus FA, Porter B, Li Y, Burroughs AM, Luo M-C, Liu L, Christopher DA, Mount SM, Moore PH, Sugimura T, Jiang J, Schuler MA, Friedman V, Mitchell-Olds T, Shippen DE, deP-amphilis CW, Palmer JD, Freeling M, Paterson AH, Gonsalves D, Wang L, Alam M (2008) The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). Nature 452(7190):991–996

Moore GA, Litz RE (1984) Biochemical markers for *Carica papaya*, *C. cauliflora*, and plants from somatic embryos of their hybrid. J Am Soc Hortic Sci 109(2):213–218

Nagarajan N, Navajas-Pérez R, Pop M, Alam M, Ming R, Paterson A, Salzberg S (2008) Genome-wide analysis of repetitive elements in papaya. Trop Plant Biol 1(3):191–201

Niroshini E, Everard J, Karunanayake EH, Tirimanne TLS (2008) Detection of sequence characterized amplified region (SCAR) markers linked to sex expression in *Carica papaya* L. J Natl Sci Found Sri Lanka 36(2):145–150

Noorda-Nguyen K, Ruizong J, Ayumi A, Qingyi Y, Nishijima W, Yun JZ (2010) Identification of disease tolerance loci to *Phytophthora palmivora* in *Carica papaya* using molecular marker approach. Acta Hortic 851:189–196

O'Brien CM, Drew RA (2009) Potential for using *Vasconcellea parviflora* as a bridging species in intergeneric hybridisation between *V. pubescens* and *Carica papaya*. Aust J Bot 57:592–601

O'Brien C, Drew RA (2010) Marker-assisted hybridisation and backcrossing between *Vasconcellea* species and *Carica papaya* for PRSV-P resistance. Acta Hortic 859:361–368

Ocampo Pérez J, d'Eeckenbrugge C, Risterucci AM, Dambier D, Ollitrault P (2007) Papaya genetic diversity assessed with microsatellite markers in germplasm from the Caribbean region. Acta Hortic 740:93–101

Oliveira EJ, Dantas JLL, Castellen MS, Machado MD (2008) Identificac¸a˜o de microssate´lites para o mamoeiro por meio da explorac¸a˜o do banco de dados de DNA. Rev Bras Frutic 30:841–845

Parasnis AS, Ramakrishna W, Chowdari KV, Gupta VS, Ranjekar PK (1999) Microsatellite (GATA)n reveals sex-specific differences in papaya. Theor Appl Genet 99(6):1047–1052

Parasnis AS, Gupta VS, Tamhankar SA, Ranjekar PK (2000) A highly reliable sex diagnostic PCR assay for mass screening of papaya seedlings. Mol Breeding 6(3):337–344

Pérez JO, Dambier D, Ollitrault P, D'eeckenbrugge GC, Brottier P, Froelicher Y, Risterucci A-M (2006) Microsatellite markers in *Carica papaya* L.: isolation, characterization and transferability to *Vasconcellea* species. Mol Ecol Notes 6(1):212–217

Porter B, Zhu Y, Christopher D (2009a) *Carica papaya* genes regulated by *Phytophthora palmivora*: a model system for genomic studies of compatible *Phytophthora*-plant interactions. Trop Plant Biol 2(2):84–97

Porter BW, Paidi M, Ming R, Alam M, Nishijima WT, Zhu YJ (2009b) Genome-wide analysis of *Carica papaya* reveals a small NBS resistance gene family. Mol Genet Genomics 281(6): 609–626

Ratchadaporn J, Sureeporn K, Khumcha U (2007) An analysis on DNA fingerprints of thirty papaya cultivars (*Carica papaya* L.), grown in Thailand with the use of amplified fragment length polymorphisms technique. Pak J Biol Sci 10(18):3072

Santos SC, Ruggiero C, Silva CLSP, Lemos EGM (2003) A microsatellite library for *Carica papay*a L. cv. Sunrise solo. Revista Brasileira De Fruticultura 25(2):263–267

Sharon D, Hillel J, Vainstein A, Lavi U (1992) Application of DNA fingerprints for identification and genetic analysis of *Carica papaya* and other *Carica* species. Euphytica 62(2):119–126

Somsri S, Bussabakornkul S (2008) Identification of certain papaya cultivars and sex identification in papaya by DNA Amplification Fingerprinting (DAF). Acta Hortic 787:197–206

Somsri S, Fletcher RJ, Jobin M, Drew R, Lawson W, Graham MW (1998) Developing molecular markers for sex prediction in papaya (*Carica papaya* L.). Acta Hortic 461:141–148

Sondur SN, Manshardt RM, Stiles JI (1996) A genetic linkage map of papaya based on randomly amplified polymorphic DNA markers. Theor Appl Genet 93(4):547–553

Stiles JI, Lemme C, Sondur S, Morshidi MB, Manshardt R (1993) Using randomly amplified polymorphic DNA for evaluating genetic relationships among papaya cultivars. Theor Appl Genet 85(6):697–701

Storey WB (1938) Segregation of sex types in Solo papaya and their application to the selection of seed. Am Soc Hortic Sci Proc 35:83–85

Tan SC, Weinheimer EA (1976) The isoenzyme patterns of developing fruit and mature leaf of papaya (*Carica papaya* L.). Sains Malays J Nat Sci 5(1):7–14

Tennant PF, Gonsalves C, Ling KS, Fitch MMM, Manshardt RM, Slightom JL, Gonsalves D (1994) Differential protection against *Papaya Ringspot Virus* isolates in coat protein gene transgenic papaya and classically cross protected papaya. Phytopathology 1359–1366

Urasaki N, Tarora K, Uehara T, Chinen I, Terauchi R, Tokumoto M (2002a) Rapid and highly reliable sex diagnostic PCR assay for papaya (*Carica papa*ya L.). Breeding Sci 52(4):333–335

Urasaki N, Tokumoto M, Tarora K, Ban Y, Kayano T, Tanaka H, Oku H, Chinen I, Terauchi R (2002b) A male and hermaphrodite specific RAPD marker for papaya (*Carica papaya* L.). Theor Appl Genet 104(2):281–285

Van Droogenbroeck B, Breyne P, Goetghebeur P, Romeijn-Peeters E, Kyndt T, Gheysen G (2002) AFLP analysis of genetic relationships among papaya and its wild relatives (*Caricaceae*) from Ecuador. Theor Appl Genet 105(2):289–297

Wai C, Ming R, Moore P, Paull R, Yu Q (2010) Development of chromosome-specific cytogenetic markers and merging of linkage fragments in papaya. Trop Plant Biol 3(3):171–181

Wang J, Chen C, Na J-K, Yu Q, Hou S, Paull RE, Moore PH, Alam M, Ming R (2008) Genome-wide comparative analyses of microsatellites in papaya. Trop Plant Biol 1(3):278–292

Yamamoto H (1964) Comparison of the carotenoids in yellow- and red-fleshed *Carica papaya*. Nature 201:1049–1050

Yu Q, Tong E, Skelton RL, Bowers JE, Jones MR, Murray JE, Hou S, Guan P, Acob RA, Luo M-C, Moore PH, Alam M, Paterson AH, Ming R (2009) A physical map of the papaya genome with integrated genetic map and genome sequence. BMC Genomics 10(1):371

Chapter 20
Papaya Nutritional Analysis

Marisa M. Wall and Savarni Tripathi

Introduction

The ultimate goal of papaya crop improvement is to provide a delicious and appealing fruit for human nourishment. This may be achieved through genetic, biological, cultural, and environmental approaches. Papayas are sweet, flavorful, brightly colored, and uniquely rich in vitamin C and carotenoids. Papaya is a very important dietary source of vitamin A in most tropical and subtropical regions of the world and is one of the fruit crops targeted to combat vitamin A deficiencies in developing countries (WHO 2009). One papaya fruit contributes a small amount to daily caloric intake while supplying a good source of the minerals magnesium (Mg), potassium (K), boron (B), and copper (Cu) (Hardisson et al. 2001; Wall 2006; USDA 2009). Consumption of 200 g papaya flesh exceeds the adult dietary reference intake (DRI) for vitamin C and imparts 10 % of the DRI for vitamin A (IOM 2000a, 2001; Wall 2006). Papayas ranked fourth in total carotenoid content and vitamin A activity among 38 fruit types (Isabelle et al. 2010). The major papaya carotenoids are β-carotene, ζ-carotene, β-cryptoxanthin (free and esterified), and lycopene (Yamamoto 1964; Philip and Chen 1988; Mutsuga et al. 2001; Chandrika et al. 2003; Wall 2006; Gayosso-Garcia Sancho et al. 2011). Lycopene is present only in red-fleshed varieties. The antioxidant capacity of papaya carotenoids, ascorbic acid, and polyphenols may reduce the risk for several chronic diseases, including cardiovascular disease and cancer (for a review, see Chap. 21; Bramley 2000; Rao and Agarwal 2000; Bazzano et al. 2002; Riboli and Norat 2003). High fruit intake showed a protective effect against lung and bladder cancers (Riboli and Norat 2003) and was associated with lower mortality from heart disease and strokes (Bazzano et al. 2002). However, to realize the full potential and value of the nutritional and

M.M. Wall (✉) • S. Tripathi
US Pacific Basin Agricultural Research Center, US Department of Agriculture, Agricultural Research Service, 64 Nowelo Street, Hilo, HI 96720, USA
e-mail: marisa.wall@ars.usda.gov

R. Ming and P.H. Moore (eds.), *Genetics and Genomics of Papaya*, Plant Genetics and Genomics: Crops and Models 10, DOI 10.1007/978-1-4614-8087-7_20,

health-promoting properties of papayas, an interplay of factors must be optimized. Cultivar selection, site location (climate and soil type), preharvest practices, harvest maturity, postharvest treatments, and storage conditions can impact the final nutritional quality of the fruit. In the future, some of this variation may be resolved through classical plant breeding, genetic manipulation, or postharvest technologies.

Papaya Nutritional Composition

The availability of precise fruit composition data is necessary to evaluate diets for nutritional adequacy and to conduct public health studies relating diet to disease prevention. However, fruit nutritional analyses seldom consider the effects of cultivar or environment on vitamin and mineral content. Also, the nutritional composition of a papaya fruit can vary widely with maturity and ripening. For these reasons, a range of values have been reported for fresh papaya fruit (Table 20.1), but most dieticians refer to the USDA National Nutrient Database as an authoritative source (USDA 2009).

Fresh papaya fruit (100 g) supplies 179 kJ of energy, 11 g carbohydrate, 1.7 g fiber, and less than 1 g protein or fat (USDA 2009). Total sugar concentrations range from 7.5 to 11 g/100 g fresh weight (FW), with sucrose, glucose, and fructose as the primary sugars (Table 20.1) (Zhou and Paull 2001; USDA 2009; Wall et al. 2010). Ripe papaya also has significant amounts of folate and minerals (Hardisson et al. 2001; Wall 2006; Saxholt et al. 2008; USDA 2009). Papayas (100 g) provide 9–12 % of the DRI for folate, 3–7 % of the DRI for K, 6–8 % of the DRI for Mg, and 9 % of the DRI for Cu (IOM 1998, 2000b, 2001; Hardisson et al. 2001; Wall 2006).

Papaya's most important nutritional benefit comes from vitamin C (ascorbic acid), followed by the provitamin A carotenoids (β-carotene and β-cryptoxanthin). Ascorbic acid is a potent, water-soluble antioxidant present in ripe papayas in reported amounts from 35 to 142 mg/100 g FW (Table 20.1) (Wall 2006; Roberts et al. 2008; Saxholt et al. 2008; USDA 2009; Isabelle et al. 2010; Tripathi et al. 2011). The average vitamin C content for Solo-type cultivars grown in Hawaii ranged from 51 to 84 mg/100 g FW (Wenkam 1990; Franke et al. 2004; Wall 2006; Wall et al. 2010; Tripathi et al. 2011). Papaya ranked second among 44 fresh fruits and vegetables for vitamin C content, exceeding oranges, grapefruit, and broccoli (Franke et al. 2004). In another study of 38 fruits, papayas ranked fifth for ascorbic acid content (Isabelle et al. 2010). Among tropical fruits, only guava has significantly more vitamin C than papaya (Wills et al. 1986). As a rich source of vitamin C, papaya may protect against coronary heart disease, reduce age-related oxidative stress, and improve cognitive function and immune responses when added to the diet (Joshipura et al. 2001; Martin et al. 2002).

Carotenoids produce the vibrant, yellow-orange color of papayas and deliver nutritional value as vitamin A precursors and antioxidants. Papaya vitamin A content ranges from 18.7 to 88.7 μg retinol activity equivalents (RAE)/100 g FW (Table 20.1), with most reports averaging 44–65 μg RAE/100 g FW (Setiawan et al. 2001; Wall 2006; USDA 2009; Schweiggert et al. 2011).

Table 20.1 Nutritional values for papaya fruit (per 100 g edible fresh weight)

Nutrient	Range of reported values[a]	USDA nutrient database[b]
Water (g)	81.0–89.4	88.1
Energy (kJ)	123–245	179
Protein (g)	0.06–0.86	0.47
Fat (g)	0.05–0.9	0.26
Carbohydrate (g)	6.17–17.2	10.82
Dietary fiber (g)	0.5–2.7	1.7
Sugars, total (g)	7.46–10.98	7.82
Sucrose (g)	1.90–6.07	NA
Glucose (g)	2.64–3.60	4.09
Fructose (g)	2.05–3.30	3.73
Vitamin C (mg)	35.5–142.0	60.9
Vitamin A (μg RAE)	18.7–88.7	47
Folate (μg)	37–63	37
Thiamin (mg)	0.021–0.036	0.023
Riboflavin (mg)	0.024–0.058	0.027
Niacin (mg)	0.227–0.338	0.357
Pantothenic acid (mg)	0.218	0.191
Vitamin B6 (mg)	0.019	0.038
Calcium (mg)	7.8–50.9	20
Iron (mg)	0.13–0.78	0.25
Magnesium (mg)	14.0–48.8	21
Phosphorus (mg)	1.7–46.7	10
Potassium (mg)	89.7–423.0	182
Sodium (mg)	2.3–67.8	8
Zinc (mg)	0.02–0.30	0.08
Copper (mg)	0.02–0.28	0.05
Manganese (mg)	0.01–0.07	0.04

[a]Values compiled from the following references: Franke et al. (2004), Hardisson et al. (2001), Isabelle et al. (2010), Morton (1987), Roberts et al. (2008), Saxholt et al. (2008), Setiawan et al. (2001), Schweiggert et al. (2011), Tripathi et al. (2011), Vinci et al. (1995), Wall (2006), Wall et al. (2010), Wenkam (1990), and Wills et al. (1986)

[b]USDA 2009, national nutrient database for standard reference

Some of the reported variation can be attributed to differences in cultivars and fruit ripeness. It is difficult to interpret vitamin A values reported prior to 2000, because new units and DRI levels for vitamin A were set by the Food and Nutrition Board of the Institute of Medicine (IOM) that assume a 50 % lower conversion of provitamin A carotenoids into active retinol by the body (IOM 2001; Murphy 2002). As a result, vitamin A values are expressed as RAE, rather than the retinol equivalents (RE) used in earlier studies. Also, analytical methods that did not separate β-carotene from β-cryptoxanthin would lead to inaccurate and overvalued measures of vitamin A because (1) β-carotene has greater provitamin A activity than β-cryptoxanthin and (2) papayas contain high amounts of the less active β-cryptoxanthin (Philip and Chen 1988; Granado et al. 1997; Chandrika et al. 2003; Wall 2006).

In addition to provitamin A activity, papaya carotenoids support human health as potent antioxidants. Lycopene, the major pigment in red-fleshed cultivars (Yamamoto 1964; Chandrika et al. 2003), has the greatest ability to scavenge free radicals, followed by β-cryptoxanthin and β-carotene (Miller et al. 1996). In epidemiological studies, lycopene-rich diets have been inversely correlated with the incidence of heart disease and specific cancers (Bramley 2000). Lycopene levels ranged from 1.4 to 5.7 mg/100 g FW in red-fleshed fruit (Yamamoto 1964; Setiawan et al. 2001; Wall 2006; Gayosso-Garcia Sancho et al. 2011; Schweiggert et al. 2011). Accordingly, total carotenoids were highest in ripe, red-fleshed cultivars (1.2 to 6.4 mg/100 g FW) due to the abundance of lycopene, whereas total carotenoid content averaged 0.8 to 2.5 mg/100 g FW for yellow-fleshed cultivars (Philip and Chen 1988; Setiawan et al. 2001; Wall 2006; Isabelle et al. 2010; Gayosso-Garcia Sancho et al. 2011; Schweiggert et al. 2011).

Papaya total antioxidant capacity is highly correlated to the concentrations of carotenoids and vitamin C ($R^2 = 0.988–0.995$) (Gayosso-Garcia Sancho et al. 2011). Ascorbic acid contributes 96.7 % of the hydrophilic antioxidant activity of fresh papayas (Isabelle et al. 2010). The fruit also contain a moderate amount of polyphenols (45–110 mg/100 g FW) which are further sources of antioxidants (Melo et al. 2006; Isabelle et al. 2010).

Fruit Development and Nutritional Composition

Papaya is a climacteric-type fruit that exhibits a striking increase in respiration and ethylene production rates coincident with ripening. Ethylene triggers or alters many of the changes in fruit color, aroma, and softening. Hence, maturity stage and degree of ripeness can greatly influence the amount of nutritional and bioactive compounds present in papayas. Sugars accumulate in the flesh between 100 and 140 days after anthesis, in tandem with a 10-fold increase in acid invertase activity and changes in seed and flesh color (Chan et al. 1979; Zhou and Paull 2001). At the earlier stages of fruit development, glucose is the principal sugar, but sucrose content increases during ripening, ranging from 2 to 6 g/100 g FW in ripe fruit (Zhou and Paull 2001; Gomez et al. 2002; Wall et al. 2010). Total sugar content remains relatively constant during postharvest ripening, suggesting that sugar accumulation in the flesh is dependent on translocation from the plant to fruit before harvest (Zhou and Paull 2001; Bron and Jacomino 2006). However, fruit ripened with ethylene had similar total soluble solid (TSS) contents (10.5–10.8° Brix) as papayas ripened on-tree (11.1° Brix) (Schweiggert et al. 2011). Minor changes in papaya mineral content occur during ripening (Tripathi et al. 2011). Ca, P, Na, and Cu declined slightly from the color break to fully ripe stages. However, fruit mineral content is affected by soil fertility, orchard location, and production practices more than fruit maturation.

Vitamin C content can increase with maturity and ripening and decline during storage (Hernandez et al. 2006; Gayosso-Garcia Sancho et al. 2011; Tripathi et al. 2011).

At harvest, papayas with greater than 15 % yellow skin color had higher ascorbic acid (AA) content than mature green fruit (Bron and Jacomino 2006). Also, papayas accumulated 25 % more vitamin C when the fruit were ripened on-tree versus off-tree (Tripathi et al. 2011). AA content was 1.5–4 times higher in ripe papayas than mature green fruit, indicating synthesis of AA occurred during ripening (Wenkam 1990; Barata-Soares et al. 2004; Gayosso-Garcia Sancho et al. 2011). Glucose, mannose, and galactose are key precursors for AA synthesis; therefore, cell wall disassembly during ripening may provide substrates for AA synthesis (Barata-Soares et al. 2004). However, AA also degrades during ripening, and the balance between synthesis and degradation is an important aspect to AA variation.

Papaya vitamin A content is directly related to the degree of carotenoid synthesis and accumulation that occurs during fruit maturation. Carotenoid biosynthesis initiates when fruit growth is complete. About 130 days after anthesis, the peel color is green and the mesocarp is white with orange spots as carotenoid synthesis begins (Schweiggert et al. 2011). Fruit at this stage had very low amounts of β-carotene, lycopene, and β-cryptoxanthin esters; total carotenoids were 0.13 mg/100 g FW. Total carotenoid content of the flesh increased to 3.6 mg/100 g FW at the color break stage, and vitamin A content was 46.2 μg RAE/100 g FW. During postharvest ripening, total carotenoids increased further and ranged from 5.4 to 6.2 mg/100 g FW in fully ripe "Pococi" fruit, with corresponding vitamin A levels of 69.7–82.3 μg RAE/100 g FW (Schweiggert et al. 2011). There was a predominant rise in lycopene and β-cryptoxanthin esters (β-cryp-laurate and β-cryp-caprate) in the red-fleshed papaya during ripening, whereas carotenoid precursors (phytoene, phytofluene, and ζ-carotene) were detected in trace amounts in ripe fruit. Lycopene and β-cryptoxanthins approximately doubled in concentration from the color break to fully ripe stages in papayas (Schweiggert et al. 2011). For another red-fleshed cultivar ("Maradol"), fruit with less than 25 % yellow skin had total carotenoids of 0.92 mg/100 g FW that increased to 3.3 mg/100 g FW in fully ripe fruit (Gayosso-Garcia Sancho et al. 2011). β-Carotene contents doubled, β-cryptoxanthin tripled, and lycopene increased ten-fold during ripening of "Maradol" fruit (Rivera-Pastrana et al. 2010; Gayosso-Garcia Sancho et al. 2011).

β-Cryptoxanthin and its esters are the principle provitamin A carotenoids in yellow-fleshed papayas, comprising 55–75 % of the total (Yamamoto 1964; Philip and Chen 1988; Kimura et al. 1991; Chandrika et al. 2003; Wall 2006; Wall et al. 2010). Although β-cryptoxanthin has half the provitamin A activity as β-carotene, non-esterified and esterified forms of β-cryptoxanthin are absorbed by the body with comparable efficiency (Breithaupt et al. 2003). Therefore, papaya β-cryptoxanthins contribute a substantial amount of the fruit's provitamin A. Vitamin A content increased 57 % in yellow-fleshed papayas (cv. Rainbow) from the mature green to fully ripe stages (Tripathi et al. 2011). When fruits were tree ripened, vitamin A content was 64 % greater than when ripened postharvest, indicating higher levels of β-carotene and β-cryptoxanthin accumulation in papayas attached to the tree (Tripathi et al. 2011).

Environmental Effects on Nutritional Composition

Papaya nutritional composition is influenced greatly by varying environmental conditions during crop production. Temperature, rainfall, wind, light intensity, day length, relative humidity, soil fertility, pH and moisture content, and disease and pest pressure impact plant growth and development, ultimately affecting fruit nutritional quality. In general, the concentrations of sugars and vitamins in mature fruit are modulated by climate (temperature, sunlight, rainfall), whereas mineral contents are closely related to soil fertility.

The availability and transport of carbohydrates from leaves to fruit determines papaya sugar content (Zhou et al. 2000). Environmental factors that improve the plant's photosynthetic capacity, and therefore sugar metabolism, can generate sweeter fruit. One mature leaf provides carbohydrates for about three fruit, although the plant is able to adapt to partial defoliation before fruit sugar content declines (Zhou et al. 2000). Papayas can maintain high carbon assimilation rates under well-watered, high light intensity conditions (Campostrini and Glenn 2007). However, as air temperature rises from 30 to 40 °C, the leaf-to-air vapor pressure deficit increases, stomata close, and net assimilation decreases (Campostrini and Glenn 2007). Differences in air temperature also can alter time intervals for fruit maturation, leading to variability in nutritional composition. A dramatic increase in sugar concentrations occurred 120 days after anthesis in summer-harvested papayas, but lagged by 10 days in winter-harvest fruit as maturity was delayed by cooler temperatures (Chan et al. 1979). Papayas have all stages of leaf and fruit development present on a single plant. Therefore, the source-sink balance in papayas can alter fruit production and sweetness, so that plants with a high fruit load may have reduced fruit sugar contents. Conversely, fruit thinning can promote sugar accumulation in the remaining fruit on the tree (Zhou et al. 2000).

Selection of genotypes with the potential for high ascorbic acid synthesis appears more important than climactic conditions in producing high amounts of vitamin C at harvest. However, among environmental factors, light intensity and temperature are important determinants of final vitamin C content (Lee and Kader 2000). As such, seasonal variations in climate can impact vitamin C content. Papayas harvested in the summer had more vitamin C (60.4–64.5 mg/100 g FW) than winter-harvest fruits (45.3–55.6 mg/100 g FW) (Wall 2006). Longer day lengths and higher light intensities in the summer enable plants to boost photosynthesis, thereby producing more ascorbic acid and glucose, a precursor to AA in fruits (Shewfelt 1990; Lee and Kader 2000). Also, fruit exposed to sunlight have higher vitamin C than shaded fruit on the same plant (Lee and Kader 2000). In some crops, soil moisture deficits, excessive nitrogen levels, and elevated temperatures lead to lowered AA contents (Shewfelt 1990; Lee and Kader 2000).

Vitamin A content is mediated by environmental factors that affect carotenoid synthesis. Carotenoid synthesis tends to increase with nitrogen fertility, and light is necessary for formation of β-carotene (Shewfelt 1990). Chilling temperatures can reduce carotenoid levels by delaying papaya ripening or by disrupting key enzymes

in the carotenoid pathway (Rivera-Pastrana et al. 2010). Climatic factors appeared to have a greater effect on papaya vitamin A content than genotypic variation (Kimura et al. 1991; Wall 2006). In Brazil, the variety "Formosa" had the lowest and highest vitamin A levels from two different geographic regions. Fruit from Bahia, where the temperatures were warmest, had higher β-carotene, β-cryptoxanthin, and lycopene concentrations than fruit from cooler São Paulo (Kimura et al. 1991). For Hawaii-grown papayas, the mean vitamin A content for 60 samples was 44.1 μg RAE/100 g FW and ranged from 18.7 to 74.0 μg RAE/100 g FW, depending on cultivar, location, and season (Wall 2006). The cultivar "Sunrise" had β-carotene levels ranging from 80 to 410 μg/100 g FW when grown in multiple locations, indicating that environmental variation contributed to a wide range in carotenoid concentrations (Wall 2006). Most likely, variable orchard temperatures can lead to slight differences in papaya maturity, thereby affecting final vitamin A levels.

Large variation exists in the mineral content of papayas among production regions (Hardisson et al. 2001). For example, Na appears highest in papayas from coastal orchards, and K and Mg tend to be greatest in fruit grown on tropical islands with volcanic soils (Hardisson et al. 2001; Wall 2006). Papaya Na content varied according to water supply for the production area. Fruit from sea-level orchards irrigated with well water had more Na than fruit from locations where rain was the sole water source (Wenkam 1990). Other minerals (Ca, Fe, Cu) may be more related to crop management practices, such as fertilization and disease control practices. In Hawaii, papayas harvested from Maui tended to have high contents of P, K, Ca, Fe, and Cu, but fruit mineral composition was not clearly related to mineral analyses of the alluvial, silty clay soil (Wall 2006).

Conceivably there are multiple interactions among preharvest environmental conditions, genetics, and physiology that determine papaya nutritional composition at harvest, although to date, research on this topic is rare. In contrast, the effects of postharvest environments and processing conditions on nutrient content are more clearly elucidated, but beyond the scope of this chapter. Among preharvest decisions, selecting a cultivar with the genetic potential for high nutrient content and choosing a production location with a favorable climate are essential to maximize the nutritional composition of papayas.

Genotypic Variation in Nutritional Composition

Papaya is a member of the Caricaceae family that includes 6 genera and 35 species. *Carica papaya* is the only species in the *Carica* genus. Genetic studies of papaya (*C. papaya*) and *Vasconcellea*, the largest genus within Caricaceae, indicate that papaya diverged early from the rest of the family (Aradhya et al. 1999; Badillo 2000; Van Droogenbroeck et al. 2004). Although papaya germplasm exhibits significant phenotypic variation for horticultural traits (including fruit size, sweetness, flesh color, and flavor), the genetic diversity within *C. papaya* is quite narrow (Kim et al. 2002). The average genetic similarity within 63 papaya accessions was

estimated as 0.880. Intergeneric hybridization with *Vasconcellea* has been challenging for introducing genetic diversity to papaya due to genome incompatibility. The use of bridge species and embryo rescue has been used to overcome these barriers (Manshardt and Wenslaff 1989; O'Brien and Drew 2010). However, genetic transformation is also a promising, but untested, alternative for improving the nutritional content of papayas.

Genetic variation in nutrient quality, though limited, does occur within the *C. papaya* species, as well as within individual cultivars or accessions. Most commonly, papaya cultivars have been developed for production characteristics (yield, fruit set and size, disease resistance) with little regard for fruit nutritional content. As a consequence, unexploited variation for nutritional quality remains and may be drawn upon through breeding programs. However, studies that distinguish genetic from environmental variation are rare, as are reports of deliberate selection within papaya germplasm for phytonutrients (except for carotenoids determining flesh color). Therefore, nutritional quality does not seem to be a high priority in papaya breeding programs.

Variation among cultivars for vitamin C content ranges from 45 to 142 mg/100 g FW in fully ripe papayas (Melo et al. 2006; Wall 2006). The highest concentrations of AA were reported for Solo type varieties (Franke et al. 2004; Melo et al. 2006). Vitamin C content (per 100 g FW fully ripe fruit) was 142 mg for cultivar "Hawaii," 100 mg for "Golden," 47–122 mg for "Sunrise," and 47–85 mg for "Rainbow" (Bron and Jacomino 2006; Melo et al. 2006; Wall 2006; Roberts et al. 2008; Tripathi et al. 2011). In contrast, cultivar "Formosa" had 64 mg AA/100 g FW (Melo et al. 2006). The Solo-type varieties also tend to have higher sugar contents than the large-fruited cultivars, and ascorbic acid and sugar contents are positively correlated. Therefore, selection for sweet fruit with high vitamin C content is possible.

Within *C. papaya*, variation in provitamin A pigments ranged from 146 to 371 μg β-carotene and 426 to 1,034 μg β-cryptoxanthin per 100 g FW in yellow-fleshed cultivars ("Kapoho," "Laie Gold," "Rainbow"). For red-fleshed papayas ("Formosa," "Golden," "Maradol," "Sunrise," "SunUp," "Tailandia"), β-carotene varied from 49 to 810 μg and β-cryptoxanthin from 288 to 1,260 μg/100 g FW (Kimura et al. 1991; Wall 2006; Rivera-Pastrana et al. 2010; Barreto et al. 2011). The red-fleshed varieties contain lycopene, which accounts for higher total carotenoids than in yellowed-fleshed types (Chandrika et al. 2003). Lycopene is the central compound in the carotenoid biosynthetic pathway and is the precursor for cyclization reactions leading to β-carotene, β-cryptoxanthin, and other xanthophylls (Barreto et al. 2011). In red-fleshed fruit, lycopene comprises 56–66 % of the total in "Formosa," "Solo," and "Tailandia" cultivars, with values ranging from 1.5 to 4.7 mg/100 g FW (Kimura et al. 1991). Lycopene varied from 1.4 to 3.7 mg/100 g FW in "Sunrise" fruit (Wall 2006). Ripe "Maradol" and "Golden" fruit contained 1.2 and 1.3 mg lycopene/100 g FW, respectively (Rivera-Pastrana et al. 2010; Barreto et al. 2011). Among the red-fleshed genotypes, "Tailandia" is distinctive, with a vivid red color and high lycopene concentration (Kimura et al. 1991). Also, total carotenoids of fully ripe,

red-fleshed papayas varied from 5.2 to 6.7 mg for "Tailandia" and from 5.4 to 6.2 mg/100 g FW for "Pococi" (Kimura et al. 1991; Schweiggert et al. 2011). Thus, these two cultivars are promising germplasm to include in a papaya breeding program aimed at increasing fruit nutritional value.

Future Prospects to Enhance Papaya Nutritional Content

The primary sources of variation for nutritional plant products are genetic, preharvest environment, maturity at harvest, and postharvest conditions. Thus at this time, the best prospects for enhancing papaya nutritional content are to maximize the genetic potential within the existing *C. papaya* germplasm through optimal varietal selection, orchard edaphic factors, crop management, harvest maturity, and postharvest treatments and storage temperatures. Nevertheless, major advances in papaya genetic tools provide an avenue for further germplasm improvement. Combining classical breeding approaches with marker-assisted selection appears to be a practical and efficient strategy for papayas. A backcross program linked with microsatellite marker selection was demonstrated for introgression of desirable genes into an elite recurrent parent (Ramos et al. 2011). A comprehensive high-density genetic linkage map for papaya is available and useful for mapping quantitative trait loci and for marker-assisted selection (Blas et al. 2010). In addition, transgenic techniques have been successful, as proven with the "Rainbow" papaya, genetically modified for resistance to papaya ringspot virus (PRSV). The transgenic process alone (for a single gene trait) does not alter papaya nutrient quality. Genetically modified (GM) papayas with resistance to PRSV had similar nutritional composition as non-GM fruit (Mutsuga et al. 2001; Roberts et al. 2008; Jiao et al. 2010; Tripathi et al. 2011).

In recent years, the significant progress made in papaya genomics, gene identification, and transcript characterization will strengthen our understanding of how these factors affect nutritional quantity and quality. Papaya genomic resources include the draft genome sequence of "SunUp" and an integrated physical and genetic map (Ming et al. 2008; Yu et al. 2009). The availability of this genomic information enables future in-depth studies directed at the nutritional aspect of papaya fruit. Differentially expressed gene transcripts associated with fruit ripening have been described, including several related to vitamin C, biotin and folate biosynthesis, and precursors for carotenoids (Fabi et al. 2010). Genes involved in papaya carotenoid biosynthesis have been identified and characterized (Skelton et al. 2006; Blas et al. 2010; Devitt et al. 2010; Yan et al. 2011). Papaya flesh color is controlled by a single genetic locus, with yellow dominant over red (Devitt et al. 2010). In yellow-fleshed cultivars, lycopene is rapidly converted into β-carotene and xanthophylls by the action of lycopene β-cyclases and carotene hydrolases. However, the conversion of lycopene in red-fleshed papayas is either lacking or less efficient, thus lycopene accumulates in these fruit (Yan et al. 2011). A papaya

chromoplast-specific lycopene β-cyclase gene was cloned from red-fleshed papaya and found responsible for fruit flesh color in papaya (Blas et al. 2010; Devitt et al. 2010). A sequence analysis of 15 wild and cultivated papaya accessions revealed a two base-pair insertion in the coding region of the recessive red-fleshed allele, resulting in a frameshift mutation and a premature stop codon. The two base-pair insertion (tt) was present in all red-fleshed genotypes and absent in all yellow-fleshed genotypes tested (Blas et al. 2010; Devitt et al. 2010). This genomic and genetic information is most obviously useful for molecular breeding to improve the nutritional quality of papayas. Also, potential modification of carotenoid biosynthesis may be possible through direct genetic manipulation of key genes of the pathway or through altering a gene's regulation and its expression in fruits; for a more complete discussion on genomics of fruit development, see Chap. 14.

Thus, the genetic characterization of ripening and color development genes has potential applications to enhance papaya nutrition through traditional breeding or transgenic methods. For example, a traditional breeding approach to break the observed linkage between the red-fleshed trait and short shelf life would include screening large, segregating populations for red-fleshed and firm texture. A PCR-based screening tool for the codominant marker, CPFC1 (*C. papaya* flesh color 1), could hasten the marker-assisted selection of fruit flesh color within populations (Blas et al. 2010). Also, despite the fact that papayas reach their maximum vitamin content when fully ripe, the fruit are harvested at earlier maturity stages to extend postharvest life. One strategy for increasing the vitamin and antioxidant levels in papayas is to combine the lycopene β-cyclase gene from red-fleshed fruit with regulatory transcripts to delay fruit softening. This would allow fruit to ripen longer on the tree, accumulating more sugars, carotenoids, and vitamins before harvest. However, there are significant challenges to inhibiting fruit softening. Gene transcripts for softening-related enzymes have been identified, but are ethylene dependent. Genetic techniques that inhibit ethylene or disrupt polygalacturonase and endoxylanase will delay softening, but the fruit will have an undesirable, rubbery texture (Manenoi and Paull 2007; Fabi et al. 2009). Other untested possibilities include increasing vitamin C levels through over expression of a gene encoding for D-galacturonic acid reductase, as has been proposed for strawberries (Agius et al. 2003), or increasing sugar content by controlling gene expression for sucrose synthase and acid invertase with fruit-specific promoters (Zhou and Paull 2001). Also, it may be possible to enhance papaya antioxidant capacity via regulation of genes encoding flavonoid biosynthesis enzymes, similar to reports for transgenic tomatoes (Shukla and Mattoo 2009).

Many nutritionists and consumers are aware of the dietary and aesthetic value of papayas. Further increases in nutritional content are possible if the germplasm base can be diversified, either through traditional or transgenic methods, including through introgression of genes from different plant sources. However, many nutritional compounds are secondary compounds and attempts to regulate key genes in biosynthetic pathways could lead to unintended consequences of lower yield, aroma, flavor, or visual quality.

References

Agius F, Gonzalez-Lamothe R, Caballero JL, Munoz-Blanco J, Botella MA, Valpuesta V (2003) Engineering increased vitamin C levels in plants by overexpression of a D-galacturonic acid reductase. Nature Biotech 21:177–181

Aradhya MK, Manshardt RM, Zee F, Morden CW (1999) A phylogenetic analysis of the genus *Carica* L. (Caricaceae) based on restriction fragment length variation in a cpDNA intergenic spacer region. Genet Resour Crop Evol 46:579–586

Badillo VM (2000) *Carica* L. vs. *Vasconcella* St. Hill. (Caricaceae) con la rehabilitacion de este ultimo. Ernstia 10:74–79

Barata-Soares AD, Gomez ML, de Mesquita CH, Lajolo FM (2004) Ascorbic acid biosynthesis: a precursor study on plants. Braz J Plant Physiol 16:147–154

Barreto G, Fabi JP, De Rosso VV, Cordenunsi BR, Lajolo FM, Nascimento J, Mercadante AZ (2011) Influence of ethylene on carotenoid biosynthesis during papaya postharvest ripening. J Food Comp Anal 24:620–624

Bazzano LA, He J, Ogden LG, Loria CM, Vupputuri S, Myers L, Whelton PK (2002) Fruit and vegetable intake and risk of cardiovascular disease in US adults: the first National Health and Nutrition Examination Survey Epidemiologic follow-up study. Am J Clin Nutr 76: 93–99

Blas AL, Ming R, Liu Z, Veatch OJ, Paull RE, Moore PH, Yu Q (2010) Cloning of the papaya chromoplast-specific lycopene β-cyclase, *CpCYC-b*, controlling fruit flesh color reveals conserved microsynteny and a recombination hot spot. Plant Physiol 152:2012–2022

Bramley PM (2000) Is lycopene beneficial to human health? Phytochemistry 54:233–236

Breithaupt D, Weller P, Wolters M, Hahn A (2003) Plasma response to a single dose of dietary esters from papaya (*Carica papaya* L.) or non-esterified β-cryptoxanthin in adult human subjects: a comparative study. Br J Nutr 90:795–801

Bron IU, Jacomino AP (2006) Ripening and quality of "Golden" papaya fruit harvested at different maturity stages. Braz J Plant Physiol 18:389–396

Campostrini E, Glenn DM (2007) Ecophysiology of papaya: a review. Braz J Plant Physiol 19:413–424

Chan HT, Hibbard KL, Goo T, Akamine EK (1979) Sugar composition of papayas during fruit development. HortScience 14:140–141

Chandrika UG, Jansz ER, Wickramasinghe S, Warnasuriya ND (2003) Carotenoids in yellow- and red-fleshed papaya (*Carica papaya* L.). J Sci Food Agric 83:1279–1282

Devitt LC, Fanning K, Dietzgen RG, Holton TA (2010) Isolation and functional characterization of a lycopene β-cyclase gene that controls fruit colour of papaya (*Carica papaya* L.). J Exp Bot 61:33–39

Fabi JP, Cordenunsi BR, Seymour GB, Lajolo FM, Oliveira do Nascimento JR (2009) Molecular cloning and characterization of a ripening-induced polygalacturonase related to papaya fruit softening. Plant Physiol Biochem 47:1075–1081

Fabi JP, Mendes LR, Lajolo FM, Nascimento J (2010) Transcript profiling of papaya fruit reveals differentially expressed genes associated with fruit ripening. Plant Sci 179:225–233

Franke AA, Custer LJ, Arakaki C, Murphy SP (2004) Vitamin C and flavonoid levels of fruits and vegetables consumed in Hawaii. J Food Comp Anal 17:1–35

Gayosso-Garcia Sancho LE, Yahia EM, Gonzalez-Aguilar GA (2011) Identification and quantification of phenols, carotenoids, and vitamin C from papaya (*Carica papaya* L., cv. Maradol) fruit determined by HPLC-DAD-MS/MS-ESI. Food Res Int 44:1284–1291

Gomez M, Lajolo B, Cordenunsi B (2002) Evolution of soluble sugars during ripening of papaya fruit and its relation to sweet taste. J Food Sci 67:442–447

Granado F, Olmedilla B, Blanco I, Gil-Martinez E, Rojas-Hidalgo E (1997) Variability in the intercomparison of food carotenoid content data: a user's point of view. Crit Rev Food Sci Nutr 37:621–633

Hardisson A, Rubio C, Baez A, Martin MM, Alvarez R (2001) Mineral composition of the papaya (*Carica papaya* variety Sunrise) from Tenerife Island. Eur Food Res Technol 212:175–181
Hernandez Y, Lobo MG, Gonzalez M (2006) Determination of vitamin C in tropical fruits: a comparative evaluation of methods. Food Chem 96:654–664
Institute of Medicine (IOM) (1998) Dietary reference intakes for thiamin, riboflavin, niacin, vitamin B6, folate, vitamin B12, pantothenic acid, biotin, and choline. National Academy Press, Washington
Institute of Medicine (IOM) (2000a) Dietary reference intakes for vitamin C, vitamin E, selenium, and carotenoids. National Academy Press, Washington
Institute of Medicine (IOM) (2000b) Dietary reference intakes for calcium, phosphorus, magnesium, vitamin D, and fluoride. National Academy Press, Washington
Institute of Medicine (IOM) (2001) Dietary reference intakes for vitamin A, vitamin K, arsenic, boron, chromium, copper, iodine, iron, manganese, molybdenum, nickel, silicon, vanadium, and zinc. National Academy Press, Washington
Isabelle M, Lee BL, Lim MT, Koh WP, Huang D, Ong CN (2010) Antioxidant activity and profiles of common fruits in Singapore. Food Chem 123:77–84
Jiao Z, Deng J, Li G, Zhang Z, Cai Z (2010) Study on the compositional differences between transgenic and non-transgenic papaya (*Carica papaya* L.). J Food Comp Anal 23: 640–647
Joshipura KJ, Hu FB, Manson JE, Stampfer MJ, Rimm EB, Speizer FE, Colditz G, Ascherio A, Rosner B, Spiegelman D, Willett WC (2001) The effect of fruit and vegetable intake on risk for coronary heart disease. Ann Intern Med 134:1106–1114
Kim MS, Moore PH, Zee F, Fitch MM, Steiger DL, Manshardt RM, Paull RE, Drew RA, Sekioka T, Ming R (2002) Genetic diversity of *Carica papaya* as revealed by AFLP markers. Genome 45:503–512
Kimura M, Rodriguez-Amaya DB, Yokoyama SM (1991) Cultivar differences and geographic effects on the carotenoid composition and vitamin A value of papaya. Leben Wissen Technol 25:415–418
Lee SK, Kader AA (2000) Preharvest and postharvest factors influencing vitamin C content of horticultural crops. Postharvest Biol Technol 20:207–220
Manenoi A, Paull RE (2007) Papaya fruit softening, endoxylanase gene expression, protein and activity. Physiol Plant 131:470–480
Manshardt RM, Wenslaff TF (1989) Interspecific hybridization of papaya with other species. J Am Soc Hortic Sci 114:689–694
Martin A, Cherubin A, Andres-Lacueva C, Paniagua M, Joseph J (2002) Effects of fruits and vegetables on levels of vitamins E and C in the brain and their association with cognitive performance. J Nutr Health Aging 6:392–404
Melo E, de Lima VL, Maciel MI (2006) Polyphenol, ascorbic acid and total carotenoid contents in common fruits and vegetables. Braz J Food Technol 9:89–94
Miller NJ, Sampson J, Candeias LP, Bramley PM, Rice-Evans CA (1996) Antioxidant activities of carotenes and xanthophylls. FEBS Lett 384:240–242
Ming R, Hou S, Feng Y, Yu Q, Dionne-Laporte A, Saw JH, Senin P, Wang W, Ly BV, Lewis KL et al (2008) The draft genome of the transgenic tropical tree fruit papaya (*Carica papaya* Linnaeas). Nature 452:991–996
Morton J (1987) Papaya. In: Morton J (ed) Fruits of warm climates. Florida Flair Books, Miami. http://www.hort.purdue.edu/newcrop/morton/papaya_ars.html. Accessed 1 Jun 2011
Murphy SP (2002) Dietary reference intakes for the U.S. and Canada: update on implications for nutrient databases. J Food Comp Anal 15:411–417
Mutsuga M, Ohta H, Toyoda M, Goda Y (2001) Comparison of carotenoid components between GM and non-GM papaya. J Food Hyg Soc Jpn 42:367–373
O'Brien C, Drew RA (2010) Marker-assisted hybridization and backcrossing between *Vasconcellea* species and *Carica papaya* for PRSV-P resistance. Acta Hortic 859:361–368
Philip T, Chen T (1988) Quantitative analyses of major carotenoid fatty acid esters in fruits by liquid chromatography: persimmon and papaya. J Food Sci 53:1720–1722

Ramos HCC, Pereira MG, Silva FF, Goncalves LSA, Pinto FO, de Souza Filho GA, Pereira TSN (2011) Genetic characterization of papaya plants (*Carica papaya* L.) derived from the first backcross generation. Genet Mol Res 10:393–403

Rao AV, Agarwal S (2000) Role of antioxidant lycopene in cancer and heart disease. J Amer Coll Nutr 19:563–569

Riboli E, Norat T (2003) Epidemiologic evidence of the protective effect of fruit and vegetables on cancer risk. Am J Clin Nutr 78:559–569

Rivera-Pastrana DM, Yahia EM, Gonzalez-Aquilar GA (2010) Phenolic and carotenoid profiles of papaya fruit (*Carica papaya* L.) and their contents under low temperature storage. J Sci Food Agric 90:2358–2365

Roberts M, Minott DA, Tennant PF, Jackson JC (2008) Assessment of compositional changes during ripening of transgenic papaya modified for protection against papaya ringspot virus. J Sci Food Agric 88:1911–1920

Saxholt E, Christensen AT, Moller A, Hartkopp HB, Hess Ygil K, Hels OH (2008) Danish food composition databank, revision 7. Department of Nutrition, National Food Institute, Technical University of Denmark. http://www.foodcomp.dk/. Accessed 17 Aug 2011

Schweiggert RM, Steingass CB, Mora E, Esquivel P, Carle R (2011) Carotenogenesis and physico-chemical characteristics during maturation of red fleshed papaya fruit (*Carica papaya* L.). Food Res Int 44:1373–1380

Setiawan B, Sulaeman A, Giraud DW, Driskell JA (2001) Carotenoid content of selected Indonesian fruits. J Food Comp Anal 14:169–176

Shewfelt RL (1990) Sources of variation in the nutrient content of agricultural commodities from the farm to the consumer. J Food Qual 13:37–54

Shukla V, Mattoo AK (2009) Potential for engineering horticultural crops with high antioxidant capacity. CAB Rev 066:1–22

Skelton RL, Yu Q, Srinivasan R, Manshardt R, Moore PH, Ming R (2006) Tissue differential expression of lycopene β-cyclase gene in papaya. Cell Res 16:731–739

Tripathi S, Suzuki JY, Carr JB, McQuate GT, Ferreira SA, Manshardt RM, Pitz KY, Wall MM, Gonsalves D (2011) Nutritional composition of Rainbow papaya, the first commercialized transgenic fruit crop. J Food Comp Anal 24:140–147

U.S. Department of Agriculture (USDA), Agricultural Research Service (2009) National nutrient database for standard reference, release 23. http://www.nal.usda.gov/fnic/foodcomp/search/. Accessed 17 Aug 2011

Van Droogenbroeck B, Kyndt T, Maertens I, Romeijn-Peeters E, Scheldeman X, Romero-Motochi JP, Van Damme P, Goetghebeur P, Gheysen G (2004) Theor Appl Genet 108:1473–1486

Vinci G, Botre F, Mele G, Ruggieri G (1995) Ascorbic acid in exotic fruits: a liquid chromatographic investigation. Food Chem 53:211–214

Wall MM (2006) Ascorbic acid, vitamin A and mineral composition of banana (*Musa* sp.) and papaya (*Carica papaya*) cultivars grown in Hawaii. J Food Comp Anal 19:434–445

Wall MM, Nishijima KA, Fitch MM, Nishijima WT (2010) Physicochemical, nutritional and microbial quality of fresh-cut and frozen papaya prepared from cultivars with varying resistance to internal yellowing disease. J Food Qual 33:131–149

Wenkam NS (1990) Food of Hawaii and the Pacific Basin, fruits and fruit products: raw, processed, and prepared. vol 4: composition. Hawaii Agric Expt Sta Res Ext Ser 110: 96 pp

Wills R, Lim J, Greenfield H (1986) Composition of Australian foods. 31. tropical and sub-tropical fruit. Food Technol Aust 38:118–123

World Health Organization (WHO) (2009) Global prevalence of vitamin A deficiency in populations at risk 1995-2005. WHO global database on vitamin A deficiency. http://www.who.int/nutrition/publications/micronutrients/vitamin_a_deficiency/9789241598019/en/index.html. Accessed 28 Nov 2011

Yamamoto HY (1964) Comparison of the carotenoids in yellow- and red-fleshed *Carica papaya*. Nature 201:1049–1050

Yan P, Gao XZ, Shen WT, Zhou P (2011) Cloning and expression analysis of phytoene desaturase and ζ-carotene desaturase genes in *Carica papaya*. Mol Biol Rep 38:785–791

Yu Q, Tong E, Skelton RL, Bowers JE, Jones MR, Murray JE, Hou S, Guan P, Acob RA, Luo MC, Moore PH, Alam M, Paterson AH, Ming R (2009) A physical map of the papaya genome with integrated genetic map and genome sequence. BMC Genomics 10:371–382

Zhou L, Paull RE (2001) Sucrose metabolism during papaya (*Carica papaya*) fruit growth and ripening. J Am Soc Hortic Sci 126:351–357

Zhou L, Christopher DA, Paull RE (2000) Defoliation and fruit removal effects on papaya fruit production, sugar accumulation, and sucrose metabolism. J Am Soc Hortic Sci 125:644–652

Chapter 21
Papaya as a Medicinal Plant

Timothy J. O'Hare and David J. Williams

Introduction

Papaya has been used for medicinal purposes for many hundreds of years in tropical and subtropical regions where it is grown, particularly central and South America, southern and Southeast Asia, as well as sub-Saharan Africa. The reasons for use are varied, with perhaps the best known being a cure for intestinal worms, as a topical application to accelerate wound repair or skin damage and as a potential abortion agent. Other ailments that have been reported to be ameliorated with papaya include diabetes, dengue fever, malaria, amoebic dysentery, cancer, and hypertension. The plant tissue used also varies, although the immature fruit, latex, seeds, and leaves are the major parts of the plant used medicinally, and they may be applied topically, ingested, or injected, depending on the ailment (Krishna et al. 2008).

In the last 20 years, there has been increasing investigation into the efficacy of "folk" remedies, providing insights into their effectiveness as a cure, the physiological mechanism by which they may be effective, and, in some circumstances, the likely compound(s) responsible for the physiological action. This chapter discusses these recent investigations.

T.J. O'Hare (✉)
QAAFI, Centre for Nutrition and Food Sciences, University of Queensland,
Locked Mailbag 7, Mail Service 437, Gatton, QLD 4343, Australia
e-mail: t.ohare@uq.edu.au

D.J. Williams
Agri-Science Queensland, Department of Agriculture, Fisheries and Forestry,
Brisbane, QLD 4108, Australia

R. Ming and P.H. Moore (eds.), *Genetics and Genomics of Papaya*, Plant Genetics and Genomics: Crops and Models 10, DOI 10.1007/978-1-4614-8087-7_21,

Anthelmintic Activity

Papaya seeds have been used for centuries as a vermifuge (a medicine that destroys parasitic worms) in many countries (Werner 1992), although early clinical trials with humans led to contradictory results. Robinson (1958) claimed that papaya seeds were effective, but Fernando (1959) stated they were not. Later laboratory studies confirmed that various preparations of papaya tissues can kill helminths effectively in vitro and in infected animals (Dar et al. 1965; Lal et al. 1976).

The anthelmintic effect of papaya seeds has been variously ascribed to the alkaloid carpaine, benzyl thiourea (Panse and Paranjpe 1943) and to benzyl isothiocyanate (BITC) (Tang 1971). There was considerable debate about the level of anthelmintic activity of benzyl thiourea as it was believed to be an artefact that arose from a reaction between ammonia and BITC during purification of the bioactive principle (Ettlinger and Hodgkins 1956). Subsequently, Dar et al. (1965) tested BITC and benzyl thiourea individually for bioactivity and showed BITC to be about 20 times more toxic than benzyl thiourea to *Ascaris lumbricoides*.

Many years later, further studies tested the efficacy of various purified isothiocyanates against plant parasitic nematodes (Lazzeri et al. 1993; Zasada and Ferris 2003). The latter researchers reported that BITC isolated from papaya seeds and leaves was the most effective of the seven isothiocyanates tested against two major plant nematodes.

Kermanshai et al. (2001) found that BITC was the major and probably the sole source of anthelmintic activity against *Caenorhabditis elegans*. While *C. elegans* is not a human parasite, it is considered to be a good model with similarities to parasitic nematodes.

Okeniyi et al. (2007) evaluated the anthelmintic effects of dried papaya seeds in 60 asymptomatic Nigerian children with intestinal parasites. They reported a clearance rate for the various types of parasites of between 70 and 100 % following treatment. The authors did not elucidate the mechanism of action or identify the active components but confirmed that the papaya seed extract clearly had anthelmintic activity.

Latex collected from young papaya fruit has also been shown to have anthelmintic activity against *Ascaridia galli* infections in chickens (Mursof and He 1991). Experiments by Satrija et al. (1995) demonstrated a marked reduction of parasite burdens in mice receiving papaya latex, but latex had no effect against the tapeworm *Hymenolepis diminuta* in rats and the trematode *Echinostoma caproni* in mice. From this, the authors concluded that latex treatment may not be effective against all types of helminths.

It has been suggested that the anthelmintic activity of papaya latex against *A. galli* may be due to the activity of proteolytic enzymes of the cysteine catalytic class (Mursof and He 1991). It is well known that papaya latex contains several proteolytic enzymes such as papain, chymopapain, and lysozyme (Winarno 1983). Papaya latex has a particularly high content of papain (Reed and Underkofler 1966; Winarno 1983), and in vitro studies have shown that adult *A. galli* are completely dissolved

following incubation in papaya latex for a period of less than 24 h (Purwati and He 1991). Likewise, incubation of *A. galli* eggs in papaya latex for 21 days destroyed the outer structure of the eggs and reduced their infectivity (Purwati and He 1991).

Although encouraging results for the anthelmintic efficacy of papaya latex were provided by these studies, the mechanism was not elucidated. The work conducted by Stepek et al. (2005) seemed to provide clear evidence of the activity of proteolytic enzymes, where the use of purified papain and ficin increased the efficacy against the rodent gastrointestinal nematode *Heligmosomoides polygyrus*. There is ample evidence that latex contains several other proven anthelmintic agents such as BITC as well as these proteolytic enzymes (Bennett et al. 1997; O'Hare et al. 2008), so the exact mechanism may be more complex than first envisaged.

Further studies are needed to elucidate the active ingredients associated with anthelmintic activity. Success in this would permit a concentrated latex extract as an alternative treatment against nematodes of both humans and livestock, especially as these compounds would provide a cheap and easily obtainable source of anthelmintic activity and would be environmentally safe. The risk of resistance developing would presumably be low as it would require the nematodes to alter their protective coating or to encode cysteine proteinase inhibitors for insertion into this coating.

Wound Healing

One of the better-known areas of medicinal use of papaya is in wound healing. Papaya pulp is currently traditionally used in a number of countries to treat burns and for wound healing (Starley et al. 1999; Hewitt et al. 2000; Pieper and Caliri 2003). Starley et al. (1999) described papaya to be effective in desloughing necrotic tissue, preventing burn wound infection, and providing a granulating wound suitable for skin grafting. It was proposed that the active papaya component responsible for activity may be the proteolytic enzymes, chymopapain and papain. Anuar et al. (2008) reported an aqueous extract from green papaya skin, in contrast to ripe papaya skin, accelerated the repair of wounds in mice. It was found that the green papaya skin was higher in proteins with a 28–29 kDa molecular weight and with a band containing 23–25 kDa proteins that was absent in ripe papaya skin. Silva et al. (2003) earlier isolated a protein fraction from a related *Carica* species (*Carica candamarcensis*), which was shown to stimulate cell proliferation in some mammalian cell types. The active protein fraction was found to be in the 10–25 kDa range, with a protein of 23 kDa molecular weight suggested as a potential candidate. Interestingly, papain has a molecular weight of 23 kDa (Dreuth et al. 1968) and chymopapain a molecular weight of 27 kDa (Watson et al. 1990). Although it is claimed that chymopapain has greater proteolytic activity than papain, the latter corresponds more closely with the weight of the *Carica* protein fraction promoting cell proliferation. It has previously been demonstrated that the main latex components from papaya leaves are chymopapains, with very little papain present

(Mezhlumyan et al. 2003). Although chymopapains are also the predominant proteases in papaya fruit, there is significantly more papain present in the fruit latex than in latex from vegetative tissues (Robinson 1975). One may therefore expect that the efficacy of papaya latex on cell proliferation would be affected by the plant tissue from which the latex was extracted, with unripe fruit being more effective than leaf extracts.

Papaya extracts are also used to remove necrotic tissue. The proteolytic enzymes of papaya have been shown to cleave non-viable denatured proteins while leaving living tissue unaffected (Buttle and Barret 1984; Gostishchev et al. 1999; Mezhlumyan et al. 2003). It is uncertain which proteases are most active, but the ability to remove injured tissue but not healthy tissue is a useful attribute. Medicinal preparations based on papaya proteinases have been used to remove damaged tissues from the ocular sclera and spinal cord (Mezhlumyan et al. 2003). Low concentrations of papaya proteinases were observed to have a strong effect on eye tissues. Papain has also been used as an intradisc injection to treat lumbar osteochondrosis patients to remove damaged tissue, with ~80 % positive results being reported (Mezhlumyan et al. 2003). More recently, Dang et al. (2007) reported that the use of chymopapain enhanced the mechanical removal of nucleus pulposis from intervertebral discs and was more effective than using rongeurs alone.

Antifertility Properties

Impact on Females

Plant preparations have been used extensively in folk medicine over many years to facilitate good menstrual flow as well as being used for their abortifacient properties (Quisumbing 1951; Saha et al. 1961). Prominent on this list of plants are the green fruit, latex and seeds of the papaya. In stark contrast to its abortifacient properties, Eno et al. (2000) reported the use of a papaya fruit extract by traditional African healers to prevent miscarriage. While these antifertility properties are founded more on popular belief based on experiences passed on by local healers, their general efficacy is hard to prove through scientific investigation, especially when it involves humans (Saha et al. 1961).

To evaluate the abortifacient properties of papaya extracts, extensive research was conducted in the 1960s and 1970s using animal models. In these studies, various formulations from papaya, such as extracts of seeds, pulp of unripe fruit, latex and even purified papain, were administered to rats. These studies suggested that papaya latex or unripe fruit extracts caused problems at implantation (Garg 1974), increased post-implantation embryo loss (Garg and Garg 1970; Gopalakrishnan and Rajasekharasetty 1978) and led to embryotoxicity (Devi and Singh 1978). However, Schmidt (1995) performed experiments that re-examined the findings of these studies and administered standardised papain in doses up to 800 mg/kg from day 0

to day 15 of gestation. Schmidt's results showed clearly that the administration of purified papain significantly decreased the implantation loss, did not adversely affect prenatal development, and exhibited no signs of maternal toxicity. The author noted that the previous studies used plant-derived extracts, a much more complex mixture of compounds than the purified papain formulations that were used in the latter study. The Schmidt results were in good agreement with two teratogenicity studies (teratogenic agents cause malformations in an embryo or foetus), performed at the request of the Food and Drug Administration, in which papain was administered orally to mice and rats during gestation. Papain had no effect on maternal or foetal survival or any teratogenic effect (Food and Drug Research Laboratories 1974).

Adebiyi et al. (2002) evaluated the safety of ripe papaya fruit extracts on pregnant and non-pregnant rats. The ripe fruit or juice did not provoke any significant uterine stimulation that could be harmful to the foetus or cause abortion. The authors concluded that normal human consumption of ripe papaya during pregnancy does not pose any significant danger. Conversely, crude papaya latex induced spasmodic contraction of the isolated uterine muscles of pregnant and non-pregnant rats comparable to low doses of the abortion drug, prostaglandin. These results are not surprising since fully ripe papaya fruit contains very little or negligible quantity of latex, but unfortunately, the authors did not offer any insight into the mechanism of action or the nature of the active compounds contained in the latex.

Chinoy et al. (1995) investigated the effects of aqueous papaya seed extracts on the metabolism of the ovary and uterus of normal rats. Changes in the biochemical profile, irregular oestrus cycles, and increased uterine contractibility all indicated that the extract manifested antifertility and abortifacient effects in the treated animals. A later study (Adebiyi et al. 2003) reported that ethanolic papaya seed extract was toxic to rat uterine tissues and resulted in spontaneous contraction. The authors suggested BITC could be the seed constituent responsible for the spontaneous contraction. Further evidence for this was provided by similar degenerative changes of uterine tissues treated with purified BITC when compared to those treated with the seed extract. It should be noted that to achieve these spontaneous uterine contractions, large doses of purified BITC (100 μM) were needed.

While evaluation of potentially toxic agents often depends on animal bioassays to predict risk in humans, additional experimental approaches may be necessary to ascertain the antifertility properties of compounds contained in papaya tissues. Inconsistencies in the methods of application and/or extraction have caused the variable results obtained so far, and there is clearly a need for a more standardised approach.

Impact on Males

Papaya seed extracts have been advocated in many countries' folklore as an antifertility agent for males (Tiwari et al. 1982; Kirtikar and Basu 1998). In recent times, there

has been significant scientific evaluation, using animal models, of the extent of this antifertility effect and for the identification of the responsible compounds.

Chinoy and Ranga (1984) noted that papaya seed extracts changed the biochemical parameters and contractile pattern, in adult male albino rats, of the vas deferens (the tube that connects the testes with the urethra). The data showed that even short-term administration manifested as an androgen deprivation (blocked the production or action of male sex hormones) in the target organs, thereby causing an antifertility effect.

A series of studies in the 1990s investigated the effect of chloroform- and water extracts of papaya seeds fed to adult male rats on sperm motility. Lohiya and Goyal (1994) reported that within 40–60 days of treatment, chloroform extracts caused suppression of sperm motility, thereby reducing fertility to zero. The same authors could induce reversible sterility in rats using aqueous seed extracts without adverse effects on libido and toxicological profile (Lohiya et al. 1994). These findings in conjunction with their previous data suggested to Chinoy et al. (1994) that papaya seed extract could be an effective male contraceptive for rodents. Even an aqueous extract of papaya bark showed contraceptive activity in male rats (Kusemiju et al. 2002). In contrast to these observations, aqueous extracts of papaya seeds failed to exhibit contraceptive effects, at any dose, when tested on male rabbits (Lohiya et al. 2000). These authors suggested that the dissimilarities in contraceptive ability may be due to species specificity, relative resistance, or factors that generally affect biological activity. Support for this last supposition came from an earlier report (Lohiya et al. 1999) where the chloroform extract was found to decrease the sperm concentration in male adult rabbits. It also affected sperm motility, attaining less than 1 % after 75 days. Recently, high doses of aqueous papaya leaf extracts produced toxic effects on the seminiferous tubule epithelium of male rats, leading to sharp decreases in the reproductive potential of these animals (Morayo and Akinloye 2010). Lohiya et al. (2002) showed that the chloroform seed extract had good contraceptive efficacy in langur monkeys. The authors felt this action was mediated through the inhibition of sperm motility.

The success of these animal trials in showing an antifertility effect has led to the evaluation of papaya seed extract as a potential safe male contraceptive (Lohiya et al. 2000).

Identifying the compounds responsible for this contraceptive effect has been more difficult, although alkaloids, steroids, triterpenoids, or flavonoids have all been suggested (Lohiya et al. 2000, 2002; Pathak et al. 2000).

Antifungal Activity

The antifungal action of components contained in papaya tissue extracts is well described in folk medicine literature (Kirtikar and Basu 1998), and according to these sources, papaya roots and stem bark are the most commonly used tissues for this purpose. However, on several of the Pacific Islands, latex from green fruit has

been recommended for the use against ringworm (Singh 1981). Most scientific investigations have concentrated on determining the active constituents involved and their mode of action.

Hine et al. (1965) observed that mature papaya fruit was susceptible to disease when inoculated with the fungal pathogen *Phytophthora parasitica*, whereas immature fruit seemed resistant. They suggested that any contribution to this resistance by papain was minor. Gas chromatographic analysis of ripening papaya fruit showed that the level of BITC decreased markedly with maturity (Tang 1971). Direct evidence that BITC was implicated in the antifungal activity was provided when it was found that purified BITC inhibited the germination of *P. parasitica* spores (Tang 1971). A later study by the same author found BITC was lethal to *P. palmivora*, an important papaya fungal pathogen, at 15 ppm (Tang 1973).

Giordani et al. (1991, 1996, 1997) attributed the deleterious effect of papaya latex on *Candida albicans* growth to glycosidases contained in the latex. They suggested these compounds, especially α-D-mannosidase and *N*-acetyl-β-D-glucosaminidase, caused degradation of the fungal cell wall. In a series of experiments, Giordani et al. (1997) reported a synergistic action against in vitro cultured *C. albicans* of a mixture of papaya latex and the antifungal agent fluconazole. This combined action allowed a significant decrease in the therapeutic dose of fluconazole, a marked benefit, as this drug has been implicated in several adverse side effects (van Cauteren et al. 1989; Guillaume et al. 1996).

Another antifungal group of proteins, the chitinases, have been demonstrated to be useful in agriculture for controlling plant fungal pathogens (Broglie et al. 1991). Chitinases are also found in abundance in the papaya latex (Chen et al. 2007). A purified recombinant papaya chitinase has recently been shown to completely inhibit spore germination of the Brassica fungal pathogen, *Alternaria brassicicola* (Chen et al. 2007). The authors also proposed to express papaya chitinase in crop plants to produce this valuable antifungal compound via a biofactory.

Antimalarial Activity

Reports of antimalarial activity of papaya are mostly anecdotal with little scientific evidence being published. Spencer et al. (1947) found little or no activity of chloroform and water extracts of papaya flowers on *Plasmodium gallinaceum* in chickens. It took many years before other scientific investigations indicated that papaya may have significant antimalarial activity. Bhat and Surolia (2001) tested aqueous and organic solvent extracts from three commonly utilised plants on the malaria strain *Plasmodium falciparum* FCK2. The rind of unripe papaya fruit showed the highest antimalarial activity of the plant extracts tested. These researchers indicated that efforts were under way to isolate the active principles and to characterise their toxicities. To date, no further results have been published.

Recent reports (Sannella et al. 2009; Pietretti et al. 2010) indicated the efficacy of dried aqueous extracts of papaya leaf on mice infested with the malaria parasites

P. falciparum or *Plasmodium berghei*. Both studies showed that the extracts had plasmodial activity and reduced parasitemia—the presence of parasites in the blood without clinical symptoms.

Even though papaya extracts have been used for many years by traditional healers to treat malaria, no clinical use appears to have been made of these bioactive compounds and there are no reports of double-blind clinical trials having been carried out to assess their antimalarial efficacy.

Antimicrobial Activity

The antibacterial effects of different tissues of papaya are well reported in folk medicine literature (Kirtikar and Basu 1998). In contrast to most therapeutic cures attributed to papaya, there are numerous scientific investigations published to explain this antimicrobial activity.

George and Pandalai (1949) investigated the antimicrobial activity of ethanol extracts of several papaya tissues. Extracts of dried root, seeds and unripe fruit were active against *Escherichia coli* and *Staphylococcus aureus*, but the dried leaf extracts and the latex were not. These findings were confirmed by dos Fernades Viera et al. (2001) and Nkuo-Akenji et al. (2001) when both groups reported that water, acetone and ethanol extracts of papaya leaves showed no microbicidal activity. Interestingly, Nkuo-Akenji et al. (2001) noted that formulations that combined papaya leaf and root extract did have antibacterial activity against *Salmonella typhi*, *Salmonella paratyphi*, and *Salmonella typhimurium*. Nickell (1959) reviewed the antimicrobial activity of vascular plant extracts including papaya plant material showing activity against gram-positive bacteria and mycobacteria.

BITC, a compound produced by the enzymatic hydrolysis of benzyl glucosinolate by myrosinase in papaya seeds, showed significant antimicrobial activity (Virtanen 1965; Tang 1973). The Virtanen (1965) study showed that BITC had the greatest antimicrobial activity of all the purified isothiocyanates tested.

Purified extracts from papaya seeds of both ripe and unripe fruit produced significant antibacterial activity on the gram-positive bacteria *S. aureus* and *Bacillus cereus* but less so on the gram-negative bacteria tested (Emeruwa 1982). The isolation of the active constituent was only partially successful in that the active parts of the seed were identified, but not the active components, although the data indicated that it was a protein.

The seed and pulp of unripe papaya fruit were shown by the agar cup plate method to be bacteriostatic against enteropathogens such as *Bacillus subtilis*, *E. coli*, *S. typhi*, *S. aureus*, *Proteus vulgaris*, *Pseudomonas aeruginosa*, and *Klebsiella pneumonia* (Osato et al. 1993). These authors suggested that BITC may be responsible for the antimicrobial activity of these tissues.

Recently, Bennett et al. (1997) found that papaya leaves and stems contained cyanogenic glycosides. This finding was surprising because plants that contain glucosinolates do not normally contain cyanogenic glycosides (Bennett et al. 1997).

These compounds offer a potential source of highly toxic cyanide. In a review of the pharmacological properties of papaya, de Oliveira and Vitoria (2011) suggested that these cyanogenic glycosides are responsible for the antibiotic properties of papaya. It would be surprising, however, if they were the sole active agent, as several studies have shown that these cyanide-producing glycosides, when present in the tissues of papaya, were at extremely low levels (Olafsdottir et al. 2002; Seigler et al. 2002). There is the possibility that in extracting the active constituent, concentration of these compounds could occur.

Bacterial resistance to antimicrobials is common. Horizontal gene transfer is widely recognised as the bacterial mechanism responsible for the widespread distribution of antibiotic resistance, with conjugation suggested to be a major pathway (Leite et al. 2005). These authors investigated the effect of papaya seed macerate on conjugal transfer from *S. typhimurium* to *E. coli* in vitro in the digestive tract of mice. The macerate caused a significant reduction of the transconjugant population, with no lethal effect of the macerate on donor or recipients in the concentrations used. The authors felt that this indicated a possible practical use for these extracts to prevent further transfer of antibiotic-resistant genes between bacteria that are targets of antibiotic chemotherapy. Furthermore, they suggested that diet may be an important factor influencing gene exchange among enteric bacteria, opening a natural route to control or minimise the dissemination of resistance genes.

Antiviral Activity (Dengue Fever)

Viral diseases such as dengue fever result in a low thrombocyte count in the blood and require a rapid response with thrombocyte and fresh-frozen plasma transfusion (World Health Organisation 1997). Crude formulations of papaya extracts have been employed for many years in folk medicine, notably in Malaysia, for the treatment of dengue infections with haemorrhagic manifestations. The popular vehicle for this treatment is a suspension of powdered leaves in palm oil taken orally (Sathasivam et al. 2009). This traditional use gave rise to the recent investigation on the effect of papaya leaf material on the thrombocyte count in an animal model (Sathasivam et al. 2009). Thrombocyte counts measured before and after treatment of mice revealed significantly higher mean counts after dosing with the papaya leaf formulation. The authors suggested that these findings call for a dose–response investigation and for extended studies into the identification and isolation of the responsible substances.

Ahmad et al. (2011) investigated the potential of papaya leaf extracts against dengue fever in a 45-year-old patient bitten by carrier mosquitoes. Aqueous extracts were administered orally, and the platelet, white blood cell, and neutrophil counts of the patient's blood were monitored. All three counts increased with treatment, with the platelet count increasing threefold, suggesting that papaya leaves have substantial activity against dengue virus. However, the authors emphasised that this was a preliminary study and much more research on isolating the active compounds was needed before papaya could be advocated as a treatment against this viral disease.

Anticancer and Chemoprotection

Few direct studies appear to have been made on the application of papaya extracts to combat cancer. Anecdotal evidence exists of patients with advanced cancers achieving remission following consumption of a papaya leaf extract. Otsuki et al. (2010) tested the effect of an aqueous extract of papaya leaf on a number of tumour cell lines and on the anti-tumour action of human lymphocytes. The authors reported antiproliferative effects on tumour cells, promotion of Th1-type cytokine production, enhancement of cytotoxicity against tumour cells, and upregulation of immunomodulatory genes, in particular those encoding the family of monocyte-chemoattractant proteins. The latter are considered to be important in the recruitment of monocytes and memory T lymphocytes to sites of inflammation. Although the active constituents of the leaf extract were not fully identified, the active components were isolated to the fraction of molecular weight less than 1,000.

Previous papaya leaf compounds reported to potentially have anti-tumour activity include α-tocopherol (Ching and Mohamed 2001), lycopene (van Breemen and Pajkovic 2008), flavonoids (Miean and Mohamed 2001), and BITC (Basu and Haldar 2008). Although the first three compounds may have potential activity, their concentration in papaya leaves is not particularly high, and lycopene is predominantly found in the pulp of ripe, red papaya rather than leaves (Schweiggert et al. 2011). BITC, or rather its precursor prior to hydrolysis, glucotropaeolin, is abundant in most papaya shoot tissues (O'Hare et al. 2008). Both compounds have a molecular weight less than 1,000 (149 and 447, respectively).

Purified BITC has been shown to induce apoptosis (cell death) in a range of cancer cells (Basu and Haldar 2008). The action appears to be through the ability of BITC to cause phosphorylation of the anti-apoptotic protein, Bc1-xL. The authors suggested that both the death receptor-mediated extrinsic pathway and the mitochondrial intrinsic pathway combined to trigger BITC-mediated cell death of cancer cells. Nakamura et al. (2007) have further shown an *n*-hexane extract of papaya seed to be highly effective in inducing apoptosis in HL-60 cells. The activity of the extract, which contained BITC, was comparable to authentic BITC at similar concentrations.

In addition to studies inducing cell death of cancer cells, considerable research has been conducted on the ability of BITC to inhibit the bioactivation of pre-carcinogens, such as nitrosamines. This is controlled by the mammalian phase 1 cytochrome P450 enzyme, CYP2E1 (Chung 1992). Interestingly, BITC has been shown to destroy CYPE1 during metabolism (Moreno et al. 1999), in contrast to other isothiocyanates such as sulforaphane, which act by inhibiting the enzyme. Destroying the enzyme is thought to potentially be a more effective mode of action in vivo, even at very low concentration.

In addition to destroying phase 1 enzymes, BITC is also an inducer of mammalian detoxification phase 2 enzymes (Zhang and Talalay 1994). These enzymes act by inducing the production of the body's own defence system, principally through the upregulation of glutathione transferase synthesis (Prestera et al. 1993). In addition, isothiocyanates, such as BITC, are rapidly conjugated to glutathione, exported into

the systemic circulation, and metabolised through the mercapturic acid pathway (Traka and Mithen 2009). Within the low glutathione environment of the plasma, the BITC-glutathione conjugate may be cleaved by *GSTM1*, leading to circulation of free BITC in the plasma. Interestingly, if BITC behaves similarly to cruciferous isothiocyanates, humans with a functional *GSTM1* gene would maintain higher levels of plasma BITC and consequently have greater benefit from this defence system (Joseph et al. 2004).

Anti-amoebic Activity

Traditional preparations from medicinal plants are used extensively throughout the world for the treatment of amoebiasis (Kambu 1990). This is the case for *Carica papaya*, where the leaves and seeds are used frequently, particularly in Africa, for this purpose. Again there is little scientific evidence published that can support or disprove these claims.

Two groups of researchers (Lozoya et al. 1994; Morales et al. 1994) investigating the anti-diarrhoeic effect of various plant extracts unanimously reported that papaya leaf extracts were effective. They both ascribed this property to quercetin and its glycosides, quercetrin and rutin.

To date, the definitive investigation of anti-amoebic activity of papaya extracts was reported by Tona et al. (1998). These researchers tested the in vitro anti-amoebic activity of plant extracts used as anti-diarrhoeic in traditional African medicine. Aqueous extracts from mature papaya seeds had one of the highest activities, although it was significantly lower than the reference anti-amoebic drug, metronidazole. It is interesting to note that quercetin or its glycosides were not detected in the analysis of the crude papaya seed extracts. This would suggest either other compounds were the active constituents or the thin layer chromatography procedures employed were inadequate to detect this flavonoid or its derivatives. However, the authors felt these empirical results constituted enough scientific evidence to support the continued use of papaya seed extracts for the treatment of amoebiasis.

Diabetes

A number of recent studies have investigated the use of papaya extracts to reduce blood glucose and blood lipids, specifically as a potential management tool of diabetes, although the active compound(s) has not been determined. Egwim (2005) reported that administration of a crude ethanolic extract from unripe papaya fruit three times daily for 7 days reduced plasma glucose as well as cholesterol and total lipids in guinea pigs and rats. In a similar study, Banerjee et al. (2006) reported unripe papaya juice and aqueous and methanolic extracts of unripe papaya fruit to significantly lower serum cholesterol and triglycerides in rats. Similarly, Oloyede

(2008) reported aqueous extracts of unripe papaya fruit administered once daily for 4 weeks reduced activity of alanine aminotransferase (ALT) in normal and diabetic rabbits. ALT generally increases in diabetic patients, being a specific marker for liver fat accumulation (Tiikkainen et al. 2003), and can be used to predict the onset of type 2 diabetes (Sattar et al. 2004). Adeneye and Olagunju (2009) further reported aqueous seed extracts of papaya to be capable of lowering blood glucose, serum triglycerides, and low-density lipoprotein cholesterol while raising high-density lipoprotein cholesterol levels in rats. By contrast, an earlier study (Chinoy et al. 1997) reported ethanolic seed extracts of papaya to have no influence on serum cholesterol, ALT, or liver function. Interactions with other hypoglycaemic agents (glimepiride and metformin) have also been reported, with an ethanolic leaf extract of papaya being reported to increase the hypoglycaemic effect of metformin, but delaying the onset of activity by glimepiride (Fakeye et al. 2007).

Anti-hypertension

Papaya has been reported as a popular treatment of hypertension in Brazil, apparently acting as a direct vasodilator (Consolini and Ragone 2008). Eno et al. (2000) administered a crude ethanol extract from unripe papaya fruit to rats and reported a significant reduction in arterial blood pressure, 28 % lower than hydralazine, a known anti-hypertension agent and relaxant of vascular muscle. The papaya extract was also found to cause relaxation of in vitro rabbit vascular muscle. Gupta et al. (1990) had earlier reported alcoholic leaf extracts also to induce central muscle relaxation in rats. Eno et al. (2000) concluded that unripe papaya juice is likely to contain an anti-hypertensive agent which exhibits alpha-adrenoreceptor activity. Runnie et al. (2004) further reported papaya leaf extracts to induce relaxation in isolated aortic muscle rings from rats. They reported that the effect could be inhibited by the nitric oxide synthase inhibitor, *N*-omega-nitro-L-arginine, and hence papaya-induced muscle relaxation could be controlled through the action of nitric oxide. Loh and Hadira (2011) recently observed that papaya extracted in dichloromethane was also a strong inhibitor of angiotensin I-converting enzyme (ACE), the enzyme which regulates conversion of angiotensin I to angiotensin II, a potent vasoconstrictor (Zhang et al. 2000). The anti-hypertensive compounds in papaya extract remains to be identified.

Future Directions

It is clear that the medicinal use of papaya is multivaried and that many of the uses appear to be supported by recent scientific studies. Although there has been increasing research into the medicinal use of papaya over the last 20 years, in many cases the compound responsible for activity still remains unknown or unproven. In some

actions, such as wound healing, it would appear likely that papain is the most likely candidate identified to date, whereas BITC is more likely to be the controlling agent for intestinal worms. This may explain why unripe pulp, high in papain, is used to treat tissue damage, while seeds, high in BITC, are more likely to be used to treat intestinal worms.

Increasing the knowledge of the active compounds responsible for curative action will help to transfer the use of papaya from folklore remedies to mainstream medicinal use. This research will provide much needed information for the commercial extraction of these compounds, the development of synthetic analogues, or, alternatively, their production through bacteria incorporating papaya synthesis genes.

References

Adebiyi A, Ganesan AP, Prasad AN (2002) Papaya (*Carica papaya*) consumption is unsafe in pregnancy: fact or fable? Scientific evaluation of a common belief in some parts of Asia using a rat model. Br J Nutr 88:199–203

Adebiyi A, Ganesan AP, Prasad AN (2003) Tocolytic and toxic activity of papaya seed extract on isolated rat uterus. Life Sci 74:581–592

Adeneye AA, Olagunju JA (2009) Preliminary hypoglycemic and hypolipidemic activities of the aqueous seed extract of *Carica papaya* Linn in Wistar rats. Biol Med 1:1–10

Ahmad N, Fazal H, Ayaz M et al (2011) Dengue fever treatment with *Carica papaya* leaves extract. Asian Pac J Trop Biomed 1:330–333

Anuar NS, Zahari SS, Taib IA et al (2008) Effect of green and ripe *Carica papaya* epicarp extracts on wound healing and during pregnancy. Food Chem Toxicol 46:2384–2389

Banerjee A, Vaghasiya R, Shrivastava N et al (2006) Anti-hyperlipidemic effect of *Carica papaya* L. in Sprague Dawley rats. Nigerian J Nat Prod Med 10:69–72

Basu A, Haldar S (2008) Dietary isothiocyanate mediated apoptosis of human cancer cells is associated with Bc1-xL phosphorylation. Int J Oncol 33:657–663

Bennett RN, Kiddle G, Wallsgrove RM (1997) Biosynthesis of benzylglucosinolate, cyanogenic glucosides and phenylpropanoids in *Carica papaya*. Phytochemistry 45:59–66

Bhat PG, Surolia N (2001) In vitro anti-malarial activity of extracts of three plants used in the traditional medicine of India. Am J Trop Med Hyg 65:304–308

Broglie K, Chet I, Holliday M et al (1991) Transgenic plants with enhanced resistance to the fungal pathogen *Rhizoctonia solani*. Science 254:1194–1197

Buttle DJ, Barret AJ (1984) Chymopapain. Chromatographic purification and immunological characterization. Biochem J 223:81–88

Chen YT, Hsu LH, Huang IP et al (2007) Gene cloning and characterization of a novel recombinant antifungal chitinase from papaya (*Carica papaya*). J Agric Food Chem 55:714–722

Ching LS, Mohamed S (2001) Alpha-tocopherol content in 62 edible tropical plants. J Agric Food Chem 49:3101–3105

Chinoy NJ, Ranga GM (1984) Effects of *Carica papaya* seed extracts on the physiology of the *vas deferens* of albino rats. Acta Eur Fertil 15:59–65

Chinoy NJ, D'Souza JM, Padman P (1994) Effects of crude aqueous extract of *Carica papaya* seeds in male albino mice. Reprod Toxicol 8:75–79

Chinoy NJ, Dilip T, Harsha J (1995) Effect of *Carica papaya* seed extract on female rat ovaries and uteri. Phytother Res 9:169–175

Chinoy NJ, Harsha J, Shilpa G et al (1997) Toxicity related response of female albino rats treated with benzene and alcoholic papaya seed extracts. Indian J Environ Toxicol 7:62–64

Chung FL (1992) Chemoprevention of lung carcinogenesis by aromatic isothiocyanates. In: Wattenberg LW (ed) Cancer chemoprevention. CRC, Boca Raton, pp 227–245
Consolini AE, Ragone MI (2008) Cardiovascular effects of some medicinal plants from the South American regions of the Rio de La Plata basin and Patagonia, Handbook of ethnopharmacology. Research Signpost, Trivandrum, pp 233–260
Dang L, Wardlaw D, Hukins DW (2007) Removal of nucleus pulposus from the intervertebral disc – the use of chymopapain enhances mechanical removal with rongeurs: a laboratory study. BMC Musculoskelet Disord 8:122
Dar RN, Garg LC, Pathak RD (1965) Anthelmintic activity of *Carica papaya* seeds. Indian J Pharm 27:335–336
de Oliveira JG, Vitoria AP (2011) Papaya: nutritional and pharmacological characterisation and quality loss due to physiological disorders. An overview. Food Res Int 44:1306–1313
Devi S, Singh S (1978) Changes in the placenta of rat foetuses induced by maternal administration of papain. Indian J Exp Biol 16:1256–1260
dos Fernades Vieira RHS, dos Prazeres Rodrigues D, Goncalves FA et al (2001) Microbial effect of medicinal plant extracts (*Psidium guajava* and *Carica papaya*) upon bacteria isolated from fish muscle and known to induce diarrhoea in children. Rev Inst Med Trop S Paulo 43:145–148
Dreuth J, Jansonius J, Koekoek R et al (1968) Structure of papain. Nature 218:929–932
Egwim E (2005) Hypoglycemic potencies of crude ethanolic extracts of cashew roots and unripe pawpaw fruits in guinea pigs and rats. J Herb Pharmacother 5:27–34
Emeruwa AC (1982) Antibacterial substance from *Carica papaya* extract. J Nat Prod 45:123–127
Eno AE, Owo OI, Itam EH et al (2000) Blood pressure depression by the fruit juice of *Carica papaya* in renal and DOCA-induced hypertension in the rat. Phytother Res 14:235–239
Ettlinger MG, Hodgkins JE (1956) The mustard oil of papaya seed. J Org Chem 21:204–205
Fakeye TO, Oladipupo TO, Showande O et al (2007) Effects of coadministration of extract of *Carica papaya* Linn (family Caricaceae) on activity of two oral hypoglycemic agents. Trop J Pharm Res 6:671–678
Fernando PVD (1959) Preliminary investigation of *Carica papaya* seeds as a vermifuge. Indian J Child Health 8:96–100
Food and Drug Research Laboratories (1974) Teratological evaluation of papain. Waverly, New York
Garg SK (1974) Antifertility effect of oil from indigenous plants on female albino rats. Placenta Medica 26:391–393
Garg SK, Garg GP (1970) Antifertility screening of plants. Part vii. Effect of five indigenous plants on early pregnancy in albino rats. Indian J Med Res 59:302–306
George M, Pandalai KM (1949) Investigations on plant antibiotics. Part iv. Further search for antibiotic substances in Indian medicinal plants. Indian J Med Res 37:169–181
Giordani R, Siepaio M, Moulin-Traffort J et al (1991) Antifungal action of *Carica papaya* latex: isolation of fungal cell wall hydrolysing enzymes. Mycoses 34:469–477
Giordani R, Cardenas ML, Moulin-Traffort J et al (1996) Fungicidal activity of latex sap from *Carica papaya* and anti-fungal effect of D(+)-glucosamine on *Candida albicans* growth. Mycoses 39:103–110
Giordani R, Gachon C, Moulin-Traffort J et al (1997) A synergistic effect of *Carica papaya* latex sap and fluconazole on *Candida albicans* growth. Mycoses 40:429–437
Gopalakrishnan M, Rajasekharasetty MR (1978) Effect of papaya (*Carica papaya*) on pregnancy and oestrous cycle in albino rats of Wistar strain. Indian J Physiol Pharmacol 22:66–70
Gostishchev VK, Afanas'ev AN, Khokhlov AM (1999) Surgical treatment of diabetic osteoarthropathy, complicated by pyogenic process in the feet. Khirurgiya 40–44 (in Russian)
Guillaume MP, De Prez C, Cogan E (1996) Subacute mitochondrial liver disease in a patient with AIDS: possible relationship to prolong fluconazole administration. Am J Gastroenterol 91:165–168
Gupta A, Wambebe CO, Parsons DL (1990) Central and cardiovascular effects of the alcoholic extract of the leaves of *Carica papaya*. Int J Crude Drug Res 28:257–266

Hewitt HS, Whittle S, Lopez S et al (2000) Topical use of papaya in chronic skin ulcer therapy in Jamaica. West Indian Med J 49:32–33
Hine RB, Aragaki M, Tokunaga J (1965) Enzymatic inactivation of the papaya blight fungus *Phytophthora parasitica* by papain and other proteolytic enzymes. Phytopathology 55: 1223–1226
Joseph MA, Moysich KB, Freudenheim JL et al (2004) Cruciferous vegetables, genetic polymorphisms in glutathione-S-transferases M1 and T1, and prostate cancer risk. Nutr Cancer 50:206–213
Kambu K (1990) Elements de Phytotherapie Comparee. Plantes Medicinales Africaines. CRP, Kinshasa
Kermanshai R, McCarry BE, Rosenfeld J et al (2001) Benzyl isothiocyanate is the chief or sole anthelmintic in papaya seed extracts. Phytochemistry 57:427–435
Kirtikar KR, Basu BD (1998) Indian medicinal plants, 2nd edn. International Book Distributors, Dehra Dun, vol II, pp 1097–1099
Krishna KL, Paridhavi M, Patel JA (2008) Review on nutritional, medicinal and pharmacological properties of papaya (*Carica papaya*). Nat Prod Radiance 7:364–373
Kusemiju O, Noronha C, Okanlawon A (2002) The effect of crude extract of the bark of *Carica papaya* on the seminiferous tubules of male Sprague–Dawley rats. Niger Postgrad Med J 9: 205–209
Lal J, Chandra S, Raviprakash V et al (1976) In vitro anthelmintic action of some indigenous medicinal plants on *Ascaridia galli* worms. Indian J Physiol Pharmacol 20:64–68
Lazzeri L, Tacconi R, Palmieri S (1993) In vitro activity of some glucosinolates and their reaction products toward a population of the nematode *Heterodera schachtii*. J Agric Food Chem 41:825–829
Leite AA, Nardi RM, Nicoli JR et al (2005) *Carica papaya* seed macerate as inhibitor of conjugative R plasmid transfer from *Salmonella typhimurium* to *Escherichia coli in vitro* and in the digestive tract of genotobiotic mice. J Gen Appl Microbiol 51:21–26
Loh SP, Hadira O (2011) In vitro inhibitory potential of selected Malaysian plants against key enzymes involved in hyperglycemia and hypertension. Malays J Nutr 17:77–86
Lohiya NK, Goyal RB (1994) Antifertility investigations on the crude chloroform extract of *Carica papaya* seeds in male albino rats. Indian J Exp Biol 30:1051–1055
Lohiya NK, Goyal RB, Jayaprakash D et al (1994) Antifertility effect of aqueous extract of *Carica papaya* seeds in male rats. Planta Med 60:400–404
Lohiya NK, Mishra PK, Pathak N et al (1999) Reversible contraception with chloroform extract of *Carica papaya* seeds in male rabbits. Reprod Toxicol 13:59–66
Lohiya NK, Pathak N, Mishra PK et al (2000) Contraceptive evaluation and toxicological study of aqueous extract of the seeds of *Carica papaya* in male rabbits. J Ethnopharmacol 70:17–27
Lohiya NK, Mishra PK, Pathak N et al (2002) Chloroform extract of *Carica papaya* seed induced long term reversible azoospermia in langur monkeys. Asian J Androl 4:17–26
Lozoya X, Meckes M, Abou Zaid M et al (1994) Quercetin glycosides in *Psidium guajava* leaves and determination of spasmolytic principle. Arch Med Res 25:11–15
Mezhlumyan LG, Kasymova TD, Yuldashev PK (2003) Proteinases from *Carica papaya* latex. Chem Nat Comp 39:223–228
Miean KH, Mohamed S (2001) Flavonoid (myricetin, quercetin, kaempferol, luteolin, and apigenin) content of edible tropical plants. J Agric Food Chem 49:3106–3112
Morales MA, Tortoriello J, Meckes M et al (1994) Calcium antagonist effect of quercetin and its relation with spasmolytic properties of *Psidium guajava*. Arch Med Res 25:17–21
Morayo O, Akinloye O (2010) Evaluation of andrological indices and testicular histology following chronic administration of aqueous extract of *Carica papaya* leaf in wistar rat. Afr J Pharm Pharmacol 7:325–333
Moreno RL, Kent UM, Hodge K et al (1999) Inactivation of cytochrome P450 2E1 by benzyl isothiocyanate. Chem Res Toxicol 12:582–587
Mursof EP, He S (1991) A potential role of papaya latex as an anthelmintic against patent *Ascaridia galli* infection in chicken. Hemera Zoa 74:11–20

Nakamura Y, Yoshimoto M, Murata Y et al (2007) Papaya seed represents a rich source of biologically active isothiocyanate. J Agric Food Chem 55:4407–4413
Nickell LG (1959) Antimicrobial activity of vascular plants. Econ Bot 13:281–318
Nkuo-Akenji T, Ndip R, McThomas A et al (2001) Anti-salmonella activity of medicinal plants from Cameroon. Cent Afr J Med 47:155–158
O'Hare TJ, Wong LS, Williams DJ et al (2008) Papaw (*Carica papaya*) as a source of glucotropaeolin and its active derivative, benzyl-isothiocyanate. In: Proceedings of the tropical fruits in human nutrition and health Conference, Gold Coast, 8–11 November, 2008, pp 197–201
Okeniyi JAO, Ogunlesi TA, Oyelami OA et al (2007) Effectiveness of dried *Carica papaya* seeds against human intestinal parasitosis: a pilot study. J Med Food 10:194–196
Olafsdottir ES, Jorgensen LB, Jaroszewski JW (2002) Cyanogenesis in glucosinolate-producing plants: *Carica papaya* and *Carica quercifolia*. Phytochemistry 60:269–273
Oloyede OI (2008) Effect of aqueous extract from unripe pulp of *Carica papaya* on transaminase activities in selected rabbit tissue of normal and alloxan induced diabetic rabbit. Biomed Pharmacol J 1:305–310
Osato JA, Santiago LA, Remo GM et al (1993) Antimicrobial and antioxidant activities of unripe papaya. Life Sci 53:1383–1389
Otsuki N, Dang NH, Kumagai E et al (2010) Aqueous extract of *Carica papaya* leaves exhibits anti-tumor activity and immunomodulatory effects. J Ethnopharmacol 127:760–767
Panse TB, Paranjpe AS (1943) Isolation of carpasemine from papaya seeds. Proc Indian Acad Sci 18A:140
Pathak N, Mishra PK, Manivannan B et al (2000) Sterility due to inhibition of sperm motility by oral administration of benzene chromatographic fraction of the chloroform extract of the seeds of *Carica papaya* in rats. Phytomedicine 7:325–333
Pieper B, Caliri MHL (2003) Non-traditional wound care: a review of the evidence for the use of sugar, papaya/papain, and fatty acids. J Wound Ostomy Continence Nurs 30:175–183
Pietretti A, Karioiti A, Sannella A et al (2010) Antiplasmodial in vivo activity of *Carica papaya* leaf decoction. Planta Med 76:P450. doi:10.1055/s-0030-1264748
Prestera T, Holtzclaw WD, Zhang Y et al (1993) Chemical and molecular regulation of enzymes that detoxify carcinogens. Proc Natl Acad Sci USA 90:2965–2969
Purwati E, He S (1991) Pengaruh getah papaya (*Carica papaya*) terhadap infektivitas telur *Ascaridia galli* pada ayam. Hemera Zoa 74:1–5
Quisumbing EA (1951) Medicinal plants of the Philippines. Dept Agric Nat Res Tech Bull 16, Philippines: Manila Bureau of Printing
Reed G, Underkofler LA (1966) Enzymes in food processing. Academic, London
Robinson P (1958) Seeds of *Carica papaya* for mass treatment against *Ascariasis*. Indian J Child Health 7:815–817
Robinson GW (1975) Isolation and characterization of papaya peptidase A from commercial chymopapain. Biochemistry 14:3695–3700
Runnie I, Salleh MN, Mohamed RJ et al (2004) Vasorelaxation induced by common edible tropical plant extracts in isolated rat aorta and mesenteric vascular bed. J Ethnopharmacol 92:311–316
Saha JC, Savini EC, Kasinathan S (1961) Ecbolic properties of Indian medicinal plants. Part 1. Indian J Med Res 49:130
Sannella AR, Karioti A, Vincieri FF et al (2009) Antiplasmodial activity of papaya leaf decoction and its synergistic effects in combination with artemisinin. Planta Med 75:PD37. doi:10.1055/s-0029-1234516
Sathasivam K, Ramanathan S, Mansor SM et al (2009) Thrombocyte counts in mice after the administration of papaya leaf suspension. Wein Klin Wochenschr 121:19–22
Satrija F, Nansen P, Murtini S et al (1995) Anthelmintic activity of papaya latex against patent *Heligmosomoides polygyrus* infections in mice. J Ethnopharmacol 48:161–164
Sattar N, Scherbakova O, Ford I et al (2004) Elevated alanine aminotransferase predicts new-onset type-2 diabetes independently of classical risk factors, metabolic syndrome, and c-reactive protein in the west of Scotland coronary prevention study. Diabetes 53:2855–2860

Schmidt H (1995) Effect of papain on different phases of prenatal ontogenesis in rats. Reprod Toxicol 9:49–55
Schweiggert RM, Steingass CB, Heller A et al (2011) Characterization of chromoplasts and carotenoids of red- and yellow-fleshed papaya (*Carica papaya* L.). Planta 234:1031–1044
Seigler DS, Pauli GF, Nahrsedt A et al (2002) Cyanogenic allosides and glycosides from *Passiflora edulis* and *Carica papaya*. Phytochemistry 60:873–882
Silva CA, Gomes MTR, Ferreira RS et al (2003) A mitogenic protein fraction in latex from *Carica candamarcensis*. Planta Med 10:926–932
Singh A (1981) Medicinal plants in Fiji and other South Pacific Islands. Unpublished manuscript School of Natural Resources, University of the South Pacific, 65 pp
Spencer CF, Koniuszy FR, Rogers EF et al (1947) Survey of plants for anti-malarial activity. Lloydia 10:145–174
Starley IF, Mohammed P, Schneider G et al (1999) The treatment of paediatric burns using topical papaya. Burns 25:636–639
Stepek G, Buttle DJ, Duce IR et al (2005) Assessment of the anthelmintic effect of natural plant cysteine proteinases against the gastrointestinal nematode *Heligmosomoides polygyrus* in vitro. Parasitology 130:203–211
Tang CS (1971) Benzyl isothiocyanate in papaya fruit. Phytochemistry 10:117–121
Tang CS (1973) Localisation of benzyl glucosinolate and thioglucosidase in *Carica papaya* fruit. Phytochemistry 12:769–773
Tiikkainen M, Bergholm R, Vehkavaara S et al (2003) Effects of identical weight loss on body composition and features of insulin resistance in obese women with high and low liver fat content. Diabetes 52:701–707
Tiwari KC, Majumdar R, Bhattacharjee S (1982) Folklore information from Assam for family planning and birth control. Int J Crude Drug Res 20:133–137
Tona L, Kambu K, Ngimbi N et al (1998) Antiamoebic and phytochemical screening of some Congolese medicinal plants. J Ethnopharmacol 61:57–65
Traka M, Mithen R (2009) Glucosinolates, isothiocyanates and human health. Phytochem Rev 8:269–282
van Breemen RB, Pajkovic N (2008) Multitargeted therapy of cancer by lycopene. Cancer Lett 269:339–351
van Cauteren H, Lampo A, Vanderberghe J et al (1989) Toxicological profile and safety evaluation of antifungal azole derivatives. Mycoses 32(Suppl 1):60–66
Virtanen AI (1965) Studies on organic sulphur compounds and other labile substances in plants. Phytochemistry 4:207–228
Watson DC, Yaguchi M, Lynn KR (1990) The amino acid sequence of chymopapain from *Carica papaya*. Biochem J 266:75–81
Werner D (1992) Where there is no doctor. Hesperian Foundation, Palo Alto
Winarno FG (1983) Enzym pangan. PT Gramedia, Jakarta, p 115
World Health Organisation (1997) Dengue haemorrhagic fever: diagnosis, treatment, prevention and control, 2nd edn. WHO, Geneva
Zasada IA, Ferris H (2003) Sensitivity of *Meloidogyne javanica* and *Tylenchulus semipenetrans* to isothiocyanates in laboratory assays. Phytopathology 93:747–750
Zhang Y, Talalay P (1994) Anticarcinogenic activities of organic isothiocyanates: chemistry and mechanisms. Cancer Res 54(Suppl 7):1976s–1981s
Zhang R, Xu X, Chen T et al (2000) An assay for angiotensin-converting enzyme using capillary zone electrophoresis. Anal Biochem 280:286–290

Chapter 22
Allele Discovery Platform (ADP) in Papaya (*Carica papaya* L.)

P.K. Anish Kumar, Anjanabha Bhattacharya, O.P. Dutta, and Manash Chatterjee

Introduction

Papaya (*Carica papaya* L.) is a soft-wooded herbaceous dicotyledonous plant that belongs to the family *Caricaceae* and is closely related to Arabidopsis, in the same order Brassicales (Bremerk et al. 1998). Papaya and Arabidopsis share a common lineage dating back to 72 million years ago (Wikstro et al. 2001). The family *Caricaceae* contains 35 latex-containing species spread over six genera. *C. papaya*, the only species in the genus *Carica*, is used for fruit production in tropical and subtropical regions worldwide.

Papaya ranks first among fruits consumed (Hui et al. 2006) and also ranks first in nutritional profile (Bari et al 2006; Ming et al. 2008; Manshardt 1992). One medium-sized yellow-fleshed papaya can provide more than the adult minimum daily requirement of vitamin A. World Health Organization recommended papaya as one of the foods in a sustainable program for combating vitamin A deficiency (VAD) in the developing nations. There are many cultivars of papaya grown worldwide (Hui et al. 2006). Besides its nutritional and medicinal properties, papaya has a number of characteristics that contribute to its being used as an experimental model for tree crops. Flowering and fruiting are continuous throughout the year with the production of one to three ripe fruit per week and hundreds of fruit over the life of the tree. It has an efficient breeding system, each fruit producing about 800 (hermaphrodite fruit) or 1,000 (female fruit) seeds and a single tree producing hundreds of fruit in its lifetime to provide an abundance of offspring for genetic studies. Hand pollination

P.K.A. Kumar • A. Bhattacharya • M. Chatterjee (✉)
Bench Bio Pvt Ltd, Biotech Building, c/o Jai Research Foundation,
Vapi, Gujarat 396108, India
e-mail: mchatterjee@benchbio.com

O.P. Dutta
Department of Research and Development, Namdhari Seeds Pvt Ltd,
Bangalore, Karnataka 562109, India

R. Ming and P.H. Moore (eds.), *Genetics and Genomics of Papaya*, Plant Genetics and Genomics: Crops and Models 10, DOI 10.1007/978-1-4614-8087-7_22,

is easily done using pollen from male or hermaphrodite flowers on stigma of female or hermaphrodite flowers. A clonal propagation system is well established, which allows testing of the same genotypes in multiple environments. Papaya has a short juvenile phase of 3–8 months and a short generation time of 9–15 months. Papaya is among the limited number of plant species that are trioecious with three sex forms—female, male, and hermaphrodite. In the two breeding systems of papaya, it is either dioecious with male and female or gynodioecious with hermaphrodite and female.

Cultivation and domestication of fruit crops for human consumption started about 10,000 BC by selecting plant genotypes for their useful traits. As a consequence, many wild-type cultivars got eliminated in the process of selection leading to narrow gene pool. As a result, crops became unable to cope with emerging diseases, pests, and changing environmental regimes. In recent times though, many biotechnological tools have been applied to improve fruit productivity, quality, and other traits and transgenic technologies being in the forefront including RNAi (Csorba et al. 2009). However, resistance to transgenic technology, particularly in Europe, has pushed scientists to look for non-transgenic methods for its improvement.

TILLING (targeting induced local lesion in genome)/ADP (allele discovery platform) is a tool for creating variability in crop species. For decades, plant breeders have used chemicals as well as fast-moving ionizing radiation (gamma/X-rays) to bring about mutagenesis, i.e., small changes in plant genome, leading to increase in variability in the breeding population. Historically, the discovery that mutation can be artificially induced came from the experiments conducted by Muller in late 1920s. It was in 1942 that Charlotte Auerbach reported that action of chemical agent—nitrogen mustard—could cause mutations in cell (Dudhe and Chikkappa 2010).

Before Muller's discovery, biologists had to rely only on naturally induced spontaneous mutations, which were typically hard to detect and correlate with gene action. Many recessive useful mutations went undetected, and the potential of mutation breeding decreased. Later on, researchers developed a method to detect SNP mutation in wide populations using a new technique, described later in this chapter. This technique reduces time to screen large number of individuals in the population, saving resources by not having to grow many plants. Furthermore, systematically phenotypic characterization could be done at every stage, which is a useful tool for biologists.

Allele Discovery Platform

ADP is a variant of TILLING method. ADP is a technique, first developed by Colbert and coworkers in late 1990s and later modified by other groups (Colbert et al. 2001; Triques et al. 2008), to detect induced and natural polymorphisms (SNP) in any population (Simsek and Kacar 2010). The term Eco-ADP is used for SNP polymorphism, detected from natural population of a species. ADP is a reverse genetic functional genomics tool (Gilchrist and Haughn 2010), related to gene

Fig. 22.1 EMS-treated papaya mutant population (M1) growing in the field

information for a particular trait. Therefore, gene sequence has to be available in order to use ADP technology. ADP gives insight into gain or loss of gene function. This method is superior to gene inactivation by ionizing radiation, as it generates a series of alleles and gives us insight into partial loss of function of any gene instead of only knockouts. The technique mostly relies on the ability of a group of enzymes (CEL1 and EndoI), Surveyor® (Li et al. 2008; Triques et al. 2008), and Keypoint™ technologies (Poel et al. 2009) to detect mismatches in a specific gene, which were induced by ethylmethanesulfonate (EMS, an alkylating agent, which causes primarily C/G to T/A transitions in turn, wrongly pairing with complementary bases), DES (diethylsulfate, modify base G/C), ENU (ethylnitrosourea), or any potential mutagenizing agents. All of these chemicals result in base change after replication. Therefore, any chemical which causes high frequency of SNP in the target gene is preferred in chemical treatments to create good mutagenized population (Till et al. 2007).

Overview of ADP Technology

The first step is to create a high-quality mutant population using EMS. A range of phenotypes are observed in the M1 population, and fruits from individual plants are harvested and M2 seeds are collected (Fig. 22.1). DNA is extracted from M2 families, then a set of labeled primers are used to amplify a specific gene from the M2 families, and enzymes are later added to the PCR products to detect mismatched heteroduplexes in the PCR products (Table 22.1). The resultant samples are run on a DNA

Table 22.1 Steps in allele discovery platform (ADP)

1. Seeds are mutagenized by chemical mutagenizing agent (EMS) to induce random point mutations in the genome. Treated seeds denoted as M1, untreated seeds as M0, also known as mutagenized population; M = mutagenized
2. M1 plants are grown and M2 seeds are produced after selfing. This population (M1) mainly consists of heterozygous plants for a particular mutation
3. Ten individuals from each M2 family are grown; DNA is collected in pools and stored in 96 plate format
4. Usually, eightfold pooling is done; 768 samples are screened in one go. Samples are subjected to gene-specific PCR amplification with IRDye-labeled primers (IR 700Red/800Green)
5. Labeled PCR products are heated and cooled resulting in heteroduplex formation between normal plants and mutants
6. Specific enzymes (CEL1/ENDO1) are used to digest at mismatches
7. Samples are denatured and run on a LiCOR DNA analyzer system *Spots are detected. The sum of spots detected should be equal to PCR amplicon sizes. Individuals are screened to detect the line carrying the mutation. Sequencing is done to identify the SNP mutation for a particular line carrying the mutation, and the effect of such mutation is evaluated in silico and by further phenotyping*
8. Homozygous lines identified are evaluated and can be released as a variety after backcrossing to the original parent or converted into hybrids

analyzer (LiCoR system) to detect mutations. A wide spectrum of mutations can be obtained from a particular mutagenized population resulting in partial or full knockouts of a gene(s), missense, and frameshift mutations leading to altered or truncated protein, i.e., increased, decreased, or altered gene function. Probability of choosing conserved regions in expressed region of genes increases the chances of obtaining product with altered gene function (McCallum et al. 2000). Programs like CODDLE (codons to optimize discovery of deleterious lesions, Website: http://www.proweb.org/coddle/) are used to predict the maximum effect of EMS on the target region screened. Such programs can be used to design primers covering mutation-prone regions. Mutation in the splice junction can result in knockout and/or altered protein coded by the respective gene.

Before starting ADP screening for any gene, it is necessary to know the copy number in any species and more importantly in a polyploid species. Multiple amplification of related sequences will likely result in excess number of fragment sizes in polyacrylamide gels, and there is a high chance that any mutation could be lost in the background. Furthermore, it is also desirable to know which copy of the gene has been altered. Therefore, ideally the first step before starting an ADP screening is to identify copy number by performing Southern hybridization or other methods were possible.

There are other techniques for detecting mutations rapidly, which are gaining prominence—Cold PCR and QMC (quick-multiplex-consensus)-assisted HSM (high resolution melting) (Fadhil et al. 2010; Milbury et al. 2010; Chen and Wilde 2011), DHPLC (denatured-HPLC), EMAIL (endonucleolytic mutation analysis by internal labeling, Cross et al. 2008), FAMA (fluorescence-assisted mismatch analysis, Tosi et al. 2001) ADP by sequencing (Tsai et al. 2011), and pyrosequencing

(Ibrahem et al. 2010). The refined methods of detecting mutations are superior than LI-COR-based systems in several ways: first, there is no need of enzymes to identify heteroduplexes; second, there is no need to work with polyacrylamide gels, which are toxic; and third, they overcome the limitation in gel-based ADP screening that mutations in first and last 50–100 bp regions are difficult to identify. However, at present from practical point of view, cost of detecting mutations by sequencing is high compared to enzyme-based method.

Papaya Genomics

In 2008, a group of researchers from University of Hawaii first sequenced the papaya genome variety "Sunup," having three times genome coverage (Rachel 2008). The variety "Sunup" is a transgenic papaya ringspot virus (PRSV)-resistant fruit crop. This was the first reported genome sequencing of a fruit tree. Genome size of papaya is about 372 Mb but reportedly has about 24,746 genes (Ming et al. 2008; Rachel 2008) because of lack of genome-wide duplication. The genome is about three times the size of model plant "Arabidopsis" and contains far less number of disease-resistant genes (Ming et al. 2008). Most gene families have fewer members in papaya, but there are few exceptions, including MADS box gene families, indicating that such genes may influence or play multiple tasks of turning on and off the other housekeeping genes (Rachel 2008). Moreover, papaya genome is highly euchromatic (Ming et al. 2008).

Since the genome of papaya has been already sequenced, specific gene-related information is available by performing intensive data mining and comparative analysis with other sequenced crops. Therefore, a large number of genes can be used for potential targets in ADP screening program. Here, we have enlisted some traits, which can be modified to improve papaya crop.

Targets for ADP

Flower Types

Papaya is polygamous and bears three types of flowers, male, female, and hermaphrodite. Out of these three types, hermaphrodite plant produces uniform fruits that are used for production in most papaya-growing regions. Specific primers are now available to distinguish one flower (sex) type from the other in the pre-flowering stages (Gangopadhyay et al. 2007). However, from practical point of view, such markers are not useful for the farmers due to their high cost. Hermaphrodite fruit produces seeds that segregate in the ratio of 2:1 (hermaphrodite: female). Therefore, true hermaphrodite plants are nonexistent. Several hypotheses were proposed to explain the sex-segregating patterns in papaya, including one gene with three alleles, M, M^h, and M

(Hofmeyr 1939) and X, Y, and Y2 chromosomes (Horovitz and Jimenez 1967). It is now clear that papaya has a pair of nascent sex chromosomes and the two sex determination genes are in the male-specific region of the Y chromosome (MSY) (Wang et al. 2012; Gschwend et al. 2012). Therefore, manipulating sex determination gene will lead to the development of true-breeding hermaphrodite varieties.

Fruit Quality

Genes involved in fruit ripening from representative of two independently generated cDNA libraries from red and yellow fruit-bearing varieties have been reported (Devitt et al. 2006). Several transcripts found in abundance were related to chitinase, ACC (aminocyclopropane-1-carboxylic acid) related to cell wall softening. Papaya flesh color is controlled by chromoplast-specific lycopene β-cyclase, *CpCYC-b*, in the carotenoid biosynthesis pathway (Blas et al. 2010). Some of these genes can be used as a target gene for improving fruit quality. Recently long-shelf-life muskmelon lines have been generated by TILLING (Dahmani-Mardas et al. 2010). Similar strategy could be applied in papaya.

Papaya Ringspot Virus

PRSV, a potyvirus, causes widespread crop loss worldwide where papaya is grown (Tennant et al. 2007; Tripathi et al. 2004). It causes loss of up to 100 % and there is no natural resistance in papaya germplasm in the world. Typical symptoms are mosaic leaves, dark green rings on fruit, and in severe cases shoestring-like leaves (Gonsalves 1998). It is transmitted by aphids and also by mechanical means (Bhargava and Khurana 1970; Gonsalves 1998). PRSV resistance in papaya has been explained in details in a review by Fermin et al. (2010). Resistance to PRSV has been achieved in other crops as described in Brotman et al. (2002). Several natural resistance genes against potyviruses, from distinct crops, were shown to encode defective forms of eIF4E and eIF4Eiso proteins (Caranta et al 2003, 2004; Gao et al 2006; Jahn and Kang 2003; Kang et al 2005; Ruffel et al 2002, 2005; Nicaise et al 2003; Nieto et al 2006). For example, pvr2 from pepper is resistant to potato virus Y (PVY) and tobacco etch virus (TEV), mo1 from lettuce is resistant to lettuce mosaic virus (LMV), sbm1 from pea is resistant to pea seed-borne mosaic virus (PSbMV), and pot-1 from tomato is resistant to PVY and TEV. In all cases, although these genes control diverse resistance phenotypes, resistance results from a small number of amino acid changes in the eIF4E and eIF4Eiso proteins encoded by the recessive resistance alleles that harbor point mutations. Recently potyvirus such as melon necrotic spot virus (MNSV)-resistant melon plants and tomato plants resistant to potato virus Y (PVY) and pepper mottle virus (PepMoV) have been successfully created by TILLING melon and tomato eIF4E and eIF4Eiso genes (Nieto et al. 2007; Piron et al 2010). Therefore, it is hypothesized that TILLING

orthologs of papaya eIF4E and eIF4Eiso genes has a good potential for generating non-transgenic potyvirus-resistant varieties.

Nematode Resistance

Nematodes cause huge loss in papaya. Significantly among the nematode species, *Meloidogyne* sp. and *Rotylenchulus reniformis* cause widespread damage in papaya (Borai and Duncan 2005). Genes specifically acting at the site of nematode-plant interaction play an important role in developing varieties resistant to nematode, which includes gene related to biosynthesis of plant hormone auxin and lignification-related genes.

Efficient Crop Nutrition

Papaya removes large amount of nutrients from the soil including micronutrients (Chang et al. 1983). Regulating plant physiology and architecture may help to regulate the nutrient requirements of the crop.

Papain Production

Papaya is also being grown for papain production, which is a proteolytic enzyme. Papain has many medicinal and prophylactic activities (Nitsawang et al. 2006). Furthermore, it is used in certain anticancer drugs. Enhancing expression of papain-coding gene may increase papain production.

Cyanide Reduction

Cyanogenic glucosides are present in leaves of papaya, which are toxic to human health. Gene knockouts or low expression for cyanogenic glucosides will produce plants with altered cyanide levels, harmless to humans and livestock (Hui et al. 2006). Recently cyanide-free sorghum lines have been generated by TILLING (Cecilia et al. 2012).

Allergen Reduction

There are reports indicating that in some people, consuming papaya fruits causes skin rashes (De Clerck et al. 2003). Genes responsible for causing allergy could be deactivated by ADP method.

Antioxidant Levels

Papaya fruits are rich source of antioxidants (Lim et al. 2007), which destroy free radicals formed in our bodies, thus, reducing aging and health-related problems (MacDonald-Wicks et al. 2006). Increasing levels of antioxidants will increase health-promoting properties of papaya.

Medicinal Use

Papaya is also used for treatment of burns (Starley et al. 1999). There are several other medicinal properties of papaya, as described extensively in Krishna et al. (2008). Enzymes responsible for healing response could be identified and engineered for higher expression.

Shortening Life Cycle of the Crop or Early-Bearing Crop

The period from sowing a seed and harvesting fruits from papaya takes about 9 months or more. During this period, the field remains occupied and it makes it difficult to produce more than one crop per year. As the genome is sequenced, it might be possible to manipulate flowering genes, which could make this plant bear fruits early than observed naturally.

Manipulating Root Architecture

Genes related to root development can be manipulated in this crop. Papaya is a shallow-rooted crop and, therefore, it has limited reach to the nutrition and water regimes available in soil. Papaya is also vulnerable to water logging and high winds. The plant can be engineered to have better root architecture.

Drought Resistance

Water requirement of papaya is enormous. Manipulating genes involved in stomata activity may result in a variety which thrives on less available water.

Disease and Pest Resistance

Bacteria, nematode, virus, fungi, and various insects attack papaya crop. Scientists are now looking into the detailed function of plant receptor and effector mechanism to arrive at a better strategy to control diseases (Glowacki et al. 2011). Insects like aphids, leafhoppers, and thrips attack papaya crop. Papaya varieties could be improved to be more tolerant to such insects. Manipulating composition of cap flowing through phloem could act as a deterrent to insects.

Source and Sink

Manipulating source and sink genes will result in diverting nutrient resources from leaves to developing flower and fruits (Website 1). Papaya is unique as it lacks starch reserves and fruits must remain attached to the mother plant to grow and accumulate sugar. Two genes, invertase and sucrose synthase (SS), are essential for accumulating sugar in developing fruits. Manipulating such genes could increase efficiency of production (Website 1).

Phenotypic Database

One of the by-products of ADP is phenotypic database. When any crop species is treated with EMS, detailed phenotyping on M2 families can be done for certain traits, for example, yield per plant, plant height, plant architecture, fruit size, and seedless fruits. Data collected are entered into a phenotypic database, and a literature survey is done to correlate with gene information wherever available for the traits. In some cases, the information can be used for ADP screening of potential candidate genes. If necessary, genes responsible for the interesting trait can be isolated by creating mapping population and positional cloning and manipulated further for crop improvement.

Conclusions

Papaya is a model fruit crop because of its small genome, short generation time, rich genomic resources, and nascent sex chromosomes. Therefore, this species can be exploited to study genetic basis of fruit quality and characteristics. Papaya also stands first in terms of nutritional properties. In order to exploit their qualitative and quantitative parameters, ADP can be utilized in a short time frame to improve this crop.

ADP technology is non-transgenic and therefore cost to develop variety is significantly lower as it requires minimal to no regulatory approval. Moreover, as there is still increasing concern against transgenic products, ADP can be exploited to create traits in elite lines, which can be directly released as a variety or used as a parent in a breeding program. A number of traits can be manipulated in a relatively short time in combination with conventional breeding. As papaya genome has been sequenced, detailed data mining and comparative genomic studies are required to identify novel genes related to certain function or traits that can be incorporated into ADP screening for trait development.

Acknowledgments We would like to thank Mr Uday Singh and Dr N. Anand (Namdhari Seeds) for financial support and encouragement. We also thank Dr A. Bendahmane (URGV lab, INRA) for advice on the TILLING approach and PRSV resistance. The contributions of all the glass house and field staff involved with the growing of the mutant populations are hereby acknowledged.

References

Bari L, Hassan P, Absar N, Haque ME, Khuda M, Perrvin MM (2006) Nutritional analysis of two local varieties of papaya (*Carica papaya* L.) at different maturation stages. Pakistan J Biol Sci 9

Bhargava KS, Khurana SMP (1970) Insect transmission of papaya virus with special reference to papaya mosaic virus. Zbl Bekt 124:688–696

Blas AL, Ming R, Liu Z, Veatch OJ, Paull RE, Moore PH, Yu Q (2010) Cloning of the papaya chromoplast-specific lycopene β-cyclas, CpCYC-b, controlling fruit flesh color reveals conserved microsynteny and a recombination hot spot. Plant Physiol 152:2013–2022

Borai FE, Duncan LW (2005) Nematode parasites of subtropical and tropical fruit tree crops. In: Luc M, Sikora RA, Bridge J (eds) Plant parasitic nematodes in subtropical and tropical agriculture, 2nd edn. CAB International, Oxfordshire, pp 467–492

Bremerk K, Chase MW, Stevens PF (1998) An ordinal classification for the families of flowering plants. Ann Mo Bot Gard 85:531–553

Brotman Y, Silberstein L, Kovalski L, Perin C, Dogiment C, Pilrat M, Klingler L, Thompson A, Perl-Treves R (2002) Resistance gene homologues in melon are linked to genetic loci conferring disease and pest resistance. TAG 104:1055–1063

Caranta C, Ruffel S, Bendahmane A, Palloix A, Robaglia C (2003) Mutations in the eIF4E gene and resistance against Potyvirus. Patent GENOPLANTE VALOR WO 03/066900

Caranta C, Ruffel S, Palloix A, Fabre N (2004) Combination of mutation in both eIF4E et eIF(iso)4E for resistance against Potyvirus. Patent GENOPLANTE VALOR FR 04/05616

Cecilia KB, Roslyn MG, Natalie O'D, Peter N, Kenneth J, Tomas L, Carl EO, Peter S, John DH, Birger LM, Alan DN (2012) A combined biochemical screen and TILLING approach identifies mutations in *Sorghum bicolor* L. Moench resulting in acyanogenic forage production. Plant Biotechnol J 10:54–66

Chang SS, Hu NH, Chen CC, Chu TF (1983) Diagnosis criteria of boron deficiency in papaya and the soil boron status of Taitung area (Taiwan). J Agric Res China 32:238–252

Chen Y, Wilde D (2011) Mutation scanning of peach floral genes. BMC Plant Biol 11:96

Colbert T, Till BJ, Tompa R, Reynolds S, Steine MN, Yeung AT, McCallum CM, Comai L, Henikoff S (2001) High throughput screening for induced point mutations. Plant Physiol 126:480–484

Cross MJ, Waters DLE, Lee LS, Henry RJ (2008) Endonucleolytic mutation analysis by internal labeling (EMAIL). Electrophoresis 29:1291–1301

Csorba T, Pantaleow V, Burgyan J (2009) RNA silencing: an antiviral mechanism. Adv Virus Res 75:35–72

Dahmani-Mardas F, Troadec C, Boualem A, Lévêque S, Alsadon AA, Aldoss AA, Dogimont C, Bendahmane A (2010) Engineering melon plants with improved fruit shelf life using the TILLING approach. PLoS One 5(12):e15776

De Clerck LS, Ebo DG, Bridts CH, Stevens WJ (2003) Angio-edema and oral allergy syndrome due to the consumption of *Carica Papaya*. J Allergy Clin Immunol 111:S103

Devitt LC, Sawbridge T, Holton TA, Mitchelson K, Dietzen R (2006) Discovery of genes associated with fruit ripening in *Carica papaya* using expressed sequence tags. Plant Sci 170:356–363

Dudhe MY, Chikkappa GK (2010) Plant sciences. New Vishal Publications, Delhi, p 96

Fadhil W, Ibrahem S, Seth R, Ilyas M (2010) Quick-multiplex-consensus (QMC)-PCR followed by high-resolution melting: a simple and robust method for mutation detection in formalin-fixed paraffin-embedded tissue. J Clin Pathol 63:134–140

Fermin GA, Castro LT, Tennant PF (2010) CP-transgenic and non-transgenic approaches for the control of papaya ringspot: current situation and challenges. Transgenic Plant J 10:1–15

Gangopadhyay G, Roy SK, Ghose K, Poddar R, Bandyopadhyay T, Basu D, Mukherjee KK (2007) Sex detection of *Carica papaya* and *Cycas circinalis* in pre-flowering stage by ISSR and RAPD. Curr Sci 92:524–526

Gao ZH, Johansen E, Eyers S, Thomas CL, Noel Ellis TH, Maule AJ (2006) The potyvirus recessive resistance gene, *sbm1*, identifies a novel role for translation initiation factor eIF4E in cell-to-cell trafficking. Plant J 40:376–385

Gilchrist E, Haughn G (2010) Reverse genetics techniques: engineering loss and gain of gene function in plants. Brief Funct Genomics 9:103–110

Glowacki S, Macioszek VK, Kenonowicz AK (2011) R proteins as fundamentals of plant innate immunity. Cell Mol Biol Lett 16:1–24

Gonsalves D (1998) Control of papaya ringspot virus in papaya: a case study. Annu Rev Phytopathol 36:415–437

Gschwend AR, Yu Q, Tong EJ, Zeng F, Han J, VanBuren R, Aryal R, Charlesworth D, Moore PH, Paterson AH, Ming R (2012) Rapid divergence and expansion of the X chromosome in papaya. Proc Natl Acad Sci USA 109:13716–13721

Hofmeyr JDJ (1939) Some genetic breeding aspects of *Carica papaya* L. Agron Trop 17:345–351

Horovitz S, Jimenez H (1967) Cruzamientos interespecificas and intergenerius on cariceas ysus implicaciones fitotechicas. Agron Trop 17:323–343

Hui YH, Cano MP, Gusek T, Sidhu JS, Sinha NK (2006) Handbook of fruits and fruit processing. Blackwell, Oxford

Ibrahem S, Seth R, O'Sullivan B, Fadhil W, Taniere P, Ilyas M (2010) Comparative analysis of pyrosequencing and QMC-PCR in conjunction with high resolution melting for KRAS/BRAF mutation detection. Int J Exp Pathol 91:500–505

Jahn MM, Kang GC (2003) Recessive plant viral resistance results from mutations in translation initiation factor eif4e. Patent WO 04/057941

Kang BC, Yeam I, Frantz JD, Murphy JF, Jahn MM (2005) The pvr1 locus in *Capsicum* encodes a translation initiation factor eIF4E that interacts with Tobacco etch virus VPg. Plant J 42:392–405

Krishna KL, Paridhavi M, Patel JA (2008) Review on nutritional, medicinal and pharmacological properties of papaya (*Carica papaya* L.). Nat Prod Radiance 7:364–373

Li J, Wang L, Mamon H, Kulke MH, Berbeco RGM, Makrigiorgos GM (2008) Replacing PCR with COLD-PCR enriches variant DNA sequences and redefines the sensitivity of genetic testing. Nat Med 14:579–584

Lim YY, Lim TT, Tee JJ (2007) Antioxidant properties of several tropical fruits: a comparative study. Food Chem 103:1003–1008

MacDonald-Wicks LK, Wood LG, Garg ML (2006) Methodology for the determination of biological antioxidant capacity in vitro: a review. J Sci Food Agric 86(13):2046–2056

Manshardt A (1992) Papaya. In: Hammerschlag FA, Ritz RE (eds) Biotechnology in agriculture, No 8. Biotechnology of perennial fruit crops. CABI, Willingford, pp 489–511

McCallum CM, Comai L, Greena EA, Henikoff S (2000) Targeting Induced Local Lesions IN Genomes (ADP) for plant functional genomics. Plant Physiol 123:439–442

Milbury CA, Li J, Makrigiorgos GM (2010) Ice-COLD-PCR enables rapid amplification and robust enrichment for low-abundance unknown DNA. Nucleic Acids Res 14:579–584

Ming R, Hou S, Feng Y, Yu Q, Dionne-Laporte A, Saw JH et al (2008) The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). Nature 452:991–996

Nicaise V, German-Retana S, Sanjuan R, Dubrana MP, Mazier M, Maisonneuve B, Candresse T, Caranta C, LeGall O (2003) The eukaryotic translation initiation factor 4E controls lettuce susceptibility to the potyvirus *Lettuce mosaic virus*. Plant Physiol 132:1272–1282

Nieto C, Morales M, Orjeda G, Clepet C, Monfort A, Sturbois B, Puigdomenech P, Pitrat M, Caboche M, Dogimont C, García-Mas J, Aranda MA, Bendahmane A (2006) An *eIF4E* allele confers resistance to an uncapped and non-polyadenylated RNA virus in melon. Plant J 48:452–462

Nieto C, Piron F, Dalmais M, Marco CF, Moriones E, Gómez-Guillamón ML, Truniger V, Gómez P, Garcia-Mas J, Aranda MA, Bendahmane A (2007) EcoTILLING for the identification of allelic variants of melon eIF4E, a factor that controls virus susceptibility. BMC Plant Biol 21:34

Nitsawang S, Hatti-Kaul R, Kanasawud P (2006) Purification of papain from *Carica papaya* latex: aqueous two-phase extraction versus two-step salt precipitation. Enzyme Microb Technol 39:1103–1107

Piron F, Nicolai M, Minoia S, Piednoir E, Moretti A, Salgues A, Zamir D, Caranta C, Bendahmane A (2010) An induced mutation in tomato eIF4E leads to immunity to two potyviruses. PLoS One 5(6):e11313

Poel HJA, van Orsouw NJ, Hogers RCJ, de Both MTJ, van Eijk MJT (2009) High-throughput detection of induced mutations and natural variation using KeyPoint™ technology. PLoS One 4:e4761

Rachel E (2008) Study decodes papaya genome. Sci News 173:16

Ruffel S, Dussault MH, Palloix A, Moury B, Bendahmane A, Robaglia C, Caranta C (2002) A natural recessive resistance gene against *Potato virus* Y in pepper corresponds to the eukaryotic initiation factor 4E (eIF4E). Plant J 32:1067–1075

Ruffel S, Gallois JL, Lesage ML, Caranta C (2005) The recessive potyvirus resistance gene pot-1 is the tomato orthologue of the pepper pvr2-eIF4E gene. Mol Genet Genomics 274:346–353

Simsek O, Kacar YA (2010) Discovery of mutations with ADP and EcoADP in plant genomics. Sci Res Essays 5:3799–3882

Starley IF, Mohammed P, Schneider G, Bickler SW (1999) The treatment of paediatric burns using topical papaya. Burns 25:636–639

Tennant PF, Fermin GA, Roye ME (2007) Viruses infecting papaya (*Carica papaya* L.): etiology, pathogenesis and molecular biology. Plant Viruses 1:178–188

Till BJ, Cooper J, Tai TH, Colowit P, Greene EA, Henikoff S, Comai L (2007) Discovery of chemically induced mutations in rice by TILLING. BMC Plant Biol 7:19

Tosi M, Verpy E, Meo T (2001) Detection of mutations by fluorescence assisted mismatch analysis (FAMA). Clin Mol Genet 119

Tripathi S, Bau HJ, Chen LF, Yeh SD (2004) The ability of papaya ringspot virus strains overcoming the transgenic resistance of papaya conferred by the coat protein gene is not correlated with higher degree of sequence divergence from the transgene. Eur J Plant Pathol 110:871–882

Triques K, Piednoir E, Dalmais M, Schmidt J, Le Signor C, Sharkey M, Caboche M, Sturbois B, Bendahmane A (2008) Mutation detection using ENDO 1: application to disease diagnostics in humans and TILLING and Eco-TILLING in plants. BMC Mol Biol 9:42

Tsai H, Howell T, Comai L (2011) Discovery of rare mutation in population: ADP by sequencing. Plant Physiol 156:1257–1268

Wang J, Na J-K, Yu Q, Gschwend AR, Han J, Zeng F, Aryal R, VanBuren R, Murray JE, Zhang W, Pérez RN, Feltus FA, Lemke C, Tong EJ, Chen C, Wai CM, Singh R, Wang M-L, Min X, Alam M, Charlesworth D, Moore PH, Jiang J, Paterson AH, Ming R (2012) Sequencing papaya X and Yh chromosomes revealed molecular basis of incipient sex chromosome evolution. PNAS 109:13710–13715

Wikstro MN, Savolainen V, Chase MW (2001) Evolution of the angiosperms: calibrating the family tree. Proc R Soc Lond B 268:2211–2220

Website

Source-sink relationship during papaya fruit growth and development. http://www.ctahr.hawaii.edu/t-star/papaya.htm

Index

R. Ming and P.H. Moore (eds.), *Genetics and Genomics of Papaya*, Plant Genetics and Genomics: Crops and Models 10, DOI 10.1007/978-1-4614-8087-7,

B

N