Compendium of Plant Genomes
Series Editor: Chittaranjan Kole

Nirala Ramchiary · Chittaranjan Kole *Editors*

The Capsicum Genome

Springer

Compendium of Plant Genomes

Series editor

Chittaranjan Kole, Raja Ramanna Fellow, Department of Atomic Energy, Government of India, ICAR-National Research Center on Plant Biotechnology, New Delhi, India

Whole-genome sequencing is at the cutting edge of life sciences in the new millennium. Since the first genome sequencing of the model plant Arabidopsis thaliana in 2000, whole genomes of over 100 plant species have been sequenced and genome sequences of several other plants are in the pipeline. Research publications on these genome initiatives are scattered on dedicated web sites and in journals with all too brief descriptions. The individual volumes elucidate the background history of the national and international genome initiatives; public and private partners involved; strategies and genomic resources and tools utilized; enumeration on the sequences and their assembly; repetitive sequences; gene annotation and genome duplication. In addition, synteny with other sequences, comparison of gene families and most importantly potential of the genome sequence information for gene pool characterization and genetic improvement of crop plants are described.

Interested in editing a volume on a crop or model plant? Please contact Dr. Kole, Series Editor, at ckoleorg@gmail.com

More information about this series at http://www.springer.com/series/11805

Nirala Ramchiary · Chittaranjan Kole
Editors

The Capsicum Genome

Springer

Editors
Nirala Ramchiary
School of Life Sciences
Jawaharlal Nehru University
New Delhi, India

Chittaranjan Kole
Raja Ramanna Fellow, Department of
Atomic Energy, Government of India
ICAR-National Research Center on Plant
Biotechnology
New Delhi, India

ISSN 2199-4781 ISSN 2199-479X (electronic)
Compendium of Plant Genomes
ISBN 978-3-319-97216-9 ISBN 978-3-319-97217-6 (eBook)
https://doi.org/10.1007/978-3-319-97217-6

Library of Congress Control Number: 2018967743

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

This book series is dedicated to my wife Phullara,
and our children Sourav, and Devleena

Chittaranjan Kole

Preface to the Series

Genome sequencing has emerged as the leading discipline in the plant sciences coinciding with the start of the new century. For much of the twentieth century, plant geneticists were only successful in delineating putative chromosomal location, function, and changes in genes indirectly through the use of a number of "markers" physically linked to them. These included visible or morphological, cytological, protein, and molecular or DNA markers. Among them, the first DNA marker, the RFLPs, introduced a revolutionary change in plant genetics and breeding in the mid-1980s, mainly because of their infinite number and thus potential to cover maximum chromosomal regions, phenotypic neutrality, absence of epistasis, and codominant nature. An array of other hybridization-based markers, PCR-based markers, and markers based on both facilitated construction of genetic linkage maps, mapping of genes controlling simply inherited traits, and even gene clusters (QTLs) controlling polygenic traits in a large number of model and crop plants. During this period, a number of new mapping populations beyond F_2 were utilized and a number of computer programs were developed for map construction, mapping of genes, and for mapping of polygenic clusters or QTLs. Molecular markers were also used in the studies of evolution and phylogenetic relationship, genetic diversity, DNA fingerprinting, and map-based cloning. Markers tightly linked to the genes were used in crop improvement employing the so-called marker-assisted selection. These strategies of molecular genetic mapping and molecular breeding made a spectacular impact during the last one and a half decades of the twentieth century. But still they remained "indirect" approaches for elucidation and utilization of plant genomes since much of the chromosomes remained unknown and the complete chemical depiction of them was yet to be unraveled.

Physical mapping of genomes was the obvious consequence that facilitated the development of the "genomic resources" including BAC and YAC libraries to develop physical maps in some plant genomes. Subsequently, integrated genetic–physical maps were also developed in many plants. This led to the concept of structural genomics. Later on, emphasis was laid on EST and transcriptome analysis to decipher the function of the active gene sequences leading to another concept defined as functional genomics. The advent of techniques of bacteriophage gene and DNA sequencing in the 1970s was extended to facilitate sequencing of these genomic resources in the last decade of the twentieth century.

As expected, sequencing of chromosomal regions would have led to too much data to store, characterize, and utilize with the-then available computer software could handle. But the development of information technology made the life of biologists easier by leading to a swift and sweet marriage of biology and informatics, and a new subject was born—bioinformatics.

Thus, the evolution of the concepts, strategies, and tools of sequencing and bioinformatics reinforced the subject of genomics—structural and functional. Today, genome sequencing has traveled much beyond biology and involves biophysics, biochemistry, and bioinformatics!

Thanks to the efforts of both public and private agencies, genome sequencing strategies are evolving very fast, leading to cheaper, quicker, and automated techniques right from clone-by-clone and whole-genome shotgun approaches to a succession of second-generation sequencing methods. The development of software of different generations facilitated this genome sequencing. At the same time, newer concepts and strategies were emerging to handle sequencing of the complex genomes, particularly the polyploids.

It became a reality to chemically—and so directly—define plant genomes, popularly called whole-genome sequencing or simply genome sequencing.

The history of plant genome sequencing will always cite the sequencing of the genome of the model plant Arabidopsis thaliana in 2000 that was followed by sequencing the genome of the crop and model plant rice in 2002. Since then, the number of sequenced genomes of higher plants has been increasing exponentially, mainly due to the development of cheaper and quicker genomic techniques and, most importantly, the development of collaborative platforms such as national and international consortia involving partners from public and/or private agencies.

As I write this preface for the first volume of the new series "Compendium of Plant Genomes," a net search tells me that complete or nearly complete whole-genome sequencing of 45 crop plants, eight crop and model plants, eight model plants, 15 crop progenitors and relatives, and 3 basal plants is accomplished, the majority of which are in the public domain. This means that we nowadays know many of our model and crop plants chemically, i.e., directly, and we may depict them and utilize them precisely better than ever. Genome sequencing has covered all groups of crop plants. Hence, information on the precise depiction of plant genomes and the scope of their utilization are growing rapidly every day. However, the information is scattered in research articles and review papers in journals and dedicated Web pages of the consortia and databases. There is no compilation of plant genomes and the opportunity of using the information in sequence-assisted breeding or further genomic studies. This is the underlying rationale for starting this book series, with each volume dedicated to a particular plant.

Plant genome science has emerged as an important subject in academia, and the present compendium of plant genomes will be highly useful both to students and teaching faculties. Most importantly, research scientists involved in genomics research will have access to systematic deliberations on the plant genomes of their interest. Elucidation of plant genomes is of interest not only for the geneticists and breeders, but also for practitioners of an array of plant science disciplines, such as taxonomy, evolution, cytology,

physiology, pathology, entomology, nematology, crop production, biochemistry, and obviously bioinformatics. It must be mentioned that information regarding each plant genome is ever-growing. The contents of the volumes of this compendium are, therefore, focusing on the basic aspects of the genomes and their utility. They include information on the academic and/or economic importance of the plants, description of their genomes from a molecular genetic and cytogenetic point of view, and the genomic resources developed. Detailed deliberations focus on the background history of the national and international genome initiatives, public and private partners involved, strategies and genomic resources and tools utilized, enumeration on the sequences and their assembly, repetitive sequences, gene annotation, and genome duplication. In addition, synteny with other sequences, comparison of gene families, and, most importantly, the potential of the genome sequence information for gene pool characterization through genotyping by sequencing (GBS) and genetic improvement of crop plants have been described. As expected, there is a lot of variation of these topics in the volumes based on the information available on the crop, model, or reference plants.

I must confess that as the series editor, it has been a daunting task for me to work on such a huge and broad knowledge base that spans so many diverse plant species. However, pioneering scientists with lifetime experience and expertise on the particular crops did excellent jobs editing the respective volumes. I myself have been a small science worker on plant genomes since the mid-1980s and that provided me the opportunity to personally know several stalwarts of plant genomics from all over the globe. Most, if not all, of the volume editors are my longtime friends and colleagues. It has been highly comfortable and enriching for me to work with them on this book series. To be honest, while working on this series I have been and will remain a student first, a science worker second, and a series editor last. And I must express my gratitude to the volume editors and the chapter authors for providing me the opportunity to work with them on this compendium.

I also wish to mention here my thanks and gratitude to the Springer staff, Dr. Christina Eckey and Dr. Jutta Lindenborn in particular, for all their constant and cordial support right from the inception of the idea.

I always had to set aside additional hours to edit books beside my professional and personal commitments—hours I could and should have given to my wife, Phullara, and our kids, Sourav, and Devleena. I must mention that they not only allowed me the freedom to take away those hours from them but also offered their support in the editing job itself. I am really not sure whether my dedication of this compendium to them will suffice to do justice to their sacrifices for the interest of science and the science community.

Kalyani, India Chittaranjan Kole

Preface to the Volume

Capsicum, also called as chili pepper, belongs to the genus *Capsicum* and Family *Solanaceae*. It is believed to be the first spice crop domesticated and cultivated in an around 6000 years ago in Central and South America. Under the genus *Capsicum*, a total of 38 different species is listed, of which six species namely *Capsicum annuum*, *C. frutescens*, *C. pubescence*, *C. chinense*, *C. baccatum*, and *C. assamicum* are cultivated. Several evidence from archaeological, genetic, and contemporary plant distributions analysis indicated that *C. annuum* was primarily domesticated in the regions of Mexico or near to Northern Central America, *C. chinense* in Amazonia, *C. frutescens* in the Caribbean region, *C. pubescens*, and *C. baccatum* in the southern Andes (Bolivia and Peru). Recently, the *C. assamicum* has been identified as a distinct domesticated species in the Northeastern part of India which is closely related to *C. frutescens* and *C. chinense* but can be differentiated due to its unique characteristics. The unique property of capsicum is due to the presence of an alkaloid complex known as Capsaicinoids which imparts pungency property to the chili fruits, and only for this reason, chili fruit extract is being used in several traditional medicinal formulations. Furthermore, several studies reported the presence of a wide variety of beneficial metabolites such as carotenoids (provitamin A), vitamins (C and E), flavonoids, and capsaicinoids. in the capsicum fruits. The chili pepper fruits also contain a wide variety of color due to the variation in carotenoids and pigments, which are also used as a coloring agent in food industries.

Despite so much of economic importance, the study toward the identification of genes or quantitative trait loci governing fruit traits (size, shape, and texture), beneficial metabolites, nutrient elements uptake, physiological traits, biotic and abiotic stress tolerance is still limited in capsicum. However, many of the genes/QTLs for those traits have been identified in tomato crop, considered as the model for fleshy fruit plants, which belongs to the same family Solanaceae as capsicum. This relatively less progress might be due to the complexity of capsicum genome, which is approximately about 3.5 Gigabase (Gb) in size, compared to the tomato genome of 900 Megabase (Mb). Nevertheless, classical breeding efforts could improve yield, fruit morphology and metabolites content, and resistance to biotic and abiotic stress tolerance in capsicum. Furthermore, advances in molecular biology, and advent of high-throughput genome and transcriptome sequencing technologies enhanced our understating of the capsicum genome structure and

function through the development of genetic maps with different molecular markers, dissection of quantitative trait loci underlying economically important traits and subsequently to identify a few genes causing trait variations. And recently, with the help of advanced high-throughput genome sequencing technologies, the whole-genome sequences (both nuclear and organellar) of *Capsicum* species have been reported. The complete sequence of chloroplast genome of *Capsicum annuum*, *C. chinense*, *C. frutescens*, *C. tovarii*, *C. chacoense*, *C. baccatum*, *C. galapagoense, C. eximium* and *C. lycianthoids*, and mitochondrial genome of *C. annuum* has been reported since 2012, while the complete nuclear genome of *C. annuum*, *C. baccatum*, and *C chinense* has been reported since 2014, respectively. Furthermore, the identification of noncoding RNAs and whole genome cytosine methylation which directly or indirectly regulates the gene expression governing the traits of interest in the capsicum genome are being reported.

This book compiles up-to-date information on research and development related to the capsicum crop. The book comprises a total of 14 chapters which starts with the introduction of capsicum crop (Chap. 1) and ends with the capsicum genome databases (Chap. 14). The first two Chaps. (1 and 2) provide a general introduction to the capsicum as a crop, its origin, available reported species, and genetic resources available around the world. Chapter 3 enumerates the history, development, and achievements of classical breeding. Chapter 4 depicts the details of cytological studies, DNA content variations, and phylogenetic relationship of different *Capsicum* species. Chapter 5 describes the development of different molecular markers and construction of genetic maps in *Capsicum* species. Chapters 6 and 7 provide information about the mapping, identification, isolation, and characterization of genes or quantitative trait loci governing economically important traits and biotic and abiotic stress tolerance. Chapters 8 and 9 summarize the sequencing efforts and the findings thereoff such as the structure of nuclear and organelle genomes, nuclear genome expansion, and the presence of repetitive elements in capsicum genome. Chapter 10 contains information on noncoding RNAs and their target genes, and Chap. 11 contains the identification of cytosine methylation in whole genome of *Capsicum annuum*. Chapter 12 gives a glimpse of phylogeny of capsicum in the recent context, and Chap. 13 provides information on application of recent advances in genomics technologies in capsicum breeding. The content of this book ends with Chap. 14 which provides the capsicum genome sequence databases and online tools. These chapters have been authored by 30 eminent scientists from 8 countries including Argentina, China, India, Israel, Italy, Japan, South Korea, and Taiwan. We express our thanks to them for their contributions and cooperation since inception until completion of this book project.

As the book contains all information from genetic resources to the gene and genome sequences, we feel this "The Capsicum Genome" book will serve as a primary resource material and will be very much useful to researchers, breeders, and students working on the capsicum crop

Dr. Nirala Ramchiary expresses his personal thanks and high gratitude to Prof. Chittaranjan Kole, Series Editor of the *Compendium of Plant Genomes*, for giving the opportunity to co-edit this book, and for his constant support and encouragement during editing of this book on "The Capsicum Genome."

Dr. Ramchiary also acknowledges the help extended by his research scholars, Abdul Rawoof and Nitin Kumar, for their assistance in editing and finalizing the chapters. The editors also acknowledge the help from all the staff of Springer Nature at all the stages.

New Delhi, India

Dr. Nirala Ramchiary
Prof. Chittaranjan Kole

Contents

Abbreviations

ABA	Abscisic acid
ABRE	ABA-responsive element
AFLP	Amplified fragment length polymorphism
AgNOR	Argyrophilic NOR
AP-PCR	Arbitrary primed PCR
ARF	Auxin-responsive factor
ASTA	American Spice Trade Association
ATG	Autophagy-related genes
AVRDC	Asian Vegetable Research and Development Center (presently World Vegetable Center)
BAC	Bacterial artificial chromosome
BC_2	Backcross second generation
BGH-UFV	Banco de Germoplasma Hortalicas
BLAST	Basic local alignment search tool
bp	Base pairs
BSA	Bulked segregant analysis
BW	Bacterial wilt
bZIP	Basic leucine zipper
CAPS	Cleaved amplified polymorphic sequence
CAS	CRISPR associated
CATIE	Tropical Agricultural Research and Higher Education Center
CDS	Coding sequence
CGIs	CpG islands
CGMS	Cytoplasmic-genetic male sterility
CGN	Centre for Genetic Resources, the Netherlands (a part of Wageningen University)
cM	CentiMorgan
CMS	Cytoplasmic male sterility
COS	Conserved ortholog set
cp	Chloroplast
CRISPR	Clustered regularly interspaced short palindromic repeats
DAA	Days after anthesis
DAF	DNA amplification fingerprinting

dCAPS	Derived CAPS
ddRAD	Double-digest restriction-site associated DNA
DEG	Differentially expressed gene
DH	Doubled haploid
dsDNA	Double-stranded DNA
EC	Enzyme Commission
EST	Expressed sequence tag
ETC	Electron transport chain
FDA	Food and Drug Administration (of USA)
FISH	Fluorescent in situ hybridization
GABA	Gamma(γ)-aminobutyric acid
Gb	Giga base
GBS	Genotyping by sequencing
GBSS	Granule-bound starch synthesis
GFP	Green fluorescent protein
GMS	Genetic/genic male sterility
GS	Genomic selection
GWAS	Genomewide association studies
HKL	Haploid karyotype length
HMG	High-mobility group
HPLC	High-performance liquid chromatography
HRM	High-resolution melting
IPK	Leibniz Institute of Plant Genetics and Crop Plant Research
IR	Inverted repeat
ISSR	Inter-simple sequence repeat
Kb	Kilobase
KEGG	Kyoto Encyclopedia of Genes and Genomes
KO	KEGG orthology
LCD	Leaf curl disease
LD	Linkage disequilibrium
lncRNA	Long noncoding RNA
LSC	Large single copy
LTR	Long terminal repeat
MAAP	Multiple arbitrary amplicon profiling
MAB	Marker-assisted breeding
MABC	Marker-assisted backcrossing
MAP	Mitogen-activated protein
MAS	Marker-assisted selection
Mb	Mega base
mCs	Methylated cytosine
ME	Mate-pair
miRNA	MicroRNA
MITEs	Miniature inverted transposable element

MLM	Mixed linear model
mncRNAs	Medium noncoding RNA
MPV	Mid-parent value
MSAP	Methylation-sensitive amplified polymorphism
MSL	Male sterile lines
Mya	Million years ago
N50	Minimum contig length
Nat-si	Natural antisense RNA
NCBI	National Center for Biotechnology Information
ncRNA	Noncoding RNA
NGS	Next-generation sequencing
NIL	Near-isogenic lines
NMSU	New Mexico State University
NOR	Nucleolar organizer region
NP	Non-pungency
nt	Nucleotide
Orf/ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PBGD	Porphobilinogen deaminase
PCD	Programmed cell death
PCR	Polymerase chain reaction
PE	Paired-end
PIP	Plasma membrane intrinsic protein
piRNA	Piwi interfering RNA
PR	Pathogenesis-related
pre-miRNA	Precursor microRNA
pri-miRNA	Primary microRNA
PWL	Postharvest water loss
QTL	Quantitative trait locus
QTLs	Quantitative trait loci
RACE	Rapid amplification of cDNA ends
RAD-seq	Restriction site-associated DNA sequencing
RAPD	Random amplified polymorphic DNA
Ra-si	Repeat associated miRNA
RdDM	RNA-directed DNA Methylation
REs	Repetitive elements
RFLP	Restriction fragment length polymorphism
RGA	Resistance genes analog
RIL	Recombinant inbred line
RNAi	RNA interference
ROS	Reactive oxygen species
rRNA	Ribosomal RNA
RuBisCo	R*ibulose 1,5-bisphosphate carboxylase/oxygenase*

SA	Salicylic acid
SCAR	Sequence-characterized amplified region
SCF	Skp–Cullin–F-box
siRNA	Short interfering RNA
SLAF-seq	Specific-locus amplified fragment sequencing
sncRNAs	Small noncoding RNA
snoRNA	Small nucleolar RNA
SNP	Single-nucleotide polymorphism
snRNA	Small nuclear RNA
SSC	Small single copy
SSCP	Single-strand conformation polymorphism
SSD	Single seed descent
SSLP	Simple sequence length polymorphism
SSR	Simple sequence repeat
ssRNA	Single-stranded RNA
STR	Short tandem repeat
TALEN	Transcription activator-like effector nucleases
Ta-si	Transacting miRNA
TCA	Tricarboxylic acid
TE	Transposable element
TF	Transcription factor
TILLING	Targeted induced local lesions in genome
tRNA	Transfer RNA
TSS	Transcription start sites
TTS	Transcription termination site
USDA	United States Department of Agriculture
Vi-si	Viral siRNA
VNTR	Variable number of tandem repeats
WGAS	Whole-genome association study
WGRS	Whole-genome resequencing
WorldVeg	World Vegetable Center, Shanhua, Taiwan

The Capsicum Crop: An Introduction

1

Pasquale Tripodi and Sanjeet Kumar

Abstract

Capsicum (*Capsicum* spp.), also called as pepper, is a main vegetable and spice crop originated in the American tropics and today cultivated all over the world for fresh, dried, and processing products. Around the genus *Capsicum* there is an increasing interest and fascination due to the considerable variation for several traits, which makes this crop extremely versatile and suitable for innumerable uses as food and non-food products. The genus *Capsicum* includes over 30 species, five of which (*C. annuum*, *C. frutescens*, *C. chinense*, *C. baccatum,* and *C. pubescens*) are domesticated and mainly grown for consumption. A large number of accessions of domesticated and wild species are stored in the world seed banks, representing a valuable resource for breeding in order to transfer traits related to resistances to various abiotic and biotic stresses as well for quality improvement. The recent advances in terms of genetic and genomic knowledge will help to unlock the potentiality of these resources. In this chapter, we provide an overview of the origin and history of the pepper, describing its economic importance, properties, and commercial market types.

1.1 Origin and Diffusion

The genus *Capsicum* is part of the large Solanaceae family, which, among the more than 90 genera and 2500 species of flowering plants, includes commercially important vegetables such as tomato, potato, and eggplant. This genus is native to tropical and subtropical America (Hunziker 2001) in a wide region comprising Mexico and northern Central America, the Caribbean, the lowland Bolivia, the northern lowland Amazonia, and the mid-elevation southern Andes, where archaeological evidence suggests use of this spice crop since 6000 BC (Davenport 1970; Basu and De 2003; Perry et al. 2007). At the beginning, fruits were exchanged for black pepper (*Piper nigrum*), a species similar in taste (though not in appearance) although not phylogenetically related to *Capsicum* (Gordo et al. 2012). For this reason, it was incorrectly named "pepper" (Walsh and Hoot 2001).

It was Fuchs, who proposed for the first time in 1543, the botanical term *Capsicum*, which was adopted later in 1753 by Linneo. The name would be the Neolithic derivation of Greek "Capsa," which refers to the peculiar shape of the

P. Tripodi (✉)
CREA Research Centre for Vegetable and Ornamental Crops, 84098 Pontecagnano Faiano, Italy
e-mail: pasquale.tripodi@crea.gov.it

S. Kumar
World Vegetable Center, Post Box 42, Shanhua, Tainan 74199, Taiwan

N. Ramchiary and C. Kole (eds.), *The Capsicum Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-319-97217-6_1

fruit. The crop was firstly introduced in Europe by Christopher Columbus during his travels after the discovery of America in the fifteenth century and later spread to Africa and Asia. Early imported varieties belong to *C. chinense* (Scotch Bonnet or Habanero) which most probably were the most consumed during that time (Walsh and Hoot 2001). The flourishing commercial exchanges of Spanish and Portuguese facilitated the spread of pepper around the globe, with an immediate success due to a well acclimatization in the regions, where they were used as a spice from that part of the population who could not afford to purchase cinnamon, nutmeg, and other spices that are widely used for seasoning and preserving food. To date, the existence of 35 *Capsicum* species is reported (Carrizo García et al. 2016), five of which, namely, *C. annuum*, *C. baccatum*, *C. chinense*, *C. frutescens*, and *C. pubescens* have been domesticated and widespread with different terms depending on the region of cultivation. In Mexico and Central America, the crop is called "chile" which was the ancient name given by local populations of the new world, in American English it becomes "chilli," in Caribbean and countries Latin American countries it is commonly referred to as "ají" and "rocoto," from which derived names of many cultivars of different species today present on the market (i.e., aji Amarillo, aji limon, aji panca, rocoto manzano, rocoto brown, and rocoto de seda). It is also known as pimiento (Spanish), red pepper and pepper (English), pepper (Italian), piment (French), paprika (German and other northern European languages). Overall, the present term "chili pepper" refers to varieties with small and spicy fruits, on the contrary, the term "sweet pepper" refers to varieties with larger fruits and little or no spicy.

1.2 Economic and Culinary Importance

World pepper production has grown considerably over 20 years (1997–2017, www.fao.org/faostat), from 2 to about 4.5 million tons of dry types and from over 17 to 36 million tons as fresh. The area harvested followed a similar trend, with an increase of the surface cultivated area of about 35% in the last 20 years, being today about 3.8 millions of hectares. Fresh pepper is cultivated in 126 countries of the world in all the continents. The world's largest producer is China with over 18 million tons annually, followed by the Mexico with about 3.5 million tons (FAOSTAT 2017). Dry pepper is cultivated in 70 countries and no relevant production is reported in Oceania. India is the largest producer with about 2.0 million tons, followed by Thailand (349.615 tons). Peppers are grown almost all over the world and are fairly easy to cultivate both in the field and in the greenhouse in a wide range of climatic and environmental conditions. Africa, Europe, and America contribute in the same proportion to the total world production (about 10–12% each) for fresh pepper; while for dry pepper, Asia and Africa are the main producers contributing to the 70.3 and 21.2%, respectively (Fig. 1.1). The economic value of pepper production has increased since 1991 becoming a good source of income for producers in many countries and giving an important role in international trading. The present worth of dry pepper is 3.8 billion dollars, while fresh pepper contributes with 30,208 billion dollars. For both, the increase observed over the past 25 years is four times higher in dry pepper and six times higher in fresh pepper.

Around the genus *Capsicum,* there is an increasing interest and fascination due to the amazing diversity in many characteristics, such as plant architecture, flower morphology, fruit typology, colors, pungency, and qualitative traits which make this crop extremely versatile and suitable for innumerable uses. As food, a variety of recipes are ensured thanks to the presence of sweet and hot types. The former are mainly widespread in temperate regions of Europe and North America where they are used freshly or cooked as vegetables. The latter are instead mainly spread in the tropical regions of America, Africa, and Asia, where they are mostly consumed fresh or dried as condiment as spice in powder or salsa in many

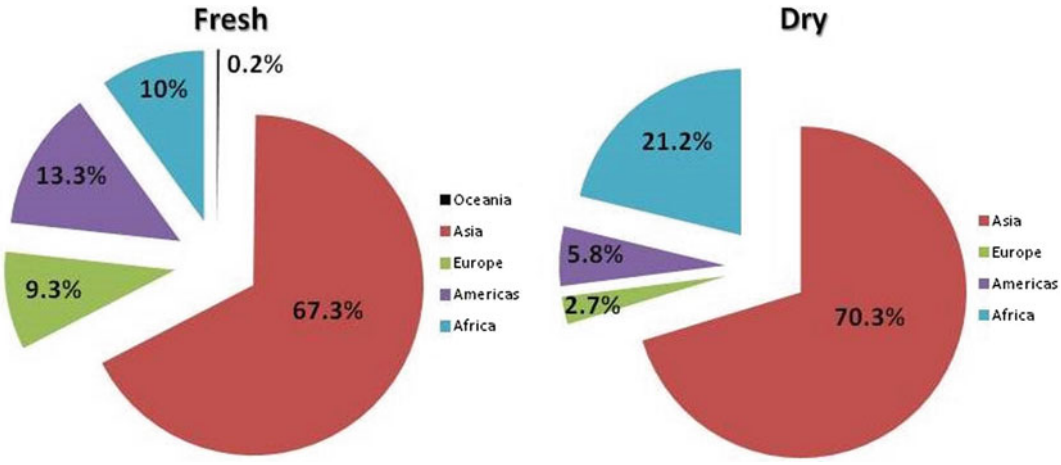


Fig. 1.1 Production share of dry and fresh pepper by region (FAOSTAT 2017)

Fig. 1.2 Examples of popular pod types of hot and sweet peppers. *Photo credit* Susan Lin, World Vegetable Center, Taiwan

dishes. Food uses of peppers could then be summarized in the following classes: (a) fresh use, of immature green fruits, mature red fruits, and leaves; (b) fresh processing, for sauces, pastes, pickles, beer etc.; (c) dried spices, from mature whole fruits and powder (Poulos 1994). Based on pod shape and size, more than 20 market types (e.g. bell, cayenne, ancho, jalapeño, pasilla, Hungarian wax, jwala, and Thai) are commercially cultivated (Fig. 1.2). Furthermore, within each of

these market types, there may be several variants; for instance, bell may have blocky, conical, or mini pods and cherry bell may have small or big pods (Fig. 1.2).

1.3 The Properties of Pepper

The uniqueness of pepper is the typical pungency due to the presence of capsaicinoids. Capsaicinoids are secondary metabolites and derivatives of phenylpropanoids produced in placental epidermis cells and accumulated in structures (blisters) located on the placenta surface (Stewart et al. 2007). The hotness sensation when consumed is given by the interaction with vanilloid receptors, supposed to be a mechanism of defense against mammalian herbivory. Capsaicin and dihydrocapsaicin are the two predominant compounds, accounting for almost 90% of total capsaicinoids. Anti-inflammatory, anticancer, and anti-obesity activities have been recognized within capsaicinoids (Luo et al. 2010). These properties are exerted by the release of substance P, a neurotransmitter involved in pain transmission by nerve (Gamse et al. 1981). Peppers are also an extremely good source of compounds exerting antioxidant properties and responsible for fruit pigmentation. Different colors are encountered in mature fruits as a result of accumulation of carotenoids in chromoplasts during ripening such as capsanthin and capsorubin (mainly in red fruits), violaxanthin and neoxanthin (mainly in yellow fruit), and lutein and β-carotene (mainly in orange fruit) (Gómez-García and Ochoa-Alejo 2013). Fruits are further well-known to have played a leading role in the discovery of vitamin C by Albert Szent-Györgyj, who extracted the first pure chemical compound from Hungarian paprika and was awarded Nobel Prize for Medicine and Physiology in 1937 (http://www.nobelprize.org). Indeed, within *Capsicum* species, a high level of ascorbic acid (vitamin C) able to satisfy the recommended daily intake (FDA 2018, attested to 60 mg for 100 g of raw pepper) is commonly found in both sweet and hot types and widely documented in the literature. High contents of other essential vitamins such vitamin A in the form of β-carotene and vitamins of group B (thiamine, riboflavin, and niacin) are recognized. All these compounds, of which content is determined by species, cultivar, environmental conditions, and maturation stage, exert their biological effects protecting cells against oxidative damage through the interaction with oxygen molecules and scavenging peroxyl radicals (Padayatty et al. 2003; Howard and Wildman 2007). Finally, antimicrobial and antivirulence properties, against *Streptococcus pyogenes*, a major human pathogen (Marini et al. 2015) and *Fusarium* infection (Tewksbury et al. 2008) a polyphagous fungus affecting many vegetables, are reported. All these properties make pepper a good candidate against diseases.

Other than food uses are recognized as active ingredient in cosmetics, pharmaceuticals, and pest management (Bosland and Votava 1999). The extractable colors from fruits due to the presence of compounds unique in pepper such as capsanthin, capsorubin, and cryptocapsin are extensively used in the food processing industry as natural colorant for a wide range of products such meats, cheeses, and other foods. Non-food uses include (a) coloring and flavoring agents, from oleoresins (carotenoids) extracts or powder, as example, paprika powder can be used to inhibit lipid oxidation of pork meat while oleoresin is used to enhance physical and sensory properties of food products (Baenas et al. 2019); (b) ethno-botanical/traditional medicine, from fruit extracts and powders (pungent fruits); (c) modern medicine/pharmaceuticals, from extracts of capsaicinoids and carotenoids which can exert analgesic, antimicrobial, antioxidant, and anti-inflammatory effects; (d) insecticides/repellents and antibacterial effect from capsaicinoids extracts and organic acids (i.e., cinnamic, coumaric, ferulic, and caffeic); (e) spiritual, using whole fruits, e.g., "ristras"; (f) ornamental, using whole plants or fruits; (g) defense/punishment, using capsaicin extracts/or powder (Kumar et al. 2006). The use in cosmesis is favored by the presence of natural compounds which allow to avoid allergies and other side effects and are addressed to protect skin oxidative and UVA-mediated damage having thanks to the anti-wrinkle action and fighting against free

radicals (Baenas et al. 2019). The industrial preparations are based on oleoresins rich of the above-mentioned bioactive compounds. Finally, among the most curious aspects of the *Capsicum* genus, there is certainly the rampant interest of many, searching and collecting, even in urban contexts, different species, characterized by a wide variety traits, as well as ornamental, aesthetically appreciated or rare varieties. This is evident in the rise of associations and websites dedicated to the subject.

1.4 Genetic Resources and Breeding

The *Capsicum* genome has an estimated size of 3.5 Gb and includes mainly diploid species with 12 chromosome ($2n = 2x = 24$). Within the genus, there are also recognized species with 13 chromosomes ($2n = 2x = 26$) as well as one tetraploid species ($2n = 4x = 48$) which is *C. annuum* var. *glabriusculum*, the wild form of the cultivated pepper. Recent investigations have grouped the *Capsicum* species in 11 clades according to main morphological features, provenance, and phylogenetic relationships (Carrizo García et al. 2016) (Table 1.1). The species of greatest interest for consumption and breeding are in three main clades namely: Annuum which includes three domesticated (*C. annuum*, *C. frutescens,* and *C. chinense*) and two wilds (*C. annuum* var. *glabriusculum* and *C. galapagoense*); Baccatum including three forms of *C. baccatum* (var. *baccatum*, var. *pendulum*, and var. *umbilicatum*) and the wilds *C. chacoense* and *C. praetermisssum*; Pubescens which only includes the homonymous domesticated species.

C. annuum is commercially most popular species worldwide. This species is characterized by pungent and non-pungent accessions with herb or sub-shrub growth and fruits having different size, shape, and colors at maturity. *C. frutescens* and *C. chinense* are mainly cultivated in American, Asian, and African countries. The former includes pungent accessions with fruits predominantly small with less than 2 cm of length, the latter instead comprise accessions highly pungent and irregular shape of fruits. The other two domesticated species (*C. baccatum* and *C. pubescens*) are cultivated in Central and South America and are distinguished by particular phenotypic characteristics such as the yellow or green spots in the corolla (*C. baccatum*) or the dark colored seeds (*C. pubescens*). Several wild species are part of the genus *Capsicum* and are principally distributed in the area of origin (Table 1.1). All of them are characterized by very small oval or spherical fruits (Fig. 1.3) with specific distinctive traits related to flower color (white, yellow, and purple with different type of spots), seed color (brownish or black), and flower shape (stellate, rotate, or campanulate) (Barboza and Bianchetti 2005). Although the uniqueness and beautiness distinguish many species of pepper, most of the breeding activities have been carried out within the Annuum clade due to the lack of interspecific barriers between *C. annuum*, *C. chinense,* and *C. frutescens* (Pickersgill 1997; Perry et al. 2007). However, the incompatibility occurring across clades could be overcome using aids such as embryo rescue. Wild and

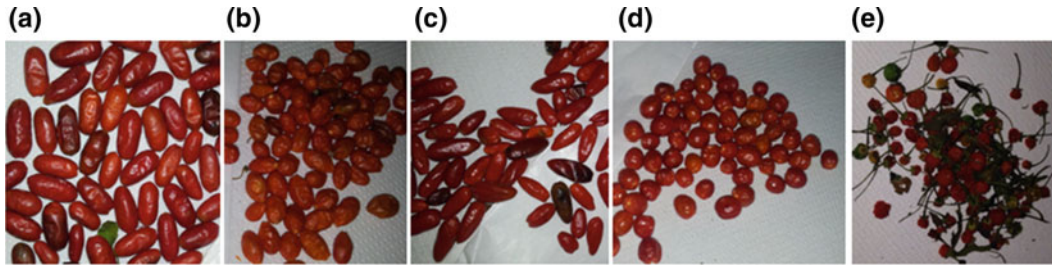


Fig. 1.3 Mature fruits of wild *Capsicum* species: **a** *C. chacoense*, **b** *C. praetermissum*, **c** *C. eximium*, **d** *C. annuum* var. *glabriusculum*, and **e** *C. flexuosum*

Table 1.1 *Capsicum* clades and related species, main features and native area

Clade[a], species name, chromosome number	Pungency	Fruit color[b]	Area of origin[a]
1. Annuum (*x*=12)			
C. annuum	Non-pungent and pungent	Variable	Central and south America regions
C. annuum var. *glabriusculum*	Pungent	Red	Venezuela, central america
C. chinense	Pungent	Variable	Central America, Colombia, Ecuador, south-eastern Brazil, Venezuela
C. frutescens	Pungent	Variable	Central America, central-eastern Brazil, Colombia, Ecuador,Venezuela
C. galapagoense	Pungent	Red	Galapagos Islands
2. Baccatum (x=12)			
C. baccatum var. *baccatum*	Non-pungent and pungent	Variable	Argentina, Bolivia Paraguay, Peru'
C. baccatum var. *pendulum*	Non-pungent and pungent	Variable	Argentina, Bolivia Paraguay, Peru'
C. baccatum var. *umbilicatum*	pungent	Variable	Argentina (north and central), Bolivia (lowlands)
C. chacoense	Pungent	Red	Argentina, Bolivia, paraguay
C. praetermissum	Pungent	Red	South-eastern Brazil
3. Tovarii (x=12)	Pungent		
C. tovarii	Pungent	Red	Perù
4. Pubescens (x=12)	Pungent		
C. pubescens	Pungent	Variable	Argentina, Bolivia, central America, Ecuador, Perù
5. Purple corolla (x=12)	Pungent		
C. cardenasii	Pungent	Red	Bolivia (highlands)
C. eximium	Pungent	Red	Argentina (north and central), Bolivia (lowlands)
C. eshbaughii[*]	Pungent	Red	Bolivia (lowlands)
6. Atlantic forest (x=13)	Pungent		
C. campylopodium	Pungent	Greenish-yellow	South-eastern Brazil
C. cornutum	Pungent	Greenish-yellow	South-eastern Brazil
C. friburgense	Pungent	Greenish-yellow	South-eastern Brazil
C. hunzikerianum	Pungent	Greenish-yellow	South-eastern Brazil
C. mirabile	Pungent	Greenish-yellow	South-eastern Brazil
C. pereirae	Pungent	Greenish-yellow	South-eastern Brazil
C. recurvatum	Pungent	Greenish-yellow	South-eastern Brazil
C. schottianum	Pungent	Greenish-yellow	South-eastern Brazil
C. villosum var. *villosum*	Pungent	Greenish-yellow	South-eastern Brazil

(continued)

Table 1.1 (continued)

Clade[a], species name, chromosome number	Pungency	Fruit color[b]	Area of origin[a]
7. Longidentatum (x=13)			
C. longidentatum	Non-pungent	Greenish-yellow	Central-eastern Brazil
8. Bolivian (x=*)			
C. caballeroi	Pungent	Red	Bolivia (lowlands)
C. minutiflorum	Pungent	Red	Bolivia (lowlands)
C. ceratocalyx	Pungent	Red	Bolivia (highlands)
C. coccineum	Pungent	Red	Bolivia, western Brazil
9. Flexuosum			
C. flexuosum	Non-pungent and pungent	Red	South-eastern Brazil, north-eastern Argentina and eastern Paraguay
10. Caatinga (x=13)			
C. caatingae	Pungent	Greenish-yellow	Central-eastern Brazil
C. parvifolium	Pungent	Greenish-yellow	Central-eastern Brazil, Colombia, Venezuela
11. Andean (x = 12)			
C. rhomboideum	Non-pungent	Red	Central America, Colombia, Ecuador, Perù, Venezuela
C. scolnikianum	Non-pungent	Red	Colombia, Ecuador, Perù
C. geminifolium	Non-pungent	Red	Ecuador, Perù
C. lanceolatum	Non-pungent	Red	Ecuador, Perù, central America
C. dimorphum	Non-pungent	Red	Colombia, Ecuador, Perù

[a]according to Carrizo García et al. (2016)
[b]at maturity stage
*chromosome number not reported

domesticated species have been used particularly for disease resistance and results are widely documented in the literature.

Breeding of pepper, has four main macro-objectives to achieve and related to: (a) main agronomic traits such as yield, fruit features such as color and shape, plant habit, and fruit set; (b) resistances to abiotic stresses such as drought and salinity which limit the cultivation in certain areas; (c) resistances to a plethora of bacterial, fungal, and viral disease causing severe damage to cultivations and loss of quality of the production; (d) quality, for which breeding objectives are mainly related to the improvement of various bioactive compounds such as capsaicinoids, isoprenoids, flavonoids, and vitamin C. The international initiatives aimed to enhance *Capsicum* genetic resources including the progress in breeding and genomics are discussed further in the chapters to be followed.

References

Baenas N, Beovíc M, Llic N, Moreno DA, García-Viguera C (2019) Industrial use of pepper (*Capsicum annum* L.) derived products: technological benefits and biological advantages. Food Chem 274:872–885

Basu SK, De AK (2003) Capsicum: historical and botanical perspectives. In: De AK (ed) Capsicum: the genus *Capsicum*. Taylor & Francis, London

Barboza GE, Bianchetti LDB (2005) Three new species of *Capsicum* (Solanaceae) and a key to the wild species from Brazil. Syst Bot 30:863–871

Bosland PW, Votava EJ (1999) Peppers. Vegetable and spice capsicums. In: Crop production science in horticulture, vol 12. 1st edn. CABI, United Kingdom

Carrizo García C, Barfuss MHJ, Sehr EM, Barboza GE, Samuel R, Moscone EA et al (2016) Phylogenetic relationships, diversification and expansion of chili peppers (*Capsicum*, Solanaceae). Ann Bot 118:35–51

Davenport WA (1970) Progress report on the domestication of Capsicum (chili peppers). Proc Assoc Am Geogr 2:46–47

Marini E, Magi G, Mingoia M, Pugnaloni A, Facinelli B (2015) Antimicrobial and Anti-Virulence activity of Capsaicin against Erythromycin-Resistant, cell-invasive group A Streptococci. Front Microbiol 6:1281

FAOSTAT (2017) www.faostat.fao.org (accessed on 06 January 2019)

FDA (2018) https://www.fda.gov/Food/GuidanceRegulation/GuidanceDocumentsRegulatoryInformation/LabelingNutrition/ucm064928.htm. Last access on 23 May 2018

Gamse R, Lackner D, Gamse G, Leeman SE (1981) Effect of capsaicin pretreatment on capsaicin-evoked release of immunoreactive somatostatin and substance-P from primary sensory neurons. Naunyn Schmiedebergs Arch Pharmacol 316:38–41

Gómez-García MR, Ochoa-Alejo N (2013) Biochemistry and molecular biology of carotenoid biosynthesis in chili peppers (*Capsicum* spp). Int J Mol Sci 14:19025–19053

Gordo SM, Pinheiro DG, Moreira EC, Rodrigues SM, Poltronieri MC, de Lemos OF, da Silva IT, Ramos RT, Silva A, Schneider H, Silva WA Jr, Sampaio I, Darnet S (2012) High-throughput sequencing of black pepper root transcriptome. BMC Plant Biol 12:168

Howard LR, Wildman REC (2007) Antioxidant vitamin and phytochemical content of fresh and processed pepper fruit (*Capsicum annuum*). In: Wildman REC (ed) Handbook of nutraceuticals and functional foods. CRC Press, Boca Raton, FL, pp 165–191

https://www.nobelprize.org/nobel_prizes/medicine/laureates/

Hunziker AT (2001) Genera Solanacearum: the genera of Solanaceae illustrated, arranged according to a new system. Gantner, Ruggell

Kumar S, Kumar R, Singh J (2006) Cayenne/American pepper (*Capsicum* species). In: Peter KV (ed) Handbook of herbs and spices, vol 3. Woodhead Publishing. Cambridge, UK, pp 299–312

Luo X-J, Peng J, Li Y-J (2010) Recent advances in the study of capsaicinoids and capsinoids. Eur J Pharmacol. https://doi.org/10.1016/j.ejphar.2010.09.074

Perry L, Dickau R, Zarrillo S, Hoist I, Pearsall DM, Piperno DR, Berman MJ, Cooke RG, Rademaker K, Ranere A, Raymond S, Sandweiss DH, Scaramelli F, Tarble K, Zeidler JA (2007) Starch fossils and the domestication and dispersal of chili peppers (*Capsicum* spp. L.) in the Americas. Science 315:986–988

Pickersgill B (1997) Genetic resources and breeding of *Capsicum* spp. Euphytica 96:129–133

Padayatty SJ, Katz A, Wang Y, Eck P, Kwon O, Lee JH, Chen S, Corpe C, Dutta A, Dutta SK, Levine M (2003) Vitamin C as an antioxidant: evaluation, of its role in disease prevention. J Am Coll Nutr 22:18–35

Poulos JM (1994) Pepper breeding (*Capsicum* spp.): achievements, challenges and possibilities. Plant Breed Abstr 64:143–155

Stewart C Jr, Mazourek M, Stellari GM, O'Connell M, Jahn M (2007) Genetic control of pungency in *C. chinense* via the *Pun1* locus. J Expt Bot 58:979–991

Tewksbury JJ, Reagan KM, Machnicki NJ, Carlo TA, Haak DC, Penaloza ALC, Levey DJ (2008) Evolutionary ecology of pungency in wild chilies. Proc Natl Acad Sci USA 105:11808–11811

Walsh BM, Hoot SB (2001) Phylogenetic relationships of *Capsicum* (Solanaceae) using DNA sequences from two noncoding regions: the chloroplast atpB-rbcL spacer region and nuclear waxy introns. Int J Plant Sci 162:1409–1418

Genetic Resources of Capsicum

2

Derek W. Barchenger, Ponnam Naresh and Sanjeet Kumar

Abstract

Peppers belong to the genus *Capsicum* of the Solanaceae family and represent plants producing fruits with variable degrees of pungency (highly pungent to nonpungent). Peppers are native to the tropical and temperate Americas. Capsaicinoids (the secondary metabolite responsible for pungency) are uniquely produced in the genus *Capsicum,* which consists of approximately 35 species. There are five widely domesticated and cultivated species (*C. annuum*, *C. frutescens*, *C. chinense, C. baccatum*, and *C. pubescens*). The wild progenitor *C. annuum* var. *glabriusculum* chili pepper (chiltepin) typically has small-round, erect, highly pungent, deciduous, and soft-fleshed fruits. *Capsicum* genetic resources have been successfully used in modern plant breeding programs to develop and commercialize sweet and hot pepper cultivars with diverse market types. Primarily these include breeding for abiotic and biotic stresses and speciality traits for industrial extraction. However, unlike the closely related tomato, the use of wild *Capsicum* germplasm in pepper improvement programs is extremely limited. There are currently no wild *Capsicum* species listed as vulnerable, threatened, or endangered by the US Endangered Species Act. However, this is likely inaccurate as tropical rainforest is being used for agriculture and other forms of habitat modification, resulting in the natural habitat of wild *Capsicum* germplasm being lost. The genetic resources against biotic stresses have the potential to be depleted, due to the rapid evolution of new pathotypes. Therefore, the search for new resistance source against specific pathogens and their deployment in commercial cultivars is a continuous process. Ensuring alignment of national and international policy regulations is needed so that unique *Capsicum* genetic resources are able to be collected, conserved, and distributed, which is critical to the overall success of ex situ conservation.

2.1 Introduction

Peppers [hot pepper (syn. chili or chilli or chile) and sweet pepper (syn. capsicum or bell pepper)] belong to the genus *Capsicum* ($2n = 24$ or $2n = 26$) of the nightshade (Solanaceae) family. There are approximately 35 species in the *Capsicum* genus of which five species have been domesticated and cultivated around the world.

D. W. Barchenger · S. Kumar (✉)
World Vegetable Center, Post Box 42, Shanhua, Tainan 74199, Taiwan
e-mail: sanjeet.kumar@worldveg.org

P. Naresh
Central Horticultural Experiment Station (ICAR-IIHR), Bhubaneswar 751019, India

N. Ramchiary and C. Kole (eds.), *The Capsicum Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-319-97217-6_2

Peppers are native to the tropical and temperate Americas. It is generally agreed that the genus *Capsicum* originated in Bolivia, but the center of domestication of five cultivated species and their dispersal patterns remain largely speculative (Perry et al. 2007), the advent of sequence-based phylogenetics is providing new insights into this area (Carrizo Garcia et al. 2016). The expansion of the genus followed a clockwise pattern around the Amazon basin, toward central and southeastern Brazil, then back to western South America, and finally northward to Central America (Carrizo Garcia et al. 2016). The origin and domestication of the genus *Capsicum*, properties and economic importance of pepper, and a very brief description of genetic resources have been described in the first chapter of this book. In light of advancement of plant molecular genetics and bioinformatics tools, the concept of genetic resources has evolved. Thus in addition to plant genetic resources, *Capsicum* genetic resources also include a variety of other forms of genetic resources such as sequence-based molecular markers, gene sequence data, linkage and physical maps, publically available reference genomes of several *Capsicum* species, organelles, and pan-genome information. However, in this chapter, our discussion is restricted to describing the classical *Capsicum* domestication events, the recent discovery of highly pungent landraces, identification, genetic characterization, and utilization of resistant sources in breeding for major biotic and abiotic stresses and conservation status and future prospects of *Capsicum* plant genetic resources.

2.2 Domestication Syndrome

Domestication syndrome refers to a set of desirable traits that were, or are, being selected by the domesticators (Gepts 2014). Usually after acquiring such favorable traits, domesticated species lose the ability to grow naturally (wild) and thus need human care to complete their life cycle. The wild *C. annuum* var. *glabriusculum* chili pepper (chiltepin) typically has small-round, erect, highly pungent, deciduous, and soft-fleshed fruits. In wild chili, secondary metabolites (capsaicinoids) in fruit have been shown to function to deter fruit consumption by mammals (Tewksbury and Nabhan 2001). In immature fruits/seeds, the function is simple—to deter consumption by granivores and herbivores—but in ripe fruits the function is complex, as consumption can be either beneficial or detrimental, depending on whether the consumer disperses or destroys seeds (Levey et al. 2006). Video monitoring of wild *C. annuum* and *C. chacoense* plants revealed fruit removal only by bird species specializing in lipid-rich fruits, and both chili species had fruit with unusually high lipids. These results supported Tewksbury and Nabhan's (2001) directed deterrence hypothesis and suggested that through secondary metabolites (in this case, capsaicinoids) fruiting plants can recognize seed predators and seed dispersers (Levey et al. 2006). Furthermore, capsaicinoids have antifungal properties and deter invertebrate pests. To this end, domestication of chili and high levels of capsaicinoids seem to be in direct contradiction. It is widely accepted that comparisons of wild and domesticated species provide a basis for elucidating crop domestication syndromes, with wild relatives generally having greater diversity as compared to the domesticated lines. The high level of variability in wild species is especially true for chemical defense strategies, which are generally selected against by domesticators. However, Luna-Ruiz et al. (2018) reported that the variability in capsaicinoid concentration is considerably lower in wild relatives of chili, indicating human-guided selection has led to greater diversity in a trait that would normally not be required in the human-controlled agricultural setting. The major domestication syndromes of chili peppers are: (i) non-deciduous fruits (remain on the plant until harvested manually); (ii) pendent fruits (associated with size increase, better protection from exposure to sun, and avoiding consumtion by birds); and (iii) fruit appearance and varied degree of pungency (associated with consumer preference) (Kumar et al. 2018).

2.3 *Capsicum* Species Complex and Gene Pools

The *Capsicum* gene pool is extensive compared to other crops. In fact, among the five domesticated species and closely related wild species, there are three genetic complexes, the *annuum*, *baccatum*, and *pubescens*, that can be accessed for each of the five domesticated species (Table 2.1). These complexes are based on the degree of genetic proximity and reproductive compatibility, and provide a basis for the primary, secondary, and tertiary gene pools, which are extremely important for plant breeders. The primary gene pool includes members of the same species or closely related species that can be directly hybridized with the species of interest to produce vigorous and fertile progeny. The secondary gene pool includes plants that belong to related species, but the progeny is often sterile or not vigorous. The tertiary gene pool includes species that can be hybridized with the species of interest, but the progeny must go through embryo rescue to be viable. Based on phylogenetic study, species belonging to these gene pools have recently been described under different clades (Garcia et al. 2016). The most widely cultivated species, *C. annuum*, is believed to be domesticated in the highlands of Mexico and includes most of the Mexican chile (syn. chili), most of the chili of Asia and Africa and sweet peppers of temperate countries. For *C. annuum*, the primary gene pool consists of breeding lines, cultivars, and landraces within the species as well as the wild progenitor chiltepin. The secondary gene pool includes *C. baccatum*, *C. chacoense*, *C. chinense*, *C. frutescens*, and *C. galapagoense*, while the tertiary gene pool consists of *C. cardenasii*, *C. eximium*, *C. lanceolatum*, *C. praetermissum*, *C. pubescens*, *C. rhomboideum*, and *C. tovarii.* Due to the non-adaptability of *C. annuum* in lowland tropics of Latin America, its cultivation was replaced by *C. frutescens* and *C. chinense* (Pickersgill 1997). The cultivation of *C. baccatum* and *C. pubescens* is mostly restricted to Latin American countries like Peru, Bolivia, Columbia, and Brazil.

2.4 Discovery of Highly Pungent Landraces

Ever since the initiation of domestication (as reviewed by Bosland and Votava 2012), human-guided selection for diversity in the unique secondary metabolite capsaicinoids continues. In the industrial, pharmaceutical, and nutraceutical markets, high levels of capsaicinoids are desirable for a diverse range of products with uses ranging from crowd control to arthritis treatment (Kumar et al. 2006). In 2000, Naga Jolokia, a landrace from Assam, India, was discovered to be India's hottest (Mathur et al. 2000). A variant of this (Bhut Jolokia) was found to be world's hottest, with a possible origin through natural hybridization between *C. chinense* and *C. frutescens* (Bosland and Baral 2007). This triggered interest in analyzing pungency levels in more germplasm; as a result, a number of highly pungent *Capsicum* landraces have been identified. Trinidad Scorpion, a landrace from Trinidad and Tobago, was then found to be even hotter than Bhut Jolokia, followed by Butch T Scorpion (from Australia) and most recently Trinidad Moruga Scorpion (Bosland

Table 2.1 Three recognized species complexes of *Capsicum*

Complex	Species representative
C. annuum complex	*C. annuum*, C. frutescens*, C. chinense*, C. chacosense, C. galapagoense*
C. baccatum complex	*C. baccatum*, C. praetermissum, C. tovarii*
C. pubescens complex	*C. pubescens*, C. cardenasii, C. eximium*

*Domesticated species

et al. 2012). A number of close variants of Trinidad Scorpion are known from Trinidad and Tobago, e.g., Trinidad 7-Pot Jonah and Douglah Trinidad Chocolate, which are genetically distinct from Bhut Jolokia (Bosland et al. 2012). Highly pungent variants (landraces) of Bhut Jolokia are also known, and they are believed to have originated from sympatric domesticated species (Rai et al. 2013). The natural variability within and between these landraces is currently maintained by local communities and warrants ex situ conservation (Kumar et al. 2018).

2.5 Discovery and Utilization of Genetic Resources for Stress Resistance

Capsicum genetic resources have been successfully used in modern plant breeding programs to develop and commercialize sweet and hot pepper cultivars with diverse market types. These cultivars are bred for resistance to one or more biotic and abiotic stresses using genes from landraces and different cultivated species. The status of germplasm screening and characterization against selected biotic and abiotic stresses and their utilization in breeding programs is discussed below.

2.5.1 Resistance to Biotic Stresses

2.5.1.1 Anthracnose

Different species of *Colletotrichum* (primarily *C. scovillei, C. truncatum,* and *C. gloeosporioides*) cause one of the most serious diseases of pepper, anthracnose. This disease usually causes pre- and post-harvest fruit rot and is likely the most devastating disease of chili pepper production under hot, humid tropical and subtropical conditions. While *C. annuum* is the most commercially important chili pepper species, it is generally considered to be an anthracnose-susceptible species with no known sources of resistance within the *C. annuum*. Broad-spectrum resistance sources were identified in *C. baccatum* (PBC80, PBC81, PI594137, PI497985-1, and PI260550) and *C. chinense* (PBC932) and used to improve resistance in *C. annuum* through introgression breeding (Gniffke et al. 2013). The variable mode of inheritance of resistance to anthracnose induced by *Colletotrichum* isolates is known. For example, resistance to *C. scovillei* was found to be controlled by a single recessive gene (Kim et al. 2008; Mahasuk et al. 2009), duplicate genes (Lin et al. 2007), and genes with quantitative effects and quantitative trait loci (QTLs; Lee et al. 2010). These variations are attributed to the use of different susceptible parents, pathogen species and pathotypes, fruit maturity stages, and screening (inoculation) methods (Suwor et al. 2015). Resistance from PBC932 was introgressed to *C. annuum* lines through conventional backcrossing (Gniffke et al. 2013). Likewise, resistance from *C. baccatum* PBC80 was introgressed into *C. annuum* by conventional hybridization via bridge cross (*C. annuum* × *C. chinense*; Yoon and Park 2005) and also by using embryo rescue to overcome interspecific hybridization barriers (Yoon et al. 2006). The Sequence Characterized Amplified Region-Indel (SCAR-Indel) and Simple Sequence Repeats (SSR-HpmsE032) markers associated with anthracnose resistance proved to be useful for MAS in *C. annuum* introgression lines derived from PBC932 and *C. baccatum* PBC80 (Suwor et al. 2017), respectively. Among a number of reports on QTLs associated with anthracnose resistance, the most recent report utilized two populations from *C. annuum* Bangchang × *C. chinense* PBC932, and *C. baccatum* PBC80 × *C. annuum* CA1316 crosses and identified flanking single-nucleotide polymorphism (SNP)-based markers that can be used to select quantitative trait loci (QTLs) through marker-assisted selection (MAS); however, primer sequences for these markers were not provided in the publication (Mahasuk et al. 2016).

2.5.1.2 Bacterial Wilt

Bacterial wilt disease is characterized by sudden death of pepper plants, particularly in the hot and humid regions (tropics). Disease is caused by

Ralstonia solanacearum, a soilborne pathogen, which is very difficult to manage using chemical and cultivation strategies. In general, among the hot and sweet peppers, sweet pepper genotypes are more susceptible to bacterial wilt and many hot pepper genotypes resistant to bacterial wilt are known (Gniffke et al. 2013). Three pathotypes (*P-1, P-2*, and *P-3*) of *R. solanacearum* and pathotype-specific resistant chili pepper germplasm are known (Lebeau et al. 2011). Broad-spectrum resistant genotypes, viz., MC-4, PBC1347, PBC066 (screened against a Malaysian isolate), PBC631 (screened against a Sri Lankan isolate), and PBC473 (screened against an Indonesian isolate), have been identified (Lopes and Boiteux 2004). In India, a number of resistant sources have been identified through All India Coordinated Vegetable Improvement Project. Genetics of bacterial wilt resistance has been found to be monogenic recessive in Anugraha (Thakur et al. 2014) and polygenic in LS2341 (Mimura et al. 2009). An SSR marker (*CAMS451*) linked to major resistant QTL '*Bw1*' is reported (Mimura et al. 2009).

2.5.1.3 Phytophthora Root Rot

Causing an estimated $100 million USD in global losses annually, the soilborne oomycete *Phytophthora capsici* is likely the most destructive pathogen to pepper production worldwide (as reviewed by Barchenger et al. 2018a). *Phytophthora capsici* exhibits a very high level of diversity due to sexual reproduction between two mating types (A1 and A2). Furthermore, *P. capsici* can also have long-lived and widespread clonal lineages in tropical and subtropical regions. Based on a pathogen surveillance and race characterization study, it is anticipated that A2 mating type is becoming more prevalent in Taiwan (Barchenger et al. 2018b), while a ratio that does not significantly deviate from a 1:1 (A1:A2) has been reported in many regions around the world (Barchenger et al. 2018a). Sources for *P. capsici* resistance have been identified in *C. annuum* landraces including Criollo de Morelos-334 (CM334), PI201232, PI201234 from Mexico, and Perennial from India. Being the most highly resistant accession identified to date, CM334 is commonly used as a resistance source, as it has broad-spectrum resistance against different isolates of the pathogen (Sy et al. 2008; Quirin et al. 2005). However, at WorldVeg, resistance breeding program preferred to use resistance derived from PI201232 and PI201234 because of undesirable nonprolific fruit-bearing habits in the backcrossed progenies derived from CM334. The inheritance of resistance in CM334 has been reported as being variable (recessive to dominant polygenic with additive and epistatic effects) depending on the use of susceptible parent and screening methodology (Lefebvre and Palloix 1996). In CM334, common QTLs, *Phyto-U* (Ogundiwin et al. 2005), *Phyto 5.2* (Quirin et al. 2005), and *Pc 5.1* (Mallard et al. 2013), are associated with the resistance to *P. capsici* and confined to chromosome *5* (Minamiyama et al. 2007; Kim et al. 2008). In PI201234, an SSR marker (*ZL6726*) linked to the resistance gene *CaPhyto* has been validated and could be used in MAS in resistance breeding against *P. capsici* race 2 (Wang et al. 2016).

2.5.1.4 Viruses

Hot and sweet peppers are prone to virus infection and can be infected by more than 60 different species of viruses in different parts of world (Gniffke et al. 2013). The symptoms induced by virus infection in pepper range from very mild leaf chlorosis to severe leaf curl, plant stunting, necrosis, dieback, and death. However, it is rarely possible to make an accurate diagnosis based on symptoms alone, because two viral species can induce very similar symptoms (Kenyon et al. 2014). Different species of viruses within the genera *Potyvirus*, *Cucumovirus*, *Tospovirus*, *Tobamovirus*, and *Begomovirus* cause significant economic losses in different pepper production systems and regions (Kenyon et al. 2014). Due to their broad host range and large number of insect vectors, complete control of these viruses is difficult. Thus, virus resistance breeding is one of the main objectives of pepper breeders worldwide.

Potyviruses

Potyviruses (ssRNA) perhaps remain the most prevalent among the viruses infecting peppers in many regions including Southeast Asia. The major potyviruses transmitted by aphid in non-persistent manner are *Chilli veinal mottle virus* (ChiVMV), *Pepper mottle virus* (PepMoV), *Potato virus Y* (PVY), and *Tobacco etch virus* (TEV). Over the past five decades, a number of resistant sources with variable modes of inheritance (including polygenic) have been reported and extensively used in breeding programs. Two resistant sources (IHR2451 and IHR4503), with monogenic recessive gene action, have recently been identified (Naresh et al. 2016). Most of these potyvirus recessive resistance genes (e.g., *pvr1, pvr2*, and *pvr6*) in peppers (and other crops) have been confirmed to encode eukaryotic translation initiation factors (*eIF4E*) or their isomers (Kang et al. 2005; Wang and Krishnaswamy 2012). Hence, potyvirus particle requires wild *eIF4E* gene (producing cap-binding protein) to complete their infection cycle in the host cells to generate susceptible reaction. A number of recessive and dominant genes conferring resistance to different potyvirus species in *C. chinense* (PI159236, PI152225, PI159236) and *C. annuum* (Yolo Y, Florida VR2, Avelar, CM334, Perennial, NW4) have been identified (Lee et al. 2013). Sequence analyses of the *Pvr7* flanking markers and the *Pvr4*-specific gene markers have shown that dominant resistant genes *Pvr7* and *Pvr4* are the same gene mapped on physical interval of 258 kb on chromosome 10 (Venkatesh et al. 2018).

Cucumovirus

Within the genus *Cucumovirus* (ssRNA), *Cucumber mosaic virus* (CMV), transmitted nonpersistently by aphids, is the most common species to infect peppers. Viral infection can cause severe systemic mosaic symptoms such as leaf distortion and fruit lesion, which result in drastic yield loss. The damage caused by CMV depends mainly on the stage of plant at infection, strains prevalent and cultivars grown in a particular region. The CMV resistance sources have been identified among several different cultivated species, including Sapporo-oonaga, Nanbu-onaga (Suzuki et al. 2003), Bukang (Kang et al. 2010), BJ0747 (Yao et al. 2013), IHR2451, and IHR4503 (Naresh et al. 2016) in *C. annuum*; Pen-3 (Nono-Womdim et al. 1991) and PI439381-1-3 (Suzuki et al. 2003) in *C. baccatum*; and BG2814-6 (Grube et al. 2000a), LS1839-2-4 (Suzuki et al. 2003), and PBC688 (Guo et al. 2017) in *C. frutescens*. The mode of inheritance of CMV resistance genes has been reported to be monogenic recessive (Pochard and Daubeze 1989), partially dominant (Lapidot et al. 1997), dominant gene cluster (Grube et al. 2000b), a single dominant in Bukang (Kang et al. 2010), and polygenic in *C. annuum* cv. Perennial (Chaim et al. 2001; Naresh et al. 2016). Major and minor QTLs (Caranta et al. 2002) and two major QTLs in BJ0747-1-3-1-1 have also been reported (Yao et al. 2013). In Bukang, a cleaved amplified polymorphic sequence (CAPS) marker has been identified that is linked to *Cmr1*, a dominant resistant gene located on centromeric region of LG2 (Kang et al. 2010). Recently in PBC688, two key QTLs, *qCmr2.1* (chromosome *2*) and *qCmr11.1* (chromosome *11*) were identified and CA02g19570 was suggested to be possible candidate gene of *qCmr2.1* for resistance to CMV (Guo et al. 2017).

Begomovirus

Over the past decade, *Begomovirus* (ssDNA and dsDNA) has emerged as one of the most serious production constraints of pepper, especially in Asia. Leaf curl disease associated with members of *Begomovirus* causes severe growth retardation with symptoms of reduced leaf size, chlorosis, and severe curling of the leaves causing failure of the crop. There are numerous species of chili-infecting begomoviruses, and every year new species are identified. The tendency for genetic recombination, the acquisition of extra DNA components, and the synergistic interaction among different begomoviruses have resulted in the rapid emergence of new viruses that can infect new hosts, cause new disease symptoms, and overcome host resistance (Varma and Malathi 2003; Chakraborty et al. 2003; Singh

et al. 2016). In order to understand the basis for and to predict epidemic outbursts and global spread of *Begomovirus*, Jabłońska-Sabuka et al. (2015) utilized a mathematical model. They found that intensive farming and breeding resistant cultivars were the major triggers for aggressive virus adaptability through mutation speedup. In fact, it has been estimated that farmers that adopt low, medium, and high integrated management strategies for *Begomovirus* could improve incomes by 17, 26, and 80%, respectively (Swaminathan et al. 2016).

Sources of resistance to *Begomovirus* in pepper have been identified, including the *C. chinense* accessions BG-3821 in Mexico (Anaya-Lopez et al. 2003) and 'Bhut Jolokia' in India (Adluri et al. 2017). The *C. annuum* accessions DLS-Sel-10, WBC-Sel-5, PBC142, PBC145, PBC345 (Srivastava et al. 2015; 2017), PBC143, PBC144, PBC149, PBC495, VI012005 (Kenyon et al. 2014), GKC-29, BS-35, EC-497636, Kalyanpur Chanchal (Singh et al. 2016), and S-343 (Thakur et al. 2019) have been reported to be resistant. However, different methods were used to identify these sources of resistance, including observations made in the field when planted with highly symptomatic susceptible lines. Inheritance of resistance in BG3821 appears to be controlled by two genes with duplicate recessive epistatic relation (Garcia-Neria and Rivera Bustamante 2011). Similarly, for 'Bhut Jolokia,' a single recessive gene has been reported to control resistance to *PepLCV* (Rai et al. 2014). Contrastingly, resistance to chili leaf curl disease (ChiLCD) in S-343 has been found to be monogenic dominant (Thakur et al. 2019).

Tospoviruses

Tospoviruses (ssRNA), transmitted by thrips in a persistent manner, are increasingly becoming a more serious threat to pepper production. *Tomato spotted wilt virus* (TSWV) is major virus infecting peppers worldwide (Momol et al. 2000). Resistant accessions are available mostly within the species *C. chinense*. For example, PI152225 and PI159236 are resistant (hypersensitive reaction) to TSWV and contain the dominant resistance gene *Tsw* (Boiteux et al. 1993; Moury et al. 1997). The *Tsw* gene has been tagged by molecular markers and mapped to chromosome *10* (Jahn et al. 2000; Moury et al. 2000). The other *C. chinense* resistant sources with hypersensitive reactions are PI159234, CNPH 275, PI-15, C00943, and ECU-973 (Jahn et al. 2000; Cebolla-Cornejo et al. 2003). Resistance in *C. chinense* AC09-207 has been reported to have monogenic dominant gene action and is nonallelic genes in PI152225 and PI159236 (Hoang et al. 2013). The *C. baccatum* accession PIM26-1 has been identified as a resistant source against a wild and resistance-breaking aggressive isolate of TSWV (Soler et al. 2015).

Tobamoviruses

Tobamoviruses (ssRNA) are also important pepper-infecting viruses that are not transmitted by an insect vector. External seed contamination, internal seed infection, and contact are the primary modes of transmission for tobamoviruses. The mixed infection of several tobamoviruses is common. Symptoms usually occur on leaves as mosaic and mottle, but systemic necrosis and necrotic spots can also occur. In pepper, resistance to tobamoviruses is conferred by the localization (*L*) locus located on chromosome *11*. The seven known pepper-infecting species of tobamoviruses have been classified in terms of pathogenicity as pathotypes (P_0, P_1, $P_{1,2}$, $P_{1,2,3}$, and $P_{1,2,3,4}$) based on their ability to overcome the resistance associated with the corresponding *L* alleles (cited in Hsu et al. 2018). *Tomato mosaic virus* (ToMV) is prevalent in Southeast Asia and Southern Europe (Moury and Verdin 2012). A recent study conducted by us revealed a unique situation of symptom-specific utilization of markers in peppers, where a L^2 allele-linked molecular marker developed in *C. frutescens* was completely effective for selection in diverse *C. annuum* sweet pepper germplasm (Hsu et al. 2018).

2.5.1.5 Root-Knot Nematode

Root-knot nematodes (genus *Meloidogyne*) are polyphagous pest causing root nodules (gall) leading to stunted growth and chlorosis of pepper plants. There are more than 70 species of *Meloidogyne*; however, *M. javanica*, *M. arenaria*,

M. incognita, and *M. hapla* are major pepper-infecting nematodes worldwide (Sanchez-Puerta and Masuelli 2011). A number of resistant sources including Punjab Tej, Japani Longi, JCA-288, Perennial, and CM334 are known (Sarath Babu et al. 2011). The accessions PI322719, PI201234, and CM334 possess highly stable and dominant gene conditioned resistance across the different species of *Meloidogyne.* Approximately, 20 resistant genes (*Me*) have been reported in peppers (Wang and Bosland 2006). The genes *Me1, Me3,* and *Me7* are stable and control a wide resistance against different species (*M. arenaria*, *M. javanica*, and *M. incognita*). These genes are clustered at 28 cM interval on chromosome P9, and linked markers (SCAR, SSCP, and CAPs) were developed (Djian-Caporalino et al. 2007). A SCAR marker linked to *N* gene (allelic to *Me7*) has also been developed (Wang et al. 2009; Fazari et al. 2012).

2.5.1.6 Insects

Arthropod pests are major production constraints for pepper worldwide. The species infecting peppers varies based on region as well as production system (greenhouse vs. open field). Thrips (*Thrips palmi*, *Frankliniella occidentalis*), broad mite (*Polyphagotarsonemus latus*), aphid (*Myzus persicae*), and spider mites are major pests of peppers. Besides causing direct damage to plant and yield loss, thrips and aphids are vectors for pathogenic viruses. The *C. annuum* accessions PBC145 and C00069 were found resistant to all three pathogens (Gniffke et al. 2013). The *C. annuum* accession AC1979 was highly resistant against thrips (Maharijaya et al. 2011, 2012). Resistant sources for mites (*P. latus)* have been reported from India, including IIHR-243-1-1-15 and Musalwadi selection (Borah 1987), Jwala, G-5, Pant C1 (Naitam et al. 1990), PMR-21, KDSC-210 (Mallapur 2000), EC378630, EC378633, EC391082, IC214991, NIC23897 (Sarath Babu et al. 2002), IC342390, IC572492, IC337281, and IC344366 (Rameash et al. 2015). The first QTL for resistance to thrips was reported on chromosome *5* (Linders et al. 2010). Additionally, in AC 1979, a major QTL for resistance to thrips was located on chromosome *6* (Maharijaya et al. 2015).

2.5.2 Tolerance to Abiotic Stress

Since the mid-twentieth century, many regions of the world have experienced considerable changes in the nature of droughts, floods, and extreme weather events (Lesk et al. 2016). Therefore, breeding for tolerance to abiotic stress is becoming an increasingly important component of many breeding programs. The first step in a successful breeding program is the identification of genetic resources with the traits of interest.

2.5.2.1 Moisture Deficit Tolerance

The moisture deficit stress (drought) causes a significant reduction in pepper plant growth and yield (Kirada et al. 2007), whereas capsaicinoid content increases under drought stress (Bosaland and Votava 2012; Sung et al. 2005). *C. chinense* IHR4502 have been identified as drought tolerant with strong and deep root system (Naresh et al. 2017).

2.5.2.2 Heat and Salt Tolerance

Pepper, especially sweet pepper, production is adversely affected by high temperatures (>32°C), humidity, and low light intensity. Under heat stress conditions, sweet pepper root and shoot growth are seriously affected (Aloni et al. 1992), leading to flower abscission and reduced pollen viability, fruit set, and finally total marketable yield. The optimum growing conditions for sweet pepper are day/night amplitude of 7–9 °C and 24 h mean temperature of 21–23 °C (Bakker and Van Uffelen 1988).

Several methods to screen heat tolerant sweet pepper lines have been developed, such as *in vitro* pollen germination and pollen tube length (Reddy and Kakani 2007), correlation between root temperature of seedlings, photosynthetic rate, stomata aperture, and intercellular CO_2 (Feng and Jiang 2000). Chilly Chili, Medusa, Thai Hot, Explosive Ember, and Treasures Red were identified as heat tolerant based on cumulative temperature response index (Gajanayake et al. 2011). The sweet pepper line AVPP9823 and the hot pepper line AVPP9905 are resistant to heat under field conditions. On the basis of heat susceptibility index, hot pepper

Pepsi-17-2 was found to be heat tolerance (Kaur et al. 2016).

The growth and yield of chili are adversely affected by salinity (Zhani et al. 2012), and it is estimated that each unit of salinity there will be 14% reduction in yield (Munns and Tester 2008). The varieties CO1, K1, Jayanthi, Arka Suphal, and accession EC497636 were reported to be highly saline tolerant (Balasankar et al. 2017). Overall, there is a general lack of genetic resources for tolerance to abiotic stress in pepper. Potentially, wild relatives can be useful sources of genetic variability in breeding for climate change adaptation (Prohens et al. 2017). One such candidate, chiltepins, evolved under marginal conditions, such as low rainfall and high temperature. However, more studies need to be conducted in this area.

2.6 Conservation of Genetic Resources

Unlike the closely related tomato, the use of interspecific hybrids in pepper is extremely limited, especially for wild *Capsicum* relatives (Mongkolporn and Taylor 2011). There are several reasons for this, but the primary reason is likely the overall lack of phenotypic data available for the species of *Capsicum*, outside of *C. annuum* (Barchenger and Bosland 2019). Another reason for the limited use of interspecific hybrids in pepper is an overall lack of access to these germplasms to many modern breeding programs. However, there are several public germplasm repositories that house both domesticated and wild *Capsicum* germplasm.

2.6.1 Ex Situ

The public germplasm repositories distribute plant material for research, education, training, and developmental purposes. Availability of seed is dependent upon quantity in the collection and national or international regulation of movement of plant material. Furthermore, concerns regarding phytosanitary issues limit or completely prevent germplasm distribution among international germplasm repositories and to plant breeders and other scientists internationally. A comprehensive understanding of gene bank coverage and gaps for *Capsicum* is lacking in part due to insufficient collaboration among public germplasm repositories. There are numerous collections of *Capsicum* germplasm at both the national and the international levels. The largest two collections are housed at the World Vegetable Center (WorldVeg) in Shanhua, Taiwan, and the United States Department of Agriculture (USDA) in Griffin, GA, USA. However, the degree of overlap in these two collections is not currently known, and it is predicted to be quite high. As of 2018, the current holdings of *Capsicum* at WorldVeg were 7178 accessions, which included 5480 (*C. annuum*), 740 (*C. frutescens*), 504 (*C. chinense*), 380 (*C. baccatum*), 30 (*C. pubescens*), 25 (*C. chacoense*), 9 (*C. praetermissum*), 4 (*C. eximium*), 3 (*C. tovarii*), 2 (*C. galapagoense*), and 1 (*C. lanceolatum*). The US National Plant Germplasm System has 1000 (*C. annuum*), 492 (*C. chinense*), 383 (*C. baccatum*), 280 (*C. frutescens*), 45 (*C. pubescens*), 19 (*C. chacoense*), 1 each of *C. eximium*, and *C. galapagoense* accessions. The Tropical Agricultural Research and Higher Education Center (CATIE) germplasm database houses 884 *Capsicum* accessions, and Chile Pepper Institute at the New Mexico State University (NMSU) maintains a large collection of both domesticated and wild *Capsicum* accessions. National gene banks in India, South Korea, China, the Netherlands, France, and Japan also maintain considerable number of collections of *Capsicum* germplasm.

2.6.2 In Situ

There are currently no wild *Capsicum* species listed as vulnerable, threatened, or endangered by the US Endangered Species Act. However, this is likely inaccurate as tropical rainforest is giving way to agriculture and other forms of habitat modification, resulting in the natural habitat of wild *Capsicum* germplasm being lost. Although

efforts have been made to collect and conserve wild *Capsicum* species ex situ, little has been done to protect the natural habitats and the native populations of these species (Tewksbury et al. 1999). Additionally, the species have economic importance, potentially increasing conscious maintenance of populations (Pagán et al. 2010). However, overexploitation has potentially resulted in the extinction of some populations of wild species (González-Jara et al. 2011; Nabhan, 1990). Similar potential threat of *Capsicum* germplasm extinction also exists in northern Himalayan region, which is considered to be secondary center of diversity for the genus *Capsicum* (Rai et al. 2013). Measures to conserve wild and managed populations of the wild *Capsicum* relatives should be implemented to maintain the source and the architecture of genetic variation (González-Jara et al. 2011).

2.7 Future Outlook

The genetic resources against biotic stresses have the potential to be depleted, due to the rapid evolution of new pathotypes. Therefore, the search for new resistance source against specific pathogens and their deployment in commercial cultivars is a continuous process. Wild *Capsicum* relatives have been underutilized in pepper breeding programs (Mongkolporn and Taylor 2011). All known examples of introgression breeding have been limited to interspecific hybridization and backcrossing between five domesticated species. For instance, in most widely cultivated *C. annuum*, resistant genes for anthracnose and tospovirus were introgressed from *C. baccatum* and *C. chinense*. Hence, it would be worthwhile to assemble and initiate screening *Capsicum* wild species against prioritized pathogens and abiotic stresses. This will possibly allow identification of novel sources of resistance for future breeding use. The sequence data information (including whole genome sequence; Kim et al. 2014; Qin et al. 2014) on *Capsicum* genetic resources is expected to be increasing at a tremendous pace. These data could be used as screening and breeding tools for various traits in pepper breeding. It is also expected that hurdle of lack of widely applicable molecular markers in pepper could be solved through bioinformatic tools and precise phenotyping. Ensuring that national and international policies and regulations are aligned so that unique *Capsicum* genetic resources are able to be collected, conserved, and openly distributed is critical to the overall success of ex situ conservation (Perramond 2005). Hence, germplasm conservation may best be promoted through better characterization and evaluation of current collections, both phenotypically and genotypically, and through building information systems that facilitate access to these data (Barchenger and Bosland 2019). Finally, greater awareness of the value and threats to these wild resources are needed in order to generate the momentum to better conserve them in situ and ex situ.

References

Adluri P, Baldoldiya GM, Nath P (2017) Screening of Bhut Jolokia (*Capsicum chinense* Jacq.) germplasm of North East India against chili leaf curl virus. Int J Pure Appl Biosci 5:1189–1196

Aloni B, Karni L, Daie J (1992) Effect of heat stress on the growth, root sugars, acid invertase and protein profile of pepper seedlings following transplanting. J Hort Sci 67:717–725

Anaya-Lopez JL, Torres-Pacheco I, Gonzalez-Chavira M, Garzon-Tiznado JA, Pons-Hernandez JL, Guevara-Gonzalez RG et al (2003) Resistance to geminivirus mixed infections in Mexican wild peppers. HortScience 38:251–255

Bakker JC, Van Uffelen JAM (1988) The effects of diurnal temperature regime on growth and yield of glasshouse sweet pepper. Neth J Agric Sci 36:201–208

Balasankar D, Praneetha S, Arumugam T, Jeyakumar P, Manivannan N, Arulmozhiselvan K (2017) Genotypic response of chilli (*Capsicum annuum* L.) on germination and seedling characters to different salinity levels. Int J Curr Microbiol Appl Sci 6:2197–2205

Barchenger DW, Bosland PW (2019) Wild chile pepper (*Capsicum sp.*) of North America. In: Greene S, Williams K, Khoury CK, Kantar MB, Marek L (eds) North American Crop Wild Relatives vol II. Springer International Publishing, New York, NY (in press)

Barchenger DW, Sheu ZM, Kumar S, Lin SW, Burlakoti R, Bosland PW (2018a) Race characterization of *Phytophthora* root rot *on Capsicum* in Taiwan as a basis for anticipatory resistance breeding. Phytopathology 108:964–971

Barchenger DW, Lamour KL, Bosland PW (2018b) Challenges and strategies for breeding resistance in *Capsicum annuum* to the multifarious pathogen, *Phytophthora capsici*. Front Plant Sci 9:628. https://doi.org/10.3389/fpls.2018.00628

Boiteux LS, Nagata T, Dutra WP, Fonseca MEN (1993) Sources of resistance to *tomato spotted wilt virus* (*TSWV*) in cultivated and wild species of *Capsicum*. Euphytica 67:89–94

Borah DC (1987) Bioecology of *Polyphagotarsonemus latus* (Banks) (Acari: Tarsonemidae) and *Scirtothrips dorsalis* Hood (Thysanoptera: Thripidae) infesting chilli and their natural enemies. Ph.D. thesis submitted to the University of Agricultural Sciences, Dharwad, Karnataka, India, p 330

Bosland P, Baral JB (2007) 'Bhut Jolokia'—the world's hottest known chile pepper is a putative naturally occurring inter specific hybrid. HortScience 42:222–224

Bosland PW, Coon D, Reeves G (2012) 'Trinidad Moruga Scorpion' pepper is the world's hottest measured chile pepper at more than two million scoville heat units. HortTechnology 22:535–538

Bosland PW, Votava EJ (2012) Peppers, vegetable and spice capsicum, 2nd edn. CABI, Wallingford, UK

Caranta C, Pxieg ES, Lefebvre V, Daubeze AM, Thabuis A, Palloix A (2002) QTLs involved in the restriction of *Cucumber mosaic virus* (CMV) long-distance movement in pepper. Theor Appl Genet 104:586–591

Caranta C, Thabuis A, Palloix A (1999) Development of a CAPS marker for the *Pvr4* locus: a tool for pyramiding potyvirus resistance genes in pepper. Genome 42:1111–1116

Carrizo Garcia C, Barfuss Michael HJ, Sehr EM, Barboza GE, Samuel R, Moscone AE, Ehrendorfer F (2016) Phylogenetic relationships, diversification and expansion of chili peppers (*Capsicum*, Solanaceae). Ann Bot 118:35–51

Cebolla-Cornejo J, Soler S, Gomar B, Soria F (2003) Screening *Capsicum* germplasm for resistance to *Tomato spotted wilt virus* (*TSWV*). Ann Appl Biol 143:143–152

Chaim BA, Grube RC, Lapidot M, Jahn M, Paran I (2001) Identification of quantitative trait loci associated with resistance to Cucumber mosaic virus in *Capsicum annuum*. Theor Appl Genet 102:1213–1220

Chakraborty S, Pandey PK, Banerjee MK, Kalloo G, Fauquet CM (2003) Tomato leaf curl Gujarat virus: a new *Begomovirus* species causing severe leaf curl disease of tomato in Varanasi, India. Phytopathology 93:1485–1495

Djian-Caporalino C, Fazari A, Arguel MJ, Vernie T, VandeCasteele C, Faure I, Brunoud G, Pijarowski L, Palloix A, Lefebvre V, Abad P (2007) Root-knot nematode (*Meloidogyne* spp.) *Me* resistance genes in pepper (*Capsicum annuum* L.) are clustered on the P9 chromosome. Theor Appl Genet 114:473–486

Fazari A, Palloix A, Wang LH, Hua MY, Sage-Palloix AM, Zhang BX, Djian-Caporalino C (2012) The root-knot nematode resistance N-gene co-localizes in the Me-genes cluster on the pepper (*Capsicum annuum* L.) P9 chromosome. Plant Breed 131:665–673

Feng YL, Jiang SM (2000) Effect of root system temperature on physiological characteristics of thick red pepper (*Capsicum annuum* var. *grossum*). Plant Physiol Commun 36:308–311

Gajanayake B, Trader BW, Raja Reddy K, Harkess RL (2011) Screening ornamental pepper cultivars for temperature tolerance using pollen and physiological Parameters. HortScience 46:878–884

Garcia-Neria MA, Rivera-Bustamante RF (2011) Characterization of *Geminivirus* resistance in an accession of *Capsicum chinense* Jacq. Mol Plant-Microbe Inter 24:172–182

Gepts P (2014) The contribution of genetic and genomic approaches to plant domestication. Curr Opin Plant Biol 18:51–59

Gniffke PA, Shieh SC, Lin SW, Sheu ZM, Chen JR, Ho FI, Tsai WS, Chou YY, Wang JF, Cho MC, Roland S, Kenyon L, Ebert AW, Srinivasan R, Kumar S (2013) Pepper research and breeding at AVRDC—the world vegetable center. In: Proceedings of XV EUCARPIA meeting on genetics and breeding of capsicum and eggplant, Turin, Italy, 2–4 September, pp 305–311

González-Jara P, Moreno-Letelier A, Fraile A, Piñero D, García-Arenal F (2011) Impact of human management on the genetic variation of wild pepper, *Capsicum annuum* var. *glabriusculum*. PLoS One 6(12):e28715. https://doi.org/10.1371/journal.pone.0028715

Grube RC, Zhang Y, Murphy JF, Loaiza-figueroa F, Lackney VK, Provvidenti R, Jahn M (2000a) New source of resistance to *Cucumber mosaic virus* in *Capsicum frutescens*. Plant Dis 84:885–891

Grube RC, Blauth JR, Andredos MSA, Caranta C, Jahn MK (2000b) Identification and comparative mapping of a dominant potyvirus resistance gene cluster in *Capsicum*. Theor Appl Genet 101:852–859

Guo G, Wang S, Liu J, Pan B, Diao W, Ge W, Gao C, Snyder JC (2017) Rapid identification of QTLs underlying resistance to Cucumber mosaic virus in pepper (*Capsicum frutescens*). Theor Appl Genet 130:41–52

Hsu CC, Lin SW, Wang YW, Chan YL, Lee LM, Barchenger DW, Kenyon L, Kumar S (2018) Resistance to *Tomato mosaic virus* (ToMV) in sweet pepper (*Capsicum annuum*). Act Hort (in press)

Hoang NH, Yang HB, Kang BC (2013) Identification and inheritance of a new source of resistance against *Tomato spotted wilt virus* (*TSWV*) in *Capsicum*. Sci Hort 161:8–14

Jabłońska-Sabuka M, Kalaria R, Kauranne T (2015) A dynamical model for epidemic outbursts by *Begomovirus* population clusters. Ecol Model 297:60–68

Jahn M, Paran I, Hoffmann K, Radwanski ER, Livingstone KD, Grube RC, Aftergoot E, Lapidot M, Moyer JW (2000) Genetic mapping of the *Tsw* locus for resistance to the tospovirus tomato spotted wilt virus in *Capsicum* spp. and its relationship to the *Sw-5* gene for resistance to the same pathogen in tomato. Mol Plant-Microbe Inter 13:673–682

Kenyon L, Kumar S, Tsai WS, Hughes J (2014) Virus diseases of peppers (*Capsicum* spp.) and their control. Adv Virus Res 90:293–350

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

Kang WH, Hoang NH, Yang HB, Kwon JK, Jo SH, SeoJK Kim KH, Choi D, Kang BC (2010) Molecular mapping and characterization of a single dominant gene controlling CMV resistance in peppers (*Capsicum annuum* L.). Theor Appl Genet 120:1587–1596

Kaur N, Dhaliwal MS, Jindal S, Singh P (2016) Evaluation of hot pepper (*Capsicum annuum* L.) genotypes for heat tolerance during reproductive phase. Int J Bio-Resour Stress Manage 7:126–129

Kim HJ, Nahm SH, Lee HR, Yoon GR, Kim KT, Kang BC, Kweon OC, Cho MC, Kwon JK, Han JH, Kim JH, Park M, Ahn JH, Choi SH, Her NH, Sung JH, Kim BD (2008a) BAC-derived markers converted from RFLP linked to *Phytophthora capsici* resistance in pepper (*Capsicum annuum* L.). Theor Appl Genet 118:15–27

Kim SH, Yoon JB, Do JW, Park HG (2008b) Inheritance of anthracnose resistance in a new genetic resource *Capsicum baccatum* PI594137. J Crop Sci Biotechnol 11:13–16

Kim S, Park M, Yeom SI, Kim YM, Lee JM, Lee HA, Seo E, Choi J, Cheong K, Kim KT, Jung K, Lee GW, Oh SK, Bae C, Kim SB, Lee HY, Kim SY, Kim MS, Kang BC, Jo YD, Yang HB, Jeong HJ, Kang WH, Kwon JK, Shin C, Lim JY, Park JH, Huh JH, Kim JS, Kim BD, Cohen O, Paran I, Suh MC, Lee SB, Kim YK, Shin Y, Noh SJ, Park J, Seo YS, Kwon SY, Kim HA, Park JM, Kim HJ, Choi SB, Bosland PW, Reeves GR, Jo SH, Lee BW, Cho HT, Choi HS, Lee MS, Yu Y, Do Choi Y, Park BS, van Deynze A, Ashrafi H, Hill T, Kim WT, Pai HS, Ahn HK, Yeam I, Giovannoni JJ, Rose JKC, Sorensen I, Lee SJ, Kim RW, Choi IY, Choi BS, Lim JS, Lee YH, Choi D (2014) Genome sequence of the hot pepper provides insights into the evolution of pungency in *Capsicum* species. Nat Genet 46:470–478

Kirada C, Topcu S, Cetin M, Dasgan HY, Kaman H, Topaloglu F, Derici MR, Ekici B (2007) Prospects of partial root zone irrigation for increasing irrigation water use efficiency of major crops in the Mediterranean region. Ann Appl Biol 150:281–291

Kumar S, Kumar R, Singh J (2006) Cayenne/American pepper (*Capsicum* species). In: Peter KV (ed) Handbook of herbs and spices, vol 3. Woodhead Publishing, Cambridge, UK, pp 299–312

Kumar S, Shieh HC, Lin SW, Schafleitner R, Kenyon L, Srinivasan R, Wang JF, Ebert AW, Chou YY (2018) Peppers (*Capsicum* spp.): domestication and breeding for global use. In: Mandal D, Shukla AC, Siddiqui MW (eds) Sustainable horticulture, vol 1: diversity, production, and crop improvement, Part III. Crop improvement and biotechnology. Apple Academic/CRC Press, Waretown, NJ, USA, pp 387–400

Lapidot M, Paran I, Ben-joseph R, Ben-harush S, Pilowsky M, Cohen S, Shifriss C (1997) Tolerance to *Cucumber mosaic virus* in pepper: development of advanced breeding lines and evaluation of virus level. Plant Dis 81:185–188

Lebeau A, Daunay MC, Frary A, Palloix A, Wang JF, Dintinger J, Chiroleu F, Wicker E, Prior P (2011) Bacterial wilt resistance in tomato, pepper, and eggplant: genetic resources respond to diverse strains in the *Ralstonia solanacearum* species complex. Phytopathology 101:154–165

Lee J, Hong JH, Do JW, Yoon JB (2010) Identification of QTLs for resistance to anthracnose to two *Colletotrichum* species in pepper. J Crop Sci Biotechnol 13:227–233

Lee HR, Jung H, Gon Y, Lee J, Kim HJ, Kang BC, Harn CH (2013) Development of a novel codominant molecular marker for chilli veinal mottle virus resistance in *Capsicum annuum* L. Euphytica 193:197–205

Lefebvre V, Palloix A (1996) Both epistatic and additive effects of QTL are involved in polygenic induced resistance to disease: a case study, the interaction pepper–*Phytophthora capsici* Leonian. TheorAppl Genet 93:503–511

Lesk C, Rowhani P, Ramankutty N (2016) Influence of extreme weather disasters on global crop production. Nature 529. https://doi.org/10.1038/nature16467

Levey DJ, Tewksbury JJ, Cipollini ML, Carlo TA (2006) Field test of the directed deterrence hypothesis in two species of wild chili. Oecologia 150:61–68

Lin SW, Gniffke PA, Wang TC (2007) Inheritance of resistance to pepper anthracnose by *Colletotrichum acutatum*. Acta Hort 760:329–334

Linders EGA, Nicolet JLME, Van Wijk HJ (2010) Insect resistant plant, vol. PCT/EP2008/055374: Syngenta Participations AG United States

Lopes CA, Boiteux LS (2004) Biovar specific and broad spectrum sources of resistance to bacterial wilt (*Ralstonia solanacearum*) in *Capsicum*. Crop Breed Appl Biot 4:350–355

Luna-Ruiz de J, Nabhan GP, Aguilar-Melendez A (2018) Shifts in plant chemical defenses of chile pepper (*Capsicum annuum* L.) due to domestication in Mesoamerica. Front Ecol Evol 6:48. https://doi.org/10.3389/fevo.2018.00048

Maharijaya A, Vosman B, Steenhuis-Broers G, Harpenas A, Purwito A, Visser RGF, Voorrips RE (2011) Screening of pepper accessions for resistance against two thrips species (*Frankliniella occidentalis* and *Thrips parvispinus*). Euphytica 177:401–410

Maharijaya A, Vosman B, Steenhuis-Broers G, Pelgrom K, Purwito A, Visser RGF, Voorrips RE (2015) QTL mapping of thrips resistance in pepper. Theor Appl Genet 128:1945–1956

Maharijaya A, Vosman B, Verstappen F, Steenhuis-Broers G, Mumm R, Purwito A, Visser RG, Voorrips RE (2012) Resistance factors in pepper inhibit larval development of thrips (*Frankliniella occidentalis*). Entomol Exp Appl 145:62–71

Mahasuk P, Struss D, Mongkolporn O (2016) QTLs for resistance to anthracnose identified in two *Capsicum* sources. Mol Breed 36:10. https://doi.org/10.1007/s11032-016-0435-5

Mahasuk P, Taylor PWJ, Mongkolporn O (2009) Identification of two new genes conferring resistance to *Colletotrichum acutatum* in *Capsicum baccatum*. Phytopathology 99:1100–1104

Mallapur CP (2000) Screening of chilli genotypes against thrips and mites. Insect Environ 5:154–155

Mallard S, Cantet M, Massire A, Bachellez A, Sophie E, Lefebvre V (2013) A key QTL cluster is conserved among accessions and exhibits broad-spectrum resistance to *Phytophthoracapsici*: a valuable locus for pepper breeding. Mol Breed 32:349–364

Mathur R, Dangi RS, Dass SC, Malhotra RC (2000) The hottest chilli variety in India. Curr Sci 79:287–288

Mimura Y, Kageyama T, Yoshikawa M, Hirai M (2009) QTL analysis for resistance to *Ralstonia solanacearum* in *Capsicum* accession LS2341. J Jpn Soc Hort Sci 78:307–313

Minamiyama Y, Tsuro M, Kubo T, Hirai M (2007) QTL analysis for resistance to *Phytophthoracapsici* in pepper using a high density SSR-based map. Breed Sci 57:129–134

Momol MT, Pappu HR, Dankers W, Rich JR, Olson SM (2000) First report of *tomato spotted wilt virus* in Habanero and Tabasco peppers in Florida. Plant Dis 84:1154

Mongkolporn D, Taylor PWJ (2011) Capsicum. In: Kole C (ed) Wild crop relatives: genomic and breeding resources. Springer, Berlin

Moury B, Palloix A, Selassie-Gebre K, Marchoux G (1997) Hypersensitive resistance to tomato spotted wilt virus in three *Capsicum chinense* accessions is controlled by a single gene and is overcome by virulent strains. Euphytica 94:45–52

Moury B, Pflieger S, Blattes A, Lefebvre V, Palloix A (2000) A CAPS marker to assist selection of tomato spotted wilt virus (TSWV) resistance in pepper. Genome 43:137–142

Moury B, Verdin E (2012) Viruses of pepper crops in the Mediterranean basin: a remarkable stasis. Adv Virus Res 84:127–162

Munns R, Tester M (2008) Mechanisms of salinity tolerance. Annu Rev Plant Biol 59:651–681

Nabhan G (1990) Conservationists and forest services join forces to save wild chiles. Diversity 6:47–48

Naitam NR, Patang Rao DA, Deshmukh SD (1990) Resistance response of chilli cultivars to leaf curl. MPKV Res J 14:206–207

Naresh P, Bhatt RM, Venkatachalapathi V, Gangadharrao P, Madhavi Reddy K (2017) Inheritance of root traits in an interspecific Cross of *Capsicum annuum C. chinense* in the presence of low moisture. Int J Veg Sci 23:575–583

Naresh P, Krishna Reddy M, Hema Chandra Reddy P, Madhavi Reddy K (2016) Screening chilli (*Capsicum spp*.) germplasm against *Cucumber mosaic virus* and *Chilli veinal mottle virus* and inheritance of resistance. Eur J Plant Pathol 146:451–464

Nono-womdim R, Marchoux G, Pochard E, Palloix A, Gebre-selassie K (1991) Resistance of pepper lines to the movement of *Cucumber mosaic virus*. J Phytopathol 132:21–32

Ogundiwin EA, Berke TF, Massoudi M, Black LL, Huestis G, Choi D, Lee S, Prince JP (2005) Construction of 2 intraspecific linkage maps and identification of resistance QTL for *Phytophthora capsici* root-rot and foliar-blight diseases of pepper (*Capsicum annuum* L.). Genome 48:698–711

Pagán I, Betancourt M, de Miguel J, Piñero D, Fraile A, García-Arenal F (2010) Genomic and biological characterization of chiltepin yellow mosaic virus, a new potyovirus infecting *Capsicum annuum* var. *aviculare* in Mexico. Arch Virol 155:675–684

Perramond E (2005) The politics of ecology: local knowledge and wild chili collection in Sonora, Mexico. J Lat Am Geogr 4:59–75. https://doi.org/10.1353/lag.2005.0025

Perry L, Dickau R, Zarrillo S, Holst I, Pearsall DM, Piperno DR, Berman MJ, Cooke RG, Rademaker K, Ranere AJ, Scott Raymond JC, Sandweiss DH, Scaramelli F, Tarble K, Zeidler JA (2007) Starch fossils and the domestication and dispersal of chili peppers (*Capsicum* spp. L.) in the Americas. Science 315:986–988

Pickersgill B (1997) Genetic resources and breeding of *Capsicum* spp. Euphytica 96:129–133

Pochard E, Daubeze AM (1989) Progressive construction of a polygenic resistance to *Cucumber mosaic virus* in the pepper. In: Proceedings of EUCARPIA meeting on genetics and breeding of capsicum and eggplant, 7th, Kragujevac, Yugoslavia, pp 189–192

Prohens J, Gramazio P, Plaza M, Dempewold H, Kilian B, Diez MJ, Fita A, Herraiz FJ, Rodriguez-Burruezo A, Soler S, Knapp S, Vilanova S (2017) Introgressiomics: a new approach for using crop wild relatives in breeding for adaptation to climate change. Euphytica 213. https://doi.org/10.1007/s10681-017-1938-9

Qin C, Yub C, Shena Y, Fang X, Chen L, Mind J, Cheng J, Zhao S, Xu M, Luo Y, Yang Y, Wu Z, Mao L, Wu H, Ling-Hu C, Zhou H, Lin H, González-Morales S, Trejo-Saavedra DL, Tian H, Tang X, Zhao M, Huang Z, Zhou A, Yao X, Cui J, Li W, Chen Z, Feng Y, Niu Y, Bi S, Yang X, Li W, Cai H, Lu X, Montes-Hernández S, Leyva-González MA, Xiong Z, He X, Bai L, Tan S, Tang X, Liu D, Liu J, Zhang S, Chen M, Zhang L, Zhang L, Zhang Y, Liao W, Zhang Y, Wang M, Lv X, Wen B, Liu H, Luan H, Zhang Y, Yang S, Wang X, Xu J, Li X, Li S,

Wang J, Palloix A, Bosland PW, Li Y, Krogh A, Rivera-Bustamante RF, Herrera-Estrella L, Yin Y, Yu J, Hu K, Zhang Z (2014) Whole-genome sequencing of cultivated and wild peppers provides insights into *Capsicum* domestication and specialization. Proc Natl Acad Sci USA 111:5135–5140

Quirin EA, Ogundiwin EA, Prince JP, Mazourek M, Briggs MO, Chlanda TS, Kim KT, Falise M, Kang BC, Jahn MM (2005) Development of sequence characterized amplified region (SCAR) primers for the detection of Phyto. 5.2, a major QTL for resistance to Phytophthora capsici Leon in pepper. Theor Appl Genet 110:605–612

Rai VP, Kumar R, Kumar S, Rai A, Kumar A, Singh M, Singh SP, Rai AB, Paliwal R (2013) Genetic diversity in *Capsicum* germplasm based on microsatellite and random amplified microsatellite polymorphism markers. Physiol Mol Biol Plants 19:575–586

Rai VP, Kumar R, Singh SP, Singh S, Kumar S, Singh M, Rai M (2014) Monogenic recessive resistance to pepper leaf curl virus in *Capsicum*. Sci Hort 172:34–38

Rameash KS, Pandravada SR, Sivaraj N, Sarath Babu B, Chakrabarty SK (2015) Screening chilli (*Capsicum annuum* L.) genotypes for resistance to broad mite (*Polyphagotarsonemus latus* Banks) and analysing the geographic distribution of resistance. Elec J Plant Breed 6:928–937

Reddy KR, Kakani VG (2007) Screening *Capsicum* species of different origins for high temperature tolerance by *in vitro* pollen germination and pollen tube length. Sci Hort 112:132–135

Ruffel S, Gallois JL, Moury B, Robaglia C, Palloix A, Caranta C (2006) Simultaneous mutations in translation initiation factors *eIF4E* and *eIF(iso)4E* are required to prevent pepper veinal mottle virus infection of pepper. J Gen Virol 87:2089–2098

Sarath Babu B, Pandravada SR, Reddy JK, Varaprasad KS, Sreekanth M (2002) Field screening of pepper germplasm for source of resistance against leaf curl caused by thrips, *Scirtothrips dorsalis* Hood and mites, *Polyphagotarsonemus latus* Banks. Indian J Plant Prot 30:7–12

Sarath Babu B, Pandravada SR, Prasada Rao RDVJ, Anitha K, Chakrabarty SK, Varaprasad KS (2011) Global sources of pepper genetic resources against arthropods, nematodes and pathogens. Crop Prot 30:389–400

Sanchez-Puerta MV, Masuelli RW (2011) Evolution of nematode-resistant Mi-1 gene homologs in three species of *Solanum*. Mol Genet Genom 285:207–218

Singh AK, Kushwaha N, Chakraborty S (2016) Synergistic interaction among begomoviruses leads to suppression of host defense-related gene expression and breakdown of resistance in chilli. Appl Microbiol Biotechnol. https://doi.org/10.1007/s00253-015-7279-5

Soler S, Debreczeni DE, Vidal E, Aramburu J, Lopez C, Galipienso L, Rubio L (2015) A new *Capsicum baccatum* accession shows tolerance to wild-type and resistance-breaking isolates of *Tomato spotted wilt virus*. Ann Appl Biol 167:343–353

Srivastava A, Mangal M, Saritha RK, Kalia P (2017) Screening of chilli pepper (*Capscium* spp.) lines for resistance to the begomovirus causing chili leaf curl disease in India. Crop Prot 100:177–185

Srivastava A, Mangal M, Saritha RK, Santosh LJ, Uttamgir Gosavy G, Kalia P (2015) Natural epiphytotic screening of chilli germplasm lines against leaf curl virus complex. Int J Trop Agric 33:3581–3586

Sung U, Chang YY, Ting NL (2005) Capsaicin biosynthesis in water stress hot pepper fruits. Bot Bull Acad Sin 46:35–42

Suwor P, Sanitchon J, Thummabenjapone P, Kumar S, Techawongstien S (2017) Inheritance analysis of anthracnose resistance and marker-assisted selection in introgression populations of chili (*Capsicum annuum* L.). Sci Hort 220:20–26

Suwor P, Thummabenjapone P, Sanitchon J, Kumar S, Techawongstien S (2015) Phenotypic and genotypic responses of chili (*Capsicum annuum* L.) progressive lines with different resistant genes against anthracnose pathogen (*Colletotrichum* spp.). Eur Plant Pathol 143:725–736

Suzuki K, Kuroda T, Miura Y, Muria J (2003) Screening and wild traits of virus resistant source in *Capsicum* spp. Plant Dis 87:779–783

Swaminathan B, Siva Balan KC, Anadaraja N, Manikanda Boopathi N, Schreinemachers P, Srinivasan R, Wu MH (2016) Profitability of begomovirus management strategies among chilli farmers in Tamil Nadu: a gross margin impact analysis. Indian J Agric Res 50:159–166

Sy O, Steiner R, Bosland PW (2008) Recombinant inbred line differential identifies race-specific resistance to phytophthora root rot in *Capsicum annuum*. Phytopathology 98:867–870

Tewksbury JJ, Nabhan G (2001) Directed deterrence by capsaicin in chillies. Nature 41:403–404

Tewksbury JJ, Nabhan GP, Norman D, Suzan H, Tuxill J, Donovan J (1999) In situ conservation of wild chiles and their biotic associations. Conserv Biol 13:98–107

Thakur PP, Mathew D, Nazeem PA, Abida PS, Indira P, Girija D, Shylaja MR, Valsala PA (2014) Identification of allele specific AFLP markers linked with bacterial wilt (*Ralstonia solanacearum* (Smith) resistance in hot peppers (*Capsicum annuum* L.). Physiol Mol Plant Pathol 87:19–24

Thakur H, Jindal SK, Sharma A, Dhaliwal MS (2019) A monogenic dominant resistance for leaf curl virus disease in chilli pepper (*Capsicum annuum* L.). Crop Prot 116:115–120

Varma A, Malathi VG (2003) Emerging geminivirus problems: a serious threat to crop production. Ann App Biol 142:145–164

Venkatesh J, An J, Kang WH, Jahn M, Kang BC (2018) Fine mapping of the dominant potyvirus resistance gene *Pvr7* reveals a relationship with *Pvr4* in *Capsicum annuum*. Phytopathology 108:142–148

Wang A, Krishnaswamy S (2012) Eukaryotic translation initiation factor 4E-mediated recessive resistance to plant viruses and its utility in crop improvement: eIF4E-mediated resistance to plant viruses. Mol Plant Pathol 13:795–803

Wang D, Bosland PW (2006) The genes of *Capsicum*. HortScience 41:1169–1187

Wang LH, Gu XH, Hu MY, Mao SL, Zhang ZH, Peng DL, Yun XF, Zhang BX (2009) A SCAR marker linked to the *N* gene for resistance to root knot nematodes (*Meloidogyne* spp.) in pepper (*Capsicum annuum* L.). Sci Hort 22:18–322

Wang P, Wang L, Guo J, Yang W, Shen H (2016) Molecular mapping of a gene conferring resistance to *Phytophthora capsici* Leonian race 2 in pepper line PI201234 (*Capsicum annuum* L.). Mol Breed 36:66. https://doi.org/10.1007/s11032-016-0464-0

Yao M, Li N, Wang F, Zhibiao YE (2013) Genetic analysis and identification of QTLs for resistance to *Cucumber mosaic virus* in chili pepper (*Capsicum annuum* L.). Euphytica 193:135–145

Yoon JK, Park HG (2005) Trispecies bridge crosses, (*Capsicum annuum* × *C. chinense*) × *C. baccatum*, as an alternative for introgression of anthracnose resistance from C. *baccatum* into C. *annuum*. Hort Environ Biotechnol 46:5–9

Yoon JB, Yang DC, Do JW, Park HG (2006) Overcoming two post-fertilization genetic barriers in interspecific hybridization between *Capsicum annuum* and *C. baccatum* for introgression of anthracnose resistance. Breed Sci 56:31–38

Zhani K, Elouer MA, Aloui H, Hannachi C (2012) Selection of a salt tolerant Tunisian cultivar of chilli (*Capsicum frutescens*). Eurasia J Bio Sci 6:47–59

Capsicum Breeding: History and Development

3

Arpita Srivastava and Manisha Mangal

Abstract

Capsicum or chili peppers were predominantly domesticated first in America and was introduced from there to rest of the world by Columbus. Capsicum breeding initially started as selection from wild species for different purposes and further improvement was based on the art of selection. With time, the breeding for crop improvement became more scientific and classical methods like mass selection, pedigree method, single-seed descent method, backcross, and hybridization are currently being utilized for capsicum improvement. Genetic diversity of capsicum is large, allowing alternatives to several new gene rearrangements. Capsicum fruits have high nutritional value, bringing benefits to consumer's health. This fact has contributed to increase the market and consumption of capsicum in the world. Search for capsicum genotypes with increased yield, disease and abiotic stress resistance and improved quality is the goal in capsicum breeding programs. Lately, new strategies for improvement like mutation breeding, polyploidy, haploid breeding, embryo rescue, and utilization of molecular markers have been used in capsicum breeding. With continuous advancement in molecular technologies, it is becoming an essential tool which when combined with traditional selection and crossing techniques can result in significant progress in already established capsicum genetic breeding program.

3.1 Introduction

Capsicum belonging to Solanaceae family are cultivated worldwide which are being utilized for different purposes with different quality and trait requirements. There is a huge amount of diversity in *capsicum* genus and so is the diversity in its usage. Capsicums contain all the important nutrients for which it has been considered as a food and used in fresh or dried form for many years. Capsicum fruits are known for its high vitamin C content which is reported to be twice that of citrus fruits. Contrary to this, dried red chilies are very high in vitamin A and are a great source of β-carotene (Shetty et al. 2013). They promote health benefits such as reducing obesity and diabetes (Kwon et al. 2007). Chilies have antibacterial qualities and also contain bioflavonoid along with antioxidants most commonly present in apple juice. It is also reported to be effective in protecting against cancer

A. Srivastava (✉) · M. Mangal
Division of Vegetable Science, ICAR-Indian Agricultural Research Institute, New Delhi, India
e-mail: asrivastava45@gmail.com

M. Mangal
e-mail: manishamangal@rediffmail.com

N. Ramchiary and C. Kole (eds.), *The Capsicum Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-319-97217-6_3

(Pramanick and Srivastava 2013). Capsaicin cream is used to relieve the sensation of pain in such conditions as arthritis and other painful chronic conditions (Bhattacharya et al. 2010). Capsicum extracts are used in cosmetics as well as pharmaceuticals. Nowadays, growing capsicum in pots or gardens for ornamental purpose is also gaining importance (Bosland and Votava 2012).

World spice trade is dominated by hot pepper while sweet pepper has become a popular vegetable in the tropics. Based on fruit shape and size and utilities, capsicum market can be grouped into five broad categories: (i) fresh market which produces green, red, or multicolor whole fruits; (ii) fresh processing market for sauce, paste, canning, pickles; (iii) dried spice market for whole fruits and pepper powder; (iv) industrial extracts like oleoresin, capsaicinoids, and carotenoids; and (v) ornamental types (plants and/or fruits) (Poulos 1994). Based on the end user's demand the focus of capsicum breeding programs will change. Capsicum breeding programs also focus to reduce the stresses imposed by pests as well as extreme environmental conditions.

3.2 History of Capsicum Breeding

Western hemisphere is the place of origin of capsicums, and they were known and used as food since 7500 BC. They are native of South America and from there they spread to Central America. Columbus is credited for introducing capsicum to Europe from where it spread to Africa and Asia. Classical studies in capsicum in early times focussed mainly on genetic inheritance of important horticultural traits, mutant forms, disease resistance traits, male sterility, and quality traits. There are various reports stating these traits to be governed by single genes having dominant or recessive mode of action and some traits by quantitative trait loci (Deshpande 1933; Daskalov 1973; Shuh and Fontenot 1990). Studies of this type have been summarized in Table 3.1.

Capsicum annuum is the most important species of the genus *capsicum* as it is cultivated widely on a commercial scale. Initially, plant breeding was mostly based on the art of selecting individuals which was rather slow and casual process. However, with the introduction of Mendelian principles on genetics and heredity, plant breeding which was considered as art became "science". Currently, different methods of selection are used in breeding plants (including chili peppers), and the choice of method depends mainly on objective(s) of the breeding program (Greenleaf 1986; Singh et al. Singh et al. 2014a, b). Basically, the strategy of capsicum breeders is to develop a single genotype with higher genetic potential as productivity, disease resistance, and content of bioactive compounds.

3.3 Current Breeding Objectives for Capsicum Improvement

Capsicum breeding objectives for both hot and bell pepper differ with the country of cultivation, purpose of cultivation, cultivation condition, end user as well as customer preference of the region. Some countries prefer hot- and pungent-type pepper while some prefer sweet types. The diseases affecting the crop also vary with climate prevalent in the respective countries. Broadly, the objectives of capsicum breeding have been summarized in Fig. 3.1. However, disease resistance breeding is one of the foremost objectives in capsicum breeding (Paran et al. 2004). Pohronezny (2003) has provided a complete illustration of various diseases affecting capsicum. Disease resistance breeding basically starts with the identification of resistant sources, understanding its genetics followed by introgression in promising genotypes. In capsicum, substantial utilization of disease and pest resistance from wild species introgressed into elite cultivars to improve disease resistance has been documented. Interspecific hybridization program for resistant gene introgression also involves understanding the level of crossability between species.

Table 3.1 Genes of major horticultural traits and disease resistance in *capsicum*

Plant character	Character type and its gene symbol	Reference
Anthocyanin coloration of leaves, stem, flower, and immature fruits	*A* (incomplete dominance) along with modifier gene *MoA*	Lippert et al. (1965), Odland (1960)
Plant height	Dwarf nature controlled by nine recessive genes *dw1* to *dw9*	Daskalov (1973a), Restaino (1989), Yazawa et al. (1991), Aniel et al. (2001)
Branchless	*Bl* (recessive gene)	Bergh and Lippert (1964)
Branching	*ct* (for plant habit), *fa* (clustered fruit habit) and *dt* (for determinate growth) along with modifiers Dominant genes Dt and Ct (indeterminate growth habit)	Mc Cammon and Honma (1984)
Leaf shape	Nl (narrow leaf), broad leaf (bl), small leaf (sl-1 & sl2), curved leaf (cl), folded leaf (fl), round tip leaf (rl-1, rl-2, rl-3),	Daskalov (1973b), Aniel et al. (2001)
Leaf pubescence	Two dominant genes *H* & *Sm* (HHSmSm: presence of pubescence, hhsmsm: glabrous leaves)	Shuh and Fontenot (1990)
Flowers	Multiple flowers by three dominant genes (Mf-1, Mf-2 and Mf-3) *ef*: early flowering *lf*: late flowering *nf*:n o flowering	Shuh and Fontenot (1990), Pathak et al. (1985)
Fruit shapes	*P*: Pointed fruit shape *fb*: Non bulging fruit shape *ce*: fruit base with enclosed calyx *O*: round fruit shape with modifiers *up-1* & *up-2*: Erect fruit *pf*: parthenocarpy	Deshpande (1933), Daskalov and Poulos (1994), Peterson (1959), Gopalkrishnan et al. (1989), Lippert et al. (1965), Pathak et al. (1983)
Immature fruit color	Three alleles of a recessive gene: *sw1, sw2 & sw3* *sw1*: sulphur white *sw2*: yellowish green *sw3*: cedar green sw1 > sw2 > sw3	Odland and Porter (1938)
Mature fruit color	y^+: red colour *y*: yellow colour *cl* & *y* + : brown colour *Ccs:* capsanthin-capsorubin synthase enzyme that synthesizes red carotenoid pigment *Psy:* locus responsible for development of fruit colour *Psy/C2*: rate limting factor in carotenoid production	Boswell (1937), Smith (1950), Hurtado-Hernandez and Smith (1985), Papovsky and Paran (2000), Lefebvre et al. (1998), Thorup et al. (2000), Huh et al. (2001)
Pungency	*Pun*: Controls acyl transferase responsible for capsaicin synthesis *lov*: non pungency due to loss of vesicles on the placental walls	Deshpande, (1935), Greenleaf (1952), Daskalov and Poulos (1994), Votava and Bosland (2002)
Beta carotene	*B, t and bc*: Confer high beta carotene contents	Chalukova et al. (1993), Daskalov et al. (1995)

(continued)

Table 3.1 (continued)

Plant character	Character type and its gene symbol	Reference
Male sterility	Genetic male sterility: Total 20 genes have been identified, *ms-1* to *ms-20* Cytoplasmic male sterility: Major gene *ms* in interaction with *S* cytoplasm *Rf*: Restorer of fertility locus	Shifriss and Frankel (1969), Shifriss and Rylski (1972), Daskalov (1973a), Daskalov and Poulos (1994), Shifriss (1973), Meshram and Narkhade (1982), Pathak et al. (1983), Peterson (1958), Novac et al. (1971) Daskalov (1973a)
Tobacco mosaic virus	L^3, L^2, L^1, L^+: series of multiple alleles L^2: localization of TMV L^1: imperfect localization of TMV L^+: mottling $L^3 > L^2 > L^1 > L^+$	Homes (1937), Boukema et al. (1980), Boukema, (1980)
Cucumber mosaic virus	*cm*: recessive geneand 4 QTLs	Singh and Thakur (1977), Gil-Ortega and Arteaga (1988), Ben Chaim et al. (2001)
Tomato spotted wilt virus	*Tsw*: Hypersensitive resistance to TSWV	Moury et al. (Moury et al. 1997a, b)
Bacterial leaf spot	*Bs1, Bs2, Bs3, Bs4*: Hypersensitive resistance *bs5* and *bs6*: nonhypersensitive recessive resistance *gds*: general defense system	Cook and Guevara (1984), Cook and Stall (1963), Hibberd et al. (1987), Kim and Hartmann (1985), Sahin and Miller, (1997), Csillery et al. (2004), Szarka and Csillery (1995), Jones et al. (2002)
Phytophthora disease	*Psr:* Stem resistance to *Phytophthora* *Pfo*: Foliar resistance *Pfr*: Fruit rot resistance	Sy et al. (2005), Walker and Bosland (1999), Saini and Sharma (1978)
Anthracnose resistance	*Anr1*: Resistance to *Colletotrichum dematium* *Anr2, Anr3, Anr4*: Resistance to *Colletotrichum gleosporoides* *Anr5*: Resistance to *Colletotrichum capsici*	Park et al. (1990), Fernandes and Ribeiro (1998), Lin et al. (2002)
Bacterial Wilt	Two genes with incomplete dominance	Matsunaga et al. (1998)
Powdery Mildew	Three genes: *lmr-1, lmr-2, lmr-3*	Shifriss et al. (1992)
Root-knot nematodes	*N:* Resistance to *M. incognita acritaMe1, Me2, Me3, Me4, Me5*: Resistance to Meloidogynae spp. *Me6*: *M. arenaria* and *M. javanica* *Mech1* and *Mech2*: Suppresses nematode resistance	Fery and Harrison (1990), Hendy et al. (1995), Pegard et al. (2005), Djian-Caporalino et al. (2004)
Bentazon herbicide resistance	*Bzt*: Tolerance	Fery and Harrison (1990)

Important diseases of attention in present-day scenario worldwide are viral diseases like Potyviruses [*Potato virus Y* (PVY), *Potato virus X* (PVX), *Pepper mottle virus* (PepMoV), *Pepper veinal mottal virus* (PVMY), *Tobacco etch virus* (TEV), *Chilli veinal mottle virus* (CVMV), *Pepper severe mosaic* (PSMV), *Pepper yellow mosaic* (PYMV)], Tospoviruses [*tomato spotted wilt virus* (TSWV), *Impatiens necrotic spot virus* (INSV), *Groundnut ringspot virus* (GRSV), *Groundnut bud-necrosis virus* (GBNV)], Cucumovirus [*Cucumber mosaic virus* (CMV)], Tobamoviruses [*Tobacco mosaic virus* (TMV), *Tomato mosaic virus* (ToMV), *Pepper mild-mosaic virus* (PMMV)], fungal diseases like powdery mildew, phytophthora root rot, anthracnose,

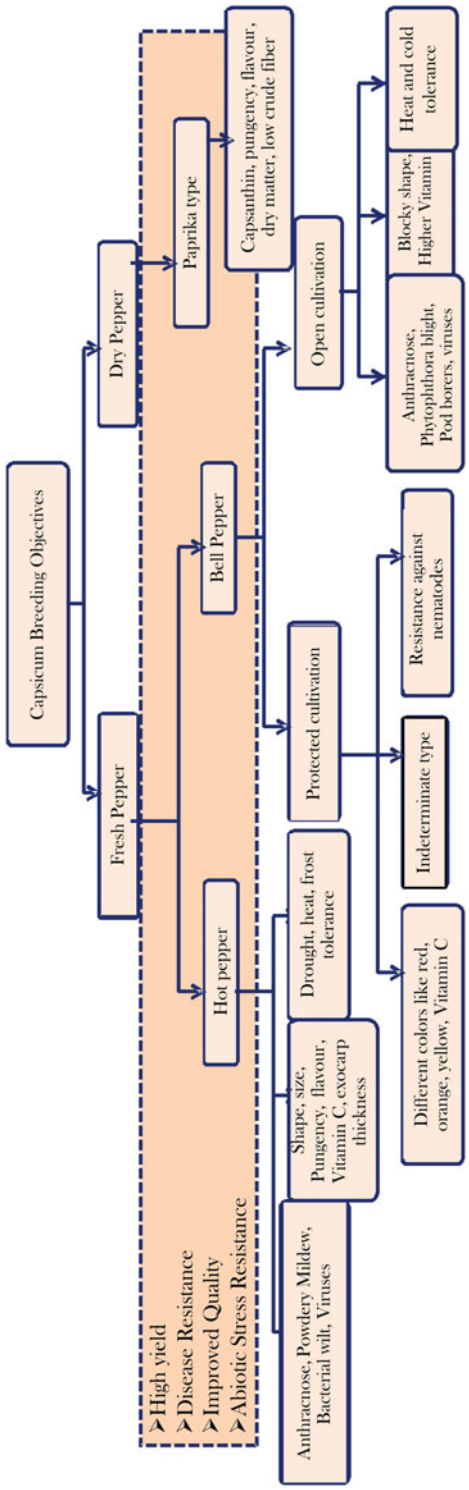


Fig. 3.1 Breeding objectives for capsicum improvement

bacterial diseases like bacterial wilt, and bacterial spot and pests like whiteflies, thrips, mites and root-knot nematodes. List of wild species utilized as disease-resistant sources is summarized in Table 3.2. Utilizing wild germplasm for introgression of disease resistance genes into promising genotypes has contributed significantly to crop improvement, particularly in terms of yield and quality improvement as well as stability in capsicum production. Introgression efforts to transfer disease resistance genes into superior genotypes have often been difficult, especially when resistance traits are under polygenic control and linked to undesirable horticultural and economic traits. With continuous evolution and emergence of new pathogen races and strains against available resistant genotypes necessitates regular search and use of new resistant sources.

The second objective for which capsicum breeders throughout the world are striving is increasing yield, thereby increasing overall productivity. In this respect, heterosis breeding program is gaining importance. Emphasis on development of hybrids based on male sterility systems is desired as it saves time and labor required for hybrid seed production. Both genetic (GMS) and cytoplasmic male sterility (CMS) systems have been utilized to produce hybrids, but CMS system is more widely exploited. With the identification of new CMS sources, their maintainers and diversification of CMS systems also become an important objective. Further for development of good hybrid, identification of restorers with good general and specific combining ability, and incorporation of resistant genes in these CMS lines and restorers should also be an area of focus.

Breeding objectives of capsicum also depend on the market demand and end utility. This includes breeding for horticultural and biochemical traits. Fresh market breeders look for traits like fruit color at unripe stage usually green (light, medium, or dark), fruit length and its width and pericarp thickness. Apart from this, the level of pungency is an important and unique aspect of capsicum breeding. Understanding people's preferences for pungency in a particular

Table 3.2 List of wild and cultivated species as source of disease resistance

Capsicum species	Resistant source	Diseases and resistant genes	References
C. baccatum	PBC 80	*Colletotrichum* spp. (anthracnose)	Montri et al. (2009), Mongkolporn and Taylor (2011)
	PBC 81	*Colletotrichum* spp. (anthracnose)	Montri et al. (2009), Mongkolporn and Taylor (2011)
	C-153	TSWV	Rosellol et al. (1996)
C. chacoense	PI260435	*Bs2* (*Xanthomonas campestris pv. vesicatoria*)	Cook and Guevara (1984)
C. chinense	7204	*Tsw* (tomato spotted wilt tospovirus)	Moury et al. (1997a, b)
	CNPH725	*Tsw* (tomato spotted wilt tospovirus)	Boiteux and de Avila (1994), Boiteux (1995)
	ECU-973	*Tsw* (tomato spotted wilt tospovirus)	Cebolla-Cornejo et al. (2003)
	PI152225	L^3(tobacco mosaic virus)*Pvr1* (TEV-C, TEV-F, PepMoV, PVY) *Tsw* (tomato spotted wilt tospovirus)	Boukema (1980, 1982, 1984), Boukema et al. (1980), Kyle and Palloix (1997), Black et al. (1991), Boiteux (1995), Jahn et al. (2000)
	PI159236	L^3 (tobacco mosaic virus) *Pvr1* (TEV-C, TEV-F, PepMoV, PVY) *Pvr7* (PepMoV) *Tsw* (tomato apotted wilt tospovirus)	Boukema (1980, 1982, 1984), Kyle and Palloix (1997), Grube et al. (2000) Black et al. (1991), Boiteux (1995), Jahn et al. (2000)
	PI315008	L^3 (tobacco mosaic virus)	Boukema (1980)
	PI315023	L^3 (tobacco mosaic virus)	Boukema (1980)
	PI315024	L^3(tobacco mosaic virus)	Boukema (1980)
	PBC932	*co1, co2, co3* (*Colletotrichum capsici*)	Pakdeevaraporn et al. (2005), Mahasuk et al. (2009a, b)
C. chacoense		*Tswv*	Boukema (1982)
C. pubescens	PI235047	Bacterial spot (*Xanthomonas campestris pv. vesicatoria*)	Sahin and Miller (1998)
C. annuum	CM334 (Serrano Criolle de Morelos-334)	*Phytophthora capsici* *Pfo* (*Phytophthora* foliar rot) *Pfr*(*Phytophthora* fruit rot) *Psr* (*Phytophthora* stem rot) *Pvr4* (PVY pathotypes 0, 1 and 2 and PepMoV) *Pvr5* (common PVY strains) *Me* (*Meloidogyne* sp.)	Saini and Sharma (1978), Gil-Ortega et al. (1991), Walker and Bosland (1999), Sy et al. (2005, 2008)
	PI264281	*Pvr2* (PVY pathotypes 0 and 1; TEV)	Kyle and Palloix (1997)
	SC46252	*pvr2* (PVY pathotypes 0 and 1; TEV)	Kyle and Palloix (1997)
	PM687 (inbred PI322719-a local Indian population)	Bacterial wilt (*Ralstonia solanacearum*) *Meloidogyne* spp.	Lafortune et al. (2005)
	PM217 (PI201234)	*Meloidogyne* spp.	Djian-Caporalino et al. (1999)
	AC2258 (PI201234)	*Phytophthora capsici*	Djian-Caporalino et al. (1999)

region is a very significant aspect. Pungency, an important attribute of capsicum commercially, is due to the presence of chemical complex or alkaloids known as capsaicinoids (Perucka et al. 2001). Capsaicin and dihydrocapsaicin are the two most abundant capsaicinoids in capsicum constituting about 90%, where capsaicin alone accounts for ~71% of the total capsaicinoids in most of the pungent varieties (Kosuge 1970). Capsaicin content of capsicum is one of the

major quality parameters that capsicum breeders look into while developing commercial varieties (Ohnuki et al. 2001; Kawabata et al. 2006; Hachiya et al. 2007). Capsaicinoid content is, nowadays, determined by high-performance liquid chromatography (HPLC), gas chromatography–mass spectrometry (GC–MS), and liquid chromatorgraphy–mass spectrometry (LC–MS) techniques. HPLC analytical technique is also used to estimate capsanthin content in capsicum. *Capsicum* genus uniquely has capsanthin–capsorubin synthase (CCS) which is an enzyme that synthesizes two red pigments—capsanthin and capsorubin (Guzman et al. 2011). Breeding for higher capsanthin is targeted in red capsicum to be used as dried spice (whole fruits and powder), and for industrial extracts (paprika oleoresin, capsaicinoids, and carotenoids). The red color in chili which is due to capsanthin and capsorubin, and yellow color due to β-carotene and violaxanthin is measured in American Spice Trade Association (ASTA) units (Englewood 1985). Generally, the higher the ASTA color value, the deeper is the red color of the genotype on ripening. The range of capsanthin content is 70 ASTA units (low), 71–100 ASTA units (medium), and 101–150 ASTA units (high). ASTA color affects the brightness of a product, while the surface color has an impact on the hue of product.

Hence, the development of paprika varieties to meet high demand of nonpungent pods with high color value for oleoresin extraction for industries is another important objective in capsicum breeding. Besides conventional nutritional uses in food, the other uses of capsicum such as in defense, spiritual, and ethnobotanical are also known (Kumar et al. 2006; Meghvansi et al. 2010). Accordingly, the breeding goals also vary.

For dry capsicum, dry matter content is an important quality character to be bred for making dry powder and whole dry fruit purpose. These are also the major characters desired for export purpose. A high dry-matter content of red chili fruit is important from commercial viewpoint in spice industry, but there is no positive relation between the dry matter content and its capsaicin content (Dhall 2008). Thin pericarp is necessary for dry capsicum as drying can be more easily accomplished. On drying, fruits with thick pericarp show wrinkled surface and dull appearance.

Increasing industrialization, risks of crop failure due to changing climate and demand (domestic and export) for more nutritious and safer foods emphasis are also being laid on breeding for genotypes with increased tolerance to high temperature, drought and wide adaptability.

3.3.1 Specific Objectives for Sweet Pepper Breeding

The main objectives of sweet pepper genetic improvement are developing varieties with blocky shape and different colors like medium or dark green at unripe stage and red, yellow or orange at ripe stage. The main objective is to select and develop new breeding lines and/or cultivars of capsicum with high levels of antioxidants and vitamins. These include: ascorbic acid (vitamin C), flavonoids (phenolics), red and yellow/orange carotenoids (including vitamin A-precursors like α- and β-carotene, β-cryptoxanthin) (Tomlekova et al. 2009a, b). Breeding efforts also include selection for high fruit set and yield under a range of growing conditions like open and protected, including the study of abiotic stresses like low temperatures, water stress, and saline stress. (Hein 2017; Negi et al. 2018). Activities also include breeding for stability of carotenoid extracts under long-term storage conditions and against photo-oxidation. Sweet pepper and hot pepper are affected by many common pathogens, but disease of importance in sweet pepper is Phytophthora fruit rot, anthracnose, viruses, powdery mildew, and bacterial wilt under open cultivation. Breeding sweet pepper genotypes with wider adaptability is another important objective as it is a cool season crop, and hence tropicalization is necessary. This will ensure availability of the crop in nontraditional areas during greater part of the year (Ferrara et al. 2011). Under protected cultivation breeding sweet pepper lines with indeterminate growth habit, amenability to training

and pruning, blocky fruit and resistance to root-knot nematode are the major breeding goals (Parker et al. 1995; de Swart 2007).

3.4 Breeding Methods for Capsicums

Conventional breeding methods like mass selection, pureline selection, pedigree breeding, single-seed descent method, backcross breeding, and heterosis breeding have been used for the crop improvement in capsicum. Other breeding methods like mutation breeding and polyploidy breeding have also been attempted to create variation and subsequently utilize in capsicum improvement programs. Mass selection, pureline selection, pedigree breeding, single-seed descent method, and backcross breeding were strategies utilized earlier for capsicum improvement when systematic plant breeding started. Mass selection is one of the simplest techniques which has been used for capsicum improvement (Table 3.3). Improvement for multiple traits of simple inheritance can be done simultaneously without any concerns about pedigree. Initially, it was used to improve landraces or open-pollinated cultivars of capsicum. In this approach, characters with high heritabilities are easily fixed and a reasonable level of variability is also maintained. Pure line selection was basically applicable to landraces/local cultivars which were being grown by farmers. In this method, superior plants are selected, then harvested separately and evaluated next year to observe plant progeny performance. Progeny showing superior performance and devoid of genetic variability, is bulk harvested and evaluated further with check cultivar(s) in replicated trials. This method has been extensively used to develop several varieties for commercial cultivation in chili capsicum (Table 3.3).

Pedigree selection is a breeding scheme where selection is affected among and within family, and the selected individuals are given a pedigree number so that any progeny in any generation can be traced back to the original plant which was first selected in F_2 generation. This has been one of the most commonly followed approaches for cultivar development in capsicum (Table 3.3). Selection of superior parental cultivars is crucial step in this method. This method is often utilized in conjunction with backcrossing to introgress important genes into advanced inbreds.

In single-seed descent (SSD) method, one seed from a single fruit is harvested from each plant in a segregating generation without applying any selection. The segregating generation are grown under greenhouse facilities to advance more generations in a year, to generate large number of inbred lines to be used in test crosses for development of hybrids and to generate recombinant inbred line populations to be utilized in mapping studies. Backcross method is the most widely used strategy in disease resistance breeding program of capsicum. This is normally used to transfer single gene/few genes from primitive cultivars/wild forms to leading cultivars. In some cases, even BC_2 families may be routed through pedigree method of breeding (modified backcross) instead of following a routine backcrossing program which needs 5–6 backcrosses with the recurrent parent.

Heterosis breeding has been advantageous for increased hot pepper or bell pepper production although open-pollinated varieties are still commonly available. Several hybrids have been developed in capsicum; however, the hybrid development program should be continuous so as to make the seeds available to the growers at affordable cost. F_1 hybrids of capsicum are gaining popularity after the initiation of research and seed production work in vegetables by a large number of private sector seed companies. To make seed production, more economic male sterility is extensively utilized in chili for hybrid seed development. The discovery of some male-sterile mutants which help to eliminate more laborious operations of emasculation combined with various marker genes further facilitate identification of undesirable types at seedling stage itself. Presently in chili, genetic male sterility (GMS) and cytoplasmic-genetic male sterility (CGMS) are being commercially exploited for the development of hybrids. Of the

Table 3.3 Achievements made by different breeding methods in capsicum improvement

Breeding approach	Significant achievements/varieties released	References
Introduction	• NuMex Centennial (Mexico), • CO 4 (introduced from Srilanka in India), • CO 3 (introduced from Srilanka in India),	TNAU Portal: http://agritech.tnau.ac.in/horticulture/horti_vegetables_chilli.html Vidhi J: http://www.biologydiscussion.com/vegetable-breeding/top-7-breeding-methods-of-capsicum-india/68448
Mass Selection	• It is still used in Mexico to select seed for Poblano, guajillo and other traditional capsicum landraces. • Heritage New Mexico 6-4: Selection made from NuMex Big Ji	Dewitt 2014; Vidhi J
Pureline Selection	• G 1, G 2, G 3, G 4, NP 46A, K 1, Co 1,CO.2, Musalwadi, Sindhur, Patna Red, Pant C 1, PLR1, KI	Gopalakrishnan (2007), Ramachandran (2013); Vidhi J
Single plant Selection	• New Mexico No. 6 • NuMex Conquistador: single plant selection from 'New Mexico 6-4 • NUMEX JOE E. PARKER: single plant selection from population of 'New Mexico 6-4 • NuMex Sweet: single plant selection from population of 'New Mexico 6-4	Dewitt 2014
Pedigree Method	• Andhra Jyoti, Pusa Jwala, Pusa Sadabahar, X 235, K2, Punjab Lal and Jawahar 218 (India) • New Mexico No. 9 Sandia: Cross of New Mexico No. 9 and California Anaheim Rio • Grande 21: Cross of New Mexico No. 6 and Anaheim • Española Improved: hybridization between 'Sandia' and a Northern New Mexico line of chilli • Numex Sunrise, Numex Sunset and Numex Eclipse: All three cultivars originated from a hybridization between 'Permagreen,' a green bell pepper, and 'New Mexico 6-4 • Numex RNaky: cross of 'Rio Grande 21,' 'New Mexico 6-4,' Bulgarian paprika, and an early-maturing native type • Numex Sunburst, Numex Sunflare and Numex Sunglo: derived by pedigree breeding from a seed source from India in Mexico • Numex Primavera, NuMex Memorial Day' and 'NuMex Thanks, Numex Garnet, NuMex Primavera' • AVPP0506, Berke's Joy, AVPP0105, AVPP0206, AVPP0303, AVPP0409, AVPP0411, AVPP0512, AVPP0514 (from World Vegetable Centre) • Bell Pepper: Spartan Garnet – California Wonder × Dwarf Pimiento Selection from the variety Santanka • Spartan Emerald – Morgold × California Wonder • Sonnette-An F_2 derived line originating from the cross (Morgold × California Wonder) × Keystone Resistant Giant	Ramachandran (2013), Dewitt 2014; Vidhi J: http://www.biologydiscussion.com/vegetable-breeding/top-7-breeding-methods-of-capsicum-india/68448; AVRDC: https://avrdc.org/seed/improved-lines/chili-capsicum/

(continued)

Table 3.3 (continued)

Breeding approach	Significant achievements/varieties released	References
Backcross method	• Pyramiding of genes conferring resistance to PMMoV, PVY and TSWV in sweet Charleston capsicum lines (BC: Backcrossing; L4: The gene conferring resistance to PMMoV pathotype 1, 2, 3. c Tsw: The gene conferring resistance to TSWV; dPVY: Potato Virus Y pathotype 1–2) • Introgression of heat shock protein (Hsp70 and sHsp) genes into the Malaysian elite chilli variety Kulai (*capsicum annuum* L.) • Resistance has been successfully introgressed into commercial capsicum cultivars, including resistance to tobamoviruses from *capsicum chacoense*and*C. Chinense* • Resistance to tomato spotted wilt virus (TSWV) from *C. chinense* and *C. baccatum* • Resistance to anthracnose fruit rot from *C. chinense* • Resistance to *Phytophthora capsici* from *C. Annuum* cv. CM334 and resistance to bacterial leaf spot disease from *C. annuum* and *C. chacoense*. • *p-AMT* and*Pun1* markers were used to develop a new fresh cultivar containing capsinoids, named 'Maru Salad'.	Vidhi J: http://www.biologydiscussion.com/vegetable-breeding/top-7-breeding-methods-of-capsicum-india/68448; Boukema (1980), Cook and Guevara (1984), Kim and Hartmann (1985), Hibberd et al. (1987), Boiteux et al. (1994), Berzal-Herranz et al. (1995), de la Cruz et al. (1997), Voorrips et al. (2004), Vallejos et al. (2010), Mallard et al. (2013), Hoang et al. (2013), Soler et al. (2017), Liu et al. (2014), Usman et al. (2008)
Heterosis breeding	• No. 1 Zao Fong: First F1 hybrid of China • Based on CGMS system: Kashi Surkh, Kashi Early, (IIVR, India) • Arka Meghana, Arka Sweta, Arka Harita (IIHR, India) • YU JING NO 2, (China) • F1 Hybrid Coral & F1 Hybrid Dara, Clover Seeds, Hong Kong China • VNR38, VNR108, VNR174, VNR Seeds, India • VNR200, VNR332 • F1 Forever Tropicasem, Senegal Many SSA countries • Remington, F1 Alpha Seeds, South Africa Many SSA countries • F1 TSS AVRDC No.4 Suntech Seeds, Taiwan Taiwan • F1 TSS AVRDC No.2 Yung Shan Seeds, Taiwan Taiwan • F1 Hsing AVRDC No.3 Suntech Seeds, Taiwan Taiwan • (sweet pepper) Yun Pepper No.2 Horticulture Research Institute, YAAS, China China • Yun High Pungency No.1 Horticulture Research Institute, YAAS, China China • Ulka F1, Masaya 315, East-West Seeds, India India • Yuvraj IN	Tong (1998), Gopalkrishnan (2007), Lin et al. (2013), Dhaliwal et al. (2015); IIHR Website: https://www.iihr.res.in/division-varities/786; IIVR website: https://www.iivr.org.in/iivr-varieties/by-crop

(continued)

Table 3.3 (continued)

Breeding approach	Significant achievements/varieties released	References
	• Super F1, Muria F1 East-West Seeds, Thailand Sri Lanka • Hybrid Indus Seeds, India • Based on GMS system: • CH-1, CH-3, CH-27 multiple disease resistant hybrid PAU, Ludhiana, India	
Mutation Breeding	• Horgoskaslatka-X-3–resistant to CMV; Yugoslavia Karasz • Albena: Attractive fruits with better flavor, Bulgaria S. Daskalov, Institute of Genetics, Sofia 113 • Krischimski ran: Hybrid variety with high yield, earliness and improved fruit quality, Bulgaria S. Daskalov and L. Milkova Institute Institute of Genetics, Sofia 113 • MDU 1: Compact plant type with higher yield and capsaicin content, Tamil Nadu Agriculture University, India • Lyulin: Hybrid variety based on induced male sterility, very early and high yield, Institute of Genetics, Sofia 113	Daskalov (1986)

two types of male sterility, CGMS has been preferred over GMS for hybrid seed production because maintenance of GMS shows segregation of male sterility and male fertility (Table 3.3).

In capsicum, CGMS was first reported by Peterson (1958) in the USDA accession PI164835. Till date, no other CMS sources have been reported. In capsicum CMS system, male sterility is caused by two abnormal mitochondrial genes—"*orf507*" and "*atp6-2*" (Kim et al. 2001a, b; Kim and Kim 2005; Gulyas et al. 2006). As the genes are present in the mitochondria, these are maternally inherited. Expression of male sterility also requires the absence of a nuclear gene for the restoration of fertility. For successful hybrid seed development, a restorer line is needed where the restoration of fertility is governed by a single dominant gene. Male sterility maintenance requires a maintainer line with a fertile cytoplasm but the absence of nuclear gene for fertility restoration. As the CGMS system of hybrid seed production requires three lines, i.e., CMS line, maintainer of male-sterile line, and a restorer of fertility in hybrids; it is called three-line system of hybrid seed production in capsicum.

GMS system has also been used for hybrid capsicum production but to a limited extent. In the GMS system, expression of male sterility is controlled by homozygous recessive genes (*ms/ms*) while homozygous dominant or heterozygous plants (*Ms/MS* or *Ms/ms*) exhibit male fertility. Maintenance of male sterility in GMS requires isogenic line with difference only at *Ms* locus is required, i.e., *MsMs* and *Msms*. Crossing between these two lines produces progeny with a mixture of male fertile (Ms/ms) and male sterile (ms/ms) in equal proportions. Male fertile plants are identified in field visually and discarded while the male-sterile lines are used for hybrid seed production (Shifriss 1997, Table 3.3).

3.5 Other Strategies Utilized for Capsicum Improvement and Achievements

Success in capsicum cultivar development initially relied considerably on the breeder's experience, his discretion to isolate promising genotypes as well as luck. Even today with the

availability of advanced breeding techniques, the breeder's experience and judgement are important factors for success in cultivar development. Therefore, plant breeding is still regarded as a combination of science and art. Apart from the established breeding methods for improvement of capsicum, many other techniques have also been attempted and success has been registered to some extent in these strategies which include mutation breeding, polyploid and haploid development, transgenics and marker-assisted breeding.

3.5.1 Mutation Breeding

Mutations are the ultimate source for creating genetic variations. Mutation breeding involves generation of new variability through chemical and physical mutagenesis followed by the development of new varieties utilizing this variability. It is now a pillar of modern plant breeding. Mutation breeding has been found to be effective and efficient breeding tool in capsicum. Daskalov (1986) has written an exhaustive review on this subject. Seeds are the most desirable parts to be treated with mutagen in capsicum. It is recommended to use seeds of uniform size, possessing 96–100% germinability and moisture content (about 13%) to obtain good reproducibility of results. When ionizing radiations are used as mutagen lethal dose should assure survival of 40–60% (Raghavan and Venkatasubban 1940) while it should be 70–80% when chemical mutagens are used (Paran et al. 2007; Hwang et al. 2014; Arisha et al. 2015, Jo et al. 2016).

Bell peppers are in general more radiosensitive than the hot peppers. Pollen grains have also been treated with gamma rays or X-rays and used for the pollination of emasculated nonirradiated flowers immediately after irradiation (Daskalov 1986). The M_1 generation (first generation after mutagen treatment) plants must be raised on isolated plots (at least 700 m apart from other capsicum plants) to prevent cross-pollination followed by bagging of the M_1 flowers to avoid outcrossing. At least 3000–5000 M_1 plants must be raised per experiment. 20–25 M_2 plants per M_1 plant or 10–15 M_2 plants per M_1 fruit (with 2–3 fruits per M_1 plant) are grown in the next generation. The size of the M_2 field population is about 70,000–100,000 plants, but it depends on the kind of selection to be performed and the number of observations to be made. The selection of desirable mutants is carried out mainly in the M_2 generation. To allow progeny testing all discovered mutants must be selfed, usually by bagging the flowers.

Mutation breeding approach has often been used in capsicum for functional gene annotation and also to create novel variability to be utilized in breeding. Sweet pepper cultivar "Maor" has been used to generate mutation population which was later utilized for isolation and characterization of genes controlling plant architecture and flowering (Daskalov 1974; Elitzur et al. 2009; Jeifetz et al. 2011; Cohen et al. 2014). Similar mutation populations have been generated in chili peppers using the cultivar "Yuwol-cho" (Hwang et al. 2014). Jeong et al. (2012) attempted targeted induced local lesions in genome (TILLING) approach in the same cultivar "Yuwol-cho" and successfully isolated a line that exhibited resistance to the *tobacco etch virus* (TEV) from this population.

Daskalov (1968, 1973b) had developed large mutation populations in capsicum using X-rays and gamma irradiation. Novel male-sterile lines were isolated from these populations and then characterized for utilization in breeding. These populations were also utilized to develop capsicum cultivars with useful characteristics such as resistance to *cucumber mosaic virus* (CMV), improved flavor, higher yield, and compact plant stature (Daskalov 1986). Honda et al. (2006) used heavy ion beams (^{12}C and ^{20}Ne ion beams) to develop a mutat population, but the screening of mutants was mainly performed in the M_1 generation. Ultraviolet irradiation has been used in capsicum to create mutants with increased level of vitamin C and E (alpha tocopherol) (Daskalov 1986).

Three male-sterile lines were isolated from capsicum mutant population created by gamma irradiation by Daskalov and Mihailov (1988) and

subsequently utilized in hybrid development. Recently, mutants with changed shoot architecture in hot pepper (Paran et al. 2007), some induced mutants in sweet pepper (Honda et al. 2006) and capsicum with increased β-carotene and orange color on maturity (Tomlecova et al. 2009a, b) have been recovered. After gamma irradiation of the dry seeds of capsicum (*C. annuum* L.), many promising mutants were obtained, the most interesting of which were induced male-sterile mutations. Male sterility is governed by single recessive genes, denoted by *ms-3, ms-4, ms-6, ms-7,* and *ms-8.* The male-sterile lines *Pazardjishka kapia* ms-3 and *Zlaten medal* ms-8 recovered after mutagen treatment were used to test their combining ability against original male-sterile line used for hybrid production. The results obtained indicate that there is no significant difference in the combining ability for early and total yields. Three male-sterile lines were crossed with a large number of lines in order to obtain hybrid combinations for different purposes. Most of the hybrid combinations exceeded the check with regard to early yield. Some hybrids were also characterized by an increase in total yield. Two hybrid combinations, named Krichimski ran and Lyulin, were released as cultivars utilizing the male-sterile lines recovered from mutat population.

3.5.2 Polyploidy Breeding

Polyploidization events are often associated with increase in vigor followed by adaptation of the newly formed polyploid to novel conditions. According to Van de Peer et al. (2009), superiority of polyploids over their diploid counterparts has been attributed to the phenomenon of transgressive segregation, i.e., formation of extreme phenotypes. Malhova's 1977 work suggests that *capsicum* may respond to changes in ploidy in the same way as *Solanum.* It is relatively easy to increase or decrease ploidy levels artificially in *capsicum*. Somatic doubling can be achieved by treating wounded leaf axils with colchicine. However, synthetic autotetraploids seem to have no agronomic or breeding advantages over diploids. Polyploid capsicum usually expresses morphological characters like stunted growth and the presence of larger, thicker, and dark green leaves (Tapadar 1963; Bose and Panigrahy 1969; Biswas and Bhattacharyya 1971; Indira and Susan 1977). The deep green color of the leaves in polyploids has been attributed to the presence of more numerous and larger chloroplasts (Raghuvanshi and Joshi 1964). Tetraploid capsicum exhibit increased dry weight in leaf, stem and root, leaf area and thickness when compared with the diploid one. The tetraploids have reported increased ability to absorb water, NO_3–N and K with a consequent increase in the photosynthetic ability; and bear smaller but uniform-size fruits, independent of fruit loading (Takizawa et al. 2008).

The tetraploid capsicum has been found to flower about one month later than the diploids. The total number of flowers produced was less; this decrease being primarily due to the non-branching nature of the polyploidy (Tapadar 1963; Biswas and Bhattacharyya 1971; Indira and Susan 1977). Raghuvanshi and Sheila (1964) observed delayed and extended flowering with larger and varied number of floral parts in the colchiploids of *capsicum frutescens*. Larger flowers with increased size of pollen grains are also characteristic of polyploids (Watts 1980). Colchicine treatment of seeds has produced tetraploid plants of *C. annuum* variety "Chigusa" (Nihon Horticultural Production Institute) (Ishikawa et al. 1997). Flow cytometric analyses of these seeds showed that ≈20% of the seeds treated with colchicine were tetraploid. In comparison with diploid flowers which typically had six petals and stamens tetraploid flowers had seven petals and stamens, 20% larger ovaries, and 25% larger diameter pollen grains (Ishikawa 2001). Polyploids have also been reported to have sterility which may be attributed to abnormalities observed in meiosis (Pal et al. 1941). Following colchicine treatment, a plant of chili pepper cv. CO-2 was found to have chromosome numbers ranging from $2n = 38$ to 96. It had 4.95% pollen fertility and set no seed, and its growth was stunted (Rao 1987). Haploids

produced through anther culture have been doubled using colchicines, but the homozygotes thus produced have not as yet been exploited to produce commercial F_1 hybrids with exhibiting heterosis but have been used to study the genetic mechanism of resistance to pests (Hendy et al. 1985) and diseases (Daubeze et al. 1990; Palloix 1992).

Malhova (1977) produced interspecific hybrid which otherwise was difficult between *capsicum. pubescens* and *C. annuum* by pollinating *C. pubescens* with pollen of autotetraploid *C. annuum*. This result gives a direction that induced autotetraploidy may be used to overcome post-fertilization barriers in other interspecific crosses of *capsicum* genus. Pochard (1970, 1977) has produced a set of trisomies for *C. annuum.* These trisomies can be utilized to identify genes present on particular chromosomes either because of distorted segregation ratios which occur in the progeny of F_1 hybrids trisomic for that chromosome (Pochard 1977) or because of dosage effects which can be detected when trisomics are compared to normal diploid individuals (Tanksley 1984). Location of gene *c* (controlling pungency) on acrocentric chromosome number '*XI*' (Pochard 1977) and its presence on long arm (Pickersgill 1977) were confirmed using these trisomics as the trait pungency segregated independently of the markers present one short arms of the acrocentric chromosomes.

3.5.3 Haploid Breeding

First haploids in the *genus capsicum* were developed through in vitro anther culture of *C. annuum* and *C. frutescens* (George and Narayanaswamy 1973; Kuo et al. 1973; Wang et al. 1973; Novak 1974). Lower recovery of haploid plants from androgenic cultures in earlier studies encouraged to design experiments which aimed to identify the factors influencing induction of androgenesis. From the various experiments conducted on haploid induction, it was concluded that androgenic response depended on growing conditions, age, and genotype of the donor plant (Ercan et al. 2006; Niklas-Nowak et al. 2012; Grozeva et al. 2013; Koleva-Gudeva et al. 2013; Alremi et al. 2014), developmental stage of microspores in the anther (Nowaczyk and Kisiała 2006; Parra-Vega et al. 2013; Barroso et al. 2015), culture medium composition, concentration and combination of growth regulators, organic and inorganic additives (Büyükalaca et al. 2004; Zhao et al. 2010; Taşkin et al. 2011; Roshany et al. 2013; Olszewska et al. 2014), and pretreatment of flower buds and/or anthers (Koleva-Gudeva 2007; Özkum and Tıpırdamaz 2007; Irikova et al. 2011; Nowaczyk et al. 2015).

Technology for development of doubled haploids is one of the fastest technique to achieve complete homozygosity in any crop species, but its application in capsicum improvement is still limited because of recalcitrant nature of capsicum (Grozeva et al. 2009; Ercan and Sensoy 2011; Olszewska et al. 2014). Capsicum breeding requires genetically stable and homozygous plants to understand genetics as well as mapping and identification of genes for various morphological traits and biotic and abiotic stress-related traits. Despite low frequency of results, several studies concerning the practical aspect of the haploid breeding in different *capsicum* species is being undertaken (Olszewska et al. 2010, 2011; Shrestha et al. 2010; Luitel et al. 2012; Luitel and Kang 2013a, b; Shmykova et al. 2014; Trajkova and Koleva-Gudeva 2014). There have been reports on development of varieties and F_1 hybrids based on parental lines developed from doubled haploid (DH) technology (Chunling and Baojun 1995; Pauk et al. 2010). DH capsicum lines with improved yield characteristics and dry matter content in fruits were also obtained (Kisiała et al. 2011). Superior DH lines with considerable variation in plant and fruit traits (Shrestha et al. 2011) and androgenic capsicum lines with positive traits have also been isolated (Koleva-Gudeva and Trajkova 2012). Capsicum DHs with improved quality aspect like fruit shape, taste, fruit firmness, dry matter content, total soluble content, phenolic content, and antioxidant activity like CUPRAC and FRAP have been developed (Luitel and Kang 2013b).

Nowaczyk et al. (2014) used DH technology for stabilization of soft-flesh *capsicum* spp. recombinants.

The DH lines obtained from anther culture of capsicum in vitro exhibited different levels of resistance to *Xanthomonas campestris* pv. *vesicatoria* (Hwang et al. 1998) and *Phytophthora capsici* (Nervo et al. 2007). These resistant DH lines can be used to develop new multiple-disease–resistant genotypes. Resistant lines to PVY and lines with important qualitative and quantitative traits have also been isolated through anther culture (Arnedo Andrés et al. 2002; Mitykó and Gémes Juhász 2006). Todorova et al. (2013) recovered capsicum lines with high productivity, improved fruit traits, and low susceptibility to Verticillium wilt were produced through haploid culture. Application of microspore embryogenesis has been used to create genotypes with improved productivity, resistance to *Verticillium dahliae* Kleb (Grozeva et al. 2009; Koleva-Gudeva and Trajkova 2012; Todorova et al. 2013; Trajkova and Koleva-Gudeva 2014) and *Tobacco mosaic virus* (TMV).

3.5.4 Embryo Rescue

Embryo rescue has most often been used to overcome the post-zygotic hybridization barriers in interspecific crosses. Incompatibility during hybridization is more common among *capsicum* species belonging to different gene pools, but incompatility has been reported within same gene pool also as between *C. annuum* and *capsicum chinense* or *C. frutescens*. Many interspecific crosses in *capsicum* spp. produce fruits with shriveled seeds which are incapable of germinating normally because endosperm and/or embryo have not developed properly. There have been reports on successful recovery of hybrid embryo of interspecific crosses in *capsicum* genus. The first attempt of embryo rescue in *capsicum* spp. was done by Fari et al. (1983) where embryo was recovered from the cross of *C. annuum* and *C. baccatum*. Another example for wide hybridization is between *C. annuum* and *C. baccatum* where immature interspecific embryos or embryo were/was rescued before abortion occurs (Shivanna and Bahadur 2015). Embryo excision and in vitro embryo culture is a technically complex process. Also, the stage at which embryo abortion occurs after hybridization may depend on the specific genotypes involved in the cross. Within *capsicum* genus, some authors could rescue interspecific embryos at the latest immature stages (Yoon et al. 2006) while there are also examples in which embryos had to be rescued at the earliest stages (Hossain et al. 2003; Manzur et al. 2015). However, embryo rescue at earlier stage is more difficult with lower efficiency of recovering interspecific hybrids (Shen et al. 2011). Anthracnose resistance found in *C. baccatum* lines has been introgressed into *C. annuum* via the rescue followed by culture of embryo obtained from interspecific crosses between these two species (Yoon et al. 2006). Genetic bridge which is the based on the use of phylogenetically closer species to the two species affected by crossability barriers is an alternative approach to overcome the above problem. In this method, the bridge species is used to which has the ability to cross with both the target species. The bridge species is crossed first with one target species, and the hybrid so obtained is then crossed with the second target species (Shivanna and Bahadur 2015). *C. chinensis* has been found to be an ideal bridge species to perform the wide hybridization between *C. annuum* and *C. baccatum* (Pickersgill 1988).

3.5.5 Transgenic Development

Genetic transformation has provided an alternative approach for capsicum improvement program. The major advantage which transgenics technology offers is that it overcomes interspecific or intergeneric barriers and enables transfer of useful genes or novel traits into capsicum. The first capsicum transformation work has been reported in 1990 (Liu et al. 1990). However, poor reproducibility of capsicum is the major limiting factor for capsicum transformation studies. Table 3.4 summarizes the major efforts

Table 3.4 Transgenic studies in capsicum using Agrobacterium mediated transformation with CAMV 35S promoters

Genotype	Explant	Agrobacterium strain	Selection marker	Gene of interest/Reporter gene	Vector	Reference
Yolo Wonder, California Wonder, NVH3051, Jupiter, Liberty Bell, Guatemalan wild accession	Leaves, cotyledons, hypocotyls	A281/C58	Kanamycin (Neomycin phosphotransferase II (*nptII*)	B-glucuronidase (GUS)	p_3-1-GUS	Liu et al. (1990)
Golden tower	Cotyledons	LBA 4404	Kanamycin, carbenicillin	Cucumber Mosaic Virs I_{17}N-satellite RNA	pRok/105	Lee et al. (1993)
Zhong Hua No. 2	Leaves, cotyledons, hypocotyls	GV3111-SE	Kanamycin	CMV-CP	pHCM40	Zhu et al. (1996)
Pusa Jwala	Cotyledons	EHA 105	Kanamycin	GUS	pBI 121	Manoharan et al. (1998)
No. 40017	Cotyledons	C58CIRifR, LBA 4404, EHA 101, A281	Kanamycin, geneticin, hygromycin (*hpt*), methorexate (*dhfr*), phosphinothricin (*bar*)	GUS	pRGG plasmid set	Mihalka et al. (2000)
Nockkwang	Hypocotyls	LBA 4404	Kanamycin	*OsMADS1*	pGA1209	Kim et al. (2001a, b)
'Pusa Jwala'	Cotyledons	C58	Kanamycin	*GUS*	pGV1040	Shivegowda et al. (2002)
VS300-1	Young embryonic tissue, Cotyedons	LBA 4404	Chlorsulfuron (*sur B*)	*CaCel1*	pWTT2132	Harpester et al. (2002)
Golden tower	Hypocotyls, Cotyledons	LBA 4404	Kanamycin	TsiI	pMBP2	Shin et al. (2002b)
Nockkwang	Hypocotyls, Cotyledon		Kanamycin	GUS	pMBP2	Shin et al. (2002a)
F1 Xiangyan 10, Zhongjiao, Zhongjiao 5, Zhongjiao 6	Cotyledons	LBA 4404	Kanamycin	GUS	pBI121	Li et al. (2003)

(continued)

Table 3.4 (continued)

Genotype	Explant	Agrobacterium strain	Selection marker	Gene of interest/Reporter gene	Vector	Reference
Feherozon	Cotyledons	ShooterGRif^{R}, GV3170Rif^{R}	Hygromycin, Kanamycin	GUS	pRGG*hpt*, pRGG*neo*	Mihalka et al. (2003)
P915, P4090, P410, P101	Hypocotyls, Cotyledon	EHA 105, LBA 4404	Kanamycin	*TMV-CP, PPII*	pCAMBIA	Lee et al. (2004)
'K1', 'K2', 'PLR1'	Shoot tips	EHA 105	Cefotaxime and Kanamycin	*GUS*	pIG121 Hm	Sobhakumari and Lalithakumari (2005)
Byadagi Kaddi	Cotyledon, Hypocotyl, Cotyledonary nodal region/shoot tip	EHA105	Kanamycin	*cryI(Ac)*	pBinBt3	Channappagoudar (2007)
Nockkwang	Hypocotyls, Cotyledon	GV3101	Hygromycin	GUS	pCAMBIA1301	Ko et al. (2007)
C. frutescens	Pollen transformation	EHA 101	Hygromycin	*GUS*	pCAMBIA 1301	Sharma et al. (2008)
Habanero Pepper 'Naranja'	Leaf	LBA 4404	Rifampicin, Streptomycin, Kanamycin	GUS	pCAMBIA2301/pCAMex, pCAMex::PR10/pCAMex::Esterase	Arcos-Ortega et al. (2010)
F1 hybrids Fiesta, Ferrari and Spirit	Cotyledons	GV3101	Kanamycin, Cefotaxime, thidiazuron (TDZ)	BnBBM	pMP90 Ti plasmid	Heidmann et al. (2011)
'Arka Gaurav', 'Arka Mohini'	Hypocotyl	EHA 105	Kanamycin	*GUS*	pBinARβC1	Kumar et al. (2012)
California Wonder		LBA 4404	Kanamycin and rifampicin	GUS	pBI121	Verma et al. (2013)
Pusa Sadabahar, Pusa Jwala	Cotyledons, Hyotcotyls	LBA 4404	Kanamycin	*GUS*	pCAMBIA2301	Mahto et al. (2018)

made in developing transgenic capsicum. Major capsicum transformation work has been done for disease resistance particularly against viruses, viz, tobacco mosaic virus (TMV), pepper mild mottle virus (PMMV) (Lee et al. 2004), tomato mosaic virus (ToMV) (Shin et al. 2002a), and cucumber mosic virus (CMV) (Shin et al. 2002b). Such transgenic virus resistance mechanism utilizing viral coat protein regions and satellite RNA is currently known as RNA silencing (Voinnet 2001). Transformation and overexpression of *TsiI*, a tobacco pathogenesis-related (PR) gene in capsicum displayed broad spectrum resistance against different pathogens like PMMV, CMV, bacterial pathogen *Xanthomonas campestris* pv. *vesicatoria,* and a fungal pathogen *P. capsici* (Shin et al. 2002a). Besides disease resistance, transformation studies in capsicum on other aspects are limited. Suppression of ripening-related endo-1.4-β-glucanase in transgenic capsicum was demonstrated by Harpester et al. (2002). Studies on transformation of foreign genes identified from other plants or organisms into capsicum are quite rare. A dwarf transgenic capsicum has been produced upon transformation with *OsMADS1* gene from rice (Kim et al. 2001a, b). RNA silencing approach has been used to identify a new gene ketoacyl-ACP reductase (*CaKR1*) in capsicum responsible for producing nonpungent fruits (Koeda et al. 2019). Capsicum transformation studies has most commonly used GUS gene (β-glucuronidase) as reporter gene, CaMV 35S as promoter, and nopaline synthase (NOS) as terminator gene (Liu et al. 1990; Zhu et al. 1996; Manoharan et al. 1998; Li et al. 2003; Mihalka et al. 2003; Kim et al. 2001a, b; Shin et al. 2002a; Lee et al. 2004). Capsicum transgenics have been most commonly developed employing agrobacterium mediated transformation using cotyledons and/or hypocotyls as explants in most studies (Manoharan et al. 1998; Pozueta-Romero et al. 2001; Kim et al. 2001a, b; Shin et al. 2002a, b; Lee et al. 2004). Direct transformation using gene gun has been attempted recently in *C. frutescens* (Chee et al. 2018).

3.5.6 Marker-Assisted Breeding

Molecular marker-assisted breeding (MAB), also called molecular-assited breeding has been now being widely utilized in improvement of capsicum. Different types of molecular markers have been developed for capsicum like isozyme markers, amplified fragment length polymorphism (AFLPs), random amplified polymorphic DNA (RAPD), restriction fragment length polymorphism (RFLPs), simple sequence repeat (SSR), and single nucleotide polymorphism (SNP and conserved ortholog set (COS) II markers. These markers have been extensively utilized to understand inheritance of important traits as well as mapping of horticultural and disease resistance genes and quantitative trait loci (QTLs).

3.6 Limitations of Traditional Breeding

Conventional plant breeding is the principle approach to crop improvement where plant genomes are manipulated within the primary gene pool of the genus. It involves methods such as hybridization followed by selection, introgression by backcross breeding, creation of new variants through induced mutagenesis and somatic hybridization to create newer combination of different genomes. Identification of commercially important new traits is based on phenotypic assessments of segregating progenies.

There are a number of issues where the applicability of conventional plant breeding in enhancement of quality and yield beyond a certain point becomes very difficult. The major distinction between modern and traditional breeding is the separation between phenotype and genotype. Genes are inherited while the phenotype is an expression of those genes within an environment. Genetic variation is induced at the DNA level, but screening and selection of genotypes is done based on phenotypic

expression. As a result, new cultivars not only contain traits of breeder's interest but also display a number of undesirable features not considered during the selection process and through conventional breeding this transfer of undesirable traits from existing to new varieties is generally inevitable.

A second issue is encountered when breeders try to capture the genetic diversity available within sexually incompatibile groups. The new traits are introgressed into cultivated varieties through wide hybridization and extensive backcrossing of generated hybrids with recipient parent. However, the targeted traits of interest do not come alone but come along with larger segments of wild chromosomes or in other words are associated with linkage drag and this linkage drag might contain genes some of which are undesirable. The third limitation results from the inability of traditional breeding to control expression of target genes in a new genetic background. The use of modern breeding strategies such as marker-assisted selection not only accelerates the introgression process but also aim to reduce linkage drag.

In short, it can be said that traditional methods in plant breeding will continue to develop new and improved varieties. However, these methods are, by themselves, not sufficient to allow complete expression of genetic potential of a genotype, where molecular breeding plays its role being precise, rapid and cost effective in comparison with conventional phenotypic selection.

3.7 Need of Molecular Breeding

Direct selection which is based on the performance of the genotype or phenotypic values of target traits is effective for qualitative traits while selection for quantitatively inherited complex character is often elusive due to environmental influence. Under such condition, indirect selection is said to be a better alternative. Conventional breeding has met with limited success due to polygenic control of resistance traits, wide range of pathogen strains distributed in different environments, complexity of host–pathogen interaction and wide variability in pathogenicity. Indirect selection can be based on some other traits which is easily measured but tightly linked to the other traits of interest which is difficult to measure or is influenced by environment. Indirect selection for yield is limited via other traits. Due to the limitation of equipment, facilities, and resources, selection for target genes often becomes impractical. Advent of molecular (DNA) markers has created a powerful and practicable tool to perform gene selection in plant breeding. Although marker-assisted gene selection is not a real gene selection, it provides the best indirect selection tool for target genes at the DNA level. So, marker-assisted selection (MAS) is an effective and reliable approach.

Kole and Gupta (2004) and Collard and Mackill (2008) have illustrated the advantages of MAS when compared with conventional phenotypic selection. Selection using molecular markers is simpler compared to phenotypic breeding. Apart from this, selection may be carried out at any stage of the plants and single plant may be selected with high reliability. Molecular markers have offered large opportunities ranging from the localization of a gene to development of newer genotype combinations having good yield with stress-resistant genes. This saves a lot of time in the breeding process. They have aided in discovering more information about the function of the genes of interest. Apart from gene location and its selection, molecular markers assist in genetic diversity assessment, quality control, and marker-assisted breeding. To meet the growing demand for increasing capsicum production as well as disease-resistant genotypes, use of molecular markers is important to hasten the pace of improvement program. Molecular markers in capsicum have been utilized for DNA fingerprinting as well as genetic diversity analysis of capsicum (Bahrami et al. 2009; Hossain et al. 2014; Rego et al. 2011; Costa et al. 2016), QTL analysis of important biotic stresses (Lee et al. 2011; Lu et al. 2012; Dwivedi et al. 2013; Han et al. 2018) and MAS (Grube et al. 1996; Tanaka et al. 2014; Jeong et al. 2015; Suwor et al. 2017). Estimating the genetic diversity among capsicum genotypes helps in reliable differentiation of

genotypes. Genetic diversity analysis and varietal identification in capsicum have been carried out using different types of marker system like isozymes (Litoriya et al. 2010); RAPD (Bhadragoudar and Patil 2011, Thul et al. 2012); AFLP (Lafebvre et al. 2001; Ibiza et al. 2012); SSR (Ibiza et al. 2012) and ISSR markers (Thul et al. 2012). Genetic diversity assessment using molecular markers is useful in selecting diverse parental combinations for hybrid development, understanding evolutionary relationship between different capsicum species and for exact varietal identification. Molecular characterization of germplasm is important for the conservation and utilization of plant genetic resources (Thul et al. 2012). MAS is a molecular breeding technique that helps to avoid the difficulties concerned with conventional plant breeding. Advances in the science of genomics has led to the identification of thousands of DNA markers in capsicum which include mapped micro-satellite markers and more recently, single nucleotide polymorphisms (SNPs) (Huang et al. 2001; Lee et al. 2013; Buso et al. 2016; Cheng et al. 2016; Taranto et al. 2016).

With the SSRs and SNPs, some genes controlling biotic and abiotic stress resistances, quality characters and various aspects of plant development have been cloned and characterized in capsicum, which are excellent assets for molecular-assisted breeding (Sect. 3.5.5). At present, SSRs are the most widely used markers by capsicum researchers due to their availability in large numbers in the public domain including their simplicity and effectiveness (Cheng et al. 2016; https://solgenomics.net/). Genes/QTLs for several important traits of capsicum such as pungency (Lee et al. 2005; Stewart et al. 2005, 2007), fertility restoration (Zhang et al. 2000; Min et al. 2008; Jo et al. 2010), soft flesh and deciduous fruits (Rao and Paran 2003), capsanthin content (Lefebvre et al. 1998), fruit size and shape (Ben Chaim et al. 2001; Rao et al. 2003), male sterility (Chen et al. 2012), parthenocarpy (Tiwari et al. 2011), resistance to CMV (Kang et al. 2010), potyviruses (Murphy et al. 1998; Kang et al. 2005), chili veinal mottle virus (Ruffel et al. 2006; Hwang et al. 2009), tobamoviruses (Berzal-Herranz et al. 1995; Tomita et al. 2011), bacterial spot (Tai et al. 1999; Mazourek et al. 2009; Pierre et al. 2000; Jordan et al. 2006), anthracnose (Voorrips et al. 2004), Phytophthora rot (Thabuis et al. 2003; Ogundiwin et al. 2005; Kim et al. 2008), powdery mildew (Lefebvre et al. 2003), and root-knot nematodes (Djian-Caporalino et al. 2001, 2007) have been mapped, and some of them have been utilized for MAS (Grube et al. 1996; Tanaka et al. 2014; Jeong et al. 2015; Suwor et al. 2017) which otherwise would have been very difficult using conventional breeding.

3.8 Future Prospects

Achievements made in capsicum breeding illustrate its possibilities for further improvement. There is considerable opportunity for further improvement of capsicum. With the availability of whole genome sequence of capsicum and single nucleotide polymorphism (SNP) discovery through genotyping by sequencing method, the genetic diversity in capsicum has been unraveled. With this genomic information in hand, we can believe that the genetic makeup of capsicum may be modified to a much greater extent than we normally appreciate. However, studies reporting association between this vast genetic diversity and the observed phenotypic variability is still poor. Finding new associations between the generated genomic resources and important traits of importance in capsicum such as fruit size, fruit production, pungency, abiotic stress tolerance, nutritional content, and disease resistance is an important research area. The exploitation of transgenic technology in capsicum is also slow as capsicum is highly recalcitrant to transformation and regeneration process. Further with the availability of capsicum genome sequence, latest genome-editing technologies and their potential applications in the genetic improvement of capsicum can be explored. However, lack of well-characterized target gene information is a major limiting factor that restricts the broad application of gene/genome-editing technologies to capsicum. Utilization of new upcoming

technologies will continue to advance which in combination with traditional techniques of selections and crosses already established in capsicum genetic breeding will become an essential tool.

References

Alremi F, Taskın H, Sönmez K, Buyukalaca S, Ellialtioglu S (2014) Effect of genotype and nutrient medium on anther culture of pepper (*Capsicum annuum* L.). Turk J Agric Nat Sci 1:108–116

Aniel Kumar O, Anitha V, Roseline Subhashini K, Raja Rao KG (2001) Induced morphological mutations in *Capsicum annuum* L. Capsicum Eggplant Newsl 20:72–75

Arcos-Ortega GF, Chan-Kuuk RA, González-Kantún WA, Souza-Perera R, Nakazawa-Ueji YE, Avilés--Berzunza E, Godoy-Hernández G, Lawton MA, Aguilar JJZ (2010) Agrobacterium tumefaciens-transient genetic transformation of Habanero pepper (*Capsicum chinense* Jacq.) leaf explants. Electron J Biotechnol 13(4). https://doi.org/10.2225/vol13-issue4-fulltext-1

Arisha MH, Shah SN, Gong ZH, Jing H, Li C, Zhang HX (2015) Ethyl methane sulfonate induced mutations in M_2 generation and physiological variations in M_1 generation of peppers (*Capsicum annuum* L.). Front Plant Sci 6:399

Arnedo-Andres MS, Gil-Ortega R, Luis-Arteaga M, Hormaza JI (2002) Development of RAPD and SCAR markers linked to the *Pvr4* locus for resistance to PVY in pepper (*Capsicum annuum* L.). Theor Appl Genet 105:1067–1074

Bahrami RM, Hassani ME, Mohammadi A, Lessan SH, Ghazi Zade S (2009) Evaluation of genetic diversity in *Capsicum* spp. as revealed by RAPD markers. Acta Hort 829:40

Barroso PA, Rego MM, Rego ER, Soares WS (2015) Embryogenesis in the anthers of different ornamental pepper (*Capsicum annuum* L.) genotypes. Genet Mol Res 14:13349–13363

Ben Chaim, Paran I, Grube RC, Jahn M, van Wijk M, Peleman J (2001) QTL mapping of fruit-related traits in pepper (*Capsicum annuum*). Theor Appl Genet 102:1016–1028

BenChaim A, Grube RC, Lapidot M, Jahn M, Paran I (2001) Identification of quantitative trait loci associated with resistance to cucumber mosaic virus in *Capsicum annuum*. Theor Appl Genet 102:1213–1220

Bergh BO, Lippert LF (1964) Six new mutant genes in the pepper, *Capsicum annuum* L. Amer Nat 99:159–166

Berzal-Herranz A, de la Cruz A, Tenllado F, Diaz-Ruiz JR, Lopez L, Sanz AI, Vaquero C, Serra MT, Garcia-Luque I (1995) The *Capsicum* L^3 gene-mediated resistance against tobamovirus is elicited by the coat protein. Virology 209:498–505

Bhadragoudar MR, Patil CG (2011) Assessment of genetic diversity among *Capsicum annuum* L. genotypes using RAPD markers. Afr J Biotechnol 10 (76):17477–17483

Bhattacharya A, Chattopadhyay A, Mazumdar D, Chakravarty A, Pal S (2010) Antioxidant constituents and enzyme activities in chilli peppers. Intl J Veg Sci 16:201–211

Biswas AK, Bhattacharyya NK (1971) Induced polyploidy in legumes I. *Cyamopsis psoraloides* D.C. Cytologia 36:469–479

Black LL, Hobbs HA, Gatti JM (1991) Tomato spotted wilt virus resistance in *Capsicum chinense* PI152225 and 159236. Plant Dis 75:863

Boiteux LS (1995) Allelic relationships between genes for resistance to tomato spotted wilt tospovirus in *Capsicum chinense*. Theor Appl Genet 90:146–149

Boiteux LS, de Avila AC (1994) Inheritance of a resistance specific to tomato spotted wilt tospovirus in *Capsicum chinense*'PI 159236'. Euphytica 75:139–142

Bose S, Panigrahi UC (1969) Studies on induced polyploidy in *Zinnia linearis* Benth. Cytologia 34:103–111

Bosland PW, Votava EJ (2012) Peppers: vegetable and spice *Capsicums*, 2nd edn. CABI Publishing, London, UK

Boswell VR (1937) Improvement and genetics of tomatoes, peppers and eggplant. Yearbook of Agriculture. U.S. Govt. Printing Office, Washington, pp 176–206

Boukema IW (1980) Allelism of genes controlling resistance to TMV in *Capsicum* L. Euphytica 29:433–439

Boukema IW (1982) Resistance to a new strain of TMV in *Capsicum chacoense* Hunz. Capsicum Newsl 1:49–51

Boukema IW (1984) Resistance to TMV in *Capsicum chacoense* Hunz. is governed by allele of the *L*-locus. Capsicum Newsl 3:47–48

Boukema IW, Jansen K, Hofman K (1980) Strains of TMV and genes for resistance in *Capsicum.* In: Proceedings of capsicum working group meeting. Wageningen, The Netherlands, pp 44–48

Buso G, Reis AMM, deSouza Amaral ZP, Ferreria ME (2016) Novel and highly informative Capsicum SSR markers and their cross-species transferability. Genet Mol Res 15(3). https://doi.org/10.4238/gmr.15038689

Buyukalaca S, Comlekcioglu N, Abak K, Ekbic E, Kilic N (2004) Effect of silver nitrate and donor plant growing conditions on production of pepper (*Capsicum annuum L.*) haploid embryos *via* anther culture. Eur J Hort Sci 69:206–209

Cebolla-Cornejo J, Soler S, Gomar B, Soria MD, Nuez F (2003) Screening Capsicum germplasm for resistance to *tomato spotted wilt virus* (TSWV). Ann Appl Biol 143(2):143–152

Chalukova M, Daskalov S, Lukarska E, Baralieva D (1993) Beta orange mutant in pepper (*Capsicum annuum* L.). Capsicum Eggplant Newsl 12:57–58

Channappagoudar SB (2007) Studies on *in vitro* regeneration and genetic transformation in chilli (*Capsicum annuum* L.). Ph.D. thesis, University of Agricultural Sciences, Dharwad, India

Chee M, Lycett GW, Foan C (2018) Development of a direct transformation method by GFP screening and *in vitro* whole plant regeneration of *Capsicum frutescens* L. Electron J Biotechnol 34:51–58

Chen C, Hao X, Chen G, Cao B, Chen Q, Liu S et al (2012) Characterization of a new male sterility-related gene *CaMF1* in *Capsicum annuum* L. Mol Biol Rep 39:737–744

Cheng J, Zhao Z, Li B, Qin C, Wu Z, Trejo-Saavedra DL et al (2016) A comprehensive characterization of simple sequence repeats in pepper genomes provides valuable resources for marker development in *Capsicum*. Sci Rep 6:18919

Chunling L, Baojun Y (1995) Successful development of new sweet (hot) pepper cultivars by anther culture. Acta Hort 402:442–444

Cohen O, Borovsky Y, David-Schwartz R, Paran I (2014) *Capsicum annuum S* (*CaS*) promotes reproductive transition and is required for flower formation in pepper (*Capsicum annuum*). New Phytol 202:1014–1023

Collard BC, Mackill DJ (2008) Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Phil Trans Roy Soc Lond B Biol Sci 363:557–572

Cook AA, Stall RE (1963) Inheritance of resistance in pepper to bacterial spot. Phytopathology 53:1060–1062

Cook AA, Guevara YG (1984) Hypersensitivity in *Capsicum chacoense* to race 1 of pepper. Plant Dis 68:329–330

Costa MPSD, Rêgo MMD, da Silva APG, do Rêgo1 ER, Barroso PA (2016) Characterization and genetic diversity of pepper (*Capsicum* spp) parents and interspecific hybrids. Genet Mol Res 15(2). https://doi.org/10.4238/gmr.15027652

Csillery G, Szarka E, Sardi E, Mityko J, Kapitany J, Nagy B, Szarka J (2004) The unity of plant defence: genetics, breeding and physiology. In: Proceedings of 12th Eucarpia meeting on genetics and breeding of capsicum and eggplant, 17–19 May 2004, Noordwijkerhout, The Netherlands, 147–153

Daskalov S (1973a) Gene list for the pepper. Genet Plant Breed 6:401–408

Daskalov S (1973b) Investigations on induced mutagemnesis in *Capsicum annuum* L. III Mutants in the variety Zlaten Medal. Genet Plant Breed 6:419–429

Daskalov S (1974) Investigation on induced mutants in sweet pepper (*Capsicum annuum* L.). In: Proceedings of 1st meeting of the capsicum breeding and genetics, 1–4 July 1974, Budapest, Hungary, pp 81–90

Daskalov S (1968) A male sterile pepper (*C.annuum*L.) mutant. Theor Appl Genet 38:370–372

Daskalov S (1973c) Investigation of induced mutants in*Capsicum annuum*L. III. Mutants in the variety *Zlaten* medal. Genet Selekts 6:419–429

Daskalov S (1986) Mutation breeding in pepper. In: Micke et al (ed) Mutation breeding review. International Atomic Energy Agency/FAO, Vienna, No. 4, p 25

Daskalov S, Chalukova M, Baralieva D, Lukarska E (1995) Biochemical investigations of an induced beta-orange mutant in sweet pepper (*Capsicum annuum* L.) and developing varieties with increased Beta carotene content. In: Proceedings of 9th Eucarpa meeting on genetics and breeding of capsicum and eggplant, 21–25 Aug 1995, Budapest, Hungry, pp 24–27

Daskalov S, Mihailov L (1988) A new method for hybrid seed production based on cytoplasmic male sterility combined with a lethal gene and a female sterile pollenizer in *Capsicum annuum* L. Theor Appl Genet 76:530–532

Daskalov S, Poulos JM (1994) Updated Capsicum gene list. Capsicum Eggplant Newsl 13:16–26

Daubeze AM, Palloix A, Pochard E (1990) Resistance of androgenic autodiploid lines of pepper to *Phytophthora capsici* and tobacco mosaic virus under high temperature. Capsicum Newsl 8(9):47–48

de la Cruz A, Lopez L, Tenllado F, Dıaz-Ruız JR, Sanz AI, Vaquero C, Serra MT, Garcıa-Luque I (1997) The coat protein is required for the elicitation of the *Capsicum* L^2 gene-mediated resistance against the tobamoviruses. Mol Plant-Microb Interact 10:107–113

de Swart EAM (2007) Potential for breeding sweet pepper adapted to cooler growing conditions—a physiological and genetic analysis of growth traits in Capsicum. Ph.D. thesis, Wageningen University, Wageningen, The Netherlands

Deshpande RB (1933) Studies in Indian chillies. 3. The inheritance of some characters in *Capsicum annuum* L. Indian J Agric Sci 3:219–300

Deshpande RB (1935) Studies in Indian chillies: 4. Inheritance of pungency in *Capsicum annuum* L. Indian J Agric Sci 5:513–516

Dhaliwal MS, Jindal SK, Cheema DS (2015) CH-27: a multiple disease resistant chilli hybrid. Agric Res J 52 (4):127–129

Dhall RK (2008) Breeding for quality traits in chilli: a review. J Res Punjab Agric Univ 45(3 & 4):156–160

Djian-Caporalino C, Berthou F, Fazari A, Lefebre V, Paloix A, Pegard A, Pijarowski L (2004) Genetic, cytological and molecular bases of resistance to root knot nematode (*Meloidogyne* spp) in pepper (*Capsicum annuum* L.). In: Voorips RE (ed) Proceedings of 12th Eucarpa meeting on genetics and breeding of capsicum and eggplant, 17–19 May, 2004. Noordwijkerhout, The Netherlands, p 180

Djian-Caporalino C, Fazari A, Arguel MJ, Vernie T, VandeCasteele C, Faure I, Brunoud G, Pijarowski L, Palloix A, Lefebvre V, Abad P (2007) Root-knot nematode (*Meloidogyne* spp.) *Me* resistance genes in pepper (*Capsicum annuum* L.) are clustered on the P9 chromosome. Theor Appl Genet 114:473–486

Djian-Caporalino C, Pijarowski L, Fazari A, Samson M, Gaveau L, O'Byrne C, Lefebvre V, Caranta C, Palloix A, Abad P (2001) High-resolution genetic mapping of the pepper (*Capsicum annuum* L.) resistance loci Me3 and Me4 conferring heat-stable resistance to root-knot nematodes (*Meloiogyne* spp.). Theor Appl Genet 103:592–600

Djian-Caporalino C, Pijarowski L, Januel A, Lefebvre V, Daubèze A, Palloix A, Dalmasso A, Abad P (1999) Spectrum of resistance to root-knot nematodes and inheritance of heat-stable resistance in pepper (*Capsicum annuum* L.). Theor Appl Genet 99:496–502

Dwivedi N, Kumar R, Paliwal R, Kumar U, Kumar S, Singh M, Singh RK (2013) QTL mapping for important horticultural traits in pepper (*Capsicum annuum*). J Plant Biochem Biotechnol. https://doi.org/10.1007/s13562-013-0247-1

Elitzur T, Nahum H, Borovsky Y, Pekker I, Eshed Y, Paran I (2009) Co-ordinated regulation of flowering time, plant architecture and growth by FASCICULATE: the pepper orthologue of SELF PRUNING. J Exp Bot 60:869–880

Englewood CNJ (1985) Pungency of capsicums and their oleoresins (HPLC method). American Spice Trade Association, Official analytical methods of the American Spice Trade Association, pp 111–114

Ercan N, Sensoy FA (2011) Androgenic responses of different (*Capsicum annuum* L.) cultivars. Biyoloji Bilimleri Arastirma Dergisi 4:59–61

Ercan N, Sensoy FA, Sensoy AS (2006) Influence of growing season and donor plant age on anther culture response of some pepper cultivars (*Capsicum annuum* L.). Sci Hort 110:16–20

Fari MG, Csillery and Zatyko L (1983) Embryo culture: an efficient technique in interspecific hybridization and in breeding of pepper (*Capsicum*). In: EUCARPIA meeting on genetics and breeding of capsicum and eggplant, Plovdiv, Bulgaria 4–7 July, pp 31-37

Fernandes de R, Ribeiro LD (1998) Mode of resistance in *Capsicum annuum* to *Colletotrichum Gloeosporioides*. In: Proceedings of 10th Eucarpa meeting on genetics and breeding of capsicum and eggplant, 7–11 Sept 1998, Avignon, France, p 170

Ferrara A, Lovelli S, Di Tommaso T, Perniola M (2011) Flowering, growth and fruit setting in greenhouse bell pepper under water stress. J Agron 10:12–19

Fery RL, Harrison HF Jr (1990) Inheritance and assessment of Betazon herbicide tolerance in Santaka pepper. J Am Soc Hort Sci 123:1008–1011

George L, Narayanaswamy S (1973) Haploid Capsicum through experimental androgenesis. Protoplasma 78:467–480

Gil-Ortega R, Palazon Espanol C, Cuartero Zueco J (1991) Genetics of resistance to *Phytophthora capsici* in the Mexican Pepper CM-334. Plant Breed 107:50–55

Gil-Ortega R, Arteaga ML (1988) Response of pepper to two Spanish isolates of CMV. Capsicum Newsl 7:65–66

Gopalakrishnan TR (2007) Chilli. In: Peter KV (ed) Vegetable crops. Horticulture science series-4. New India Publishing, pp 77–86

Gopalkrishnan TR, Gopalkrishnan PK, Peter KV (1989) Inheritance of clusterness and fruit orientation in Chilli (*Capsicum annuum* L.). Indian J Genet 49:219–222

Greenleaf WH (1952) Inheritance of Pungency and of deciduous fruit character in peppers (*Capsicum annuum*). Proc Assoc South Agric Workers 49:110–111

Greenleaf WH (1986) Pepper breeding. In: Basset MJ (ed) Breeding vegetable crops. The AVI Publishing Company, Westport, Connecticut, pp 67–134. ISBN-13: 9780870554995

Grozeva S, Rodeva V, Todorova V, Pundeva R (2009) Obtaining of pepper plants via anther culture. Genet Breed 38:25–31

Grozeva S, Todorova V, Cholakov T, Rodeva V (2013) Effect of temperature and growth period of donor plants on pepper anther culture. In: 3rd international conference on research people and actual tasks on multidisciplinary sciences, June, Lozenec, Bulgaria, vol 1, pp 60–64

Grube R, Radwanski ER, Jahn M (2000) Comparative genetics of disease resistance within the Solanaceae. Genetics 155:873–887

Grube RC, Zhang Y, Huang B, Kyle MM (1996) Phenotypic and marker assisted breeding of *Capsicum* for *Cucumber Mosaic Virus* resistance. Hort Sci 31 (4):595

Gulyas GK, PakozdiJS Lee Y (2006) Analysis of fertility restoration by using cytoplasmic male-sterile red pepper (*Capsicum annuum*L.) lines. Breed Sci 56:331–334

Guzman I, Hamby S, Romero J, Bosland PW, Connell MAO (2011) Variability of carotenoid biosynthesis in orange colored *Capsicum* spp. Plant Sci 179(1–2):49–59

Hachiya S, Kawabata F, Ohnuki K, Inoue N, Yoneda H, Yazawa S, Fushiki T (2007) Effects of CH-19 Sweet, a non-pungent cultivar of red pepper, on sympathetic nervous activity, body temperature, heart rate, and blood pressure in humans. Biosci Biotechnol Biochem 71:671–676

HanK, LeeHY, Ro NY, Hur OS, Lee JH, Kwon JK, KangBC (2018) QTL mapping and GWAS reveal candidate genes controlling capsaicinoid content in *Capsicum*. Plant Biotechnol J. https://doi.org/10.1111/pbi.12894

Harpster MH, Brummell DA, Dunsmuir P (2002) Suppression of a ripening-related endo-1,4-β-glucanase in transgenic pepper fruit does not prevent depolymerization of cell wall polysaccharides during ripening. Plant Mol Biol 50:345–355

Heidmann I, de Lange B, Lambalk J, Angenent JC, Boutilier K (2011) Efficient sweet pepper transformation mediated by the BABY BOOM transcription factor. Plant Cell Rep 30(6):1107–1115

Hein T (2017) https://european-seed.com/2017/11/closer-look-sweet-pepper-breeding-challenges

Hendy H, Pochard E, Dalmasso A (1985) Transmission héréditaire de la résistance aux *Meloidogyne* portée par deux lignées de *Capsicum annuum*: études de descendances d'homozygotes issues d'androgénèse. Agronomie 593–100. https://doi.org/10.1051/agro:19850201

Hibberd AM, Bassett MJ, Stall RE (1987) Allelism tests of 3 dominant genes for hypersensitive resistance to bacterial spot of pepper. Phytopathology 77:1304–1307

Hoang NH, Yang HB, Kang BC (2013) Identification and inheritance of a new source of resistance against *Tomato spotted wilt virus* (TSWV) in *Capsicum*. Sci Hort 161:8–14

Homes FO (1937) Inheritance of ability to mobilize tobacco mosaic disease in pepper. Phytopathology 24:984–1002

Honda I, Kikuchi K, Matsuno S, Fukuda M, Santo H, Ryuto N, Fukumshi N, Tomoko A (2006) Effect of heavy ion bombardment on mutagenesis in sweet pepper isolated by M1 plant selection. Euphytica 15 (1):61–66

Hossain MA, Minami M, Nemoto K (2003) Immature embryo culture and interspecific hybridization between *Capsicum annuum* L. and *C. frutescens* L. via embryo rescue. Jpn J Trop Agric 47:9–16

Hossain SM, Habiba U, Bhuyan SI, Haque MS, Begum SN, Hossain DM (2014) DNA fingerprinting and genetic diversity analysis of chilli germplasm using microsatellite markers. Biotechnology 13:174–180

Huang S, Baoxi Z, Milbourne D, Cardle L (2001) Development pepper SSR markers from sequence databases. Euphytica 117(2):163–167

Huh JH, Kang BC, Nahm SH, Kim S, Ha KS, Lee MH, Kim BD (2001) A candidate gene approach identified phytoene synthase as the locus for mature fruit color in red pepper (*Capsicum* spp.). Theor Appl Genet 102:524–530

Hurtado-Hernandez H, Smith PG (1985) Inheritance of mature fruit colour in *Capsicum annuum* L. J Hered 76:211–213

Hwang D, Jeong HJ, Kwon JK, Kim H, Kang SY, Kang BC (2014) Phenotypic variants among ethyl methanesulfonate M_2 mutant lines in *Capsicum annuum*. Plant Genet Resour Charact Util 12:141–145

Hwang J, Li J, Liu WY, An SJ, Cho H, Her NH, Yeam I, Kim D, Kang BC (2009) Double mutations in *eIF4E* and *eIFiso4E* confer recessive resistance to *Chilli veinal mottle virus* in pepper. Mol Cells 27:329–336

Hwang JK, Paek KY, Cho DH (1998) Breeding of resistant pepper lines (*Capsicum annuum* L.) to bacterial spot (*Xanthomonas campestris* pv. vesicatoria) through anther culture. Acta Hort 461:301–310

Ibiza VP, Blanca J, Canizares J, Nuez F (2012) Taxonomy and genetic diversity of domesticated *Capsicum* species in the Andean region. Genet Resour Crop Evol 59(6):1077–1088

Indira C, Susan A (1977) Morphological and cytological studies on a radiation induced polyploid in *Capsicum annum* Linn. Cytologia 42:371–375

Irikova T, Grozeva S, Popov P, Rodeva V, Todorovska E (2011) *In vitro* response of pepper anther culture (*Capsicum annuum* L.) depending on genotype, culture medium and duration of cultivation. Biotechnol Biotech Eq 25:2604–2609

Ishikawa K (2001) Tetraploid bell pepper shows high in vitro pollen germination at 15°C. HortScience 36 (7):1336

Ishikawa K, Mishiba K, Yoshida H, Nunomura O (1997) Establishment of tetraploid plants of *Capsicum annuum* L. by colchicine treatment with the analysis of flow cytometry. Capsicum Eggplant Nwsl 16:44–47

Jahn M, Paran I, Hoffmann K, Radwanski ER, Livingstone KD, Grube RC, Aftergoot E, Lapidot M, Moyer J (2000) Genetic mapping of the Tsw locus for resistance to the Tospovirus Tomato spotted wilt virus in Capsicum spp. and its relationship to the Sw-5 gene for resistance to the same pathogen in tomato. Mol Plant-Microbe Interact 13:673–682

Jeifetz D, David-Schwartz R, Borovsky Y, Paran I (2011) CaBLIND regulates axillary meristem initiation and transition to flowering in pepper. Planta 234:1227–1236

Jeong HJ, Kwon JK, Pandeya D, Hwang J, Hoang NH, Bae JH, Kang BC (2012) A survey of natural and ethyl methane sulfonate-induced variations of *eIF4E* using high-resolution melting analysis in *Capsicum*. Mol breed 29:349–360

Jeong H-S, Jang S, Han K, Kwon J-K, Kang B-C (2015) Marker-assisted backcross breeding for development of pepper varieties (*Capsicum annuum*) containing capsinoids. Mol Breed, 226–235

Jo YD, Kim SH, Hwang JE, Kim YS, Kang HS, Kim SW, Kwon SJ, Ryu JH, Kim JB et al (2016) Construction of mutation populations by gamma-ray and carbon beam irradiation in chili pepper (*Capsicum annuum* L.). Hort Environ Biotechnol 57:606–614

Jo YD, Kim YM, Park MN, Yoo JH, Park M, Kim BD, Kang BC (2010) Development and evaluation of broadly applicable markers for Restorer-of-fertility in pepper. Mol Breed 25:187–201

Jones JB, Minsavage GV, Roberts PD, Johnson RR, Kousik CS, Subramanian S, Stall RE (2002) A non-hypersensitive resistance in pepper to the bacterial spot pathogen is associated with two recessive genes. Phytopathology 92(3):273–277

Jordan T, Romer P, Meyer A, Szczesny R, Pierre M, Piffanelli P, Bendahmane A, Bonas U, Lahaye T (2006) Physical delimitation of the pepper *Bs3* resistance gene specifying recognition of the *AvrBs3* protein from *Xanthomonas campestris* pv. *vesicatoria*. Theor Appl Genet 113:895–905

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

Kang WH, Huy NH, Yang HB, Kwon JK, Jo SH, Seo JK, Kim KH, Choi D, Kang BC (2010) Molecular mapping and characterization of a single dominant gene controlling CMV resistance in peppers (*Capsicum annuum* L.). Theor Appl Genet 120:1587–1596

Kawabata F, Inoue N, Yazawa S, Kawada T, Inoue K, Fushiki T (2006) Effects of CH-19 sweet, a non-pungent cultivar of red pepper, in decreasing the body weight and suppressing body fat accumulation by sympathetic nerve activation in humans. Biosci Biotechnol Biochem 70:2824–2835

Kim BS, Hartmann RW (1985) Inheritance of a gene (Bs3) conferring hypersensitive resistance to *Xanthomonas campestris* pv. *vesicatoria* in pepper (*Capsicum annuum*). Plant Dis 69:233–235

Kim DH, Kang JG, Kim S, Kim BD (2001a) Identification of *coxII* and *atp6* region as associated to CMS in *Capsicum annuum* by using RFLP and long and accurate PCR. J Kor Soc Hort Sci 42:121–127

Kim DH, Kim BD (2005) The organization of mitochondrial *atp6* gene region in male fertile and CMS lines of pepper (*Capsicum annuum* L.). Curr Genet 49(1):59–67

Kim HJ, Nahm SH, Lee HR, Yoon GB, Kim KT, Kang BC, Choi D, Kweon OY, Cho MC, Kwon JK, Han JH, Kim JH, Park M, Ahn JH, Choi SH, Her NH, Sung JH, Kim BD (2008) BAC-derived markers converted from RFLP linked to *Phytophthora capsici* resistance in pepper (*Capsicum annuum* L.). Theor Appl Genet 118:15–27

Kim S, Kim SR, Chung Sun AN, Hong YN, Lee KW (2001) Constitutive expression of rice MADS box gene using seed explants in hot pepper (*Capsicum annuum* L.). Mol Cells 12(2):221–226

Kisiała A, Olszewska D, Niklas-Nowak A, Nowaczyk P (2011) Biometrical characteristics of R2 generation of anther-derived pepper (*Capsicum* spp.) plants. Acta Agrobot 64:53–58

Koeda S, Sato K, Saito H, Nagano AJ, Yasugi M, Kudoh H, Tanaka Y (2019) Mutation in the putative ketoacyl-ACP reductase CaKR1 induces loss of pungency in Capsicum. Theor Appl Genet 132(1):65–80

Kole C, Gupta PK (2004) Genome mapping and map based cloning. In: Jain HK, Kharkwal MC (eds) Plant breeding: mendelian to molecular approaches, Narosa Publishing House, New Delhi, India, p 811

Ko MK, Soh H, Kim KM, Kim YS, Im K (2007) Stable production of transgenic pepper plants mediated by *Agrobacterium tumefaciens*. HortScience 42(6):1425–1430

Koleva-Gudeva L, Gulaboski R, Janevik-Ivanovska E, Trajkova F, Maksimova V (2013) Capsaicin—inhibitory factor for somatic embryogenesis in pepper anther culture. Elect J Biol 9:29–36

Koleva-Gudeva L, Trajkova F (2012) Anther culture of pepper: morphological characteristics of fruits of androgenetic pepper lines (*Capsicum annuum* L.). J Res Agric 1:136–145

Koleva-Gudeva L (2007) Somatic embryogenesis in pepper anther culture: the effect of incubation treatments and different media. Sci Hort 111:114–119

Kumar RV, Sharma VK, Chattopadhyay B, Chakraborty S (2012) An improved plant regeneration and *Agrobacterium*—mediated transformation of red pepper (*Capsicum annuum* L.). Physiol Mol Biol Plants 18(4):357–364

Kumar S, Kumar R, Singh J (2006) Cayenne/American pepper. In: Peter KV (ed) Handbook of herbs and spices. Woodhead Publishing, Cambridge, UK, pp 299–312

Kuo JS, Wang YY, Chien NF, Ku SJ, Kung ML, Hsu HC (1973) Investigations on the anther culture in vitro of *Nicotiana tabacum* and *Capsicum annuum*. Acta Bot Sin 15:36–52

Kwon YI, Apostolidis E, Shetty K (2007) Evaluation of pepper (*Capsicum annuum*) for management of diabetes and hypertension. J Food Biochem 31:370–385

Kyle MM, Palloix A (1997) Proposed revision of nomenclature for potyvirus resistance genes in *Capsicum*. Euphytica 97:183–188

Lafortune D, Béramis M, Daubèze AM, Boissot N, Palloix A (2005) Partial 4 resistance of pepper to bacterial wilt is oligogenic and stable under tropical conditions. Plant Dis 89:501–506

Lee CJ, Yoo EU, Shin JH, Lee J, Hwang H, Kim B (2005) Non-pungent *Capsicum* contains a deletion in the capsaicinoid synthetase gene which allows early detection of pungency with SCAR markers. Mol Cells 19(2):262–267

Lee HR, Kim KT, Kim HJ, Han JH, Yeoum SI et al (2011) QTL analysis of fruit length using rRAMP, WRKY, and AFLP markers in chili pepper. Hort Environ Biotechnol 52:602–613

Lee J, Park SJ, DoJ Wohng, Choi D, Han JH, Yoon J (2013) Development of a genetic map of chilli pepper using single nucleotide polymorphism markers generated from next generation resquencing of parents. Kor J Hort Sci Technol 31(4):473–482

Lee SJ, Kim BD, Paek KH (1993) *In vitro* plant regeneration and *Agrobacterium* mediated transformation from cotyledon explants of hot pepper (*Capsicum annuum* cv. Golden Tower). Korean J Plant Tissue Culture 20:289–294

Lee YH, Kim HS, Kim JY, Jung M, Park YS, Lee JS, Choi SH, Her NH, Lee JH, Hyung NI, Lee CH, Yang SG, Harn CH (2004) A new selection method for pepper transformation: callus-mediated shoot formation. Plant Cell Rep 23:50–58

Lefebvre V, Daubeze AM, Voort RVJ, Peleman J, Bardin M, Palloix A (2003) QTLs for resistance to powdery mildew in pepper under natural and artificial infections. Theor Appl Genet 107:661–666

Lefebvre V, Kuntz M, Camara B, Palloix A (1998) The capsanthin-capsorubin synthase gene: a candidate gene for the y locus controlling the red fruit colour in pepper. Plant Mol Biol 36:785–789

Lefebvre V, Goffinet B, Chauvet JC, Caromel B, Signoret P, Brand R, Palloix A (2001) Evaluation of genetic distances between pepper inbred lines for cultivar protection purposes: comparison of AFLP, RAPD and phenotypic data. Theor Appl Genet 102 (5):741–750

Li D, Zhao K, Xie B, Zhang B, Luo K (2003) Establishment of a highly efficient transformation system for pepper (*Capsicum annuum* L.). Plant Cell Rep 21(8):785–788

Lin S, Chou Y, Shieh H, Ebert A, Kumar S et al (2013) Germplasm dissemination by AVRDC—The World Vegetable Centre: an overview and introspection. Chron Hort 53(3):21–27

Lippert LF, Bergh BO, Smith PG (1965) Gene list for the pepper. J Hered 56(1):30–34

Litoriya N, Kaur D, Patel NJ, Talati JG (2010) Varietal identification of chilli (*Capsicum annuum* L.) by electrophoretic technique. Indian J Agric Biochem 23 (1):36–40

Liu W, Parrott WA, Hildebrand DF, Collins GB, Williams EG (1990) *Agrobacterium* induced gall formation in bell pepper (*Capsicum annuum* L.) and formation of shoot-like structures expressing introduced genes. Plant Cell Rep 9:360–364

Liu WY, Kang JH, Jeong HS, Choi HJ, Yang HB et al (2014) Combined use of bulked segregant analysis and microarrays reveals SNP markers pinpointing a major QTL for resistance to *Phytophthora capsici* in pepper. Theor Appl Genet 127 (11):2503–2513. https://link.springer.com/journal/122

Lu FH, Kwon S-W, Yoon M-Y et al (2012) SNP marker integration and QTL analysis of 12 agronomic and morphological traits in F_8 RILs of pepper (*Capsicum annuum* L.). Mol Cells 34(1):25–34

Luitel B, Adhikari P, Shrestha S, Kang W (2012) Morphological characterization of anther derived plants in minipaprika (*Capsicum annuum* L.). Kor J Breed Sci 44:450–461

Luitel B, Kang W (2013a) *In vitro* androgenic response of minipaprika *(Capsicum annuum* L.) genotypes in different culture media. Hort Environ Biotechnol 54:162–171

Luitel B, Kang W (2013b) Assessment of fruit quality variation in doubled haploids of minipaprika (*Capsicum annuum* L.). Hort Environ Biotechnol 54:257–265

Mahasuk P, Khumpeng N, Wasee S, Taylor PWJ, Mongkolporn O (2009a) Inheritance of resistance to anthracnose *(Colletotrichum capsici*) at seedling and fruiting stages in chili pepper (*Capsicum* spp.). Plant Breed 128:701–706

Mahasuk P, Taylor PWJ, Mongkolporn O (2009b) Identification of two new genes conferring resistance to *Colletotrichum acutatum* in *Capsicum baccatum*. Phytopathology 99:1100–1104

Mahto BK, Sharma P, Rajam M, Reddy P, Dhar-Ray S (2018) An efficient method for Agrobacterium-mediated genetic transformation of chilli pepper (*Capsicum annuum* L.). Indian J Plant Physiol 23:1–9

Malhova E (1977) Cytoembryology du genre *Capsicum*. Eucarpia Capsicum 77:191–197

Mallard S, Cantet M, Massire A, Bachellez A, Ewert S, Lefebvre V (2013) A key QTL cluster is conserved among accessions and exhibits broad-spectrum resistance to*Phytophthora capsici*: a valuable locus for pepper breeding. Mol Breed 32:349–364

Manoharan M, Vidya CSS, Sita GL (1998) Agrobacterium-mediated genetic transformation in hot chilli (*Capsicum annuum* L. var. *Pusa jwala*). Plant Sci 131:77–83

Manzur JB, Fita A, Prohens J, Rodríguez-Burruezo A (2015) Successful wide hybridization and introgression breeding in a diverse set of common peppers (*Capsicum annuum*) using different cultivated *Ají* (*C. baccatum*) accessions as donor parents. PLoS One 10 (12):1–18

Matsunaga H, Sato T, Monma S (1998) Inheritance of bacterial wilt resistance in the sweet pepper cv. Mie–Midori. In: Proceedings of 10th Eucarpia meeting on genetics and breeding of capsicum and eggplant, 7–11 Sept 1998, Avignon, France, pp 172

Mazourek M, Cirulli ET, Collier SM, Landry LG, Kang BC, Quirin EA, Bradeen JM, Moffett P, Jahn MM (2009) The fractionated orthology of Bs2 and Rx/Gpa2 supports shared synteny of disease resistance in the solanaceae. Genetics 182:1351–1364

Mc Cammon KR, Honma S (1984) Genetics of the umbrella branching habit in *Capsicum* L. Theor Appl Genet 68:541–545

Meghvansi MK, Siddiqui S, Khan H, Gupta VK, Vairale MG, Gogo HK, Singh L (2010) Naga Chilli: a potential source of capsaicinoids with broad-spectrum ethnopharmacological applications. J Ethnopharmacol 132:1–14

Meshram LD, Narkhade MN (1982) Natural male sterile mutant in hot Chilli (*Capsicum annuum* L.). Euphytica 31:1003–1005

Mihálka V, Balázs E, Nagy I (2003) Binary transformation systems based on 'shooter' mutants of *Agrobacterium* tumefaciens: a simple, efficient and universal gene transfer technology that permits marker gene elimination. Plant Cell Rep 21:778–784

Mihalka V, Fari M, Szasz A, Balazs E, Nagy I (2000) Optimized protocols for efficient regeneration and gene transfer in pepper (Capsicum annuum L.). J Plant Biotechnol 2:143–149

Min WK, Lim H, Lee YP, Sung SK, Kim BD, Kim S (2008) Identification of a third haplotype of the sequence linked to the Restorer-of-fertility (*Rf*) gene and its implications for male sterility phenotypes in peppers (*Capsicum annuum* L.). Mol Cells 25:20–29

Mitykó J, Gémes Juhász A (2006) Improvement in the haploid technique routinely used for breeding sweet and spice peppers in Hungary. Acta Agron Hung 54:203–219

Mongkolporn O, Taylor PWJ (2011) *Capsicum*. In: Kole C (ed) Wild crop relatives: genomic and breeding resources. Vol 7: Vegetables. Springer, Berlin, pp 43–57

Montri P, Taylor PWJ, Mongkolporn O (2009) Pathotypes of *Colletotrichum capsici* the causal agent of chilli anthracnose in Thailand. Plant Dis 93:17–20

Moury B, Palloix A, Selassie KG, Marchoux G (1997a) Hypersensitive resistance to tomato spotted wilt virus in three *Capsicum chinenese* accessions is controlled by a single gene and is overcome by virulent strains. Euphytica 94:45–52

Moury B, Selassie-Gebre K, Marchoux G, Daubeze AM, Palloix A (1997b) High temperature effects on hypersensitive resistance to *tomato spotted wilt tospovirus* (TSWV) in pepper (*Capsicum chinense* Jacq.). Eur J Plant Pathol 104:489–498

Murphy JF, Blauth JR, Livingstone KD, Lackney VK, Jahn MM (1998) Genetic mapping of the *pvr1* locus in *Capsicum* spp. and evidence that distinct potyvirus resistance loci control responses that differ at the whole plant and cellular levels. Mol Plant-Microbe Interact 11:943–951

Negi R, Thakur S, Sharma P (2018) Advances in the breeding of bell pepper—a review. Intl J Curr Microbiol App Sci 7(4):2272–2281

Nervo G, Azzimonti MT, Bonelli A, Tamietti G (2007) Application of in vitro anther culture methods to a pepper breeding program for disease resistance. In: Proceedings of 51st Italian society of agricultural genetics annual congress, Riva del Granda, Italy, Poster Abstract S 2.03

Niklas-Nowak A, Olszewska D, Kisiała A, Nowaczyk P (2012) Study of individual plant responsiveness in anther cultures of selected pepper (*Capsicum* spp.) genotypes. Folia Hort 24:141–146

Novac F, Betlach J, Dubovshi J (1971) Cytoplasmic male sterility in sweet pepper (*Capsicum annuum* L.). Phenotype and male sterile character. Z Pflanzenzucht 65:129–140

Novak F (1974) Capsicum haploids. Z Pflanzenzeucht 72:46–54

Nowaczyk L, Banach-Szott M, Olszewska D, Nowaczyk P (2014) Androgenic response of Capsicum interspecific hybrids and capsaicinoid characteristics of DH lines. Herba Polon 60:50–59

Nowaczyk L, Nowaczyk P, Olszewska D, Niklas-Nowak A (2015) Effect of 2,4-dichlorophenoxyacetic acid pretreatment of *Capsicum* spp. donor plants on the anther culture efficiency of lines selected by capsaicinoid content. Biotechnologia 2(2):179–183

Nowaczyk P, Kisiała A (2006) Effect of selected factors on the effectiveness of *Capsicum annuum* L. anther culture. J Appl Genet 47:113–117

Odland ML (1960) Inheritance of flower colour in *Capsicum annuum* L. Proc Am Soc Hort Sci 76:475–481

Odland ML, Porter AM (1938) Inheritance of the immature fruit colour of peppers. Proc Am Soc Hort Sci 36:647–651

Ogundiwin EA, Berke TF, Massoudi M, Black LL, Huestis G, Choi D, Lee S, Prince JP (2005) Construction of 2 intraspecific linkage maps and identification of resistance QTLs for *Phytophthora capsici* root-rot and foliar-blight diseases of pepper (*Capsicum annuum* L.). Genome 48:698–711

Ohnuki K, Moritani T, Ishihara K, Fushiki T (2001) Capsaicin increases modulation of sympathetic nerve activity in rats: measurement using power spectral analysis of heart rate fluctuations. Biosci Biotechnol Biochem 65:638–643

Olszewska D, Kisiala A, Nowaczyk P (2011) The assessment of doubled haploid lines obtained in pepper (*Capsicum annuum* L.) anther culture. Folia Hort 23(2):93–99

Olszewska D, Kisiała A, Niklas-Nowak A, Nowaczyk P (2014) Study of *in vitro* anther culture in selected genotypes of genus *Capsicum*. Turk J Biol 38:118–124

Olszewska D, Niklas-Nowak A, Nowaczyk P (2010) Variation in the quantitative characters of androgenic pepper lines derived from hybrid *Capsicum frutescens* × *Capsicum chinense* jacq. Veget Crops Res Bul 73:5–11

Özkum D, Tıpırdamaz R (2007) Effects of silver nitrate, activated charcoal and cold treatment on the in vitro androgenesis of pepper (*Capsicum annuum* L.). Acta Hort 729:133–136

Pakdeevaraporn P, Wasee S, Taylor PWJ, Mongkolporn O (2005) Inheritance of resistance to anthracnose caused by *Colletotrichum capsici* in *Capsicum*. Plant Breed 124:206–208

Pal BP, Ramanujan S, Joshi AB (1941) Colchicine induced polyploidy in crop-plants. II Chilli (*Capsicum annuum* L.). Indian J Genet 1:28–40

Palloix A (1992) Diseases of pepper and perspectives for genetic control. In: Capsicum newsletter special issue-Proc VIIIth EUCARPIA meeting on genetics and breeding of capsicum and eggplant, 7–10 Sept 1992 Rome, Italy, pp 120–126

Paran I, Borovsky Y, Nahon S, Cohen O (2007) The use of induced mutations to study shoot architecture in*Capsicum*. Isr J Plant Sci 55:125–131

Paran I, VanderVoort JR, Lefebvre V, Jahn M, Landry L, van Schriek M, Tanyolac B, Caranta C, Ben-Chaim A. Livingstone K, et al (2004) An integrated genetic linkage map of pepper (*Capsicum* spp.). Mol Breed 13:251–261

Park HK, Kim BS, Lee WS (1990) Inheritance of resistance to anthracnose (*Colletotrichum*spp.) in pepper (*Capsicum annuum*L.) I. Genetic analysis of anthracnose resistance by diallele crosses. J Korean Soc Hort Sci 31:91–105

Parker BL, Talekar NS, Skinner M (1995) Field guide: insect pests of selected vegetables in tropical and sub-tropical Asia. Asian Vegetable Research and Development Center, Shanhua, Tainan, Taiwan, R.O. C. Publication no. 94–427

Parra-Vega V, González-García B, Seguí-Simarro JM (2013) Morphological markers to correlate bud and

anther development with microsporogenesis and microgametogenesis in pepper (*Capsicum annuum* L.). Acta Physiol Plant 35:627–633

Pathak CS, Deshpande AA, Singh DP (1985) Non flowering mutant in chillies (*Capsicum annuum* L.). Capsicum Newsl 4:41–42

Pathak CS, Singh DP, Deshpande AA (1983) Closed flower mutant in *Capsicum annuum* L. Capsicum Newsl 2:99–100

Pauk J, Lantos C, Somogyi G, Vági P, Ábrahám Táborosi Z, Gémes Juhász A, Mihály R, Kristóf Z, Somogyi N, Tímár Z (2010) Tradition, quality and biotechnology in Hungarian spice pepper (*Capsicum annuum* L.) breeding. Acta Agron Hung 58:259–266

Pegard A, Brizzard G, Fazari A, Soucaze O, Abad P, Djian-Caporalino C (2005) Histological characterization of resistance to different root-knot nematode species related to phenolics accumulation in *Capsicum annuum*. Phytopathology 95(2):158–165

Perucka I, Materska M (2001) Phenylalanine ammonia-lyase and antioxidant activities of lipophilic fraction of fresh pepper fruits *Capsicum annuum* L. Innovat Food Sci Emerg Technol 2:189–192

Peterson PA (1958) Cytoplasmically inherited male sterility in *Capsicum*. Amer Nat 92:111–119

Peterson PA (1959) Linkage of fruit shape and colour genes in *Capsicum*. Genetics 44:407–419

Pickersgill B (1977) Chromosomes and evolution in Capsicum. In: Pochard E (ed) Capsicum 77. Comptes Rendues 3'eme Congre's Eucarpia Piment. INRA, Montfavet, Avignon, France, pp 27–37

Pickersgill B (1988) The genus *Capsicum*: a multidisciplinary approach to the taxonomy of cultivated and wild plants. Biol Zent BI 107:381–389

Pierre M, Noël L, Lahaye T, Ballvora A, Veuskens J, Ganal M, Bonas U (2000) Highresolution genetic mapping of the pepper resistance locus Bs3 governing recognition of the *Xanthomonas campestris* pv vesicatora AvrBs3 protein. Theor Appl Genet 101:255–263

Pochard E (1970) Description of trisomic individuals of *Capsicum annuum* L. obtained in progeny of a haploid plant. Ann Amelior Plantes 20:233–256

Pochard E (1977) Localisation of genes in *Capsicum annuum* L. by trisomic analysis. Ann Amélior Plantes 27:255–266

Pohronezny K (2003) Compendium of pepper diseases. American Phytopathological Society, Minnesota, USA, p 63

Popovsky S, Paran I (2000) Molecular genetics of the y locus in pepper: its relation to capsanthin-capsorubin synthase and to fruit colour. Theor Appl Genet 101:86–89

Poulos JM (1994) Pepper breeding (*Capsicum* spp.): achievements, challenges and possibilities. Plant Breed Abstr 64:143–155

Pozueta-Romero L, Houlne G, Canas L, Schantz R, Chamarro L (2001) Enhanced regeneration of tomato and pepper seedlings explants for *Agrobacterium*-mediated transformation. Plant Cell, Tissue Organ Cult 67:173–180

Pramanick K, Srivastava SK (2013) Role of capsaicin in cancer prevention. In: Sanjay K Srivastava (ed) *Role of capsaicin in oxidative stress and cancer*. Springer, Netherlands, pp 1–18

Raghavan TS, Venkatasubban KR (1940) Studies in the south Indian chillies. Proc Plant Sci 12:29–46

Raghuvanshi SS, Sheila J (1964) Cytomorphological studies on the colchiploids of *Capsicum frutescens* L. Cytologia 29:61–78

Ramachandran RK (2013) Breeding of chilli and Capsicum. In: Ramachandra RK (auth) Breeding of vegetable crops. Narendra Publishing House, pp 205

Rao GU, Ben Chaim A, Borovsky Y, Paran I (2003) Mapping of yield -related QTLs in pepper in an interspecific cross of *Capsicum annuum* and *C. frutescens*. Theor Appl Genet 106:1457–1466

Rao GU, Paran I (2003) Polygalacturonase: a candidate gene for the soft flesh and deciduous fruit mutation in Capsicum. Plant Mol Biol 51:135–141

Rao KG (1987) Colchicine induced chromosome mosaicism in chili pepper (*Capsicum annuum* L.). Proc Indian Acad Sci: Plant-Sci 97(1):55–61

Rego ER, Rego MM, Farias-Filho LP (2011) Genetic diversity in pepper (*Capsicum* spp.) by RAPD marker. Acta Hort 918:341–347

Restaino F (1989) New dwarf pepper (*Capsicum annum* L.) cv. developed through mutation induction. In: Proceedings of 7th EUCARPIA meeting on genetics and breeding of *Capsicum* and eggplant, 27–30 June 1989, Palanka, Yugoslavia, pp 55–59

Rosellol S, Diezel MJ, Jorda C, Nuezl F (1996) Screening of Capsicum chacoense accessions for TSWV resistance by mechanical inoculation. Capsicum Eggplant Newsl 16:68–78

Roshany G, Kalantarai S, Naderi R, Hassani ME (2013) Callus formation *via* anther culture in *Capsicum annuum* L. with differences in genotypes, media and incubation temperature. Tech J Engin App Sci 3:3847–3853

Ruffel S, Gallois JL, Moury B, Robaglia C, Palloix A, Caranta C (2006) Simultaneous mutations in translation initiation factors *eIF4E* and *eIF(iso)4E* are required to prevent pepper veinal mottle virus infection of pepper. J Gen Virol 87:2089–2098

Sahin F, Miller SA (1997) A source of resistance in *Capsicum* spp. Accessions to Capsicum race 6 of *Xanthomonas campestris pv. Vesictoria.* Phytopathology 87:84

Sahin F, Miller SA (1998) Resistance in *Capsicum pubescens* to *Xanthomonas campestris* pv. *Vesicatoria*. pepper Race 6. Plant Dis 82:794–799

Saini SS, Sharma PP (1978) Inheritance of resistance to fruit rot (*Phytophthora capsici* Leon.) and induction of resistance in bell pepper (*Capsicum annuum* L.). Euphytica 27:721–723

Sharma A, Kumar V, Giridhar P, Ravishankar GA (2008) Induction of *in vitro* flowering in *Capsicum frutescens* under the influence of silver nitrate and cobalt chloride and pollen transformation. Electron J Biotechnol 11:1–6

Shen X, Gmitter FG Jr, Grosser JW (2011) Immature embryo rescue and culture. In: Thorpe TA, Young EC (eds) Plant embryo culture. Humana Press, New York, pp 75–92

Shetty AA, Magadum S, Managanvi K (2013) Vegetables as sources of antioxidants. J Food Nutr Disor 2(1):1–5

Shifriss C (1973) Additional spontaneous male sterile mutant in *Capsicum annuum* L. Euphytica 22:527–529

Shifriss C (1997) Male sterility in pepper (*Capsicum annuum* L.). Euphytica 93(1):83–88

Shifriss C, Frankel R (1969) A new male sterility gene in *Capsicum annuum* L. J Amer Soc Hort Sci 94:385–387

Shifriss C, Pilovsky M, Zack JM (1992) Resistance to *Leveillula* mildew (*Oidiopsis taurica*) in *Capsicum annuum* L. In: Proceedings of 8th EUCARPIA meeting on genetics and breeding of capsicum and eggplant, Rome, Italy, 7–10 Sep 1992, pp 172-177

Shifriss C, Rylski I (1972) A male sterile *(ms-2)* gene in 'California wonder' pepper (*C. annuum*). Hort Sci 7:36

Shin R, Park JM, An JM, Park KH (2002a) Ectopic expression of *TsiI* in transgenic hot pepper plants enhances host resistance to viral, bacterial and oomycete pathogens. Mol Plant Microbe Interact 15:983–989

Shin R, Han JH, Lee GJ, Peak KH (2002b) The potential use of a viral coat protein gene as a transgene screening marker and multiple virus resistance of pepper plants coexpressing coat proteins of cucumber mosaic virus and tomato mosaic virus. Transgenic Res 11:215–219

Shivanna KR, Bahadur B (2015) Efficacy of biotechnological approaches to raise wide sexual hybrids. In: Bahadur B, Rajam MV, Sahijram L, Krishnamurthy KV (eds) Plant biology and biotechnology, vol II. Plant genomics and biotechnology. Springer India, New Delhi, pp 347–362

Shivegowda ST, Mythili JB, Lalitha A, Saiprasad GVS, Gowda R, Gowda TKS (2002) *In vitro* regeneration and transformation in chilli pepper (*Capsicum annuum* L.). J Hortic Sci Biotechnol 77:629–634

Shmykova NA, Pyshnaya ON, Shumilina DV, Dzhos EA (2014) Morphological characteristics of doubled haploid plants of pepper produced using microspore/anther in vitro culture of the interspecies hybrids of *Capsicum annum* L. and *C. chinense* Jacq. Russ Agric Sci 40:417–421

Shrestha LS, Luitel BP, Kang WH (2011) Agro-morphological characterization of anther derived plants in sweet pepper (*Capsicum annuum* L. cv. Boogie). Hort Environ Biotechnol 52:196–203

Shrestha LS, Luitel BP, Lee TJ, Kang WH (2010) Cytological and morphological characterization of anther derived plants from sweet pepper (*Capsicum annuum* L.) cv. 'Special'. Korean J Breed Sci 42:431–438

Shuh DM, Fontenot JF (1990) Gene transfer of multiple flowers and pubescent leaf from *Capsicum chinense* into *Capsicum annuum* backgrounds. J Am Soc Hortic Sci 115:499–502

Singh J, Thakur MR (1977) Genetics of resistance to *Tobacco Mosaic Virus*, *Cucumber Mosaic Virus* and leaf curl virus in hot pepper. In: Proceedings of 3rd EUCARPIA meeting on Capsicum and working group, Montfavet, Avignon, France, 5–8 July 1977, pp 119–126

Singh P, Cheema DS, Dhaliwal MS, Garg N (2014a) Heterosis and combining ability for earliness, plant growth, yield and fruit attributes in hot pepper (*Capsicum annuum* L.) involving genetic and cytoplasmic-genetic male sterile lines. Sci Hort 168:175–188

Singh R, Giri SK, Kotwaliwale N (2014) Shelf-life enhancement of green bell pepper (*Capsicum annuum* L.) under active modified atmosphere storage. Food Pack Shelf Life 1:101:112

Smith PG (1950) Inheritance of brown and green mature colour in peppers. J Hered 41:138–140

Sobhakumari VP, Lalithakumari D (2005) High frequency shoot regeneration and *Agrobacterium* mediated DNA transfer in red chilli (*Capsicum annuum* L.). PCBMB 6:9–16

Soler S, Debreczeni DE, Vidal E, Aramburu J, López C, Galipienso L, Rubio L (2017) New Capsicum baccatum accession shows tolerance to wild-type and resistance-breaking isolates of Tomato spotted wilt virus. Ann Appl Biol 170(2):286

Stewart C Jr, Kang BC, Liu K, Mazourek M, Moore SL, Eun YY, Kim BD, Paran I, Jahn MM (2005) The *Pun1* gene for pungency in pepper encodes a putative acyltransferase. Plant J 42:675–688

Stewart C Jr, Mazourek M, Stellari GM, O'Connell M, Jahn M (2007) Genetic control of pungency in *C. chinensevia* the *Pun1* locus. J Exp Bot 58:979–991

Suwor P, Sanitchon J, Thummabenjapone P, Kumar S, Techawongstien S (2017) Inheritance analysis of anthracnose resistance and marker-assisted selection in introgression populations of chilli (*Capsicum annuum* L.). Sci Hort 220:20–26

Sy O, Bosland PW, Steiner R (2005) Inheritance of *Phytophthora* stem blight resistance as compared to *Phytophthora* root rot and *Phytophthora* foliar blight resistance in *Capsicum annuum* L. J Amer Soc Hort Sci 130:75–78

Sy O, Steiner R, Bosland PW (2008) Recombinant inbred line differential identifies race-specific resistance to *Phytophthora* root rot in *Capsicum annuum*. Phytopathology 98:867–870

Szarka J, Csillery G (1995) Defence systems against *Xanthomonas campestris* pv. *vesicatoria* in pepper. In:

Proceedings of 9th Eucarpia meeting on genetics and breeding of capsicum and eggplant, Budapest, Hungary, 21–25 Aug 1995, pp 184–197

Tai TH, Dahlbeck D, Stall RE, Peleman J, Staskawicz BJ (1999) High-resolution genetic physical mapping of the region containing the *Bs2* resistance gene of pepper. Theor Appl Genet 99:1201–1206

Takizawa K, Ishikawa K, Nunomura O, Ito T (2008) Ploidy level effect on physiology of pepper plant as affected by fruit loading. Acta Hort 779:689–697

Tanaka Y, Yoneda H, Hosokawa M, Miwa T, Yazawa S (2014) Application of marker-assisted selection in breeding of a new fresh pepper cultivar (*Capsicum annuum*) containing capsinoids, low-pungent capsaicinoid analogs. Sci Hort 165:242–245

Tanksley SD (1984) Linkage relationships and chromosomal locations of enzyme-coding genes in pepper (*Capsicum annuum* L.). Chromosoma 89:353–360

Tapadar NN (1963) Studies in induced tetraploids of the family Apocynaceae I. *Rauvolfia serpentina* Benth. Cytologia 28:229–234

Taranto F, D'Agostino N, Greco B, Cardi T, Tripodi P (2016) Genome-wide SNP discovery and population structure analysis in pepper (*C. annuum*) using genotyping by sequencing. BMC Genomics 17:943

Taskin H, Buyukalaca S, Keles D, Ekbic E (2011) Induction of microspore-derived embryos by anther culture in selected pepper genotypes. Afr J Biotechnol 10:17116–17121

Thabuis A, Palloix A, Pfl ieger S, Daubeze AM, Caranta C, Lefebvre V (2003) Comparative mapping of Phytophthora resistance loci in pepper germplasm: evidence for conserved resistance loci across Solanaceae and for a large genetic diversity. Theor Appl Genet 106:1473–1485

Thorup T, Tanyolac B, Livingstone K, Popovsky S, Paran I, Jahn M (2000) Candidate gene analysis of organ pigmentation loci in the Solanaceae. Proc Natl Acad Sci USA 97:11192–11197

Thul ST, Darokar MP, Shasany AK, Khanuja SP (2012) Molecular profiling for genetic variability in *Capsicum* species based on ISSR and RAPD markers. Mol Biotechnol 51(2):137–147

Tiwari A, Vivian-Smith A, Voorrips RE, Habets MEJ, Xue LB, Offringa R, Heuvelink E (2011) Parthenocarpic potential in *Capsicum annu*um L. is enhanced by carpelloid structures and controlled by a single recessive gene. BMC Plant Biol 11:143

Todorova V, Grozeva S, Rodeva V, Masheva S (2013) Breeding evaluation of pepper lines obtained by in vitro anther culture. Genetika 45:601–610

Tomita R, Sekine KT, Mizumoto H, Sakamoto M, Murai J, Kiba A, Hikichi Y, Suzuki K, Kobayashi K (2011) Genetic basis for the hierarchical interaction between tobamovirus spp. and L resistance gene alleles from different Capsicum species. Mol Plant-Microbe Interact 24:108–177

Tomlekova N, Timina OO, Timin OY (2009a) Achievement and perspectives of sweet pepper breeding towards high Beta Carotene. Acta Hort 1:205–209

Tomlekova NB, Timina OO, Timin OY (2009b) Achievements and perspectives of sweet pepper breeding towards high beta-carotene. Acta Hort 830:205–212

Tong N (1998) Chile peppers in China. Chile Pepper Inst Newsl 7(3):1–3

Trajkova F, Koleva-Gudeva L (2014). Fruit analysis of pepper androgenic lines P3 and P4 (*Capsicum annuum* L. cv. Piran) in different maturation stages. Yearbook 2014, Goce Delcev University—Stip, Faculty of Agriculture, pp 51–66

Usman MG, Rafii MY, Martini MY, Yusuff OA, Ismail MR, Miah G (2008) Introgression of heat shock protein (Hsp70 and sHsp) genes into the Malaysian elite chilli variety Kulai (*Capsicum annuum* L.) through the application of marker-assisted backcrossing (MAB). Cell Stress Chaperones 23(2):223–234

Vallejos CE, Jones V, Stall RE, Jones JB, Minsavage GV, Shultz DC, Rodrigues R, Olsen LE, Mazourek M (2010) Characterization of two recessive genes controlling resistance to all races of bacterial spot in peppers. Theor Appl Genet 121:37–46

Van de Peer Y, Fawcett JA, Proost S, Sterck L, Vandepoele K (2009) The flowering world: a tale of duplications. Trends Plant Sci 14:680–688

Verma S, Dhiman K, Srivastava DK (2013) *Agrobacterium*-mediated genetic transformation of bell pepper (*Capsicum annuum* L. cv. California wonderWonder) with *gus* and *npt*-ii genes. Int J Adv Biotechnol Res 4:397–403

Voinnet O (2001) RNA silencing as a plant immunity system against viruses. Trends Genet 17:449–459

Voorrips RE, Finkers R, Sanjaya L, Groenwold R (2004) QTL mapping of anthracnose (*Colletotrichum* spp.) resistance in a cross between *Capsicum annuum* and *C. chinense*. Theor Appl Genet 109:1275–1282

Votava EJ, Bosland PW (2002) Novel sources of non-pungency in *Capsicum* species. Capsicum Eggplant Newsl 21:66–68

Walker SJ, Bosland PW (1999) Inheritance of *Phytophthora* root rot and foliar blight resistance in pepper. J Am Soc Hortic Sci 124:14–18

Wang YY, Sun CS, Wang CC, Chien NJ (1973) The induction of pollen plantlets of Triticale and *Capsicum annuum* anther culture. Sci Sin 16:147–151

Watts L (1980) Polyploidy. In: Watts L (ed) Flower and vegetable plant breeding. Grower Books, London, pp 22

Yazawa S, Sao T, Namiki T (1991) Interspecific hybrid dwarfism and geographical distribution of the dwarfness gene in *Capsicum*. J Jpn Soc Hort Sci 58:609–618

Yoon JB, Yang DC, Do JW, Park HG (2006) Overcoming two post-fertilization genetic barriers in interspecific hybridization between *Capsicum annuum* and *C. Baccatum* for introgression of anthracnose resistance. Breed Sci 56:31–38

Zhang B, Huang S, Yang G, Guo J (2000) Two RAPD markers linked to a major fertility restorer gene in pepper. Euphytica 113:155–161

Zhao J, Zhou X, Zhang Z, Yang B, Zhou S (2010) Effects of culture media on anther culture of chili pepper (*Capsicum annuum* L.). J Hunan Agric Univ (Nat Sci) 36:181–184

Zhu YX, Wen-Jun OY, Yi-Feng Z, Zhang-Liang C (1996) Transgenic sweet pepper plants from *Agrobacterium* mediated transformation. Plant Cell Rep 16:71–75

Cytology and DNA Content Variation of *Capsicum* Genomes

4

M. A. Scaldaferro and E. A. Moscone

Abstract

Chromosome data and characterization by fluorescent banding, silver nucleolar organizer region staining (AgNOR), and fluorescence in situ hybridization (FISH) are compiled in this chapter, together with estimations of nuclear DNA content of *Capsicum* species. To date, the diploid chromosome number of 77.8% of the species in the genus has been recorded. The chromosome number distinguishes two groups of species, one with $2n = 2x = 24$ and the other with $2n = 2x = 26$. Only two clades, Andean and Atlantic Forest, possess the chromosome number of $2n = 26$. A physical chromosome map with heterochromatin distribution besides *5S* and active and inactive *45S* ribosomal genes (rDNA) of 12 *Capsicum* taxa was constructed using fluorescent banding, AgNOR and FISH. The chromosome banding pattern with fluorochromes chromomycin A3 and 4′-6-diamidino-2-phenylindole (CMA/DAPI) reveals number of bands, distribution and content of heterochromatin, and FISH reports the localization of *5S* and active and inactive *45S* rDNA. Both methods are specific and, together with morphological characters, are instrumental for identifying taxa in *Capsicum*. AgNOR method informs the number, size, and position of just active NORs. Additionally, nuclear DNA content was estimated for nine diploid species of *Capsicum* by flow cytometry. Genome size displays significant variation between but not within species and contributes to their taxonomic grouping.

4.1 Introduction

The genus *Capsicum* (Solanaceae, subtribe Capsiceae; Olmstead et al. 2008; Särkinen et al. 2013), with 36 variable species (Carrizo García et al. 2016), is a small increasing genus from tropical and temperate areas in America, distributed from southern Mexico to central Argentina. Cultivation of sweet and hot chili peppers has great economic implication, since these vegetables and spices are highly consumed worldwide. The most important cultivated species grow around the world and belong to the *Capsicum annuum* complex (*C. annuum*, *C. chinense*, and *C. frutescens*), and two other cultivated species are predominantly regionally consumed in Latin America, *C. pubescens* and

The Author E. A. Moscone was deceased.

M. A. Scaldaferro (✉) · E. A. Moscone
Instituto Multidisciplinario de Biología Vegetal (IMBIV), CONICET-Universidad Nacional de Córdoba, Córdoba, Argentina
e-mail: mariselscaldaferro@gmail.com

M. A. Scaldaferro
Facultad de Ciencias Exactas, Físicas y Naturales, Universidad Nacional de Córdoba, Córdoba, Argentina

N. Ramchiary and C. Kole (eds.), *The Capsicum Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-319-97217-6_4

C. baccatum (Pickersgill 1997; Scaldaferro et al. 2018). Their wild relatives originated from Central and South America and were domesticated by American natives at least 6000 years before present (Pickersgill 1984; Eshbaugh 1993; Perry et al. 2007; Piperno 2011; Scaldaferro et al. 2018).

Cultivated and wild *Capsicum* species have been usually characterized by corolla color, conforming 'white- and purple-flowered groups.' However, this could not be considered to describe the flower color of more distantly wild species and the genus as a whole, since some species exhibit single-colored flowers, i.e., white, cream, yellow, ocher, pink, lilac, or purple–violet (e.g., *Capsicum chacoense*, *C. friburgense*, *C. rhomboideum*), whereas others have different color combinations in lobules, throat, and tube, often including spots of various colors which makes species delimitation complex (e.g., *Capsicum coccineum*, *C. hunzikerianum*, *C. parvifolium*; Hunziker 2001; Barboza and Bianchetti 2005) (see Table 4.1).

Fruit pungency is a characteristic of the genus due to a group of compounds called 'capsaicinoids' which produce the organoleptic heat sensation, and are exclusive for *Capsicum* (Bosland 1996). More than 20 capsaicinoids found in chili peppers (Bosland and Votava 2000) are synthesized in the epidermis of the placenta substances (Stewart et al. 2007). Although capsaicinoid contents and the intensity of pungency are quantitative traits, in the *C. annuum* complex the presence of capsaicinoids was found to be regulated by the *Pun1* gene; however, its expression is distinct among species and cultivars, since it is based on other modifier genes epistatically affected by *Pun1* and environmental conditions (Lippert et al. 1966; IBPGR 1983; Bosland and Votava 2000; Hunziker 2001; Lefebvre et al. 2002; Stewart et al. 2005). Sometimes, pungency is missing, as in all species from the $x = 13$ 'yellow-flowered group' or Andean clade (Carrizo García et al. 2016; Scaldaferro et al. 2016); e.g., *Capsicum dimorphum*, *C. geminifolium*, *C. hookerianum*, *C. lanceolatum*, *C. lycianthoides*, and *C. rhomboideum* are reported to be completely free of pungency (Barboza pers. com.). This peculiarity of the genus is also absent in *Capsicum longidentatum* Agra and Barboza (Barboza pers.com.), some accessions of *C. chacoense* (Eshbaugh 1980; Tewksbury et al. 2006), in one of *Capsicum cornutum* (as *Capsicum dusenii* Bitter; Hunziker 1971), and in cultivars of *C. annuum* var. *annuum* after human selection.

Meanwhile, chromosome markers have been very important tools to elucidate the evolution and diversification of the genus (Moscone et al. 1993, 1995, 1996a, b, 1999, 2003, 2007; Park et al. 2000; Scaldaferro et al. 2006, 2013, 2016; Romero-da Cruz et al. 2016). The most used markers for chromosome identification in the *Capsicum* species studied up to now are chromosome banding methods with fluorochromes to reveal heterochromatic regions, the use of silver impregnation to show the exact position of active NORs, and the application of FISH with rDNA probes. These techniques have provided a more defined karyo-systematic analysis, contributing to the comprehension of the diversification and evolution of the genus (Moscone et al. 1993, 1995, 1996a, b, 1999, 2003, 2007; Park et al. 2000; Scaldaferro et al. 2006, 2011, 2013, 2016; Barboza et al. 2011; Romero-da Cruz et al. 2016).

This chapter describes all *Capsicum* chromosome features studied to date. A list of the recognized *Capsicum* species and some taxonomically relevant cytogenetic traits are presented in Table 4.1.

4.2 Phylogeny of Chili Peppers

Carrizo García et al. (2016) proposed an informal classification of *Capsicum* into 11 clades according to three molecular markers, two plastid DNA markers, the *maturase K* gene (*matK*) and the *psbA-trnH* intergenic spacer, and one nuclear gene *waxy* (GBSSI, *granule-bound starch synthase*). Based on this grouping, there are only two clades where species possess the chromosome number of $2n = 26$, i.e., **Andean** and **Atlantic Forest** clades.

C. annuum var. *annuum* and var. *glabriusculum*, *C. chinense*, *C. frutescens*, and *Capsicum*

Table 4.1 Recognized *Capsicum* species and some relevant taxonomic traits

Taxon and voucher number	Corolla shape and color	Pungency	2*n*	Karyotype formula	Chromosomes with active NOR	No. and position of rDNA sites		Hc amount (HKL in μm)	1C DNA content in pg
						45S	*5S*		
Annuum clade									
1. *C. annuum* L. var. *annuum*									
Cytotype 1 (EAM 193, 251, 203)[a]	Stellate; white or cream (exceptionally violet)	No	24	10 m + 1 sm + 1 st	11 sm	3 [1 major, 2 small]	1 (6p)	1.80 (68.51)	3.41
Cytotype 2 (EAM 204, 252; NMCA 10544, 10272)[a]		Yes	24	10 m + 1 sm + 1 st	11 sm, 12 st	4 [2 major, 2 small]	1 (6p)	2.88 (70.40)	3.32
var. *glabriusculum* (Dunal) Heiser and Pickersgill									
Cytotype 1 (NMCA 10955)[a]	Stellate; white or cream	Yes	24	10 m + 1 sm + 1 st	11 sm	1 [major]	1 (6p)	2.26 (59.53)	?
Cytotype 2 (NMCA 10983)[a]		Yes	24	11 m + 1 st	1 m, 5 m	2 [major]	1 (6p)	3.54 (51.95)	?
Cytotype 3 (LQ w. no.)[e]		Yes	24	11 m + 1 st	11 m	5 [1 major, 4 small]	1 (5p)	2.33 (55.13)	?
Cytotype 4 (YSG w. no.)[e]		Yes	24	11 m + 1 st	5 m, 12 st	4 [2 major, 2 small]	1 (4p)	6.33 (53.56)	?
Cytotype 5 (Neth 804750009)[e]		Yes	24	11 m + 1 sm	12 sm	3 [1 major, 2 small]	1 (2p)	3.37 (55.43)	?
Cytotype 6 (PI 511885)[e]		Yes	24	11 m + 1 st	1 m, 5 m, 6 m, 12 st	6 [4 major, 2 small]	1 (4p)	2.97 (80.38)	?
Cytotype 7 (PI 511886)[e]		Yes	24	11 m + 1 st	1 m, 2 m, 5 m, 8m	4 [major]	1 (4p)	3.83 (70.05)	?

(continued)

Table 4.1 (continued)

Taxon and voucher number	Corolla shape and color	Pungency	2*n*	Karyotype formula	Chromosomes with active NOR	No. and position of rDNA sites		Hc amount (HKL in μm)	1C DNA content in pg
						45S	*5S*		
2. *C. chinense* Jacq.									
Cytotype 1(GEB, FC, MM 797; AA, GEB 807)[a]	Stellate; white or cream	Yes	24	11 m + 1 st	12 st	5 [2 major, 3 small]	1 (6p)	3.91 (61.31)	3.43
Cytotype 2 (EAM 201)[a]		Yes	24	11 m + 1 st	12 st	?	?	5.52 (61.36)	3.41
3. *C. frutescens* L.									
(GEB, FC, MM 795)[a]	Stellate; white or cream	Yes	24	11 m + 1 st	1 m, 12 st	9 [2 major, 7 small]	1 (5p)[s]	5.55 (66.63)	3.40
4. *C. galapagoense* Hunz.									
(PI 639682)[a]	Stellate; white	Yes	24	11 m + 1 st	12 st	?	?	2.24 (48.66)	?
Baccatum clade									
5. *C. baccatum* L.									
var. *baccatum* (GEB 163)[a]	Stellate; white with greenish spots in the throat	Yes	24	11 m + 1 st	1 m, 3 m, 10 m, 12 st	15 [4 major, 11 small]	1 (5p)[s]	7.45 (66.84)	3.71
var. *pendulum* (Willd.) Eshbaugh									
Cytotype 1 (EAM 192, 209)[a]	Stellate; white with greenish spots in the throat	Yes	24	11 m + 1 st	1 m, 3 m, 12 st	?	?	7.30 (75.53)	3.71
Cytotype 2 (EAM 205, 206, 247; ATH 25382; EAM & RN 211)[a]		Yes	24	11 m + 1 st	1 m, 3 m, 10 m, 12 st	14 [4 major, 10 small]	1 (5p)[s]	7.56 (74.31)	3.68

(continued)

Table 4.1 (continued)

Taxon and voucher number	Corolla shape and color	Pungency	2*n*	Karyotype formula	Chromosomes with active NOR	No. and position of rDNA sites		Hc amount (HKL in μm)	1C DNA content in pg
						45S	*5S*		
var. *umbilicatum* (Vellozo) Hunz. and Barboza									
(EAM 197, 253)[a]	Stellate; white with greenish spots in the throat	Yes	24	11 m + 1 st	1 m, 3 m, 10 m, 12 st	13 [4 major, 9 small]	1 (5p)[s]	9.06 (74.27)	3.76
6. *C. chacoense* Hunz.									
Cytotype 1 (EAM 104, 195, 207, 250; AAC, EAM & FE 973)[a]	Stellate; white	Yes	24	11 m + 1 st	1 m, 12 st	4 [2 major, 2 small]	1 (6p)	2.94 (65.02)	3.34
Cytotype 2 (LB & LG 525)[a]		Yes	24	11 m + 1 st	11 m, 12 st	?	?	2.44 (71.25)	3.36
7. *C. praetermissum* Heiser and Smith									
Cytotype 1 (PI 441654)[a]	Rotate; white with purple lobule margins and greenish spots in the throat	Yes	24	11 m + 1 st	7 m, 12 st	?	?	10.96 (72.55)	?
Cytotype 2 (EFM 05-17)[e]		Yes	24	11 m + 1 sm	6 m, 12 sm	11–13 [2 major, 9–11 small]	1 (7p)[s]	14.92 (76.20)	?
Purple corolla clade									
8. *C. cardenasii* Heiser and Smith									
Cytotype 1 (Neth 904750136)[e]	Campanulate; violet lobules with azure throat	Yes	24	11 m + 1 sm	7 m, 12 sm	8–11 [4 major, 4–7 small]	1 (9p)[s]	9.42 (80.76)	?
Cytotype 2 (AAC w. no; GEB w. no)[e]		Yes	24	11 m + 1 sm	7 m, 12 sm	18 [6 major, 12 small]	1 (9p)[s]	12.53 (69.04)	?
9. *C. eshbaughii* Barboza	Stellate; corolla color unknown	Yes	?	?	?	?	?	?	?

(continued)

Table 4.1 (continued)

Taxon and voucher number	Corolla shape and color	Pungency	2*n*	Karyotype formula	Chromosomes with active NOR	No. and position of rDNA sites		Hc amount (HKL in µm)	1C DNA content in pg
						45S	*5S*		
10. *C. eximium* Hunz.									
Cytotype 1 (EAM 254)[a]	Stellate; white with violet lobules, greenish in the tube	Yes	24	11 m + 1 sm	7 m, 12 sm	?	?	4.90 (68.89)	4.06
Cytotype 2 (EAM 255)[a]		Yes	24	11 m + 1 sm	7 m, 12 sm	6 [2 major, 4 small]	1 (9p)	2.10 (69.65)	?
Pubescens clade									
11. *C. pubescens* Ruiz et Pav									
(GEB 79; EAM 198, 202, 208, 256, 257)[a]	Rotate; purple or violet in the lobules, white or yellowish in the tube	Yes	24	11 m + 1 st	10 m, 12 st	14 [12 major, 2 small]	1 (3p)[s]	18.95 (80.53)	4.47
Tovarii clade									
12. *C. tovarii* Eshbaugh, Smith & Nickerent									
Cytotype 1 (ATH & GEB 25653)[a]	Stellate; variable color (purple or cream, cream with greenish spots in the lobules)	Yes	24	11 m + 1 sm	10 m, 12 sm	?	?	38.91 (70.32)	?
Cytotype 2 (NMCA 90008)[e]		Yes	24	11 m + 1 sm	6 m, 7 m, 12 sm	8 [3 major, 5 small]	1 (9q)	4.89 (67.02)	?
Caatingae clade									
13. *C. caatingae* Barboza and Agra [subnom. *C. parvifolium* (Moscone et al. 2003, 2007)]									
Cytotype 1 (ATH 25233)[a]	Rotate; white with purple spots in the lobules, greenish in the throat and tube	Yes	24	11 m + 1 st	12 st	?	?	5.52 (82.40)	5.77
Cytotype 2 (ATH 25233 bis)[a]		Yes	24	12 m	12 m	?	?	7.47 (77.60)	?

(continued)

Table 4.1 (continued)

Taxon and voucher number	Corolla shape and color	Pungency	2*n*	Karyotype formula	Chromosomes with active NOR	No. and position of rDNA sites		Hc amount (HKL in μm)	1C DNA content in pg
						45S	*5S*		
14. *C. parvifolium* Sendtn.									
(MFA & GEB 7075)[d]	Stellate; whitish green lobes and part of the limb purple adaxially, the tube cream or yellowish white	Yes	24	11 m + 1 sm	12 sm	?	?	?	?
15. *C. longidentatum* Agra and Barboza (MFA & GEB 7086)[d]	Stellate; white with two pale yellow or greenish yellow spots in the throat	No	24	12 m	12 m	?	?	?	?
Flexuosum clade									
16. *C. flexuosum* Sendtn.									
(GEB, FC & EMa 1034; JD & AIH 599)[a]	Stellate; white with greenish spots in the throat	Yes	24	11 m + 1 st	2 m, 5 m	14–15 [12–13 major, 2 small]	1 (9p)[s]	16.82 (103.69)	?
Andean clade									
17. *C. dimorphum* (Miers) Kuntze									
(AOR 2685)	Stellate; yellow, sometimes with violet spots in the throat	No	?	?	?	?	?	?	?
18. *C. geminifolium* (Dammer) Hunz. subnom. *C. scolnikianum* Hunz.									
(GEB & SLG 4819)	Campanulate; yellow with purple or brown spots in the throat	No	?	?	?	?	?	?	?
19. *C. hookerianum* (Miers) Kuntze									
(GEB & SLG 4826, 4831)	Stellate; ocher	No	?	?	?	?	?	?	?
20. *C. rhomboideum* (Dunal) Kuntze									
(YSG 19; 20)[a]	Rotate; yellow	No	26	10 m + 1 sm + 2 st	9 m	1 [major]	1 (3p)	5.07 (41.91)	?

(continued)

Table 4.1 (continued)

Taxon and voucher number	Corolla shape and color	Pungency	2*n*	Karyotype formula	Chromosomes with active NOR	No. and position of rDNA sites		Hc amount (HKL in µm)	1C DNA content in pg
						45S	*5S*		
21. *C. lanceolatum* (Greenm.) C. V. Morton and Standl.									
(NMCA90016)[b]	Stellate campanulate; white or yellowish	No	26	?	?	?	?	?	?
22. *C. lycianthoides* Bitter									
(GDB 85)	Rotate; white or yellowish with violet spots in the throat	No	26	9 m + 3 sm + 1 st	10 sm	?	?	?	?
Atlantic Forest clade									
23. *C. campylopodium* Sendtn.									
Cytotype 1 (ATH 25116)[a]	Stellate; white with golden spots in the throat	Yes	26	5 m + 6 sm + 1 st + 1 t	7 sm	?	?	32.49 (88.30)	5.74
Cytotype 2 (ATH 25128, 25130, 25136)[a]		Yes	26	10 m + 2 sm + 1 st	11 sm	?	?	20.41 (87.95)	4.53
24. *C. cornutum* (Hiern) Hunz.									
(LBB 1527, 1542, 1546)[c]	Rotate; white with violet or brownish spots in the throat, green in the tube	Yes	26	?	?	?	?	?	?
25. *C. friburgense* Bianchetti and Barboza									
(GEB 2048; LBB 1565)[c]	Campanulate urceolate; pink or lilac	Yes	26	?	?	?	?	?	?
26. *C. hunzikerianum* Barboza and Bianchetti									
(GEB, EFi, AG, GB 1648, 1649)	Stellate; white with purple spots in lobules and throat, yellowish in the tube	Yes	?	?	?	?	?	?	?
27. *C. mirabile* Mart.									
Cytotype 1 (ATH 25238, 25251)[a]	Stellate; white with purple spots in the lobules, greenish in the throat and tube	Yes	26	9 m + 2 sm + 1 st + 1 t	7 m	?	?	29.64 (83.81)	?
Cytotype 2 (ATH 25238, 25255)[a]		Yes	26	8 m + 3 sm + 1 st + 1 t	7 m	?	?	29.25 (93.72)	?
Cytotype 3 (ATH 25238, 25267)[a]		Yes	26	9 m + 3 sm + 1 t	9 m	?	?	30.93 (103.4)	?

(continued)

Table 4.1 (continued)

Taxon and voucher number	Corolla shape and color	Pungency	2*n*	Karyotype formula	Chromosomes with active NOR	No. and position of rDNA sites		Hc amount (HKL in μm)	1C DNA content in pg
						45S	*5S*		
28. *C. pereirae* Barboza and Bianchetti									
Cytotype 1 (ATH 26137)[a]	Stellate; white with purple spots in the lobules, yellowish in the throat and tube	Yes	26	9 m + 1 sm + 2 st + 1 t	4 m, 11 st	?	?	11.42 (74.52)	?
Cytotype 2 (ATH 25249)[a]		Yes	26	10 m + 2 st + 1 t	6 m, 11 st	?	?	16.04 (75.85)	?
29. *C. recurvatum* Witas.									
(GEB, MM, RS & RM 915[a]; GEB, EFi, AG, GB 1629, 1632)	Stellate; white with greenish spots in the throat	Yes	26	10 m + 2 sm + 1 st	12 sm, 13 st	4 [2 major, 2 small]	1 (3q)	5.45 (75.52)	?
30. *C. schottianum* Sendtn.									
(ATH 25160)[a]	Stellate; white with violet or brownish spots in the throat, greenish in the tube	Yes	26	9 m + 2 sm + 1 st + 1 t	11 sm	?	?	23.28 (93.71)	?
31. *C. villosum* Sendtn.									
(ATH 25169; GEB, EFi, AG, GB 1653)[a]	Stellate; white with violet or brownish spots in the throat, greenish in the tube	Yes	26	9 m + 3 sm + 1 t	10 sm, 12 sm	30 [4 major, 26 small]	1 (1p)[s]	9.74 (75.89)	?
32. *C. buforum* Hunz.									
(NMCA50029)[b]	Stellate; white with green spots at the base of the lobe. Distal half of each corolla lobe white; the proximal portion and the corolla tube purple	Yes	26	?	?	?	?	?	?
Bolivian clade									
33. *C. caballeroi* M. Nee									
(GEB & CCG 3655)	Campanulate; lemon yellow	Yes	?	?	?	?	?	?	?
34. *C. ceratocalyx* M. Nee	Rotate; yellow with green spots in the throat	Yes	?	?	?	?	?	?	?

(continued)

Table 4.1 (continued)

Taxon and voucher number	Corolla shape and color	Pungency	2*n*	Karyotype formula	Chromosomes with active NOR	No. and position of rDNA sites		Hc amount (HKL in µm)	1C DNA content in pg
						45S	*5S*		
35. *C. coccineum* (Rusby) Hunz.									
(GEB 4921)	Stellate; yellowish white with purplish spots in the throat	Yes	?	?	?	?	?	?	?
36. *C. minutiflorum* (Rusby) Hunz.									
(GEB, MM, AR, GB 1861; GEB & CCG 3644)	Stellate; yellow with greenish spots in the lobules	Yes	?	?	?	?	?	?	?

[a]Extracted from Moscone et al. (2007)
[b]Extracted from Tong and Bosland (2003)
[c]Extracted from Pozzobon et al. (2006)
[d]Taken from Barboza et al. (2011)
[e]Taken from Scaldaferro et al. (2016)
Abbreviations of collector's name: *MFA* MF Agra; *AA* A Anton; *GEB* GE Barboza; *LBB* L de Bem Bianchetti; *GDB* GD Berltran; *GB* G Bertone; *LB* L Bernardello; *CCG* C Carrizo García; *FC* F Chiarini; *AAC* AA Cocucci; *JD* J Daviña; *FE* F Ehrendorfer; *EFi* E Filippa; *EFM* E Forni Martins; *LG* L Galetto; *AG* A Gutiérrez; *AIH* AI Honfi; *ATH* AT Hunziker; *SLG* S Leiva González; *LQ* Llatas Quiroz; *EMa* E Marini; *MM* M Matesevach; *RM* R Minhot; *EAM* EA Moscone; *RN* R Neumann; *AR* A Romanutti; *AOR* A Orejuela; *YSG* Y Sanchez García; *RS* R Scrivanti; Neth collection number of Nijmegen University, the Netherlands; NMCA collection number of the College of Agriculture, New Mexico State University; PI collection number of the United States Department of Agriculture (Griffin, GA)
Abbreviations: *?* unknown data; *Hc* heterochromatin amount; *HKL* haploid karyotype length; *var.* variety; *w. no.* without number; *m* metacentric; *sm* submetacentric; *st* subtelocentric; *t* telocentric; $()^s$ synteny

galapagoense previously belonged to the 'white-flowered group' and were grouped together again in **Annuum** clade based on the phylogeny study. Nevertheless, *C. chacoense* which was also a member of that group but with controversial positions was now nested with *C. baccatum* and *Capsicum praetermissum* in **Baccatum** clade. In Moscone et al. (2007) and Scaldaferro et al. (2013), *C. praetermissum* was specifically ranked in an intermediate position

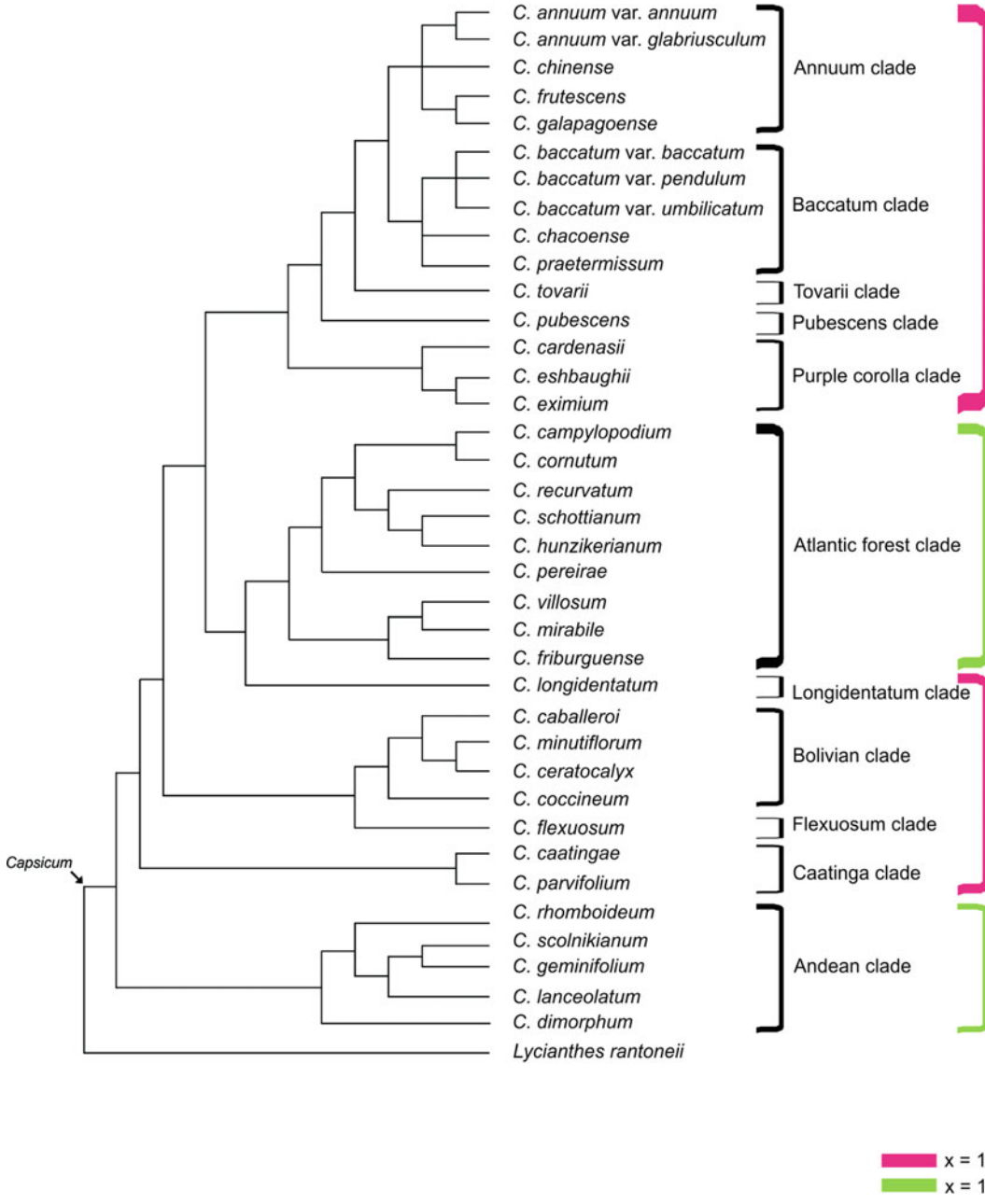


Fig. 4.1 Phylogenetic tree of *Capsicum* species, based on three molecular markers: *matK*, *psbA-trnH*, and *waxy* (Carrizo García et al. 2016), and the relationships with chromosome numbers

between **Baccatum** and **Purple corolla** clade (Moscone et al. 2007; Scaldaferro et al. 2013).

In the phylogeny, *Capsicum flexuosum* was suggested to be close to **Bolivian** clade and was recognized as the monotypic **Flexuosum** clade, although Carrizo García et al. (2016) have included *C.* aff. *flexuosum* (unknown chromosome number) in the clade, considered as a local variation because results did not support a strong specific separation from typical *C. flexuosum*.

Capsicum cardenasii, *C. eshbaughii*, *C. eximium, C. pubescens*, and *C. tovarii* were members of the traditional 'purple-flowered group.' Then, they were assigned to **Purple corolla** (*C. cardenasii*, *C. eshbaughii*, and *C. eximium*), **Pubescens,** and **Tovarii** clades, respectively.

Another group with *Capsicum campylopodium*, *C. cornutum*, *C. friburgense*, *C. hunzikerianum*, *C. mirabile*, *C. pereirae*, *C. recurvatum, C. schottianum*, and *C. villosum* belong to **Atlantic Forest** clade ($x = 13$), phylogenetically distant from the above-mentioned groups. This clade has corollas mostly white, with some variations in the throat (golden, violet, brownish, greenish, or purple spots) and in the tube (yellowish or greenish). In *C. friburgense*, corolla appears completely pink or lilac.

On the other hand, *C. rhomboideum, C. lanceolatum, C. geminifolium, C. lycianthoides, C. dimorphum*, and *C. hookerianum* are the most distant taxa, belonging to **Andean** 'yellow-flowered group,' sometimes with violet spots in the throat and with $x = 13$.

Caatingae clade includes *Capsicum caatingae* and *C. parvifolium*, both with $x = 12$ and with similar karyotype formulas. **Longidentatum** clade, with a single species, *C. longidentatum*, possesses $x = 12$ with a karyotype very similar to that of **Caatingae** clade.

Finally, **Bolivian** clade, with *Capsicum caballeroi*, *C. ceratocalyx*, *C. coccineum*, and *C. minutiflorum*, presents lemon yellow flowers, sometimes with violet spots in the throat. This clade has not been studied cytogenetically until now, and therefore, its chromosome number is still unknown; however, its position in the phylogeny suggests that Bolivian species have $2n = 24$, as the sister clades **Longidentatum**, **Flexuosum,** and **Caatingae** (Fig. 4.1).

4.3 Basic Chromosome Number

The genus *Capsicum* has two universal chromosome numbers: $2n = 2x = 24$ and $2n = 2x = 26$, the latter only found in wild species (Pickersgill 1971, 1991; Moscone 1990, 1993, 1999; Moscone et al. 1996a, 2007; Tong and Bosland 2003; Pozzobon et al. 2006; Scaldaferro et al. 2011, 2013, 2016).

To date, the diploid chromosome number for 77.8% of the recorded *Capsicum* species (28/36) is known. Among them, 15 species (15/28) have $2n = 2x = 24$, whereas 13 species (13/28) possess $2n = 2x = 26$ (Table 4.1). The phenomenon of polyploidy has never been significant in *Capsicum* and was only found in one accession of *C. annuum* var. *glabriusculum* with $2n = 4x = 48$ (Pickersgill 1977).

The chromosome numbers of the following eight species have not been reported yet: the whole Bolivian clade (*C. caballeroi*, *C. ceratocalyx*, *C. coccineum*, and *C. minutiflorum*), *C. eshbaughii* (Purple corolla clade), *C. dimorphum*, *C. hookerianum* (Andean clade), and *C. hunzikerianum* (Atlantic Forest clade).

4.4 Karyotyping of *Capsicum* Species

Since 1971, *Capsicum* species have been assessed using different methodological chromosome approaches, including classic and silver staining, fluorescent banding, FISH, and nuclear DNA content estimation (Pickersgill 1971, 1991; Moscone 1990, 1993, 1999; Moscone et al. 1993, 1995, 1996a, b, 1999, 2003, 2007; Park et al. 2000; Scaldaferro et al. 2006, 2011, 2013, 2016; Barboza et al. 2011; Romero-da Cruz et al. 2016).

Although chromosome number has been studied in 28 species, their karyotypes have been obtained only from 24, since the number of chromosomes proceeds from meiosis in some

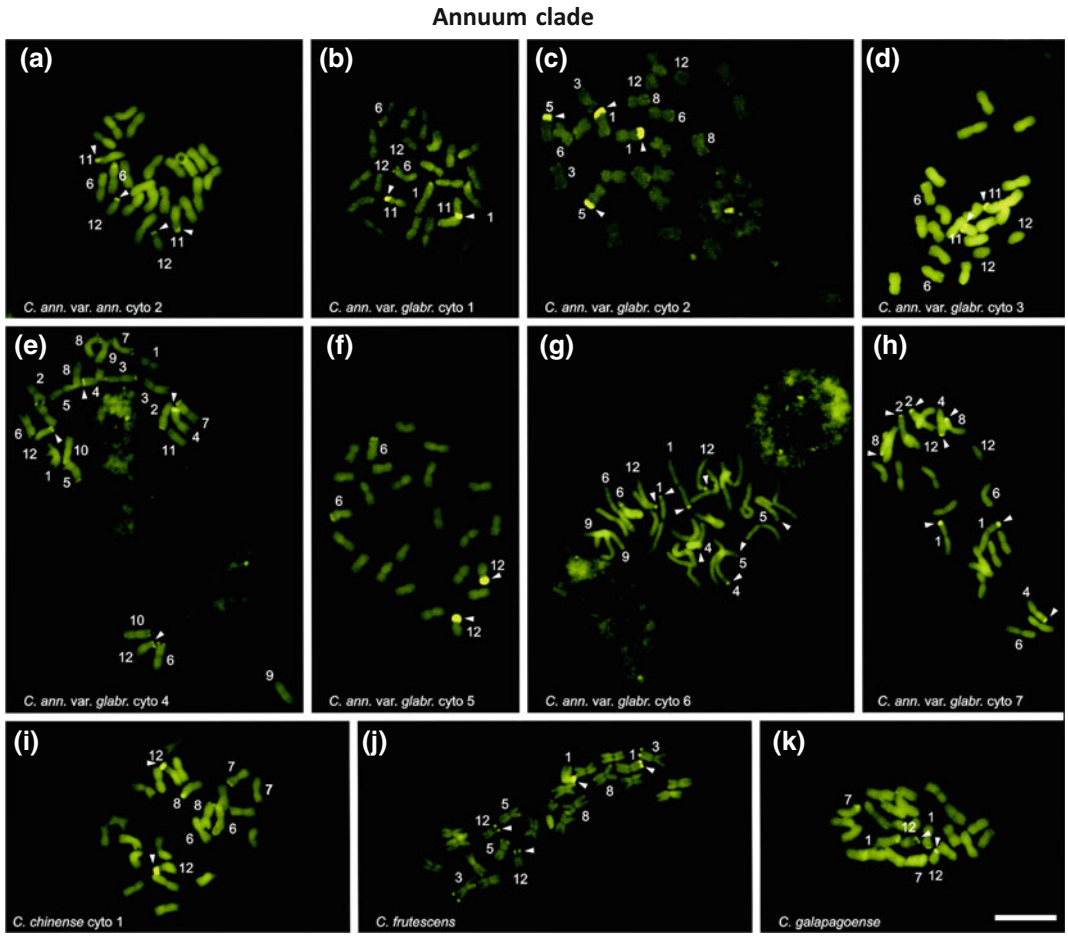


Fig. 4.2 Somatic metaphases of **Annuum** clade ($2n = 24$) stained with CMA. **a** *C. annuum* var. *annuum* cytotype 2 (NMCA 10272). **b** *C. annuum* var. *glabriusculum* cytotype 1 (NMCA 10955). **c** *C. annuum* var. *glabriusculum* cytotype 2 (NMCA 10983). **d** *C. annuum* var. *glabriusculum* cytotype 3 (Ll. Q. w. no.). **e** *C. annuum* var. *glabriusculum* cytotype 4 (YSG w. no.). **f** *C. annuum* var. *glabriusculum* cytotype 5 (Netherlands 804750009). **g** *C. annuum* var. *glabriusculum* cytotype 6 (PI 511885). **h** *C. annuum* var. *glabriusculum* cytotype 7 (PI 511886). **i** *C. chinense* cytotype 1 (GEB et al. 797). **j** *C. frutescens* (GEB et al. 795). **k** *C. galapagoense* (PI 639682). Identified homologous chromosomes are indicated with the same numbers as those used in the respective ideogram (Fig. 4.11). Arrowheads indicate CMA+ NOR-associated heterochromatin. Scale bar = 10 μm

cases (i.e., *Capsicum buforum*, *C. cornutum*, *C. friburgense*, and *C. lanceolatum*; Tong and Bosland 2003; Pozzobon et al. 2006). Half of cytogenetically studied taxa present intraspecific karyotype variation, differing in karyotype formulas, number and location of active NORs, heterochromatin content (Hc), and banding pattern (Moscone et al. 2007; Scaldaferro et al. 2013, 2016) (Figs. 4.2, 4.3, 4.4, 4.5, 4.6, 4.11, 4.12; Table 4.1).

According to the base-specific fluorochromes used for chromosome banding, there are four types of constitutive heterochromatin in *Capsicum*, which depend on composition of the satellite DNA: (1) highly GC-rich heterochromatin CMA+/DAPI−, CMA homogeneously bright and DAPI dull, occurring in NORs of every *Capsicum* species; (2) highly AT-rich heterochromatin CMA−/DAPI+, CMA dull and DAPI bright, only present in *C. campylopodium*,

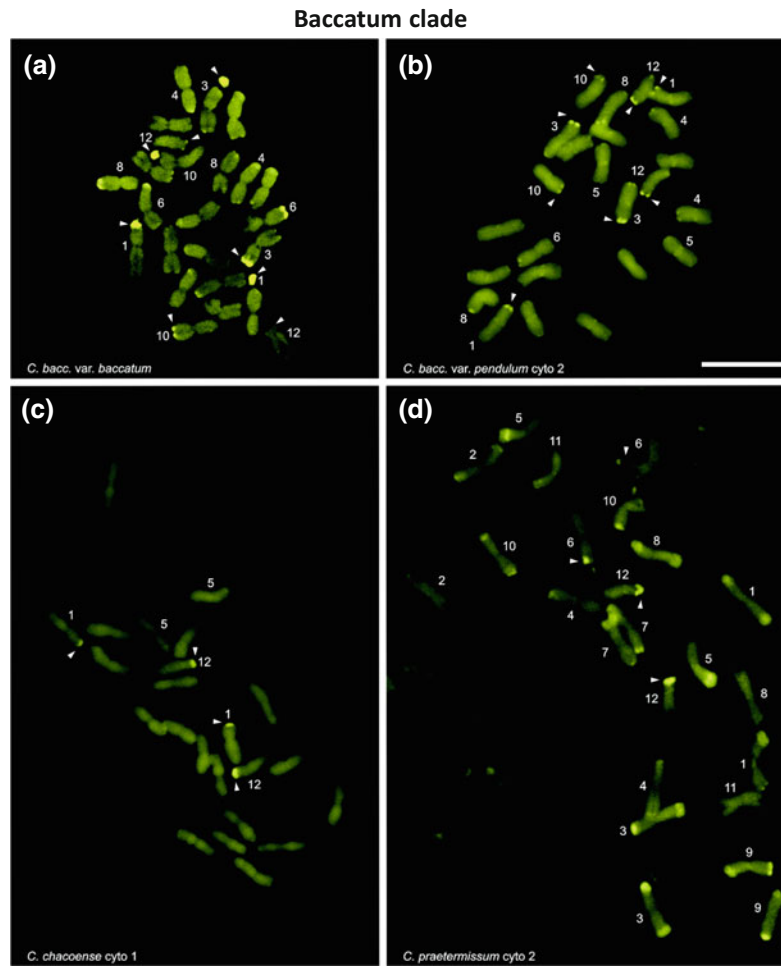


Fig. 4.3 Somatic metaphases of **Baccatum** clade ($2n = 24$) stained with CMA. **a** *C. baccatum* var. *baccatum* (GEB 163). **b** *C. baccatum* var. *pendulum* cytotype 2 (EAM & RN 211). **c** *C. chacoense* cytotype 1 (EAM 250). **d** *C. praetermissum* cytotype 2 (EFM 05-17). Identified homologous chromosomes are indicated with the same numbers as those used in the respective ideogram (Fig. 4.11). Arrowheads indicate CMA+ NOR-associated heterochromatin. Scale bar = 10 μm

C. pereirae, *C. praetermissum*, and *C. pubescens*; (3) moderately GC-rich heterochromatin CMA+/DAPIo, CMA bright and DAPI indifferent, and occurs in a variable number of distal and intercalary bands; and (4) CMA+/DAPI+ mixed distal bands CMA and DAPI bright, only observed in *C. campylopodium* and *C. praetermissum* (Fig. 4.11).

The species of Purple corolla clade share the karyotype formula but differ in heterochromatin amount: *C. eximium* exhibits a slightly different chromosome banding pattern with lower Hc than *C. cardenasii*. In *Pubescens*, *Tovarii*, *Baccatum*, and *Annuum* clades, chromosome number and karyotype formula of every species have been yet reported, but in *C. eshbaughii* from Purple corolla clade the chromosome number is still unknown.

Chromosome data from Andean clade species only has been reported for *C. lycianthoides* and *C. rhomboideum*. Both karyotype formulas are quite similar: They have small chromosomes compared to other clades, little Hc, and only one pair of NORs.

Atlantic Forest clade comprises 11 species; to date, chromosomes of only eight species in this group have been cytogenetically studied. All of them possess $x = 13$, with more asymmetrical karyotypes than those in Andean clade and with higher frequencies of *sm*, *st*, or *t* chromosomes. Their chromosome complements show longer haploid karyotype length (HKL) than Andean

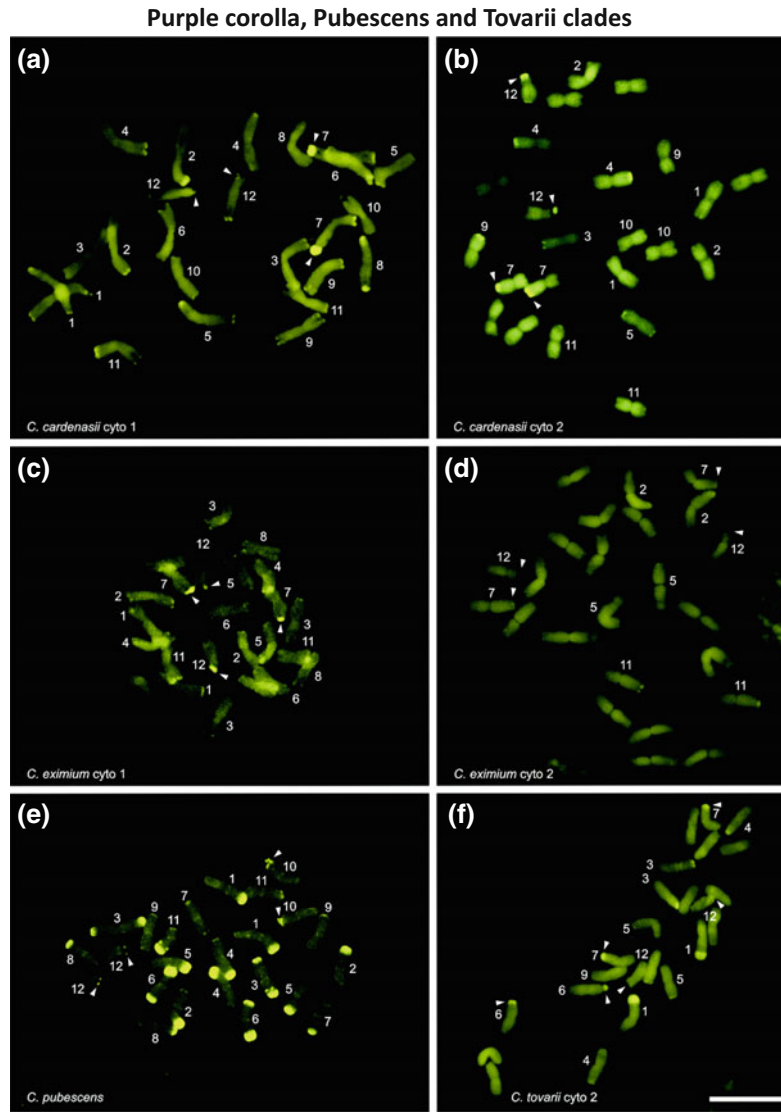


Fig. 4.4 Somatic metaphases of **Purple corolla, Pubescens,** and **Tovarii** clades ($2n = 24$) stained with CMA. **a** *C. cardenasii* cytotype 1 (GEB w. no.). **b** *C. cardenasii* cytotype 2 (Netherlands 904750136). **c** *C. eximium* cytotype 1 (EAM 254). **d** *C. eximium* cytotype 2 (EAM 255). **e** *C. pubescens* (EAM 257). **f** *C. tovarii* cytotype 2 (NMCA 90008). Identified homologous chromosomes are indicated with the same numbers as those used in the respective ideogram (Fig. 4.11). Arrowheads indicate CMA+ NOR-associated heterochromatin. Scale bar = 10 μm

clade, being twice the complement length than in *C. rhomboideum* in every case (Table 4.1).

The NOR-associated heterochromatin is CMA +/DAPI−; however, it sometimes appears as CMA+/DAPIo and includes the distal satellite and a small portion of the respective arm adjacent to the NOR. Nucleolar organizer regions are situated on short or long arms, although more often on short arms, and appear as constrictions or gaps in fluorochrome-banded chromosomes. Sometimes, the centromeric heterochromatin is visible only as faint CMA+/DAPIo paired dots, as previously reported for the cultivated taxa of *Capsicum* (Moscone et al. 1996a). The amplitude ranges of Hc (indicated as percentage of HKL) vary broadly (from 1.80 to 38.91) in the genus and correlate positively with the HKL in most of the taxa examined (Fig. 4.13). *C. annuum* and *C. tovarii* have the lowest and highest Hc, respectively, but among clades, *Annuum* remains the one with lowest Hc, whereas the species with the highest Hc prevail in Atlantic Forest (Table 4.1).

Species with $2n = 24$ show rather uniform and comparatively most symmetrical karyotypes,

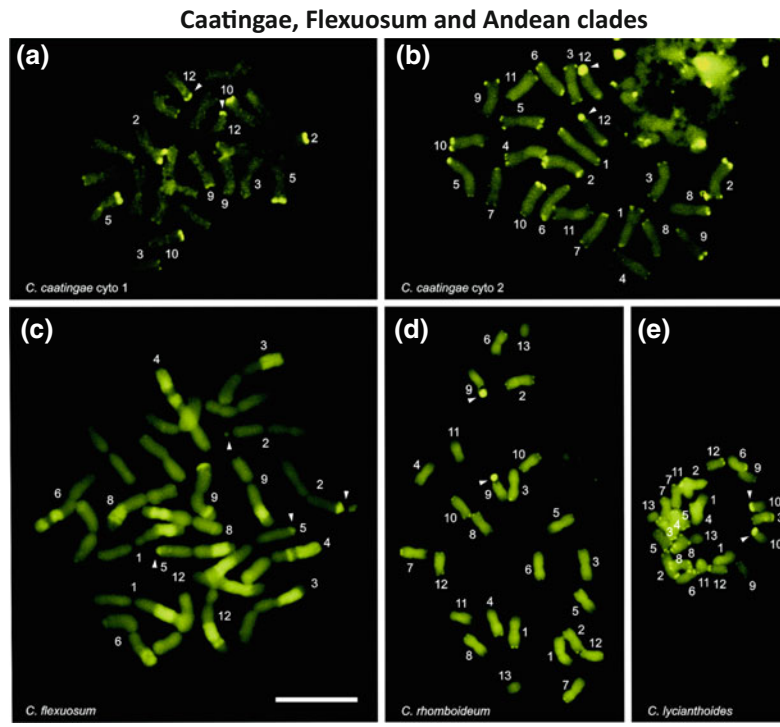


Fig. 4.5 Somatic metaphases of **Caatingae, Flexuosum** ($2n = 24$), and **Andean** clades ($2n = 26$) stained with CMA. **a** *C. caatingae* cytotype 1 (ATH 25233). **b** *C. caatingae* cytotype 2 (ATH 25233 bis). **c** *C. flexuosum* (GEB et al. 1034). **d** *C. rhomboideum* (YSG 20). **e** *C. lycianthoides* (GDB 85). Identified homologous chromosomes are indicated with the same numbers as those used in the respective ideogram (Fig. 4.11). Arrowheads indicate CMA+ NOR-associated heterochromatin. Scale bar = 10 µm

since most of them have the 11 *m* + 1 *st* karyotype formula, although 11 *m* + 1 *sm* is also frequent. In contrast, among $2n = 26$ species karyotype formulas are more asymmetrical, having nine different karyotypes among nine taxa.

4.5 Mapping of the *45S* and *5S* Ribosomal RNA Genes

Cytotaxonomy commonly uses number and distribution of secondary constrictions, AgNOR bands, satellites, and *45S* rDNA loci as morphological karyotype characters (Baeza and Schrader 2005; Xu et al. 2007; García et al. 2009, among hundreds of studied species). All of those characters are particularly associated with the highly preserved ribosomal *45S* RNA genes, described as markers of transcriptional or active *45S* rDNA genes (e.g., secondary constrictions, AgNOR bands, and satellites) and of non-functional rDNA sites or inactive *45S* rDNA loci (Kovarik et al. 2008).

Recently, a physical chromosome map of 12 *Capsicum* taxa was constructed employing AgNOR banding, which reports the number, size, and position of active NORs and FISH; used together, both methods inform about *5S* and active and inactive *45S* rRNA genes, revealing the functional *45S* rRNA genes in most species of the genus (Scaldaferro et al. 2016).

In *Capsicum*, AgNORs are frequently associated with satellites that not always differentially dye with silver nitrate. Nucleolar organizer regions appear as constrictions in chromosomes stained with fluorescent dyes in every case (Scaldaferro et al. 2013). The NORs and their associated heterochromatin are rich in GC base pairs (Moscone et al. 1996a, 2007; Scaldaferro et al. 2013) as is the prevalence in plants (Sinclair and Brown 1971). Therefore, in the genus all NORs are considered descendants of the same initial NOR due to an identical base pair constitution (Berg and Greilhuber 1993).

FISH method has resulted in an essential tool for physical gene mapping. Ribosomal genes are highly repetitive sequences or tandem

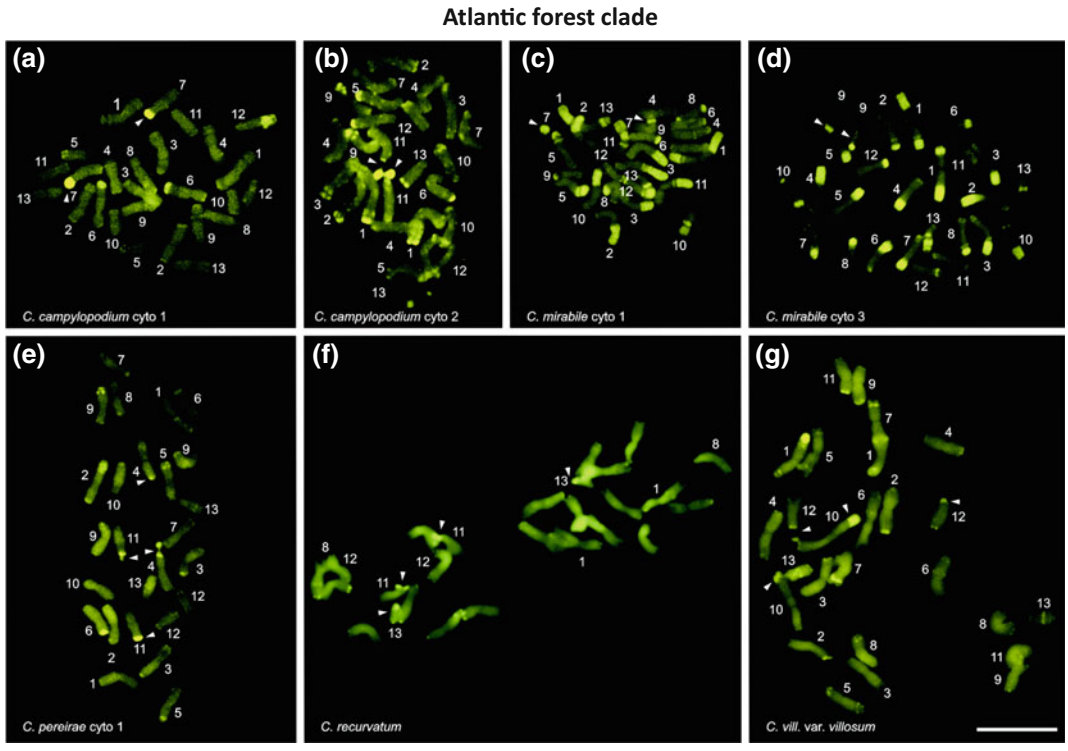
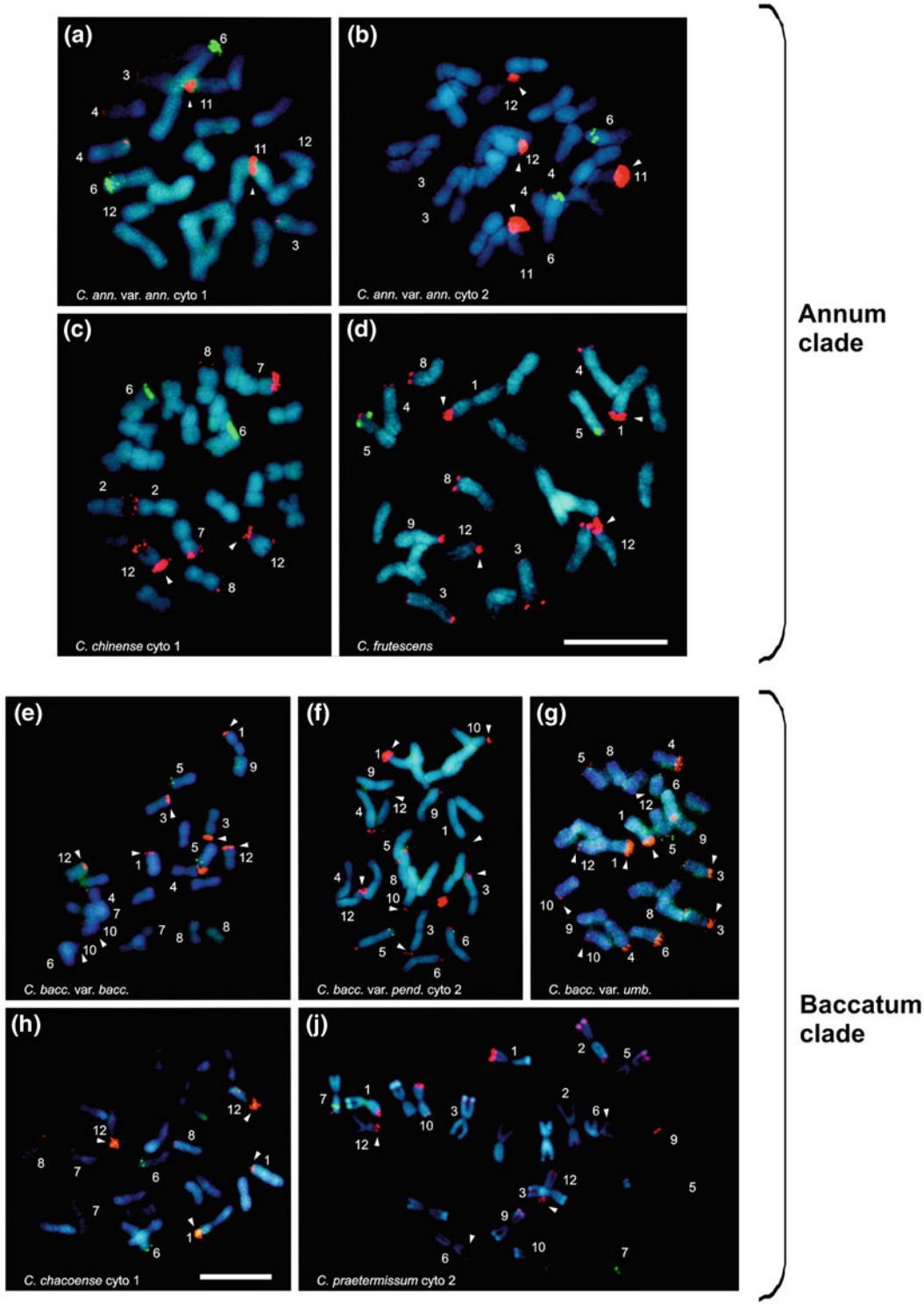


Fig. 4.6 Somatic metaphases of **Atlantic forest** clade ($2n = 26$) stained with CMA. **a** *C. campylopodium* cytotype 1 (ATH 25116). **b** *C. campylopodium* cytotype 2 (ATH 25128). **c** *C. mirabile* cytotype 1 (ATH 25251). **d** *C. mirabile* cytotype 3 (ATH 25238). **e** *C. pereirae* cytotype 1 (ATH 26137). **f** *C. recurvatum* (GEB et al. 1629). **g** *C. villosum* (GEB et al. 1653). Identified homologous chromosomes are indicated with the same numbers as those used in the respective ideogram (Fig. 4.11). Arrowheads indicate CMA+ NOR-associated heterochromatin. Scale bar = 10 μm

arrangements found in a small number of sites (loci) in the species genome. In the genus *Capsicum*, FISH of *5S* and *45S* rRNA genes shows disparity in number, size, and location among the species (Park et al. 2000; Scaldaferro et al. 2006, 2016). Physically, *5S* locus maps in a single preserved position, principally intercalary in a metacentric median chromosome. This *5S* rDNA distribution could be parsimoniously explained if the common ancestor of the genus was bearer of a single intercalary *5S* locus on a medium-sized to large chromosome. Until now, established linkage maps in *Capsicum* have not included the *5S* rRNA gene (Livingstone et al. 1999; Lefebvre et al. 2002; Paran et al. 2004).

The number and position of *45S* rDNA loci are useful characters for morphological identification of similar chromosome sites and operate as evolutionary markers between species. In *Capsicum*, the number of 45S rDNA sites is remarkably variable, ranging widely from a unique pair in *C. rhomboideum* up to 30 pairs in *C. villosum* (Figs. 4.7, 4.8, 4.12; Table 4.1), and both number and position of *45S* loci remain constant within each species, with some exceptions; e.g., in *C. annuum*, there are from 1 to 6 sites, 14 to 15 sites in *C. baccatum*, and from 8 to 18 sites in *C. cardenasii*. Although a relative constancy is observed in the *45S* loci as a whole, the smaller landmarks are more variable, as the major sites hold number and position constant within each species and cytotype. These last sites are principally concomitant with that NORs that have been previously identified by AgNOR banding, and therefore are the active sites (Scaldaferro et al. 2006, 2016). In general,


(a)
C. ann. var. ann. cyto 1
(b)
C. ann. var. ann. cyto 2
(c)
C. chinense cyto 1
(d)
C. frutescens
Annum clade
(e)
C. bacc. var. bacc.
(f)
C. bacc. var. pend. cyto 2
(g)
C. bacc. var. umb.
(h)
C. chacoense cyto 1
(j)
C. praetermissum cyto 2
Baccatum clade

◀ **Fig. 4.7** Localization of *45S* (red) and *5S* (green) rDNA loci inferred via FISH in **Annuum** and **Baccatum** clades. **a** *C. annuum var. annuum* cytotype 1 (EAM 251). **b** *C. annuum var. annuum* cytotype 2 (EAM 204). **c** *C. chinense* cytotype 1 (GEB et al. 807). **d** *C. frutescens* (GEB et al. 795). **e** *C. baccatum* var. *baccatum* (GEB 163). **f** *C. baccatum* var. *pendulum* cytotype 2 (EAM 247). **g** *C. baccatum* var. *umbilicatum* (EAM 253). **h** *C. chacoense* cytotype 1 (EAM 250). **i** *C. praetermissum* cytotype 2 (EFM 05-17). Identified homologous chromosomes are indicated with the same numbers as those used in the respective ideogram (Fig. 4.12). Arrowheads indicate active NOR. Scale bar = 10 μm

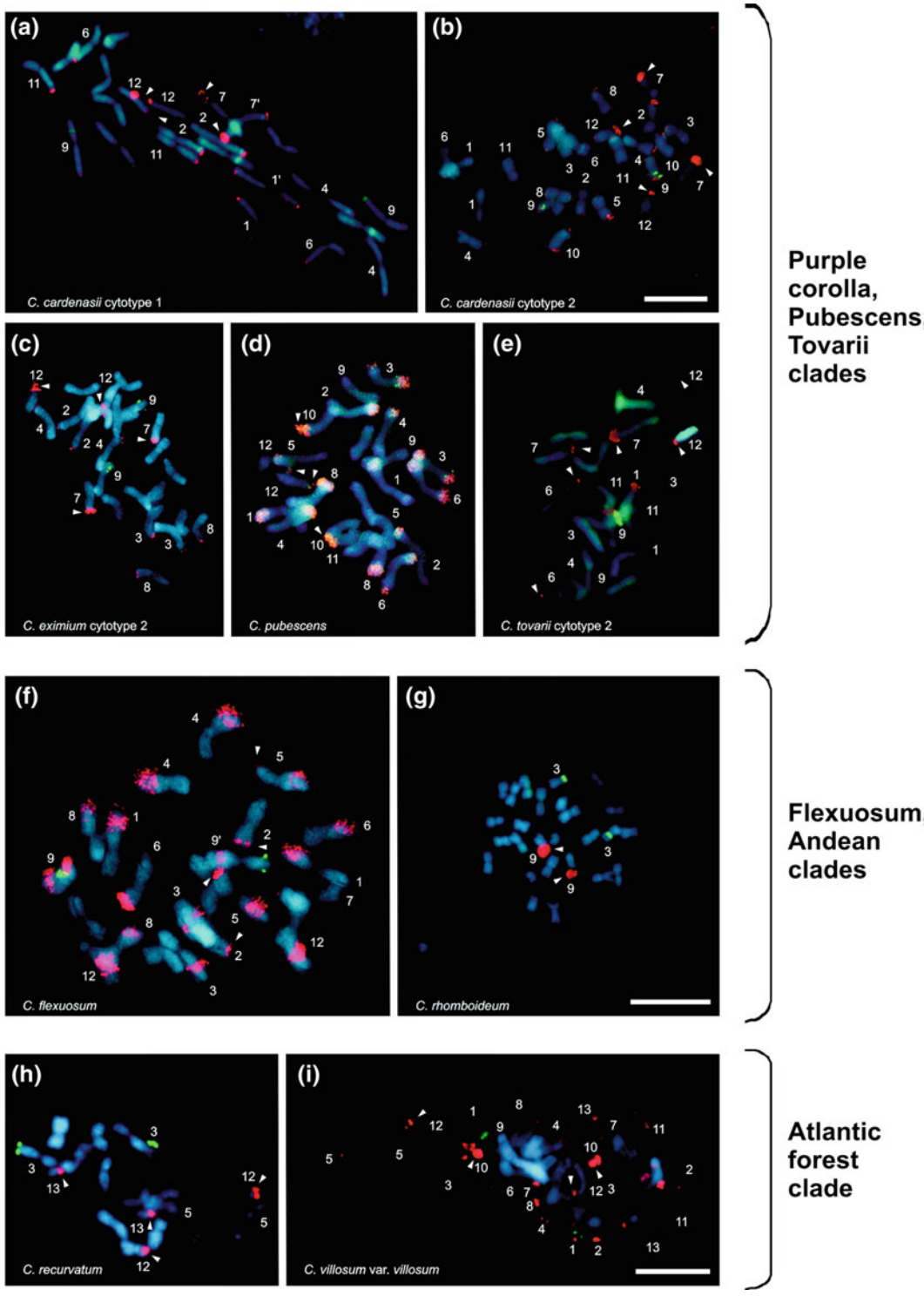
diploid plant genera species bear one pair of NOR (Raina and Khoshoo 1971), but in very few cases diploid taxa contain more than one pair of NOR. In situ, hybridization studies have identified several other rDNA loci but on chromosomes that are devoid of NORs. Hence, the signals at those sites are generally considered to be inactive sites that do not synthesize ribosomal RNA. Even in those diploid species with more than two NORs, only two remain active, as generally found using FISH (Raina and Mukai 1999).

Genomic evolution in *Capsicum* has involved considerable changes in number and distribution of the *45S* gene family, including locus loss and gain, and sequence spreading. Other mechanisms that generate variations in size, number, and position of rDNA sites are structural rearrangements, such as inversions and translocations, homologous and non-homologous unequal crossing over, gene conversion, and transpositional events (Hall and Parker 1995; Sharma and Raina 2005). Evidence suggests that positioning and remodeling of rDNA sites could be related to the rDNA gene shuffling or transposable elements playing an important role in plant genome evolution (Dubcovsky and Dvorák 1995; Raskina et al. 2004; Datson and Murray 2006).

In *Capsicum*, *45S* FISH signals mostly correspond to specific fluorescent banding, although they do not coincide absolutely in number, location, or size (Moscone et al. 2007; Scaldaferro et al. 2013, 2016). Accordingly, there would be a relationship between *45S* rDNA probes and GC-rich heterochromatic regions. Park et al. (2012) have studied in detail the evolution of constitutive heterochromatin in *Capsicum*. They showed an expansion of this genome structure 20.0–7.5 million years ago in pepper through a massive accumulation of single-type *Ty3/Gypsy*-like elements from the Del subgroup. Interestingly, derivatives of the Del elements played important roles in the expansion of constitutive heterochromatic regions. This process represents a characteristic mechanism for genome amplification in plant species through expansion of constitutive heterochromatic regions, which does not involve a genome-wide duplication event. Most recently, Qin and Yu (2014) explained that LTR expansion promoted the large genome size in *Capsicum*. Our findings about the localization of *45S* probes and their relationship with heterochromatic regions and active NORs also suggest their additional role in *Capsicum* genome diversity.

4.6 AgNOR Mapping

Silver impregnation is used to reliably detect active NORs (Ag-I; Bloom and Goodpasture 1976; Kodama et al. 1980). Active NORs vary in number from one to four pairs among *Capsicum* species. The maximum number of NORs (four pairs) only appears in species with $2n = 24$; instead, taxa with $2n = 26$ present 1–2 pairs maximum. Few species have only one pair: *C. annuum* var. *glabriusculum* (cytotypes 1, 3, and 5) and *C. rhomboideum*; instead, nine species present two pairs: *C. annuum* var. *annuum* (cytotype 2) and *C. annuum* var. *glabriusculum* (cytotypes 2 and 4), *C. chinense*, *C. frutescens*, *C. eximium* (cytotype 2), *C. cardenasii* (cytotypes 1 and 2), *C. flexuosum, C. praetermissum*, *C. recurvatum*, and *C. villosum*. The only species exhibiting three pairs is *C. tovarii* (cytotype 2). Finally, *C. annuum* var. *glabriusculum* (cytotypes 6 and 7; Fig. 4.9) and *C. baccatum* var. *baccatum* and var. *pendulum* show four pairs of NORs in their diploid complements (Fig. 4.9).


(a)
C. cardenasii cytotype 1
(b)
C. cardenasii cytotype 2
(c)
C. eximium cytotype 2
(d)
C. pubescens
(e)
C. tovarii cytotype 2
Purple corolla, Pubescens, Tovarii clades
(f)
C. flexuosum
(g)
C. rhomboideum
Flexuosum, Andean clades
(h)
C. recurvatum
(i)
C. villosum var. villosum
Atlantic forest clade

◀ **Fig. 4.8** Localization of *45S* (red) and *5S* (green) rDNA loci inferred via FISH in **Purple corolla, Pubescens, Tovarii, Flexuosum, Andean,** and **Atlantic Forest** clades. **a** *C. cardenasii* cytotype 1 (GEB w. no.). **b** *C. cardenasii* cytotype 2 (Netherlands 904750136). **c** *C. eximium* cytotype 2 (EAM 255). **d** *C. pubescens* (EAM 257). **e** *C. tovarii* cytotype 2 (NMCA 90008). **f** *C. flexuosum* (GEB et al. 1034). **g** *C. rhomboideum* (YSG 20). **h** *C. recurvatum* (GEB et al. 915). **i** *C. villosum* (GEB et al. 1653). Identified homologous chromosomes are indicated with the same numbers as those used in the respective ideogram (Fig. 4.12). Arrowheads indicate active NOR. Scale bar = 10 μm

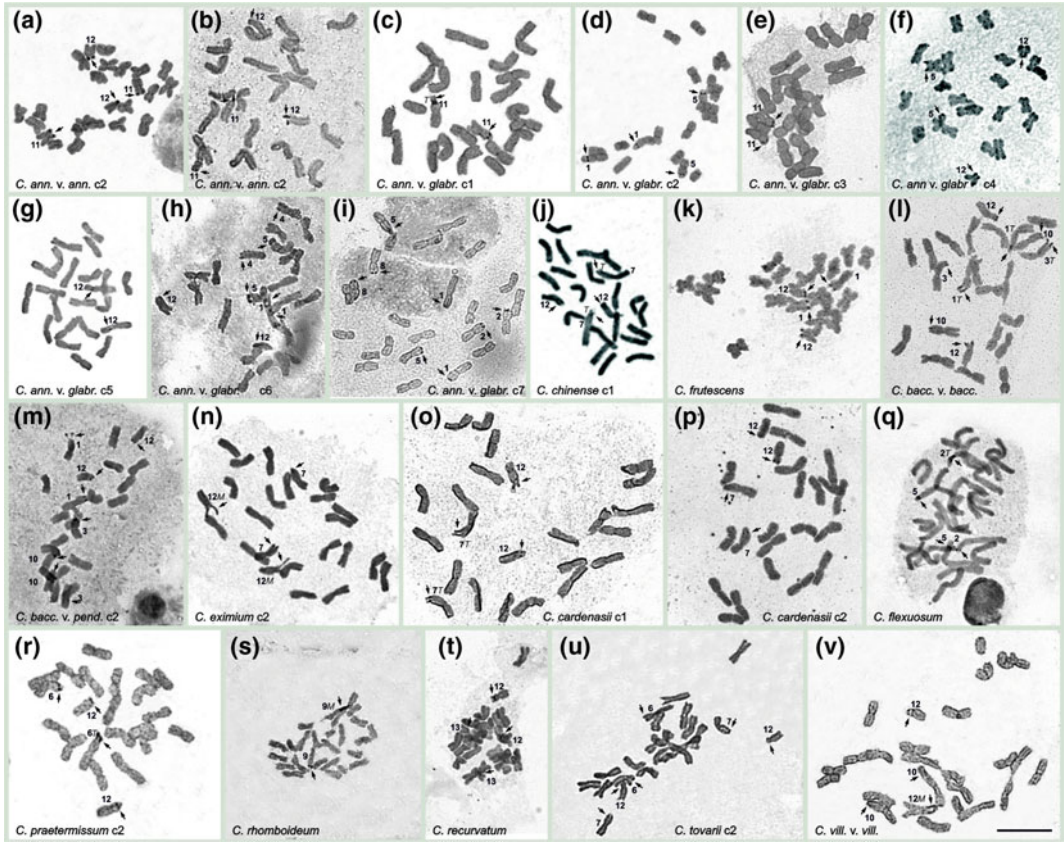


Fig. 4.9 Silver-stained somatic metaphases of *Capsicum*. **a–b** *C. annuum* var. *annuum*: **a** NMCA 10272 cytotype 2 and **b** NMCA 10544 cytotype 2. **c–i** *C. annuum* var. *glabriusculum*: **c** NMCA 10955 cytotype 1, **d** NMCA 10983 cytotype 2, **e** LQ w. no. cytotype 3, **f** YSG w. no. cytotype 4, **g** Netherlands 804750009 cytotype 5, **h** PI 511885 cytotype 6, **i** PI 511886 cytotype 7. **j** *C. chinense* cytotype 1 GEB 807. **k** *C. frutescens* GEB, FC, MM 795. **l** *C. baccatum* var. *baccatum* GEB 163. **m** *C. baccatum* var. *pendulum* EAM and RN 211 cytotype 2. **n** *C.eximium* EAM 255 cytotype 2. **o–p** *C. cardenasii*: **o** Netherlands 904750136 cytotype 1, **p** AAC w. no. cytotype 2. **q** *C. flexuosum* GEB, FC, EMa 1034. **r** *C. praetermissum* EFM 05-17 cytotype 2. **s** *C. rhomboideum* YSG 20. **t** *C. recurvatum* GEB, MM, RSc, RM 915. **u** *C. tovarii* NMCA 90008 cytotype 2. **v** *C. villosum* GEB, EFi, AG, GB 1653.*M*, macrosatellite; *T*, tandem satellite. Arrows indicate AgNOR. Scale bar = 10 μm

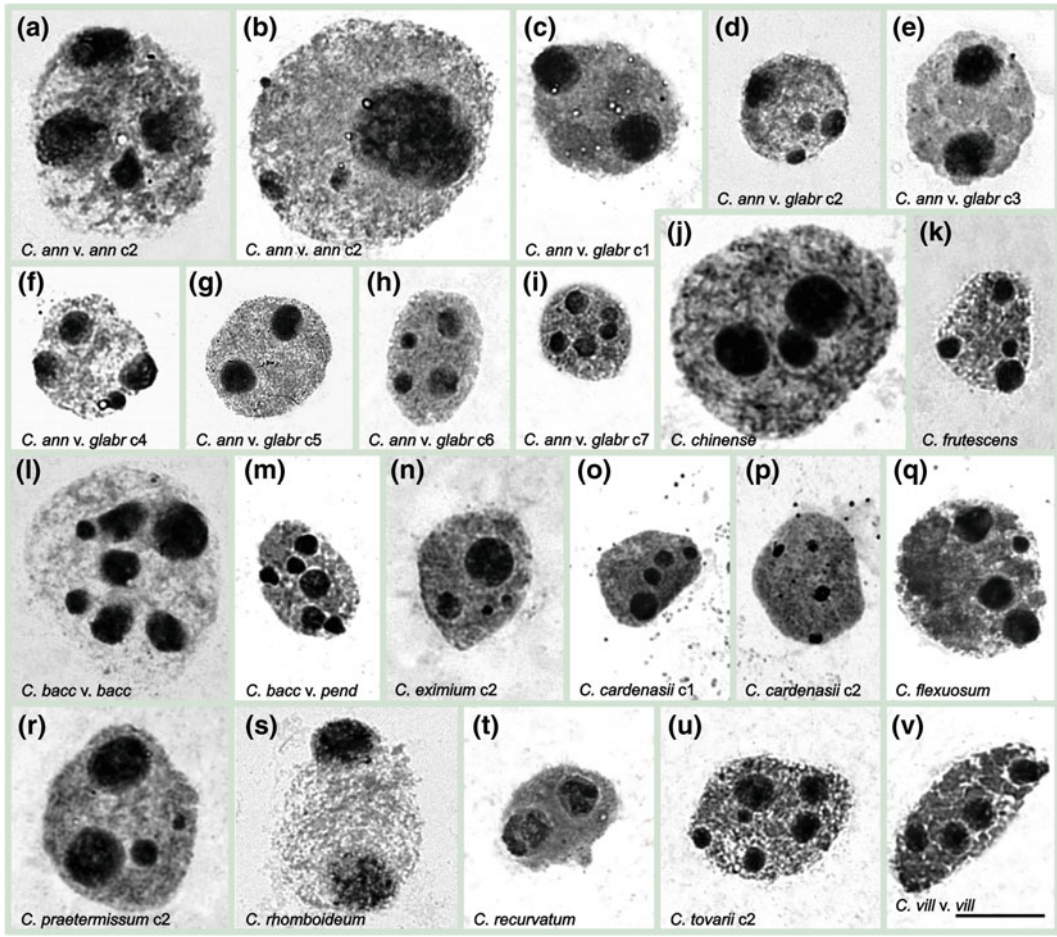


Fig. 4.10 Silver-stained interphase nucleus of *Capsicum*. **a–b** *C. annuum* var. *annuum*: **a** NMCA 10272 cytotype 2 and **b** NMCA 10544 cytotype 2. **c–i** *C. annuum* var. *glabriusculum*: **c** NMCA 10955 cytotype 1, **d** NMCA 10938 cytotype 2, **e** LQ w. no. cytotype 3, **f** YSG w. no. cytotype 4, **g** Netherlands 804750009 cytotype 5, **h** PI 511885 cytotype 6, **i** PI 511886 cytotype 7. **j** *C. chinense* GEB 807. **k** *C. frutescens* GEB, FC, MM 795. **l** *C. baccatum* var. *baccatum* GEB 163. **m** *C. baccatum* var. *pendulum* EAM and RN 211. **n** *C. eximium* EAM 255 cytotype 2. **o–p** *C. cardenasii*: **o** Netherlands 904750136 cytotype 1, **p** AAC w. no. cytotype 2. **q** *C. flexuosum* GEB, FC, EMa 1034. **r** *C. praetermissum* EFM 05-17 cytotype 2. **s** *C. rhomboideum* YSG 20. **t** *C. recurvatum* GEB, MM, RSc, RM 915. **u** *C. tovarii* NMCA 90008 cytotype 2. **v** *C. villosum* GEB, EFi, AG, GB 1653. Scale bar = 10 m

Mostly, the NORs in *Capsicum* are positioned on the short arm of the respective chromosomes, although some taxa exhibit one NOR on the long arm of different chromosome pairs, e.g., *C. annuum* var. *glabriusculum* (cytotypes 2, 4, 6, and 7), *C. eximium* (cytotype 2), *C. cardenasii* (cytotypes 1 and 2), *C. tovarii* (cytotype 2), and *C. villosum*. The sizes of NOR-associated satellites are diverse among species, individuals, and frequently among cells from the same plant. According to Battaglia terminology (Battaglia 1955), microsatellites, macrosatellites, and

tandem satellites are registered in varying proportions among species (Fig. 4.9; see *M* NOR-bearing chromosomes and *T* NOR-bearing chromosomes). Also in some cases, size of NORs varies between homologues.

In all cases, the maximum number of nucleoli seen in interphase is coincident with the maximum number of NORs found in metaphase (Fig. 4.10). Although the correspondence between the size of NORs in metaphase and the size of nucleoli in the interphase nuclei is a well-established phenomenon in plants (Burger and Knälmann 1980; Hizume et al. 1982; Linde-Laursen 1984), no size correlation has been observed in *Capsicum* (Moscone et al. 1995; Scaldaferro et al. 2016).

4.7 DNA Content of *Capsicum* Species

Moscone et al. (2003) estimated nuclear DNA content in nine diploid species of *Capsicum* by flow cytometry, using ethidium bromide to stain the DNA (internal standard, *Hordeum vulgare*, *1C* = 5.063 pg) (Table 4.1). Additionally, two samples were analyzed using Feulgen densitometry (*C. annuum* var. *annuum* and *C. pubescens*; standard, *Allium cepa*, *1C* = 16.75 pg). Very similar relative values were obtained from both staining methods. The 1C values ranged from 3.34 to 3.43 pg (3273–3361 Mbp) in *C. chacoense* and the *C. annuum* complex to 4.53–5.77 pg (4439–5655 Mbp) in *C. campylopodium* and *C. caatingae*. Genome size displayed significant variation between but not within species (except in *C. campylopodium*) and contributed to their taxonomic grouping (Moscone et al. 2003) (Figs. 4.11 and 4.12).

Quantity and distribution of heterochromatin in *Capsicum* suggested that this type of chromatin may have been gained by addition rather than by euchromatin transformation. As a consequence of the proportion change of repeated DNA sequences in the nuclear genome (particularly tandem repeats or satellite DNAs that make up heterochromatic C-bands on the chromosomes), the DNA content in angiosperms varies (Flavell 1986; Raina and Bisht 1988; Bennett and Leitch 1995; Greilhuber 1995). Another previous work showed a strong positive correlation between genome size and Hc in *Capsicum*, with higher range of variation in the latter parameter (Scaldaferro et al. 2013) (Fig. 4.13).

4.8 Concluding Remarks and Future Prospects

This chapter compiles all the chromosome features that have been studied in the American genus *Capsicum* until now, also included DNA content data. This is an innovative approach since the genus is treated based on the last phylogeny from Carrizo García et al. (2016), which allowed to relate chromosome similarities within each clade, and chromosome diversity among different clades. Our group continues working on the genus, and new sequencing technologies (Harrison and Kidner 2011; Macas et al. 2011; Buggs et al. 2012; Egan et al. 2012) would facilitate comprehensive studies of *Capsicum* genome. (Harrison and Kidner 2011; Macas et al. 2011; Buggs et al. 2012; Egan et al. 2012).

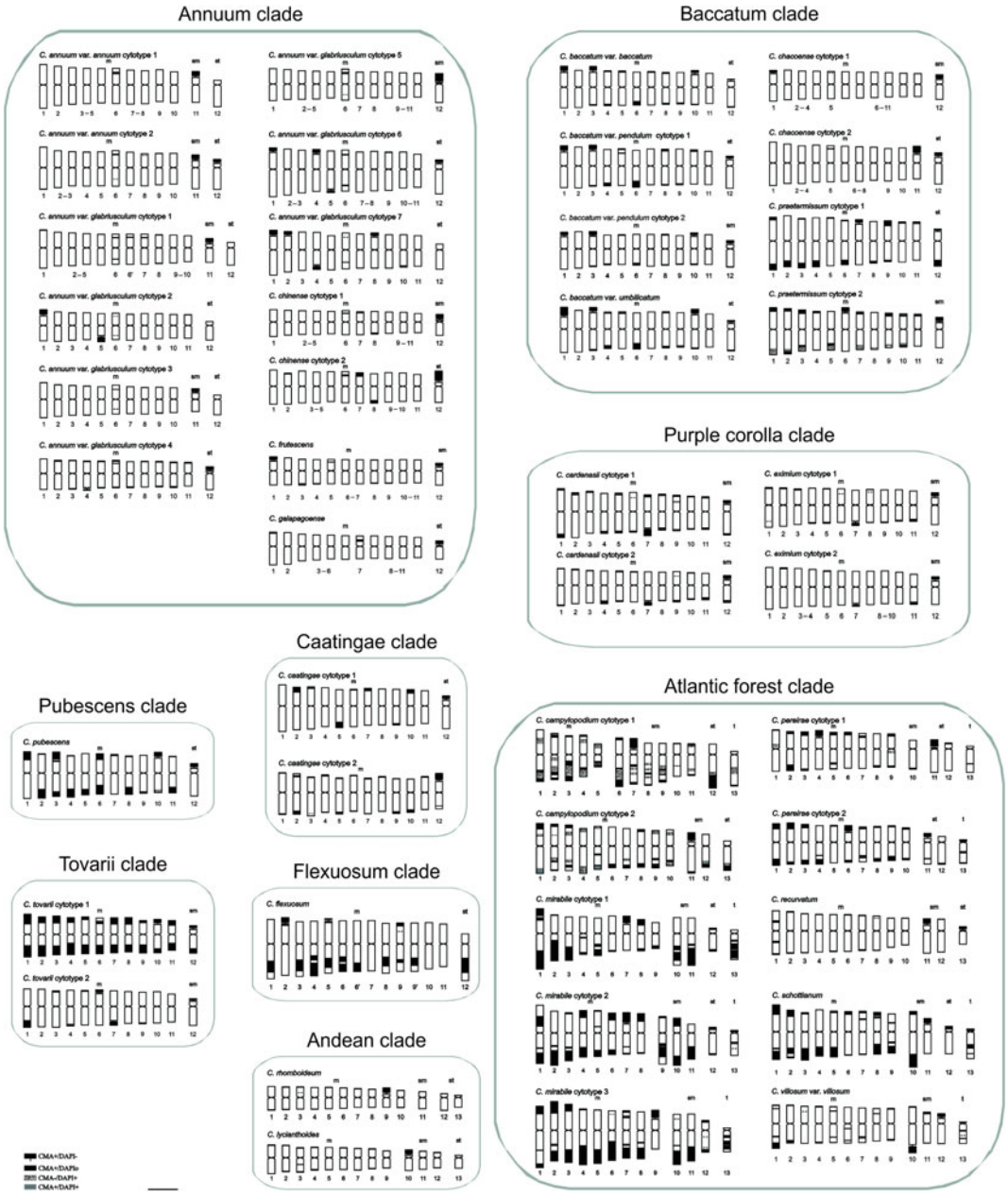


Fig. 4.11 Ideograms of *Capsicum* taxa showing heterochromatic fluorochrome banding pattern after staining with CMA/DAPI. Solid black blocks indicate CMA+/DAPIo or bands of homogeneous aspect; solid gray blocks indicate DAPI+/CMA− bands of homogeneous aspect; spotted blocks indicate CMA+/DAPI+ bands of mottled appearance. Active *45S* rDNA sites are indicated by a constriction. Chromosomes that have the same number on the ideogram are not necessarily homologous for the different taxa. In each ideogram, chromosomes with similar measures without markers were grouped. *m*, metacentric; *sm*, submetacentric; *st*, subtelocentric; *t*, telocentric. Scale bar = 5 μm

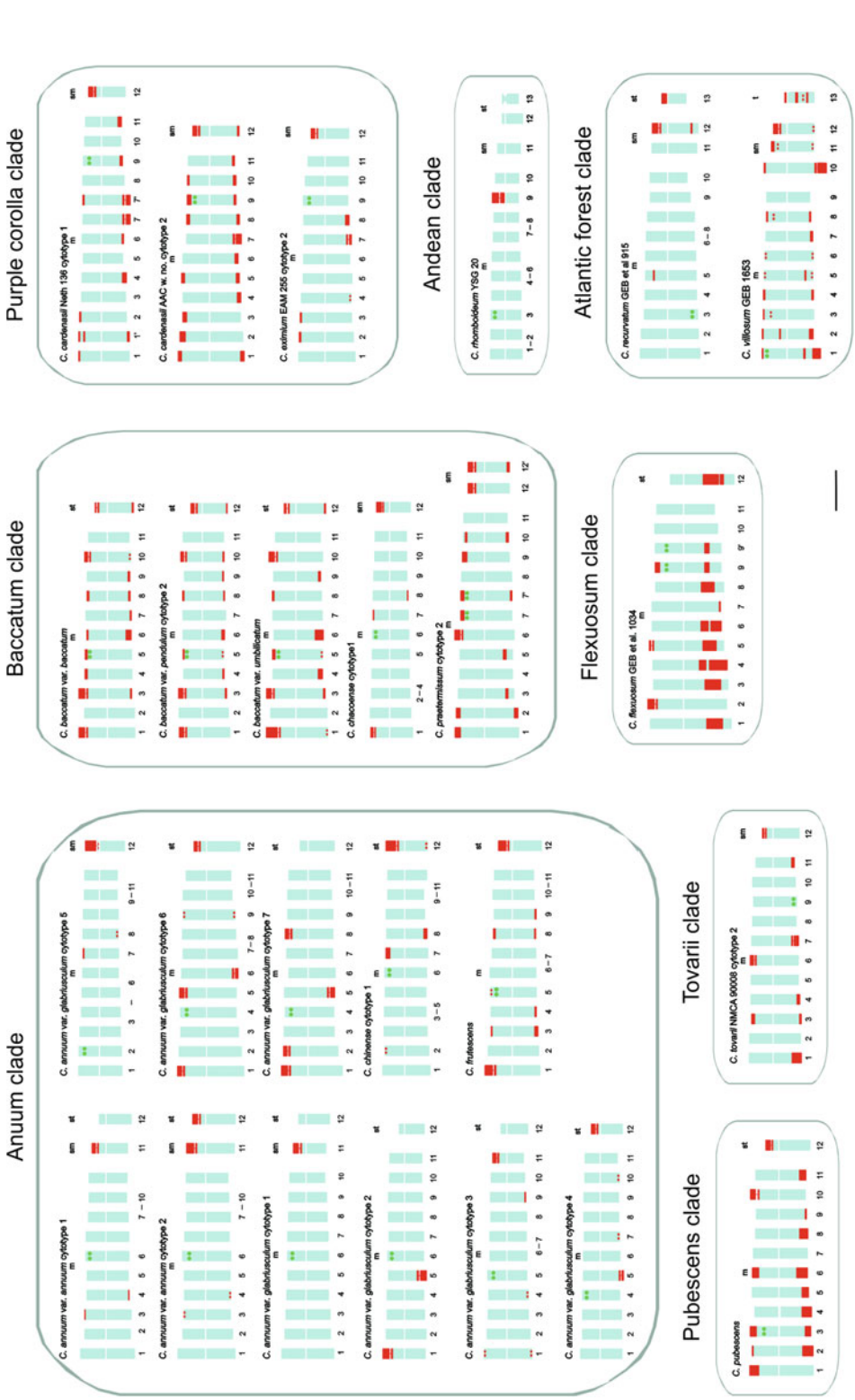


Fig. 4.12 Ideograms of *Capsicum* taxa showing the distribution of *45S* (red blocks) and *5S* (green circles) rDNA loci. Euchromatic regions appear in light blue. Active *45S* rDNA sites are indicated by a constriction. Chromosomes that have the same number on the ideogram are not necessarily homologous for the different taxa. In each ideogram, chromosomes with similar measures without markers were grouped. *m*, metacentric; *sm*, submetacentric; *st*, subtelocentric; *t*, telocentric. Some ideograms are extracted from Scaldaferro et al. (2016). Scale bar = 5 μm

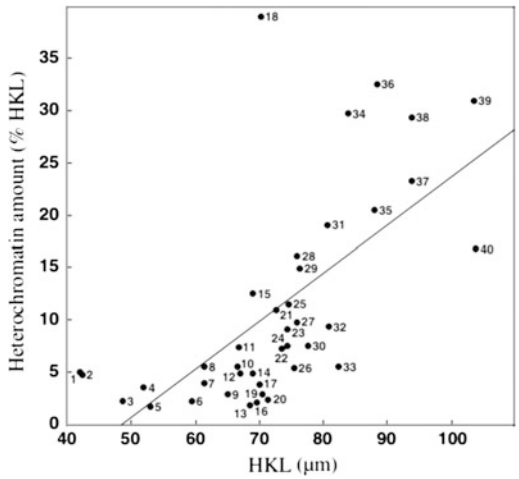


Fig. 4.13 Relationship between haploid karyotype length (HKL) and heterochromatin amount (expressed as percentage of HKL) in *Capsicum*. **1** *C. rhomboideum*; **2** *C. rhomboideum*; **3** *C. galapagoense*; **4** *C. annuum* var. *glabriusculum* cytotype 2; **5** *C. chacoense* cytotype 1; **6** *C. annuum* var. *glabriusculum* cytotype 1; **7** and **8** *C. chinense* cytotype 1 and 2, respectively; **9** *C. chacoense* cytotype 1; **10** *C. frutescens*; **11** *C. baccatum* var. *baccatum*; **12** *C. tovarii* cytotype 2; **13** *C. annuum* var. *annuum* cytotype 1; **14** *C. eximium* cytotype 1; **15** *C. cardenasii* cytotype 1; **16** *C. eximium* cytotype 2; **17** *C. annuum* var. *glabriusculum* cytotype 7; **18** *C. tovarii* cytotype 1; **19** *C. annuum* var. *annuum* cytotype 2; **20** *C. chacoense* cytotype 2; **21** *C. praetermissum* cytotype 1; **22** and **24** *C. baccatum* var. *pendulum* cytotype 1 and 2; **23** *C. baccatum* var. *umbilicatum*; **25** *C. pereirae* cytotype 1; **26** *C. recurvatum*; **27** *C. villosum*; **28** *C. pereirae* cytotype 2; **29** *C. praetermissum* cytotype 2; **30** *C. caatingae* cytotype 2; **31** *C. pubescens*; **32** *C. cardenasii* cytotype 2; **33** *C. caatingae* cytotype 1; **34**, **38,** and **39** *C. mirabile* cytotype 1, 2 and 3, respectively; **35** and **36** *C. campylopodium* cytotype 1 and 2, respectively; **37** *C. schottianum*; **40** *C. flexuosum.* The line (r = 0.65; P\0.0001) represents the correlation among data. Data from Moscone et al. (2007) and Scaldaferro et al. (2013)

Therefore, with the rapidly advancing sequencing technology and cytogenetic analysis, we will gain knowledge that could be compared with data from this chapter.

References

Baeza C, Schrader O (2005) Comparative karyotype analysis in *Haplopappus* Cass. and *Grindelia* Willd. (Asteraceae) by double FISH with rRNA specific genes. Plant Syst Evol 251:161–172

Barboza GE, Bianchetti LB (2005) Three new species of *Capsicum* (Solanaceae) and a key to the wild species from Brazil. Syst Bot 30(4):863–871

Barboza GE, Agra MF, Romero MV, Scaldaferro MA, Moscone EA (2011) New endemic species of *Capsicum* (Solanaceae) from the Brazilian caatinga: comparison with the re-circumscribed *C. parvifolium*. Syst Bot 36(3):768–781

Battaglia E (1955) Chromosome morphology and terminology. Caryologia 8:179–187

Bennett MD, Leitch IJ (1995) Nuclear DNA amounts in angiosperms. Ann Bot 76:85–91

Berg C, Greilhuber J (1993) Cold-sensitive chromosome regions and heterochromatin in *Cestrum* (Solanaceae): *C. srtigillatum*, *C. fasciculatum*, and *C. elegans*. Plant Syst Evol 185:133–151

Bloom SE, Goodpasture C (1976) An improved technique for selective silver staining of nucleolar organizer regions in human chromosomes. Hum Genet 34:199–206. https://doi.org/10.1007/BF00278889 PMID:63440

Bosland PW (1996) *Capsicums*: innovative use of an ancient crop. In: Janick J (ed) Progress in new crops. ASHS Press, Arlington, pp 479–487

Bosland PW, Votava EJ (2000) Peppers: vegetable and spice *Capsicums*. Crops production science in horticulture 12. CAB Intl Publishing, Wallingford, UK, p 204

Buggs RJA, Chamala S, Wu W, Tate JA, Schnable PS, Soltis DE, Soltis PS, Barbazuk WB (2012) Rapid, repeated, and clustered loss of duplicate genes in allopolyploid plant populations of independent origin. Curr Biol 22:248–252

Burger EC, Knälmann M (1980) Koinzidenz von Feulgen-Achromasie, in situ hybridisierung und silberbandenfärbung in vier nukleolusorganisatoren von *Vicia sativa*. Eur J Cell Biol 21:313–318

Carrizo García C, Barfuss MHJ, Sehr EM Barboza GE, Samuel R, Moscone EA, Ehrendorfer F (2016) Phylogenetic relationships, diversification and expansion of chili peppers (*Capsicum*, Solanaceae). Ann Bot 118:35–51

Datson PM, Murray BG (2006) Ribosomal DNA locus evolution in *Nemesia*: transposition rather than structural rearrangement as the key mechanism? Chromosome Res 14:845–857

Dubcovsky J, Dvorák J (1995) Ribosomal RNA multigene loci: nomads of the triticeae genomes. Genetics 140:1367–1377

Egan AN, Schlueter J, Spooner DM (2012) Applications of next-generation sequencing in plant biology. Am J Bot 99:175–185

Eshbaugh WH (1980) The taxonomy of the genus *Capsicum* (Solanaceae). Phytologia 47(3):153–165

Eshbaugh WH (1993) History and exploitation of a serendipitous new crop discovery. In: Janick J, Simon JE (eds) New crops. Wiley, New York, pp 132–139

Flavell RB (1986) Repetitive DNA and chromosome evolution in plants. Philos T R Soc Lon B 312: 227–242

García S, Garnatje T, McArthur ED, Pellicer J, Siljak-Yakovlev S, Vallès J (2009) Ribosomal DNA,

heterochromatin, and correlation with genome size in diploid and polyploid North American endemic sagebrushes (*Artemisia*, Asteraceae). Genome 52:1012–1024
Greilhuber J (1995) Chromosomes of the monocotyledons (general aspects). In: Rudall PJ, Cribb PJ, Cutler DF, Humphries CJ (eds) Monocotyledons: systematics and evolution. Royal Botanic Gardens, Kew, pp 379–414
Hall KJ, Parker JS (1995) Stable chromosome fission associated with rDNA mobility. Chromosome Res 3:417–422
Harrison N, Kidner CA (2011) Next-generation sequencing and systematics: what can a billion base pairs of DNA sequence data do for you? Taxon 60:1552–1566
Hizume M, Tanaka A, Shigematsu H (1982) Detection of nucleolar organizing regions in the chromosomes of *Nigella damacena* [i.e. *damascena*]. Experientia 38:238–239
Hunziker AT (1971) Estudios sobre Solanaceae. VII. Contribución al conocimiento de *Capsicum* y géneros afines (*Witheringia; Acnistus, Aihenaea,* etc.). Tercera parte. Kurtziana 6:241–259
Hunziker AT (2001) *Genera solanacearum*: the genera of Solanaceae illustrated, arranged according to a new system. ARG Gantner Verlag K-G, Liechtenstein
International Board for Plant Genetic Resources (IBPGR) (1983) Genetic resources of *Capsicum*, a global plan of action. IBPGR Executive Secretariat, Rome, Italy
Kodama Y, Yoshida MC, Sasaki M (1980) An improved silver staining technique for nucleolus organizer regions by using nylon cloth. Jap J Hum Genet 25:229–233. https://doi.org/10.1007/BF01997700
Kovarik A, Dadejova M, Lim YK, Chase MW, Clarkson JJ, Knapp S, Leitch AR (2008) Evolution of rDNA in Nicotiana allopolyploids: a potential link between rDNA homogenization and epigenetics. Ann Bot 101:815–823
Lefebvre V, Pflieger S, Thabuis A, Caranta C, Blattes A, Chauvet J-C, Daubeze AM, Palloix A (2002) Towards the saturation of the pepper linkage map by alignment of three intraspecific maps including known-function genes. Genome 45:839–854
Linde-Laursen IB (1984) Nucleolus organizer polymorphism in barley, *Hordeurn vulgare* L. Hereditas 100:33–43
Lippert LF, Smith PG, Bergh BO (1966) Cytogenetics of the vegetable crops. Garden pepper, *Capsicum* sp. Bot Rev (Lancaster) 32:24–55
Livingstone KD, Lackney VK, Blauth JR, van Wijk R, Jahn MK (1999) Genome mapping in *Capsicum* and the evolution of genome structure in the Solanaceae. Genetics 152:1183–1202
Macas J, Kejnovský E, Neumann P, Novak P, Koblížková A, Vyskot B (2011) Next generation sequencing-based analysis of repetitive DNA in the model dioceous plant *Silene latifolia*. PLoS ONE 6:e27335
Moscone EA (1990) Chromosome studies on *Capsicum* (*Solanaceae*) I. Karyotype analysis in *C. chacoënse*. Brittonia 42:147–154
Moscone EA (1993) Estudios cromosómicos en *Capsicum* (*Solanaceae*) II. Análisis cariotípico en *C. parvifolium* y *C. annuum* var. *annuum*. Kurtziana 22:9–18
Moscone EA (1999) Análisis cariotípico en *Capsicum baccatum* var. *umbilicatum* (*Solanaceae*) mediante bandeos AgNOR y de fluorescencia. Kurtziana 27:225–232
Moscone EA, Lambrou M, Hunziker AT, Ehrendorfer F (1993) Giemsa C-banded karyotypes in *Capsicum* (*Solanaceae*). Plant Syst Evol 186:213–229
Moscone EA, Loidl J, Ehrendorfer F, Hunziker AT (1995) Analysis of active nucleolus organizing regions in *Capsicum* (*Solanaceae*) by silver staining. Am J Bot 82:276–287
Moscone EA, Lambrou M, Ehrendorfer F (1996a) Fluorescent chromosome banding in the cultivated species of *Capsicum* (*Solanaceae*). Plant Syst Evol 202:37–63
Moscone EA, Matzke MA, Matzke AJM (1996b) The use of combined FISH/GISH in conjunction with DAPI counterstaining to identify chromosomes containing transgene inserts in amphidiploid tobacco. Chromosoma 105:231–236
Moscone EA, Klein F, Lambrou M, Fuchs J, Schweizer D (1999) Quantitative karyotyping and dual-color FISH mapping of *5S* and *18S-25S* rDNA probes in the cultivated *Phaseolus* species (*Leguminosae*). Genome 42:1224–1233
Moscone EA, Baranyi M, Ebert I, Greilhuber J, Ehrendorfer F, Hunziker AT (2003) Analysis of nuclear DNA content in *Capsicum* (*Solanaceae*) by flow cytometry and Feulgen densitometry. Ann Bot 92:21–29
Moscone EA, Scaldaferro MA, Grabiele M, Cecchini NM, Sanchez García Y, Jarret R, Daviña JR, Ducasse DA, Barboza GE, Ehrendorfer F (2007) The evolution of chili peppers (*Capsicum* – Solanaceae): a cytogenetic perspective. Proc PAA/Solanaceae Conf Acta Hortic 745:137–169 (International Society for Horticultural Science, Leuvem, Belgium, ISSN 0567-7572)
Olmstead RG, Bohs L, Migid HA, Santiago-Valentin E, Garcia VF, Collier SM (2008) A molecular phylogeny of Solanaceae. Taxon 57:1159–1181
Paran I, van der Voort JR, Lefebvre V, Jahn M, Landry L, van Schriek M, Tanyolac B, Caranta C, Ben Chaim A, Livingstone K, Palloix A, Peleman J (2004). An integrated genetic linkage map of pepper (*Capsicum* spp.). Mol Breed 13:251–261
Park Y-K, Park K-C, Park C-H, Kim N-S (2000) Chromosomal localization and sequence variation of *5S* rRNA gene in five *Capsicum* species. Mol Cells 10:18–24
Park M, Park J, Kim S, Kwon J-K, Park HM, Bae IH, Yang T-J, Lee Y-H, Kang B-C, Choi D (2012) Evolution of the large genome in *Capsicum annuum* occurred through accumulation of single-type long terminal repeat retrotransposons and their derivatives. Plant J https://doi.org/10.1111/j.1365-313x.2011.04851.x
Perry DR, Zarrillo S, Holst I, Pearsall DM, Piperno DR, Berman MJ, Cooke RG, Rademaker K, Ranere AJ, Raymond JS, Sandweiss DH, Scaramelli F, Tarble K, Zeidler JA (2007) Starch fossils and the domestication and dispersal of chili peppers (*Capsicum* spp. L.) in the Americas. Science 315:986–988

Pickersgill B (1971) Relationships between weedy and cultivated forms in some species of chili peppers (genus *Capsicum*). Evolution 25:683–691

Pickersgill B (1977) Taxonomy and the origin and evolution of cultivated plants in the New World. Nature 268:591–595

Pickersgill B (1984) Migration of chili peppers, *Capsicum* spp, in the Americas. In: Store P (ed) Pré-Columbian plant migration. Papers of the Peabody Museum of archaelogy and ethnology 76. Harvard University Press, Cambridge, pp 105–123

Pickersgill B (1991) Cytogenetics and evolution of *Capsicum* L. In: Tsuchiya T, Gupta PK (eds) Chromosome engineering in plants: genetics, breeding, evolution. Part B. Elsevier, Amsterdam, pp 139–160

Pickersgill B (1997) Genetic resources and breeding of *Capsicum* spp. Euphytica 96:129–133

Piperno DR (2011) The origins of plant cultivation and domestication in the new world tropics. Patterns, process, and new developments. Curr Anthropol 52:S453–470 https://doi.org/10.1086/659998 (The Wenner-Gren Foundation for Anthropological Research)

Pozzobon MT, Schifino-Wittmann MT, Bianchetti LB (2006) Chromosome numbers in wild and semidomesticated Brazilian *Capsicum* L. (Solanaceae) species: do $x = 12$ and $x = 13$ represent two evolutionary lines? Bot J Linn Soc 151:259–269

Qin C, Yu (2014) Whole-genome sequencing of cultivated and wild peppers provides insights into *Capsicum* domestication and specialization. Proc Nat Acad Sci USA 111(14):5135–5140

Raina SN, Bisht MS (1988) DNA amounts and chromatin compactness in *Vicia*. Genetica 77:65–77

Raina SN, Khoshoo TN (1971) Cytogenetics of tropical bulbous ornamentals III: mitotic mosaicism in 3x *Crinum augustum*. Theor Appl Genet 41:375–378

Raina SN, Mukai Y (1999) Detection of a variable number of *18S–5.8S–26S* and *5S* ribosomal DNA loci by fluorescence in situ hybridisation in diploid and tetraploid *Arachis* species. Genome 42:52–59

Raskina O, Belyayev A, Nevo E (2004) Activity of the En/Spm-like transposons in meiosis as a base for chromosome repatterning in a small, isolated, peripheral population of *Aegilops speltoides* Tausch. Chromosome Res 12:153–161

Romero-da Cruz MV, Urdampilleta JD, Forni Martins ER, Moscone EA (2016) Cytogenetic markers for the characterization of *Capsicum annuum* L. cultivars. Plant Biosyst https://doi.org/10.1080/11263504.2015.1103798

Särkinen T, Bohs L, Olmstead RG, Knapp S (2013) A phylogenetic framework for evolutionary study of the nightshades (Solanaceae): a dated 1000-tip tree. BMC Evol Biol 13:214

Scaldaferro MA, Seijo GJ, Acosta MC, Barboza GE, Ducasse DA, Moscone EA (2006) Genomic characterization of the germplasm in peppers (*Capsicum*—Solanaceae) by fluorescent in situ hybridization. Plant Sci 43(4):291–297 (National Centre for Agrarian Sciences, Sofia, Bulgaria; ISSN 0568-465X)

Scaldaferro MA, Barboza GE, Moscone EA (2011) Estudios citogenéticos en especies andinas silvestres de *Capsicum* (Solanaceae). In: Sociedad Argentina de Genética (ed) XL Congreso Argentino de Genética, III Simposio Latinoamericano de Citogenética y Evolución, I Jornadas Regionales SAG-NEA. Corrientes, Argentina, Sept 2011. J Basic Appl Genet XLI:31

Scaldaferro MA, Grabiele M, Moscone EA (2013) Heterochromatin type, amount and distribution in wild species of chili peppers (*Capsicum*-Solanaceae). Genet Res Crop Evol 60(2):693–709. https://doi.org/10.1007/s10722-012-9867-x

Scaldaferro MA, Romero da Cruz MV, Cecchini NM, Moscone EA (2016) FISH and AgNor-mapping of the *45S* and *5S* rRNA genes in wild and cultivated *Capsicum* species (Solananceae). Genome 59:95–113. https://doi.org/10.1139/gen-2015-0099

Scaldaferro MA, Barboza GE, Acosta MC (2018) Evolutionary history of the chili pepper *Capsicum baccatum* L. (Solanaceae): domestication in South America and natural diversification in the Seasonally Dry Tropical Forests. Biol J Linn Soc 124(3):466–478 https://doi.org/10.1093/biolinnean/bly062

Sharma S, Raina SN (2005) Organization and evolution of highly repeated satellite DNA sequences in plant chromosomes. Cytogenet Genome Res 109:15–26

Sinclair JH, Brown DD (1971) Retention of common nucleotide sequences in the ribosomal deoxyribonucleic acid of eukaryotes and some of their physical characteristics. Biochemistry 10:2761–2769

Stewart C Jr, Mazourek M, Stellari GM, O'Connell M, Jahn M (2007) Genetic control of pungency in *C. chinense* via the *Pun1* locus. J Exp Bot 58(5):979–991

Stewart C Jr, Kang B-C, Liu K, Mazourek M, Moore SL, Yoo EY, Kim B-D, Paran I, Jahn MM (2005) The *Pun1* gene for pungency in pepper encodes a putative acyltransferase. Plant J 42:675–688

Tewksbury JJ, Manchego C, Haak DC, Levey DJ (2006) Where did the chili get its spice? Biogeography of capsaicinoid production in ancestral wild chili species. J Chem Ecol 32:547–564

Tong N, Bosland PW (2003) Observations on interspecific compatibility and meiotic chromosome behavior of *Capsicum buforum* and *C. lanceolatum*. Genet Res Crop Evol 50:193–199

Xu YH, Yang F, Cheng YL, Ma L, Wang JB, Li LJ (2007) Comparative analysis of rDNA distribution in metaphase chromosomes of Cucurbitaceae species. Hereditas 29:614–620

5 Development and Evolution of Molecular Markers and Genetic Maps in *Capsicum* Species

Jundae Lee

Abstract

Capsicum species including six main cultivated species, *C. annuum*, *C. frutescens*, *C. chinense*, *C. baccatum*, *C. pubescens,* and *C. assamicum*, commonly known as chili peppers, are economically important crops in the world. Peppers are used as fresh vegetables, spices, pigments, and medical supplies. Over the four decades, a number of molecular marker techniques have been developed to analyze variations on DNA sequences of the genome in biological organisms. For plant breeding, molecular markers can substantially improve selection efficiency and reduce breeding time compared to conventional breeding. Genetic linkage mapping is an important basic tool for localizing gene(s) that are associated with important horticultural traits, marker-assisted selection, comparative mapping, physical mapping, and map-based cloning of the gene of interest. Recently, genetic linkage mapping has become easier owing to the advent of next-generation sequencing technology and its various applicative technologies. Here in this chapter, we reviewed the development and evolution of molecular markers and genetic maps of *Capsicum* spp. in which pepper researchers are interested.

J. Lee (✉)
Department of Horticulture, Chonbuk National University, Jeonju 54896, Republic of Korea
e-mail: ajfall@jbnu.ac.kr

5.1 Introduction

Molecular markers, which can be classified into biochemical and DNA markers, indicate a visible phenotype or fragment of DNA that is associated with a certain location within the genome (Kumar 1999). Of them, DNA marker is the most important marker because it is the most widely used. The well-known techniques for DNA markers include restriction fragment length polymorphisms (RFLPs; Botstein et al. 1980), random amplified polymorphic DNAs (RAPDs; Williams et al. 1990), cleaved amplified polymorphic sequences (CAPSs; Akopyanz et al. 1992), sequence-characterized amplified regions (SCARs; Paran and Michelmore 1993), amplified fragment length polymorphisms (AFLPs; Vos et al. 1995), simple sequence repeats (SSRs; Hearne et al. 1992), and single-nucleotide polymorphisms (SNPs; Wang et al. 1998). DNA markers can be applied in construction of genetic linkage maps, comparative mapping analysis, understanding germplasm relationships, tagging economically important genes, marker-assisted selection (MAS; Mohan et al. 1997; Ribaut and Hoisington 1998), and map-based cloning of genes (Kumar 1999). Several articles are well reviewed on molecular markers (Mohan et al. 1997; Kumar 1999; Kesawat and Das 2009; Jiang 2013).

Genetic mapping is an important method for positioning genes of interest in genome as well as identifying quantitative trait loci (QTLs)

N. Ramchiary and C. Kole (eds.), *The Capsicum Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-319-97217-6_5

responsible for natural phenotypic variation (Xu et al. 2017). Recently, rapid genome-wide SNP detection techniques using next-generation sequencing (NGS; Shendure and Ji 2008), including genotyping-by-sequencing (GBS; Huang et al. 2009), restriction site-associated DNA sequencing (RAD-seq; Baird et al. 2008), and specific-locus amplified fragment sequencing (SLAF-seq; Sun et al. 2013), shortened the time required for genetic map construction.

In this chapter, I will review the development and evolution of various molecular marker techniques, their use for *Capsicum* researches, development of the interspecific and intraspecific genetic maps of *Capsicum* spp., and their application to QTL detection and marker-assisted backcrossing (MABC).

5.2 Development and Evolution of Markers in *Capsicum* Species

5.2.1 Biochemical Marker

Biochemical markers, also known as protein markers or isozyme markers, can be examined by protein electrophoresis to identify the alleles producing isozymes (Kumar 1999). The most commonly used biochemical markers are isozymes which are variant forms of the same enzyme (Vodenicharova 1989). Isozyme markers reveal differences in the amino acid sequence and function as codominant markers. However, their use is limited due to their limited number and various posttranslational modifications.

Three isozyme markers, *Gpi-2*, *Idh-1*, and *Pgm-2*, were mapped on an interspecific genetic map of *Capsicum annuum* 'NuMex RNaky' × *Capsicum chinense* 'PI159234,' and two isozyme markers, *Idh-1* and *Pgm*, were used to compare the genetic maps between pepper and tomato (Livingstone et al. 1999). The isozyme marker *Mnr-1* corresponding to the first region of the menadione reductase (1.6.99.2) was mapped on an intraspecific genetic map of *C. annuum* 'Perennial' × 'Yolo Wonder' (Lefebvre et al. 2002).

5.2.2 DNA Marker

DNA markers, which are based on the difference of DNA sequences, can be classified into two categories, hybridization-based and polymerase chain reaction (PCR)-based markers, depending on the method to detect polymorphisms (Kumar 1999). Hybridization-based markers include restriction fragment length polymorphism (RFLP; Botstein et al. 1980) and variable number of tandem repeats (VNTRs; Nakamura et al. 1987) that generally use probes and Southern blot analysis. PCR-based markers can be more classified into random and specific, depending on the type of primer used. Random PCR-based markers include random amplified polymorphic DNA (RAPD; Williams et al. 1990), arbitrarily primed PCR (AP-PCR; Welsh and McClelland 1990), DNA amplification fingerprinting (DAF; Caetano-Anollés et al. 1991), inter-simple sequence repeats (ISSRs; Gupta et al. 1994; Zietkiewicz et al. 1994), and amplified fragment length polymorphism (AFLP; Vos et al. 1995). Specific PCR-based markers include simple sequence repeat (SSR; Hearne et al. 1992), cleaved amplified polymorphic sequence (CAPS; Akopyanz et al. 1992; Konieczny and Ausubel 1993; Lyamichev et al. 1993), sequence-characterized amplified region (SCAR; Paran and Michelmore 1993), and single-nucleotide polymorphism (SNP; Wang et al. 1998). The characteristics of these markers are given in Table 5.1.

5.2.2.1 Restriction Fragment Length Polymorphism (RFLP)

RFLP analysis follows next steps: DNA extraction, restriction enzyme digestion, agarose gel electrophoresis, Southern blotting, hybridization with radioactive probes such as random genomic clones and cDNA clones, and autoradiography (Botstein et al. 1980). RFLP reveals the differences of fragment length hybridized with probes, which are resulted from the presence or the absence of a cleavage site and insertion or deletion of DNA sequences within a fragment (Botstein et al. 1980). The major strength of RFLP markers is their high reproducibility, codominant

Table 5.1 Overview of the characteristics of some important DNA markers

Classification	Marker	Genomic abundance[a]	Locus specificity[b]	Codominance[b]	Reproducibility[a]	Required DNA quantity[a]	Labor intensity[a]	Cost per analysis[c]	First reported references
Hybridization-based marker									
	RFLP	H	Y	Y	H	H	H	H	Botstein et al. (1980)
	VNTR	M	N/Y	N/Y	H	H	H	H	Nakamura et al. (1987)
PCR-based marker									
Random primers	RAPD	H	N	N	L	L	L	L	Williams et al. (1990)
	ISSR	M-H	N	N	M-H	L	L	L	Gupta et al. (1994), Zietkiewicz et al. (1994)
	AFLP	H	N	N/Y	M-H	M	M	M	Vos et al. (1995)
Specific primers	SSR	H	Y	Y	H	L	L/M	L	Hearne et al. (1992)
	CAPS	L	Y	Y	H	L	L-M	L	Akopyanz et al. (1992)
	SCAR	L	Y	N/Y	H	L	L	L	Paran and Michelmore (1993)
	SNP	VH	Y	Y	H	L	L	L/VL	Wang et al. (1998)
	InDel	H	Y	Y	H	L	L	L/VL	Bhattramakki et al. (2002), Weber et al. (2002)

[a]*VH* very high, *H* high, *M* medium, *L* low
[b]*Y* yes, *N* no
[c]*H* high, *M* moderate, *L* low, *VL* very low

inheritance, and good transferability which can allow synteny studies (Kesawat and Das 2009).

A total of 85 RFLP markers were developed using tomato cDNA probes and mapped by using 46 F_2 individuals derived from the F_1 of *C. annuum* 'CA50' and *C. chinense* 'CA4' to compare the genetic maps between pepper and tomato (Tanksley et al. 1988). A total of 192 molecular markers including RFLPs and isozymes were used for constructing an interspecific pepper genetic map (Prince et al. 1993). A total of 85 markers including RFLP and RAPD covered approximately 820 cM of the integrated pepper linkage map (Lefebvre et al. 1995). Using pepper-derived probes, total 150 RFLP markers were developed and positioned on an interspecific F_2 linkage map of *C. annuum* 'TF68' *C. chinense* 'Habanero' (Kang et al. 2001). To analyze yield-related quantitative trait loci (QTLs) in pepper, 92 RFLP markers were used, resulting in detection of a total of 58 QTLs (Rao et al. 2003). To detect QTLs associated with *Phytophthora capsici* resistance, a RFLP-based linkage map was constructed using 100 F_2 individuals from a cross between *C. annuum* 'CM334' and *C. annuum* 'Chilsungcho' and bacterial artificial chromosome (BAC)-derived markers were developed from RFLP linked to the resistant trait (Kim et al. 2008c).

5.2.2.2 Minisatellites: Variable Number of Tandem Repeat (VNTR)

VNTR analysis is almost the same with RFLP analysis except using probes with minisatellite sequences (Kumar 1999). The polymorphism of VNTR is due to the differences in the number of repeats (Nakamura et al. 1987). No VNTR markers were used for the *Capsicum* studies.

5.2.2.3 Random Amplified Polymorphic DNA (RAPD)

Multiple arbitrary amplicon profiling (MAAP) techniques include random amplified polymorphic DNA (RAPD; Williams et al. 1990), arbitrary primed PCR (AP-PCR; Welsh and McClelland 1990), and DNA amplification fingerprinting (DAF; Caetano-Anollés et al. 1991), which are random PCR markers (Kumar 1999). RAPD markers use generally 10 bp synthetic primers of random sequence, while AP-PCR uses longer arbitrary primers than RAPDs and DAF uses shorter 5–8 bp primers to generate a larger number of fragments (Kesawat and Das 2009).

An integrated linkage map of pepper, including RFLP and RAPD markers, was constructed by alignment of three intraspecific linkage maps (*C. annuum*, Lefebvre et al. 1995). Screening with 400 RAPD primers along with bulked segregant analysis (BSA; Michelmore et al. 1991) allowed the identification of three QTLs for capsaicinoid content in *Capsicum* (Blum et al. 2003). A total of 122 RAPD markers were used for constructing an intraspecific linkage map of *C. annuum* (Sugita et al. 2005).

5.2.2.4 Inter-simple Sequence Repeat (ISSR)

ISSR, which uses microsatellites as primers, involves amplification of DNA segments present at a close distance in between two identical microsatellite repeat regions oriented in opposite directions (Gupta et al. 1994; Zietkiewicz et al. 1994). The primers can be either unanchored or more usually anchored at 3′ or 5′ end with 1–4 degenerate bases extended into the flanking sequence (Reddy et al. 2002). ISSRs have higher reproducibility due to the use of longer primers (16–25 bp) as compared to RAPD primers (10 bp, Reddy et al. 2002).

A total of 17 ISSR markers were used for differentiating the four disputed chili pepper samples (Kumar et al. 2001). Five ISSR primers amplified 204 reproducible bands of which 139 were polymorphic and they were used for assessing the genetic relation to 13 *C. annuum* cultivars (Patel et al. 2011). Eight ISSR primers were used for analyzing genetic variability in six *Capsicum* species (Thul et al. 2012). A total of 219 ISSR clear and reproducible fragments generated with 13 ISSR primers were used to evaluate the effects of in vitro culture on genetic variation in Habanero pepper (*C. chinense* Jacq.; Bello-Bello et al. 2014). Using eight ISSR anchored primers, a total of 38 bands were obtained for assessment of inter- and intraspecific differentiation in two Serrano and two Jalapeno

cultivars of *C. annuum* and one cultivar of *Capsicum pubescens* (Ibarra-Torres et al. 2015). Total 85 ISSR markers were used to construct a genetic map of *Capsicum baccatum* (Moulin et al. 2015).

5.2.2.5 Amplified Fragment Length Polymorphism (AFLP)

The AFLP technique is based on the selective PCR amplification of restriction fragments from a total digested genomic DNA (Vos et al. 1995). AFLP analysis follows next steps: DNA extraction, double digestion with two different restriction enzymes (generally *Eco*RI and *Mse*I), adaptor ligation, pre-selective amplification, selective amplification, polyacrylamide gel electrophoresis (PAGE), and silver staining (Vos et al. 1995; Blears et al. 1998). The banding profiles result from variations in the restriction sites or in the intervening regions (Kesawat and Das 2009).

A total of 430 AFLP markers were used to construct a linkage map of pepper in an interspecific F_1 population derived from a cross between *C. annuum* 'TF68' and *C. chinense* 'Habanero' (Kang et al. 2001). A genetic linkage map of the sweet pepper was constructed by 382 AFLP markers in an intraspecific doubled haploid (DH) population (Sugita et al. 2005). A total of 175 AFLP markers were used for identifying QTLs associated with anthracnose resistance in an intraspecific F_2 population of a cross between *C. baccatum* var. *pendulum* and *C. baccatum* 'Golden-aji' (Kim et al. 2010). Using an introgression BC_1F_2 population made by interspecific crosses between *C. annuum* 'SP26' and *C. baccatum* 'PBC81,' a total of 197 AFLP markers were developed to identify QTLs for resistance to anthracnose caused by *Colletotrichum scovillei* and *C. dematium* (Lee et al. 2010b).

5.2.2.6 Simple Sequence Repeat (SSR)

Microsatellites are known as simple sequence repeat (SSR; Hearne et al. 1992), short tandem repeat (STR), and simple sequence length polymorphism (SSLP) and are the smallest class of simple repetitive DNA sequences (1–6 bp) (Kesawat and Das 2009). SSR markers result from variations on the number of tandem repeats of microsatellites (Hearne et al. 1992). Specific primers (generally 20–25 bp) in the flanking regions of microsatellite can be designed to amplify the microsatellite by PCR (Kesawat and Das 2009). SSR markers can be developed from genomic or genic microsatellite sequences. Expressed sequence tag-SSR (EST-SSR) markers can be easily developed by data mining for SSRs in EST databases (Kantety et al. 2002). Genomic SSRs are more polymorphic compared to genic SSRs (EST-SSRs) and are superior to fingerprinting or varietal identification studies, while genic SSRs are useful for assessment of functional diversity (Varshney et al. 2005).

Forty-six SSR markers were first placed on an interspecific linkage map of *C. annuum* 'TF68' and *C. chinense* 'Habanero' (Lee et al. 2004). In the same population, a total of 150 EST-SSRs were developed through in silico analysis of 10,232 non-redundant EST sequences (Yi et al. 2006). By sequencing 1873 clones derived from the genomic DNA libraries of *C. annuum* 'Manganji,' 106 new SSR markers were developed and mapped on an intraspecific linkage map of *C. annuum* 'Manganji' × 'Tongari' (Minamiyama et al. 2006). In an intraspecific genetic map of *C. annuum* 'YCM334' and 'Tean,' 101 EST-SSR markers were located after screening of total 1667 EST-SSR markers (Truong et al. 2010). Total 151 SSR markers were used to construct an intraspecific linkage map of *C. annuum* 'California Wonder' and 'LS2341' (Mimura et al. 2012). To construct a high-density linkage map of *C. annuum*, 1736 genomic SSR markers and 1344 EST-SSR markers were developed from 6528 clones and 13,003 sequences, respectively (Sugita et al. 2013). To map QTLs affecting the initiation of flower primordia in pepper, 95 SSR markers were validated and integrated into a genetic map of *C. annuum* 'BA3' and *Capsicum frutescens* 'YNXML' (Tan et al. 2015). A total of 113,500 in silico unique SSR loci were identified in nuclear genome of pepper using a homemade bioinformatics workflow, and as a preliminary study, 65 SSR markers were validated among a wide collection of 21 *Capsicum* genotypes

(Cheng et al. 2016b). To map QTLs for fruit length, 400 SSR markers were surveyed, but only 28 markers were mapped in an F_2 population derived from a cross of *C. annuum* 'FL201' and *Capsicum galapagoense* 'TC07245' (Arjun et al. 2018).

5.2.2.7 Cleaved Amplified Polymorphic Sequences (CAPSs)

CAPS analysis, referred to as PCR-restriction fragment length polymorphism (PCR-RFLP), follows next steps: PCR using specific 20–25 bp primers, digestion of the PCR products with a restriction enzyme, and agarose gel electrophoresis of the digested products (Akopyanz et al. 1992; Konieczny and Ausubel 1993; Lyamichev et al. 1993). The DNA fragment length polymorphisms of CAPS result from variation in the occurrence of restriction sites (Kesawat and Das 2009). Advantages of CAPS include the requirement of only low quantities of template DNA (50–100 ng per reaction) for PCR, the codominance of alleles, the high reproducibility, and easier procedure compared to RFLP due to no requiring Southern blot hybridization and radioactive detection steps (Kesawat and Das 2009). A derived cleaved amplified polymorphic sequence (dCAPS) marker is a modified method of CAPS technique where mismatches in a PCR primer are used to create a polymorphism based on the target mutation (Neff et al. 1998).

A CAPS marker was converted from the AFLP marker E41/M49-645 linked to the *Pvr4* resistance gene in *C. annuum* (Caranta et al. 1999). A CAPS marker $SCAC_{568}$ was developed from the $OPAC10_{593}$ RAPD marker linked to *Tsw* gene to assist selection of TSWV resistance in pepper (Moury et al. 2000). Three CAPS markers, Pvr1-S, pvr1-R1, and pvr1-R2, were developed to discriminate between $Pvr1^+$, *pvr1*, $pvr1^1$, and $pvr1^2$ alleles in *Capsicum* spp. (Yeam et al. 2005). The *Rf* locus was mapped by using eight AFLP markers, and of them, the AFRF8 marker was successfully converted to a CAPS marker AFLP8CAPS which was closest to *Rf* with a genetic distance of 1.8 cM (Kim et al. 2006). A CAPS marker PR-CAPS for partial restoration (*pr*) locus was developed from the AFLP marker E-AGC/M-GCA_{112} estimated at about 1.8 cM in genetic distance (Lee et al. 2008). Two CAPS markers, PmsM1-CAPS and PmsM2-CAPS, linked to the ms_1 gene on pepper chromosome 5 were developed (Lee et al. 2010a, 2011a). Two CAPS markers, GMSK-CAPS and GMS3-CAPS, were identified to cosegregate with the ms_k and ms_3 genes, respectively (Lee et al. 2010c, d). A major QTL *CaR12.2* for the resistance was found in an introgression BC_1F_2 population made by interspecific crosses between *C. annuum* 'SP26' (susceptible) and *C. baccatum* 'PBC81' (resistant), and the CaR12.2M1-CAPS marker closely linked to the major QTL *CaR12.2* was developed (Lee et al. 2010b, 2011b). The M3-CAPS marker tightly linked to the major QTL *Phyto.5.2* for resistance to *Phytophthora* root rot was developed using two segregating F_2 populations from a cross of 'Subicho' × 'CM334' and self-pollination of a commercial cultivar 'Dokyacheongcheong' (Lee et al. 2012b). A set of allele-specific markers of *L* locus, including L2-CAPS and L0nu-CAPS markers, was developed using five pepper differential hosts including *C. annuum* 'ECW' (L^0/L^0), *C. annuum* 'Tisana' (L^1/L^1), *C. annuum* 'CM334' (L^2/L^2), *C. chinense* 'PI159236' (L^3/L^3), and *Capsicum chacoense* 'PI260429' (L^4/L^4) (Lee et al. 2012a). A codominant CAPS marker, CL000081-0555, located 1.13 cM away from the *Me1* gene, was developed using an F_2 population of a cross between *C. annuum* 'AZN-1' (susceptible line) and 'PM217' (resistant inbred line derived from 'PI201234'; Uncu et al. 2015). Recently, a CAPS marker 16,830-CAPS, tightly linked to the *Me1* gene, was developed through a fine mapping approach and the *CA09g16830* gene was identified as a candidate gene for *Me1* (Wang et al. 2018).

5.2.2.8 Sequence-Characterized Amplified Region (SCAR)

SCARs are PCR-based markers that are identified by PCR amplification of genomic DNA with a pair of specific primers (Paran and Michelmore 1993). SCARs can be classified into two types, dominant and codominant, depending on inheritance pattern: Dominant SCARs result from the

presence or the absence of amplification of the band; codominant SCARs result from the length polymorphisms caused by insertion or deletion (Paran and Michelmore 1993). SCARs have higher reproducibility compared to RAPD due to the use of longer primers (22–24 bp) designed for specific amplification of a particular locus (Kesawat and Das 2009).

A SCAR marker $SCUBC19_{1432}$ linked to the *Pvr4* locus was developed using segregating progenies obtained by crossing a homozygous resistant ('Serrano Criollo de Morelos-334') with a homozygous susceptible ('Yolo Wonder') (Arnedo-Andrés et al. 2002). Three SCAR markers, $PMFR11_{269}$, $PMFR11_{283}$, and $PMFR21_{200}$, positioned at a distance of 4.0 cM from the L^3 locus, were converted from two RAPD markers, $E18_{272}$ and $E18_{286}$, which were developed by applying the bulked segregant analysis (BSA) method to two DH populations, K9-DH and K9/AC-DH, derived from F_1 hybrid 'K9' that harbors the L^3 gene derived from 'PI159236' (Sugita et al. 2004). Two markers, *atp6*-SCAR and *coxII* SCAR, have been developed to identify the CMS cytoplasm (Kim and Kim 2005). The D4 SCAR marker for the detection of *Phyto.5.2* and a major QTL for resistance to *P. capsici* were developed (Quirin et al. 2005). A SCAR marker L4SC340, which mapped 1.8 cM from the L^4 locus, was developed from an AFLP marker L4-c, which was identified by applying BSA-AFLP method to a near-isogenic BC_4F_2 population generated by using *C. chacoense* 'PI260429' (carrying the L^4 allele) as a resistant parent (Kim et al. 2008a). The presence of a third haplotype ($Rfls^{7701}$) of the sequence linked to the *Rf* gene was reported, and two codominant SCAR markers CaRf-M1 and CaRf-M2 were developed for discriminating between $Rfls^A$, $Rfls^B$, and $Rfls^{7701}$ (Min et al. 2008b). A codominant SCAR marker AFRF4 linked to the *Rf* locus with a genetic distance of 0.1 cM was developed (Min et al. 2009). A newly developed *Rf*-linked marker BAC13T7-SCAR was developed from the sequence of a tomato BAC clone containing three genes which are homologous to petunia *Rf* gene encoding a pentatricopeptide repeat (PPR) protein (Jo et al. 2010). A codominant SCAR marker PR-Bs3 was developed by designing primers to amplify the InDel region of *Bs3* and *bs3* promoters (Römer et al. 2010). A major QTL *CcR9* for the resistance of 'PBC81' to *Colletotrichum truncatum* was identified, and the CcR9M1-SCAR marker closely linked to the QTL *CcR9* was developed (Lee et al. 2010b, 2011b). A marker SCAR_P2 linked to the ms_8 locus on the lower arm of the pepper chromosome 4 was identified (Bartoszewski et al. 2012). A set of allele-specific markers of *L* locus, including L1-SCAR, L3-SCAR, L4-SCAR, and L0c-SCAR markers, was developed (Lee et al. 2012a). A codominant SCAR marker SA133_4 linked to the QTL P5 for *Phytophthora* resistance was developed (Truong et al. 2013). The SCAR_PM54 marker was identified to be fully consistent with artificial nematode (*Meloidogyne incognita* race 2) testing, correctly predicting resistant ('PM687', 'PM217,' and 'Carolina Cayenne') and susceptible ('Yolo Wonder B,' 'California Wonder 300,' and 'CM331') genotypes (Pinar et al. 2016). Two markers SCAR-InDel and SSR-HpmsE032 associated with resistance to *C. scovillei* were validated in two *C. annuum* anthracnose-resistant introgression lines, P_{R1} derived from 'PBC932' and P_{R2} derived from 'PBC80,' showing the selection efficiency of 77% when both markers were used together (Suwor et al. 2017). Recently, a novel powdery mildew resistance locus, *PMR1*, was identified on pepper chromosome 4 using two populations consisting of 102 'VK515' $F_{2:3}$ families and 80 'PM Singang' F_2 plants, and one SCAR marker (ZL1_1826) was developed to cosegregate with the *PMR1* locus (Jo et al. 2017).

5.2.2.9 Single-Nucleotide Polymorphism (SNP) and Insertion/Deletion (InDel)

A single-nucleotide polymorphism (SNP) is a single-nucleotide difference between two DNA sequences or individuals (Wang et al. 1998), and an insertion/deletion (InDel) refers to an insertion or deletion of bases in the genome of an organism (Bhattramakki et al. 2002; Weber et al. 2002).

SNPs and InDels are highly abundant and distributed throughout the genome in plants (Kesawat and Das 2009). They are very useful tool for genetic mapping, marker-assisted breeding, and map-based cloning (Rafalski 2002; Kesawat and Das 2009). Over the past two decades, a number of different SNP genotyping methods have been developed. Various SNP genotyping assays can be classified by a combination of one of the sample preparation techniques (allele-specific hybridization, primer extension, oligonucleotide ligation, and nuclease cleavage) and one of the analysis techniques (gel separation, array, mass spectrometry, and plate reader; Gut 2001). The well-known SNP genotyping methods include TaqMan assay (Livak 1999), allele-specific PCR (AS-PCR) with universal energy-transfer-labeled primers (Myakishev et al. 2001), and high-resolution melting (HRM) analysis (Wittwer et al. 2003; Liew et al. 2004).

TaqMan assay, also known as 5′ nuclease assay, can be used to discriminate alleles that differ from a single-nucleotide substitution, using a fluorogenic probe consisting of an oligonucleotide labeled with both a fluorescent reporter dye (generally FAM or TET) and a quencher dye (Livak 1999). Amplification of the probe-specific product causes cleavage of the probe, generating an increase in specific reporter fluorescence (Livak 1999). However, the biggest problem of this method is production cost of the specific probe required for each TaqMan assay.

AS-PCR with universal energy-transfer-labeled primers was developed for high-throughput SNP genotyping (Myakishev et al. 2001). The technique involves PCR amplification with two different tailed allele-specific primers that contain priming sites for universal energy-transfer-labeled primers. This method can solve the problem of TaqMan assay by using the same universal primers for all analyses (Myakishev et al. 2001). SNP-type assay (Wang et al. 2009) and kompetitive allele-specific PCR (KASP; Semagn et al. 2014) for SNP genotyping adopt this method and can be applied to high-throughput SNP genotyping analysis.

HRM analysis, a method that allows detecting polymorphism in double-stranded DNA by comparing profiles of melting curves, can be used for genotyping SNP, SSR, and InDel markers (Liew et al. 2004; Simko 2016). HRM markers are faster, simpler, and less expensive than other marker systems requiring gel separation or labeled probes because it is directly analyzed within a closed tube with the addition of fluorescent dyes such as LCGreen® Plus, SYTO® 9, EvaGreen®, LCGreen™ I, or SYBR® Green I before PCR amplification (Wittwer 2009).

A total of 40 SNP markers using AS-PCR analysis were developed for cultivar identification in *Capsicum* (Jung et al. 2010). To develop a SNP-based genetic map in an F_2 population derived from a cross of *C. annuum* 'NB1' × *C. chinense* 'Jolokia,' 116 SNP markers using HRM analysis were developed from SNPs identified from next-generation resequencing of parents (Lee et al. 2013). To construct an EST-based linkage map in the F_2 population (*C. annuum* 'NuMex RNaky' × *C. chinense* 'PI159234'), 48 EST-based SNPs markers were developed (Park et al. 2014). To develop an InDel-based linkage map of hot pepper (*C. annuum*), 251 InDel markers were developed (Li et al. 2015). To construct a SNP-based genetic linkage map of *C. baccatum*, a total of 395 HRM markers were developed based on SNPs identified by comparing genome sequences generated through next-generation resequencing of the parents, *C. baccatum* 'Golden-aji' and 'PI594137' (Lee et al. 2016).

5.2.2.10 High-Throughput SNP Genotyping Systems

Next-generation sequencing (NGS) technologies, including 454, Solexa, SOLiD, Polonator, and HeliScope, have had a great influence on biological studies by enabling faster and less expensive analysis of genomes and transcriptomes (Shendure and Ji 2008). NGS technologies have made it easy to detect genetic variations including SNPs and InDels and to develop DNA-based molecular markers in plant genetics and breeding (Varshney et al. 2009). Indeed,

SNP markers are increasingly becoming the go-to marker system because SNPs can be identified so easily through NGS technologies (Ganal et al. 2009; Kumar et al. 2012).

High-throughput SNP genotyping systems, including Illumina Infinium iSelect HD array (International HapMap Consortium 2005), Affymetrix Axiom array (International HapMap Consortium 2005), Douglas Array Tape (www.douglasscientific.com), Fluidigm dynamic arrays (Wang et al. 2009), restriction enzyme-based genotyping-by-sequencing (GBS; Huang et al. 2009), and amplicon sequencing (Bybee et al. 2011), are very useful for plant breeding (Thomson 2014).

Fluidigm dynamic arrays, a flexible, PCR-based SNP genotyping platform, include three formats for nanofluidic integrated fluid circuits (IFCs): 96 samples × 96 SNPs, 48 samples × 48 SNPs, and 192 samples × 24 SNPs (Wang et al. 2009). A 48.48 dynamic array yields 2304 data points with 48 samples and 48 markers, and 96.96 and 192.24 dynamic arrays yield 9216 and 4608 data points, respectively (Wang et al. 2009; Thomson 2014). The dynamic arrays can be used with three types of assays: TaqMan, KASP, or SNP-type assays (Wang et al. 2009; Thomson 2014).

Genotyping-by-sequencing (GBS), a genome-wide genotyping method that enables a rapid and inexpensive analysis of the whole genome using a multiplexed NGS technology, can be applied to various areas of plant genetics and breeding, including SNP discovery, high-density genetic mapping, QTL analysis, genome-wide association studies (GWASs), genomic selection (GS), and low-cost genomics-assisted breeding (GAB) (Deschamps et al. 2012; Poland and Rife 2012). Representative GBS methods include various following protocols: restriction association DNA sequencing (RAD-seq; Baird et al. 2008), genotyping-by-sequencing (GBS; Huang et al. 2009), multiplex shotgun genotyping (MSG; Andolfatto et al. 2011), double-digested RAD-seq (Peterson et al. 2012), double-digested GBS (Poland et al. 2012), sequence-based genotyping (SBG; Truong et al. 2012), restriction enzyme sequence comparative analysis (Monson-Miller et al. 2012), ion torrent GBS (Mascher et al. 2013), restriction fragment sequencing (REST-seq; Stolle and Moritz 2013), and specific-locus amplified fragment sequencing (SLAF-seq; Sun et al. 2013) (Kim et al. 2016).

These techniques can be used in *Capsicum* spp. due to reports of whole-genome sequences of *C. annuum* 'CM334' (Kim et al. 2014), 'Zunla-1' (Qin et al. 2014), *C. annuum* var. *glabriusculum* 'Chiltepin' (Qin et al. 2014), *C. chinense* 'PI159236' (Kim et al. 2014, 2017b), and *C. baccatum* 'PBC81' (Kim et al. 2017b).

For marker-assisted backcrossing (MABC) in *Capsicum*, 412 SNPs evenly distributed on each chromosome were used to develop locus-specific markers for the Fluidigm® EP1™ genotyping system as a high-throughput SNP genotyping method (Kang et al. 2014). GBS analysis was used to detect QTLs conferring resistance to the *cucumber mosaic virus* P1 (CMV_{P1}) strain in pepper (Eun et al. 2016). An ultra-high-density bin map containing 2578 bins was constructed to identify QTLs for horticultural traits in *C. annuum* through next-generation resequencing analysis (Han et al. 2016). With an Illumina Infinium iSelect SNP array (pepper *Cap*SNP15K array), a high-density interspecific genetic map containing 5569 SNPs was constructed to analyze genetic diversity of 339 pepper elite/landrace lines (Cheng et al. 2016a). The PepperSNP16K array, which simultaneously genotyped 16,405 SNPs, was developed using the pepper haplotype map (HapMap) completed through resequencing of inbred lines (Hulse-Kemp et al. 2016). A total of 20 SNP-type assays for Fluidigm dynamic array, which were associated with several disease resistances and high capsaicinoid content, were developed for marker-assisted selection (MAS) of chili pepper (Kim et al. 2017a). Whole-genome resequencing and GBS were used for high-resolution mapping of QTLs controlling capsaicinoid content in *Capsicum* spp. (Han et al. 2018). A high-density genetic map containing 12,727 SNP markers was

constructed to identify QTLs for *cucumber mosaic virus* resistance in pepper using SLAF-seq (Li et al. 2018).

5.3 Genetic Maps in *Capsicum* Species

5.3.1 Interspecific Genetic Linkage Maps

Interspecific genetic linkage maps of *Capsicum* spp. were constructed using crosses including *C. annuum* × *C. chinense*, *C. annuum* × *C. frutescens*, and *C. annuum* × *C. baccatum* (Table 5.2). These maps are overviewed in Table 5.2.

5.3.1.1 *Capsicum annuum* × *Capsicum chinense*

The first pepper genetic linkage map was constructed by using 84 RFLP markers based on a common set of cDNA clones and selected single-copy genomic clones and by using 46 individuals derived from a cross between *C. annuum* 'Doux des Landes (CA50)' and *C. chinense* 'CA4' (Tanksley et al. 1988). A molecular genetic map of pepper covering 720 cM was constructed in an interspecific F_2 population with a total of 192 RFLP and isozyme markers (Prince et al. 1993). A genetic map of pepper consisting of 13 linkage groups that cover a total of 1245.7 cM was created from an interspecific F_2 population (*C. annuum* 'NuMex RNaky' × *C. chinense* 'PI159234'; Livingstone et al. 1999). The SNU pepper genetic map, consisting of 16 linkage groups and covering 1320 cM, was constructed in an interspecific F_2 population (*C. annuum* 'TF68' × *C. chinense* 'Habanero') with 150 RFLP and 430 AFLP markers (Kang et al. 2001). The SNU2 pepper map with 333 markers (46 SSR and 287 RFLP) in 15 linkage groups covering 1761.5 cM was generated in the same population with the SNU map (Lee et al. 2004). The SNU3 pepper map, forming 14 linkage groups and spanning 2201.5 cM, was constructed by adding 139 SSR markers based on expressed sequence tags (ESTs) (Yi et al. 2006). A SNP-based genetic map of pepper was developed in an F_2 population derived from a cross of *C. annuum* 'NB1' × *C. chinense* 'Jolokia' by using 116 SNP (HRM) markers generated from next-generation resequencing of parents (Lee et al. 2013). An EST-based linkage map of pepper (the AC2 map) was constructed in the AC99 F_2 population (*C. annuum* 'NuMex RNaky' × *C. chinense* 'PI159234') by using a total of 512 markers, comprising 214 intron-based polymorphic markers (IBPs), 143 conserved ortholog sets (COSIIs), 48 EST-SNPs (eSNPs), and 107 previously reported markers (Park et al. 2014). QTL mapping for capsaicinoid content was conducted in an interspecific population of 85 RILs derived from *C. annuum* 'TF68' × *C. chinense* 'Habanero' through a genotyping-by-sequencing (GBS) analysis (Han et al. 2018).

5.3.1.2 *Capsicum annuum* × *Capsicum frutescens*

A pepper genetic map was constructed for identifying yield-related QTLs using 248 BC_2 plants derived from a cross between *C. annuum* 'Maor' and *C. frutescens* 'BG2816' (Rao et al. 2003). In the same population, QTLs for capsaicinoid content were analyzed (Blum et al. 2003). An interspecific genetic map (*C. annuum* 'BA3' × *C. frutescens* 'YNXML') containing 129 InDel and 95 SSR markers was constructed for mapping the QTLs affecting the initiation of flower primordia (Tan et al. 2015). A linkage map with 5546 markers separated into 1361 bins on 12 linkage groups representing 1392.3 cM was produced using an interspecific population created between *C. frutescens* 'Tabasco' and *C. annuum* 'P4' and using the PepperSNP16K Infinium array (Hulse-Kemp et al. 2016). A high-density interspecific SNP genetic map of pepper was constructed in 297 F_2 individuals of *C. annuum* × *C. frutescens* using an Illumina Infinium iSelect SNP array (pepper *Cap*SNP15K array) (Cheng et al. 2016a).

Table 5.2 Overview of the interspecific genetic linkage maps of *Capsicum* spp.

Interspecific cross	Parents	Population size and type[a]	Number and type of markers[b]	Number of linkage groups	Total map length (cM)	References
C. annuum × *C. chinense*	'CA50' × 'CA4'	46 F_2	84 RFLPs	19	229	Tanksley et al. (1988)
	'CA50' × 'CA4'	46 F_2	192 RFLPs and isozymes	19	720	Prince et al. (1993)
	'NuMex RNaky' × 'PI159234'	75 F_2	350 AFLPs, 303 RFLPs, 17 RAPDs, 2 isozymes	13	1246	Livingstone et al. (1999)
	'TF68' × 'Habanero'	107 F_2	150 RFLPs, 430 AFLPs	16	1320	Kang et al. (2001)
	'TF68' × 'Habanero'	107 F_2	46 SSRs, 287 RFLPs	15	1762	Lee et al. (2004)
	'TF68' × 'Habanero'	107 F_2	139 EST-SSRs	14	2202	Yi et al. (2006)
	'NB1' × 'Jolokia'	96 F_2	116 HRMs	12	1168	Lee et al. (2013)
	'NuMex RNaky' × 'PI159234'	75 F_2	214 IBPs, 143 COSIIs, 48 eSNPs, 107 other markers	12	2336	Park et al. (2014)
	'TF68' × 'Habanero'	85 RILs	1089 bins (GBS)	12	1127	Han et al. (2018)
C. annuum × *C. frutescens*	'Maor' × 'BG2816'	248 BC_2	92 RFLPs	12	1100	Rao et al. (2003)
	'BA3' × 'YNXML'	154 and 147 F_2	129 InDels, 95 SSRs	13	1250	Tan et al. (2015)
	'P4' × 'Tabasco'	90 F_2	1361 bins (array)	12	1392	Hulse-Kemp et al. (2016)
	'BA3' × 'YNXML'	297 F_2	3826 bins (array)	12	1629	Cheng et al. (2016a)
C. annuum × *C. baccatum*	'SP26' × 'PBC81'	87 BC_1F_2	197 AFLPs, 21 SSRs	13	325	Lee et al. (2010b)

[a]*RILs* recombinant inbred lines

[b]*RFLPs* restriction fragment length polymorphisms, *AFLPs* amplified fragment length polymorphisms, *RAPDs* random amplified polymorphic DNAs, *SSRs* simple sequence repeats, *EST*-SSRs expressed sequence tag-SSRs, *HRMs* high-resolution melting markers, *IBPs* intron-based polymorphic markers, *COSIIs* conserved ortholog sets II, *eSNPs* EST-SNPs, *GBS* genotyping-by-sequencing, *InDels* insertion/deletion markers, *array* Illumina Infinium iSelect SNP array

5.3.1.3 *Capsicum annuum* × *Capsicum baccatum*

An introgression BC_1F_2 population was generated by interspecific crosses between *C. annuum* 'SP26' (susceptible) and *C. baccatum* 'PBC81' (resistant) for QTL mapping analyses of anthracnose resistance, and the introgression map consisting of 13 linkage groups with a total of 218 markers (197 AFLPs and 21 SSRs), covering 325 cM, was constructed (Lee et al. 2010b).

5.3.2 Intraspecific Genetic Linkage Maps

Intraspecific genetic linkage maps of *Capsicum* spp. have been reported in two species, *C. annuum* and *C. baccatum* (Table 5.3). The overview of these maps is given in Table 5.3.

5.3.2.1 *Capsicum annuum*

The first functional detailed map of pepper, containing 100 known-function gene markers and 9 loci of agronomic interest (*L*, *pvr2*, *Pvr4*, *C*, *up*, *Tsw*, *Me3*, *Bs3*, and *y*), was generated using three intraspecific populations including two DH populations of 'H3' × 'Vania' and 'Perennial' × 'Yolo Wonder' and one F_2 population of 'Yolo Wonder' × 'CM334' (Lefebvre et al. 2002). A genetic linkage map of the sweet pepper using an intraspecific DH population, consisting of 382 AFLP, 122 RAPD, 3 RFLP, 7 SCAR, and 4 CAPS markers, was constructed by AFLP using the high-efficiency genome scanning (HEGS) system and RAPD (Sugita et al. 2005). An SSR-based linkage map of *C. annuum*, including 106 new SSR markers distributed across 13 linkage groups and covering 1042 cM, was constructed in an intraspecific DH population derived from 'Manganji' × 'Tongari' (Minamiyama et al. 2006). A RFLP-based pepper linkage map, consisting of 202 RFLPs, 6 WRKYs, and 1 SSR and covering 1482.3 cM, was constructed to detect QTL associated with *P. capsici* resistance using 100 F_2 individuals from a cross between 'CM334' (resistant) and 'Chilsungcho' (susceptible) (Kim et al. 2008c). In the same population, 60 WRKY-based and 71 reverse random amplified microsatellite polymorphism (rRAMP)-based markers were developed and added (Kim et al. 2008b; Min et al. 2008a). A saturated intraspecific genetic map of pepper, containing 281 AFLPs, 101 EST-SSRs, 37 consensus SSRs, and 1 CAPS, was generated for studying QTLs associated with *Phytophthora* root rot resistance using a population of 126 F_8 RILs derived from a cross between 'YCM334' (resistant) and 'Tean' (susceptible) (Truong et al. 2010). The first SSR-based intraspecific genetic map of *C. annuum*, containing 151 SSRs, 90 AFLPs, 10 CAPSs, and 2 STSs and spanning 1336 cM, was constructed using a DH population derived from a cross between 'California Wonder' and 'LS2341' (Mimura et al. 2012). An SSR-based high-density linkage map of *C. annuum*, consisting of 597 SSR markers and covering 2028 cM, was developed by using DH lines derived from an intraspecific cross of 'K9-11' × 'MZC-180' (Sugita et al. 2013). The first InDel-based linkage map of hot pepper (BB-InDel map), containing 251 InDel markers and covering 1178 cM, was made using an F_2 population derived from the intraspecific cross 'BA3' × 'B702' through whole-genome resequencing of two parents (Li et al. 2015). An ultra-high-density bin map of *C. annuum*, containing 2578 bins and spanning 1372 cM, was developed for QTL mapping of horticultural traits using 120 RILs derived from a cross between 'Perennial' and 'Dempsey' (Han et al. 2016). A high-density genetic map of *C. annuum*, containing 12,727 markers on 12 chromosomes and spanning 1785 cM, was constructed using 195 F_2 individuals derived from a cross between 'BJ0747' (resistant) and 'XJ0630' (susceptible) to identify QTLs for CMV resistance using SLAF-seq (Li et al. 2018).

5.3.2.2 *Capsicum baccatum*

An intraspecific genetic map of *C. baccatum*, containing 52 SSRs, 175 AFLPs, and 100 SRAPs, and covering 1896 cM, was developed using 126 F_2 plants derived from a cross between 'Cbp' (resistant) and 'Golden-aji' (susceptible) to identify QTLs associated with anthracnose resistance (Kim et al. 2010). A reference map of *C. baccatum* based on 42 SSRs, 85 ISSRs, and 56 RAPDs, consisting of 16 linkage groups and covering 2547 cM, was constructed using 203 F_2 individuals originated from a cross of 'UENF1616' and 'UENF1732' (Moulin et al. 2015). A SNP-based genetic linkage map of *C. baccatum*, containing 395 HRM markers and covering 1056.2 cM, was generated using an F_2 population from a cross between 'Golden-aji' and 'PI594137' and was compared to *C. annuum* reference physical map (Lee et al. 2016).

Table 5.3 Overview of the intraspecific genetic linkage maps of *Capsicum* spp.

Intraspecific cross	Parents	Population size and type[a]	Number and type of markers[b]	Number of linkage groups	Total map length (cM)	References
C. annuum × *C. annuum*	'H3' × 'Vania'	101 DH	434 AFLPs, 56 RAPDs, 50 RFLPs, 3 morphological markers (*C*, *L*, *pvr2*)	12	1513	Lefebvre et al. (2002)
	'Perennial' × 'Yolo Wonder'	114 DH	325 AFLPs, 164 RAPDs, 133 RFLPs, 1 isozyme, 4 SCARs, 3 morphological markers (*C*, *L*, *up*)	26	1.668	Lefebvre et al. (2002)
	'Yolo Wonder' × 'CM334'	151 F_2	109 AFLPs, 67 RFLPs, 28 RAPDs, 2 SCARs, 2 morphological markers (*C*, *Pvr4*)	18	685	Lefebvre et al. (2002)
	'K9-11' × 'AC2258'	176 DH	382 AFLPs, 112 RAPDs, 3 RFLPs, 7 SCARs, 4 CAPSs	16	1043	Sugita et al. (2005)
	'Manganji' × 'Tongari'	117 DH	123 SSRs, 228 AFLPs, 60 RAPDs, 1 CAPS	13	1042	Minamiyama et al. (2006)
	'CM334' × 'Chilsungcho'	100 F_2	202 RFLPs, 6 WRKYs, 1 SSR	14	1482	Kim et al. (2008c)
	'CM334' × 'Chilsungcho'	100 F_2	163 rRAMPs, 134 AFLPs, 29 SSRs, 9 RFLPs, 2 RAPDs	16	1854	Min et al. (2008a)
	'CM334' × 'Chilsungcho'	100 F_2	41 WRKYs, 199 AFLPs, 26 SSRs, 8 RFLPs, 97 rRAMPs	20	2051	Kim et al. (2008b)
	'YCM334' × 'Tean'	126 RILs	281 AFLPs, 101 EST-SSRs, 37 SSRs, 1 CAPS	14	2178	Truong et al. (2010)
	'California Wonder' × 'LS2341'	94 DH	151 SSRs, 90 AFLPs, 10 CAPSs, 2 STS	12	1336	Mimura et al. (2012)
	'K9-11' × 'MZC-180'	184 DH	597 SSRs	12	2028	Sugita et al. (2013)
	'BA3' × 'B702'	178 F_2	251 InDels	12	1178	Li et al. (2015)
	'Perennial' × 'Dempsey'	120 RILs	2578 bins (WGS)	12	1372	Han et al. (2016)
	'A1' × '2602'	96 SSD F_3	906 SNPs (GBS)	12	1273	Eun et al. (2016)
	'BJ0747' × 'XJ0630'	195 F_2	12,727 SNPs (SLAF-seq)	12	1785	Li et al. (2018)
	'Perennial' × 'Dempsey'	56 RILs	2578 bins (WGS)	12	1372	Han et al. (2018)
C. baccatum × *C. baccatum*	'Cbp' × 'Golden-aji'	126 F_2	52 SSRs, 175 AFLPs, 100 SRAPs	13	1896	Kim et al. (2010)
	'UENF1616' × 'UENF1732'	203 F_2	42 SSRs, 85 ISSRs, 56 RAPDs	16	2547	Moulin et al. (2015)
	'Golden-aji' × 'PI594137'	93 F_2	395 HRMs	12	1056	Lee et al. (2016)

[a]*DH* doubled haploids, *RILs* recombinant inbred lines, *SSD* single seed descent

[b]*AFLPs* amplified fragment length polymorphisms, *RAPDs* random amplified polymorphic DNAs, *RFLPs* restriction fragment length polymorphisms, *SCARs* sequence-characterized amplified regions, *CAPSs* cleaved amplified polymorphic sequences, *SSRs* simple sequence repeats, *WRKYs* WRKY gene-based markers, *rRAMPs* reverse random amplified microsatellite polymorphisms, *EST*-SSRs expressed sequence tag-SSRs, *InDels* insertion/deletion markers, *WGS* whole-genome resequencing, *SNPs* single-nucleotide polymorphisms, *GBS* genotyping-by-sequencing, *SLAF-seq* specific-locus amplified fragment sequencing, *SRAPs* sequence-related amplified polymorphisms, *ISSRs* inter-simple sequence repeats, *HRMs* high-resolution melting markers

5.3.3 Integrated Genetic Linkage Maps

The first integrated linkage map of *C. annuum*, including mainly RFLP and RAPD markers, was constructed by alignment of three intraspecific linkage maps generated by segregating DH progenies (Lefebvre et al. 1995). An integrated genetic linkage map of pepper, consisting of 1528 AFLP, 440 RFLP, 288 RAPD, several known gene sequences, isozymes, and morphological markers and covering 1832 cM, was generated by using pooled data from six individual maps (Paran et al. 2004). An integrated pepper map, containing 169 SSR, 354 RFLP, 23 STS from BAC end sequences, 6 STS from RFLP, 152 AFLP, 51 WRKY, and 99 rRAMP markers on 12 chromosomes, was constructed using four genetic maps of two interspecific (*C. annuum* 'TF68' × *C. chinense* 'Habanero') and two intraspecific (*C. annuum* 'CM334' × 'Chilsungcho') populations (Lee et al. 2009).

5.3.4 Comparative Mapping Between Solanaceous Crops

The first RFLP-based pepper linkage map was compared to the RFLP-based tomato map, suggesting that gene repertoire is conserved but gene order is not (Tanksley et al. 1988). Comparison of the pepper, tomato, and potato genetic maps revealed a total of 30 breaks as part of 5 translocations, 10 paracentric inversions, 2 pericentric inversions, and 4 disassociations or associations of genomic regions (Livingstone et al. 1999). Disease resistance genes (*R* genes) and *R* gene homologues were compared between three solanaceous crops including tomato, potato, and pepper (Grube et al. 2000). Pepper genome was compared to tomato genome using a total of 299 orthologous markers including 263 conserved ortholog set II (COSII) markers, suggesting that the two genomes have become differentiated by a minimum number of 19 inversions and 6 translocations, as well as numerous putative single gene transpositions but share 35 conserved syntenic segments within which gene/marker order is well preserved (Wu et al. 2009). In addition, the genome of cultivated *C. annuum* and wild *C. annuum* (as well as *C. chinense*, *C. frutescens*) was found to differ by a reciprocal translocation between chromosomes 1 and 8 (Wu et al. 2009). Comparative mapping studies were performed in tomato, potato, eggplant, pepper, and diploid *Nicotiana* species (*Nicotiana tomentosiformis* and *Nicotiana acuminata*) using COSII markers, providing the first broad overview of chromosomal evolution in the family Solanaceae (Wu and Tanksley 2010). The eggplant/pepper syntenic map confirmed 10 translocations and 8 inversions already detected, and a set of 151 pepper QTL were located as well as 212 eggplant QTL, including 76 major QTLs (phenotypic variance explained, PVE $\geq$ 10%) affecting key agronomic traits (Rinaldi et al. 2016). Recently, two high-quality de novo genomes (*C. baccatum* 'PBC81' and *C. chinense* 'PI159236') and an improved reference genome (*C. annuum* 'CM334') were reported, showing dynamic genome rearrangements involving translocations among chromosomes 3, 5, and 9 between *C. baccatum* and the two other peppers and suggesting the process of speciation and evolution of the *Capsicum* species (Kim et al. 2017b).

5.3.5 Marker-Assisted Backcrossing

Marker-assisted backcrossing (MABC) is a new breeding approach that can substantially reduce breeding time and cost by using highly polymorphic markers with known positions in each chromosome (Frisch et al. 1999; Herzog and Frisch 2011). A total of 412 SNP markers were developed from EST sequences generated by large-scale transcriptome sequencing of eight accessions (*C. annuum* 'Jeju,' 'LAM32,' 'Tean,' 'CM334,' 'Yuwolcho,' 'PI201234,' and 'YCM334' and *C. chinense* 'SNU-001') using the Illumina Genome Analyzer IIx platform to facilitate MABC in hot pepper (Kang et al. 2014). Moreover, by analyzing the SNP makers via a high-throughput SNP genotyping system

(Fluidigm® EP1™ system), a genetic linkage map of *C. frutescens* 'BG2814-6' × *C. annuum* 'NuMex RNaky' was constructed and a genetic diversity of 27 *Capsicum* accessions was tested (Kang et al. 2014).

5.4 Future Prospects

Various molecular marker techniques and many genetic linkage maps can be used to develop the trait-linked markers or gene-based markers as well as to identify a gene or QTLs for important horticultural traits including male sterility (CMS and GMS), various disease resistances (anthracnose, powdery mildew, phytophthora, bacterial wilt, bacterial spot, CMV, TSWV, PMMoV, PepMoV, and nematode), and fruit traits (color, shape, size, capsaicinoid content, carotenoid content, and sugar content). To date, a few function-known pepper genes including *Bs2*, *pvr1* (*pvr2*), *pun1*, *Bs3*, *pAMT*, *L*, *Pvr4*, *Tsw*, and ms_1 were only identified through map-based cloning or candidate gene approach. Recent genome-wide genotyping technologies such as GBS, Rad-seq, and SLAF-seq will accelerate development of whole-genome genetic maps and trait-linked DNA markers and identification of genes controlling important horticultural traits.

References

Akopyanz N, Bukanov NO, Westblom TU, Berg DE (1992) PCR-based RFLP analysis of DNA sequence diversity in the gastric pathogen *Helicobacter pylori*. Nucl Acids Res 20:6221–6225

Andolfatto P, Davison D, Erezyilmaz D, Hu TT, Mast J et al (2011) Multiplexed shotgun genotyping for rapid and efficient genetic mapping. Genome Res 21:610–617

Arjun K, Dhaliwal MS, Jindal SK, Fakrudin B (2018) Mapping of fruit length related QTLs in interspecific cross (*Capsicum annuum* L. × *Capsicum galapagoense* Hunz.) of chilli. Breed Sci 68:219–226

Arnedo-Andrés MS, Gil-Ortega R, Luis-Arteaga M, Hormaza JI (2002) Development of RAPD and SCAR markers linked to the *Pvr4* locus for resistance to PVY in pepper (*Capsicum annuum* L.). Theor Appl Genet 105:1067–1074

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL et al (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS ONE 3:e3376

Bartoszewski G, Waszczak C, Gawroński P, Stępień I, Bolibok-Bragoszewska H et al (2012) Mapping of the ms_8 male sterility gene in sweet pepper (*Capsicum annuum* L.) on the chromosome P4 using PCR-based markers useful for breeding programmes. Euphytica 186:453–461

Bello-Bello J, Iglesias-Andreu LG, Avilés-Viñas SA, Gómez-Uc E, Canto-Flick A et al (2014) Somaclonal variation in Habanero pepper (*Capsicum chinense* Jacq.) as assessed ISSR molecular markers. HortScience 49:481–485

Bhattramakki D, Dolan M, Hanafey M, Wineland R, Vaske D et al (2002) Insertion-deletion polymorphisms in 3' regions of maize genes occur frequently and can be used as highly informative genetic markers. Plant Mol Biol 48:539–547

Blears MJ, De Grandis SA, Lee H, Trevors JT (1998) Amplified fragment length polymorphism (AFLP): a review of the procedure and its applications. J Ind Microbiol Biotechnol 21:99–114

Blum E, Mazourek M, O'Connell M, Curry J, Thorup T et al (2003) Molecular mapping of capsaicinoid biosynthesis genes and quantitative trait loci analysis for capsaicinoid content in *Capsicum*. Theor Appl Genet 108:79–86

Botstein B, White RL, Sholnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Amer J Hum Genet 32:314–331

Bybee SM, Bracken-Grissom H, Haynes BD, Hermansen RA, Byers RL et al (2011) Targeted amplicon sequencing (TAS): a scalable next-gen approach to multilocus, multitaxa phylogenetics. Genome Biol Evol 3:1312–1323

Caetano-Anollés G, Bassam BJ, Gresshoff PM (1991) DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. Nat Biotechnol 9:553–557

Caranta C, Thabuis A, Palloix (1999) Development of a CAPS marker for the *Pvr4* locus: a tool for pyramiding potyvirus resistance genes in pepper. Genome 42:1111–1116

Cheng J, Qin C, Tang X, Zhou H, Hu Y et al (2016a) Development of a SNP array and its application to genetic mapping and diversity assessment in pepper (*Capsicum* spp.). Sci Rep 6:33293

Cheng J, Zhao Z, Li B, Qin C, Wu Z et al (2016b) A comprehensive characterization of simple sequence repeats in pepper genomes provides valuable resources for marker development in *Capsicum*. Sci Rep 6:18919

Deschamps S, Llaca V, May GD (2012) Genotyping-by-sequencing in plants. Biology 1:460–483

Eun MH, Han JH, Yoon JB, Lee J (2016) QTL mapping of resistance to the *Cucumber mosaic virus* P1 strain

in pepper using a genotyping-by-sequencing analysis. Hort Environ Biotechnol 57:589–597
Frisch M, Bohn M, Melchinger AE (1999) Comparison of selection strategies for marker-assisted backcrossing of a gene. Crop Sci 39:1295–1301
Ganal MW, Altmann T, Röder MS (2009) SNP identification in crop plants. Curr Opin Plant Biol 12:211–217
Grube RC, Radwanski ER, Jahn M (2000) Comparative genetics of disease resistance within the solanaceae. Genetics 155:873–887
Gupta M, Chyi YS, Romero-Severson J, Owen JL (1994) Amplification of DNA markers from evolutionarily diverse genomes using single primers of simple-sequence repeats. Theor Appl Genet 89:998–1006
Gut IG (2001) Automation in genotyping of single nucleotide polymorphisms. Hum Mut 17:475–492
Han K, Jeong HJ, Yang HB, Kang SM, Kwon JK et al (2016) An ultra-high-density bin map facilitates high-throughput QTL mapping of horticultural traits in pepper (*Capsicum annuum*). DNA Res 23:81–91
Han K, Lee HY, Ro NY, Hur OS, Lee JH et al (2018) QTL mapping and GWAS reveal candidate genes controlling capsaicinoid content in *Capsicum*. Plant Biotechnol J. https://doi.org/10.1111/pbi.12894
Hearne CM, Ghosh S, Todd JA (1992) Microsatellites for linkage analysis of genetic traits. Trends Genet 8: 288–294
Herzog E, Frisch M (2011) Selection strategies for marker-assisted backcrossing with high-throughput marker systems. Theor Appl Genet 123:251–260
Huang X, Feng Q, Qian Q, Zhao Q, Wang L et al (2009) High-throughput genotyping by whole-genome resequencing. Genome Res 19:1068–1076
Hulse-Kemp AM, Ashrafi H, Plieske J, Lemm J, Stoffel K et al (2016) A HapMap leads to a *Capsicum annuum* SNP infinium array: a new tool for pepper breeding. Hort Res 3:16036
Ibarra-Torres P, Valadez-Moctezuma E, Pérez-Grajales M, Rodríguez-Campos J, Jaramillo-Flores ME (2015) Inter- and intraspecific differentiation of *Capsicum annuum* and *Capsicum pubescens* using ISSR and SSR markers. Sci Hort 181:137–146
International HapMap Consortium (2005) A haplotype map of the human genome. Nature 437:1299–1320
Jiang GL (2013) Molecular markers and marker-assisted breeding in plants. In: Andersen SB (ed) Plant breeding from laboratories to fields. Intech, pp 45–83
Jo YD, Kim YM, Park MN, Yoo JH, Park M et al (2010) Development and evaluation of broadly applicable markers for *restorer-of-fertility* in pepper. Mol Breed 25:187–201
Jo J, Venkatesh J, Han K, Lee HY, Choi GJ et al (2017) Molecular mapping of *PMR1*, a novel locus conferring resistance to powdery mildew in pepper (*Capsicum annuum*). Front Plant Sci 8:2090
Jung JK, Park SW, Liu WY, Kang BC (2010) Discovery of single nucleotide polymorphism in *Capsicum* and SNP markers for cultivar identification. Euphytica 175:91–107
Kang BC, Nahm SH, Huh JH, Yoo HS, Yu JW et al (2001) An interspecific (*Capsicum annuum* × *C. chinense*) F_2 linkage map in pepper using RFLP and AFLP markers. Theor Appl Genet 102:531–539
Kang JH, Yang HB, Jeong HS, Choe P, Kwon JK et al (2014) Single nucleotide polymorphism marker discovery from transcriptome sequencing for marker-assisted backcrossing in *Capsicum*. Korean J Hort Sci Technol 32:535–543
Kantety RV, Rota ML, Matthews DE, Sorrells ME (2002) Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. Plant Mol Biol 48:501–510
Kesawat MS, Das BK (2009) Molecular markers: it's application in crop improvement. J Crop Sci Biotechnol 12:169–181
Kim DH, Kim BD (2005) Development of SCAR markers for early identification of cytoplasmic male sterility genotype in chili pepper (*Capsicum annuum* L.). Mol Cells 20:416–422
Kim DS, Kim DH, Yoo JH, Kim BD (2006) Cleaved amplified polymorphic sequence and amplified fragment length polymorphism markers linked to the fertility restorer gene in chili pepper (*Capsicum annuum* L.). Mol Cells 21:135–140
Kim HJ, Han JH, Yoo JH, Cho HJ, Kim BD (2008a) Development of a sequence characteristic amplified region marker linked to the L^4 locus conferring broad spectrum resistance to tobamoviruses in pepper plants. Mol Cells 25:205–210
Kim HJ, Lee HR, Han JH, Yeom SI, Harn CH et al (2008b) Marker production by PCR amplification with primer pairs from conserved sequences of WRKY genes in chili pepper. Mol Cells 25:196–204
Kim HJ, Nahm SH, Lee HR, Yoon GB, Kim KT et al (2008c) BAC-derived markers converted from RFLP linked to *Phytophthora capsici* resistance in pepper (*Capsicum annuum* L.). Theor Appl Genet 118:15–27
Kim S, Kim KT, Kim DH, Yang EY, Cho MC et al (2010) Identification of quantitative trait loci associated with anthracnose resistance in chili pepper (*Capsicum* spp.). Korean J Hort Sci Technol 28:1014–1024
Kim S, Park M, Yeom SI, Kim YM, Lee JM et al (2014) Genome sequence of the hot pepper provides insights into the evolution of pungency in *Capsicum* species. Nat Genet 46:270–279
Kim C, Guo H, Kong W, Chandnani R, Shuang LS et al (2016) Application of genotyping by sequencing technology to a variety of crop breeding programs. Plant Sci 242:14–22
Kim H, Yoon JB, Lee J (2017a) Development of Fluidigm SNP type genotyping assays for marker-assisted breeding of chili pepper (*Capsicum annuum* L.). Hort Sci Technol 35:465–479
Kim S, Park J, Yeom SI, Kim YM, Seo E et al (2017b) New reference genome sequences of hot pepper reveal the massive evolution of plant disease-resistance genes by retroduplication. Genome Biol 18:210

Konieczny A, Ausubel FM (1993) A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. Plant J 4: 403–410

Kumar LS (1999) DNA markers in plant improvement: an overview. Biotechnol Adv 17:143–182

Kumar LD, Kathirvel M, Rao GV, Nagaraju J (2001) DNA profiling of disputed chilli samples (*Capsicum annuum*) using ISSR-PCR and FISSR-PCR marker assays. Forens Sci Intl 116:63–68

Kumar S, Banks TW, Cloutier S (2012) SNP discovery through next-generation sequencing and its applications. Int J Plant Genom. http://dx.doi.org/10.1155/2012/831460

Lee JM, Nahm SH, Kim YM, Kim BD (2004) Characterization and molecular genetic mapping of microsatellite loci in pepper. Theor Appl Genet 108:619–627

Lee J, Yoon JB, Park HG (2008) A CAPS marker associated with the partial restoration of cytoplasmic male sterility in chili pepper (*Capsicum annuum* L.). Mol Breed 21:95–104

Lee HR, Bae IH, Park SW, Kim HJ, Min WK et al (2009) Construction of an integrated pepper map using RFLP, SSR, CAPS, AFLP, WRKY, rRAMP, and BAC end sequences. Mol Cells 27:21–37

Lee J, Han JH, An CG, Lee WP, Yoon JB (2010a) A CAPS marker linked to a genic male-sterile gene in the colored sweet pepper, 'Paprika' (*Capsicum annuum* L.). Breed Sci 60:93–98

Lee J, Hong JH, Do JW, Yoon JB (2010b) Identification of QTLs for resistance to anthracnose to two *Colletotrichum* species in pepper. J Crop Sci Biotechnol 13:227–233

Lee J, Lee WP, Han JH, Yoon JB (2010c) Development of molecular marker linked to a genic male-sterile gene, ms_k in chili pepper. Korean J Hortic Sci Technol 28:270–274

Lee J, Yoon JB, Han JH, Lee WP, Kim SH et al (2010d) Three AFLP markers tightly linked to the genic male sterility ms_3 gene in chili pepper (*Capsicum annuum* L.) and conversion to a CAPS marker. Euphytica 173:55–61

Lee J, Do JW, Han JH, An CG, Kweon OY et al (2011a) Allelism and molecular marker tests for genic male sterility in paprika cultivars. Korean J Hort Sci Technol 29:130–134

Lee J, Do JW, Yoon JB (2011b) Development of STS markers linked to the major QTLs for resistance to the pepper anthracnose caused by *Colletotrichum acutatum* and *C. capsici*. Hort Environ Biotechnol 52: 596–601

Lee J, Han JH, Yoon JB (2012a) A set of allele-specific markers linked to *L* locus resistant to *Tobamovirus* in *Capsicum* spp. Korean J Hort Sci Technol 30:286–293

Lee WP, Lee J, Han JH, Kang BC, Yoon JB (2012b) Validity test for molecular markers associated with resistance to *Phytophthora* root rot in chili pepper (*Capsicum annuum* L.). Kor J Hort Sci Technol 30:64–72

Lee J, Park SJ, Do JW, Han JH, Choi D et al (2013) Development of a genetic map of chili pepper using single nucleotide polymorphism markers generated from next generation resequencing of parents. Korean J Hort Sci Technol 31:473–482

Lee YR, Yoon JB, Lee J (2016) A SNP-based genetic linkage map of *Capsicum baccatum* and its comparison to the *Capsicum annuum* reference physical map. Mol Breed 36:61

Lefebvre V, Palloix A, Caranta C, Pochard E (1995) Construction of an intraspecific integrated linkage map of pepper using molecular markers and doubled-haploid progenies. Genome 38:112–121

Lefebvre V, Pflieger S, Thabuis A, Caranta C, Blattes A et al (2002) Towards the saturation of the pepper linkage map by alignment of three intraspecific maps including known-function genes. Genome 45:839–854

Li W, Cheng J, Wu Z, Qin C, Tan S et al (2015) An InDel-based linkage map of hot pepper (*Capsicum annuum*). Mol Breed 35:32

Li N, Yin Y, Wang F, Yao M (2018) Construction of a high-density genetic map and identification of QTLs for cucumber mosaic virus resistance in pepper (*Capsicum annuum* L.) using specific length amplified fragment sequencing (SLAF-seq). Breed Sci 68: 233–241

Liew M, Pryor R, Palais R, Meadows C, Erali M et al (2004) Genotyping of single-nucleotide polymorphisms by high-resolution melting of small amplicons. Clin Chem 50:1156–1164

Livak KJ (1999) Allelic discrimination using fluorogenic probes and the 5' nuclease assay. Genet Anal-Biomol E 14:143–149

Livingstone KD, Lackney VK, Blauth JR, van Wijk R, Jahn MK (1999) Genome mapping in *Capsicum* and the evolution of genome structure in the Solanaceae. Genetics 152:1183–1202

Lyamichev V, Brow MAD, Dahlberg JE (1993) Structure-specific endonucleolytic cleavage of nucleic acids by eubacterial DNA polymerases. Science 260:778–783

Mascher M, Wu S, St. Amand P, Stein N, Poland J (2013) Application of genotyping-by-sequencing on semiconductor sequencing platforms: a comparison of genetic and reference-based marker ordering in barley. PLoS ONE 8:e76925

Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci USA 88:9828–9832

Mimura Y, Inoue T, Minamiyama Y, Kubo N (2012) An SSR-based genetic map of pepper (*Capsicum annuum* L.) serves as an anchor for the alignment of major pepper maps. Breed Sci 62:93–98

Min WK, Han JH, Kang WH, Lee HR, Kim BD (2008a) Reverse random amplified microsatellite polymorphism reveals enhanced polymorphisms in the 3' end of simple sequence repeats in the pepper genome. Mol Cells 26:250–257

Min WK, Lim H, Lee YP, Sung SK, Kim BD, Kim S (2008b) Identification of a third haplotype of the sequence linked to the *restorer-of-fertility* (*Rf*) gene and its implications for male-sterility phenotypes in peppers (*Capsicum annuum* L.). Mol Cells 25:20–29

Min WK, Kim S, Sung SK, Kim BD, Lee S (2009) Allelic discrimination of the *restorer-of-fertility* gene and its inheritance in peppers (*Capsicum annuum* L.). Theor Appl Genet 119:1289–1299

Minamiyama Y, Tsuro M, Hirai M (2006) An SSR-based linkage map of *Capsicum annuum*. Mol Breed 18:157–169

Mohan M, Nair S, Bhagwat A, Krishna TG, Yano M et al (1997) Genome mapping, molecular markers and marker-assisted selection in crop plants. Mol Breed 3:87–103

Monson-Miller J, Sanchez-Mendez DC, Fass J, Henry IM, Tai TH et al (2012) Reference genome-independent assessment of mutation density using restriction enzyme-phased sequencing. BMC Genom 13:72

Moulin MM, Rodrigues R, Ramos HCC, Bento CS, Sudré CP et al (2015) Construction of an integrated genetic map for *Capsicum baccatum* L. Genet Mol Res 14:6683–6694

Moury B, Pflieger S, Blattes A, Lefebvre V, Palloix A (2000) A CAPS marker to assist selection of tomato spotted wilt virus (TSWV) resistance in pepper. Genome 43:137–142

Myakishev MV, Khripin Y, Hu S, Hamer DH (2001) High-throughput SNP genotyping by allele-specific PCR with universal energy-transfer-labeled primers. Genome Res 11:163–169

Nakamura Y, Leppert M, O'Connell P, Wolff R, Holm T et al (1987) Variable number of tandem repeat (VNTR) markers for human gene mapping. Science 235:1616–1622

Neff MM, Neff JD, Chory J, Pepper AE (1998) dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: experimental applications in *Arabidopsis thaliana* genetics. Plant J 14:387–392

Paran I, Michelmore RW (1993) Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. Theor Appl Genet 85:985–993

Paran I, van der Voort JR, Lefebvre V, Jahn M, Landry L et al (2004) An integrated genetic linkage map of pepper (*Capsicum* spp.). Mol Breed 13:251–261

Park SW, Jung JK, Choi EA, Kwon JK, Kang JH et al (2014) An EST-based linkage map reveals chromosomal translocation in *Capsicum*. Mol Breed 34:963–975

Patel AS, Sasidharan N, Vala AG, Kumar V (2011) Genetic relation in *Capsicum annuum* [L.] cultivars through microsatellite markers: SSR and ISSR. Elec J Plant Breed 2:67–76

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. PLoS ONE 7:e37135

Pinar H, Mutlu N, Ozaslandan A, Argun D, Keles D et al (2016) Reliability assessment of molecular markers linked to resistance genes against *Meloidogyne* spp. in diverse peppers genotypes. Egypt J Biol Pest Con 26:515–521

Poland JA, Rife TW (2012) Genotyping-by-sequencing for plant breeding and genetics. Plant Genome 5:92–102

Poland JA, Brown PJ, Sorrells ME, Jannink JL (2012) Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS ONE 7:e32253

Prince JP, Pochard E, Tanksley SD (1993) Construction of a molecular linkage map of pepper and a comparison of synteny with tomato. Genome 36:404–417

Qin C, Yu C, Shen Y, Fang X, Chen L et al (2014) Whole-genome sequencing of cultivated and wild peppers provides insights into *Capsicum* domestication and specialization. Proc Nat Acad Sci USA 111:5135–5140

Quirin EA, Ogundiwin EA, Prince JP, Mazourek M, Briggs MO et al (2005) Development of sequence characterized amplified region (SCAR) primers for the detection of Phyto. 5.2, a major QTL for resistance to Phytophthora capsici Leon. In pepper. Theor Appl Genet 110:605–612

Rafalski A (2002) Applications of single nucleotide polymorphisms in crop genetics. Curr Opin Plant Biol 5:94–100

Rao GU, Ben Chaim A, Borovsky Y, Paran I (2003) Mapping of yield-related QTLs in pepper in an interspecific cross of *Capsicum annuum* and *C. frutescens*. Theor Appl Genet 106:1457–1466

Reddy MP, Sarla N, Siddiq EA (2002) Inter simple sequence repeat (ISSR) polymorphism and its application in plant breeding. Euphytica 128:9–17

Ribaut JM, Hoisington D (1998) Marker-assisted selection: new tools and strategies. Trends Plant Sci 3: 236–239

Rinaldi R, Deynze AV, Portis E, Rotino GL, Toppino L et al (2016) New insights on eggplant/tomato/pepper synteny and identification of eggplant and pepper orthologous QTL. Front Plant Sci 7:1031

Römer P, Jordan T, Lahaye T (2010) Identification and application of a DNA-based marker that is diagnostic for the pepper (*Capsicum annuum*) bacterial spot resistance gene *Bs3*. Plant Breed 129:737–740

Semagn K, Babu R, Hearne S, Olsen M (2014) Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. Mol Breed 33:1–14

Shendure J, Ji H (2008) Next-generation DNA sequencing. Nat Biotechnol 26:1135–1145

Simko I (2016) High-resolution DNA melting analysis in plant research. Trends Plant Sci 21:528–537

Stolle E, Moritz RFA (2013) RESTseq—efficient benchtop population genomics with RESTriction Fragment SEQuencing. PLoS ONE 8:e63960

Sugita T, Yamaguchi K, Sugimura Y, Nagata R, Yuji K et al (2004) Development of SCAR markers linked to L^3 gene in *Capsicum*. Breed Sci 54:111–115

Sugita T, Kinoshita T, Kawano T, Yuji K, Yamaguchi K et al (2005) Rapid construction of a linkage map using high-efficiency genome scanning/AFLP and RAPD, based on an intraspecific, doubled-haploid population of *Capsicum annuum*. Breed Sci 55:287–295

Sugita T, Semi Y, Sawada H, Utoyama Y, Hosomi Y et al (2013) Development of simple sequence repeat markers and construction of a high-density linkage map of *Capsicum annuum*. Mol Breed 31:909–920

Sun X, Liu D, Zhang X, Li W, Liu H et al (2013) SLAF-seq: an efficient method of large-scale *de novo* SNP discovery and genotyping using high-throughput sequencing. PLoS ONE 8:e58700

Suwor P, Sanitchon J, Thummabenjapone P, Kumar S, Techawongstien S (2017) Inheritance analysis of anthracnose resistance and marker-assisted selection in introgression populations of chili (*Capsicum annuum* L.). Sci Hort 220:20–26

Tan S, Cheng JW, Zhang L, Qin C, Nong DG et al (2015) Construction of an interspecific genetic map based on InDel and SSR for mapping the QTLs affecting the initiation of flower primordia in pepper (*Capsicum* spp.). PLoS ONE 10:e0119389

Tanksley SD, Bernatzky R, Lapitan NL, Prince JP (1988) Conservation of gene repertoire but not gene order in pepper and tomato. Proc Natl Acad Sci USA 85:6419–6423

Thomson MJ (2014) High-throughput SNP genotyping to accelerate crop improvement. Plant Breed Biotech 2:195–212

Thul ST, Darokar MP, Shasany AK, Khanuja SPS (2012) Molecular profiling for genetic variability in *Capsicum* species based on ISSR and RAPD markers. Mol Biotechnol 51:137–147

Truong HTH, Kim KT, Kim S, Chae Y, Park JH et al (2010) Comparative mapping of consensus SSR markers in an intraspecific F_8 recombinant inbred line population in *Capsicum*. Hort Environ Biotechnol 51:193–206

Truong HT, Ramos AM, Yalcin F, de Ruiter M, van der Poel HJA et al (2012) Sequence-based genotyping for marker discovery and co-dominant scoring in germplasm and populations. PLoS ONE 7:e37565

Truong HTH, Kim JH, Cho MC, Chae SY, Lee HE (2013) Identification and development of molecular markers linked to *Phytophthora* root rot resistance in pepper (*Capsicum annuum* L.). Eur J Plant Pathol 135:289–297

Uncu AT, Celik I, Devran Z, Ozkaynak E, Frary A et al (2015) Development of a SNP-based CAPS assay for the *Me1* gene conferring resistance to root knot nematode in pepper. Euphytica 206:393–399

Varshney RK, Graner A, Sorrells ME (2005) Genic microsatellite markers in plants: features and applications. Trends Biotechnol 23:48–55

Varshney RK, Nayak SN, May GD, Jackson SA (2009) Next-generation sequencing technologies and their implications for crop genetics and breeding. Trends Biotechnol 27:522–530

Vodenicharova M (1989) Use of proteins as molecular-genetic markers in plants. Genet Sel 22:269–277

Vos P, Hogers R, Bleeker M, Reijans M, van der Lee T et al (1995) AFLP: a new technique for DNA fingerprinting. Nucl Acids Res 23:4407–4414

Wang DG, Fan JB, Siao CH, Berno A, Young P et al (1998) Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science 280:1077–1082

Wang J, Lin M, Crenshaw A, Hutchinson A, Kicks B et al (2009) High-throughput single nucleotide polymorphism genotyping using nanofluidic Dynamic Arrays. BMC Genom 10:561

Wang X, Fazari A, Cao Y, Zhang Z, Palloix A et al (2018) Fine mapping of the root-knot nematode resistance gene *Me1* in pepper (*Capsicum annuum* L.) and development of markers tightly linked to *Me1*. Mol Breed 38:39

Weber JL, David D, Heil J, Fan Y, Zhao C et al (2002) Human diallelic insertion/deletion polymorphisms. Amer J Hum Genet 71:854–862

Welsh J, McClelland (1990) Fingerprinting genomes using PCR with arbitrary primers. Nucl Acids Res 18:7213–7218

Williams JGK, Kubelik AR, Livak KL, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucl Acids Res 18:6531–6535

Wittwer CT (2009) High-resolution DNA melting analysis: advancements and limitations. Hum Mut 30:857–859

Wittwer CT, Reed GH, Gundry CN, Vandersteen JG, Pryor RJ (2003) High-resolution genotyping by amplicon melting analysis using LCGreen. Clin Chem 49:853–860

Wu F, Tanksley SD (2010) Chromosomal evolution in the plant family Solanaceae. BMC Genom 11:182

Wu F, Eannetta NT, Xu Y, Durrett R, Mazourek M et al (2009) A COSII genetic map of the pepper genome provides a detailed picture of synteny with tomato and new insights into recent chromosome evolution in the genus *Capsicum*. Theor Appl Genet 118:1279–1293

Xu Y, Li P, Yang Z, Xu C (2017) Genetic mapping of quantitative trait loci in crops. Crop J 5:175–184

Yeam I, Kang BC, Lindeman W, Frantz JD, Faber N et al (2005) Allele-specific CAPS markers based on point mutations in resistance alleles at the *pvr1* locus encoding eIF4E in *Capsicum*. Theor Appl Genet 112:178–186

Yi G, Lee JM, Lee S, Choi D, Kim BD (2006) Exploitation of pepper EST-SSRs and an SSR-based linkage map. Theor Appl Genet 114:113–130

Zietkiewicz E, Rafalski A, Labuda D (1994) Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. Genomics 20:176–183

6 Molecular Mapping and Identification of QTLs and Genes for Economically Important Traits in the *Capsicum* Genome

Vijee Mohan and Ilan Paran

Abstract

Pepper exhibits large phenotypic variation for economically important traits that are mostly quantitatively inherited. In this chapter, we review the quantitative trait locus (QTL) mapping studies focused on plant growth and fruit yield and quality traits. We further review recent developments of genomic resources and genotyping techniques and their utilization for construction of ultra-high-density maps of pepper including newly developed maps established for the less explored *Capsicum* species *Capsicum baccatum*. These studies allowed a comprehensive understanding of the genetic basis for regulation of these traits in pepper and the development of molecular markers linked to favorable genes and their introgression to elite backgrounds.

6.1 Introduction

Pepper consists of a vast variation in morphological traits such as fruit color, size, and shape (Fig. 6.1), fruit quality traits such as metabolic contents of phytonutrients, yield-related traits such as response to biotic and abiotic stresses and shoot growth traits such as flowering time and plant architecture. Mapping the loci governing this variation and identification of the causative genes has been done in the last thirty years by exploiting natural variation that exists within several *Capsicum* species and in a more limited scale by induced variation mostly for flowering and shoot architecture traits. Early mapping studies in pepper have been summarized by Paran et al. (2006) and by Paran (2013). More recent review focused on mapping of economically important traits in the perspective of translational research (Ramchiary et al. 2014). The present chapter will focus on recent mapping studies that utilized newly developed genomic tools as well as on QTL studies for mapping major plant growth and fruit-related traits.

6.2 Ultra-High-Density Maps

With the advent of reference genome sequences of pepper (Kim et al. 2014; Qin et al. 2014) and affordable genomic tools such as single-nucleotide polymorphism (SNP) arrays and low-coverage next-generation sequencing (NGS)-based methods,

V. Mohan · I. Paran (✉)
Institute of Plant Sciences, Agricultural Research Organization, The Volcani Center, Rishon LeZion 7528809, Israel
e-mail: iparan@volcani.agri.gov.il

V. Mohan
e-mail: vijeebiot@gmail.com

N. Ramchiary and C. Kole (eds.), *The Capsicum Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-319-97217-6_6

Fig. 6.1 Natural variation of fruit morphology in pepper

several high-density maps have been constructed and used for mapping in pepper in recent years which are described below.

In the study of Han et al. (2016), an ultra-high-density bin map was constructed in a *Capsicum annuum* intraspecific recombinant inbred line (RIL) population from the cross of Perennial and Dempsey. SNP markers were detected by low-coverage (1X) resequencing of the RILs. Over 1 million SNPs were detected between the parents and used to construct a bin map in which all SNPs within a window of defined size were regarded as a bin using the sliding window approach. Using this approach, 2578 bins were used to construct the map that spans 1372 cM. Each of the 12 chromosomes consisted of 154–370 bins per chromosome in an average density of 1 bin/0.5 cM, thus providing a highly saturated map for efficient QTL mapping. A total of 18 plant architecture, leaf, flower, and fruit traits were measured, and a total of 86 QTLs were detected in multiple environments. This study confirmed the results from other QTL analyses for the occurrence of major fruit weight and fruit shape QTLs in chromosomes 2 and 3 (Ben Chaim et al. 2001; Barchi et al. 2009; Chunthawodtiporn et al. 2018).

Two RIL populations, an intraspecific cross of *C. annuum* Early Jalapeno × CM 334 (NM) and an interspecific *Capsicum frutescens* BG2814-6 × *C. annuum* NuMex RNaky (FA) cross, were used to construct ultra-high-density maps (Hill et al. 2015). Polymorphism detection was done using a pepper GeneChip containing 31,196 unigene expressed sequence tags (EST; Ashrafi et al. 2012). In total, 3878 and 16,167 EST markers were mapped in the NM and FA populations, respectively. The markers in the two maps were clustered into 783 and 2105 bins (markers with zero recombination were considered as a single bin) in the NM and FA populations, respectively. Because the maps were based on gene-based markers, they allowed syntenic comparison between pepper and other Solanaceae species and comparative mapping of common traits. Similar to other pepper interspecific

crosses, translocation between chromosomes 1 and 8 was observed in the FA map and the high-density map allowed to precisely locate the translocation breakpoint to a specific bin in the chromosome. The use of a high number of common markers in both maps allowed to compare recombination rate and markers' distortion in an intraspecific cross compared to an interspecific cross.

A more recent SNP Illumina Infinium array consisting of 16,000 SNPs was developed as a public tool to aid pepper breeding and mapping (Hulse-Kemp et al. 2016). The SNPs were selected based on resequencing of 22 pepper lines representing chili and bell-fruited types. The utility of the array was tested by constructing a high-density map from an interspecific cross of *C. frutescens* Tabasco × *C. annuum* blocky-type P4. A total of 5546 markers were mapped into 12 linkage groups and arranged in 1362 genetic bins. The present map and the above-mentioned FA map were compared using a common set of 822 markers and found to be highly similar. Important advantages of the Infinium SNP array are low rate of missing data, accurate calling of heterozygotes, and rapid downstream processing of the raw data.

A second Illumina Infinium SNP array was developed and utilized for mapping and diversity analysis in pepper (Cheng et al. 2016). A set of 15,000 SNPs was selected based on resequencing of the cultivars, BA3 and B702, of which approximately 8200 loci were anchored to the Zunla genome assembly (Qin et al. 2014) and scored in various populations. An interspecific cross of BA3 (*C. annuum*) × YNXML (*C. frutescens*) was used to construct an F_2 mapping population. The population was genotyped with 5828 SNPs and phenotypically scored for erect/pendant fruit orientation controlled by the *up* locus which has been previously assigned to chromosome 12 (Lefebvre et al. 1995). A major locus, *Up12.1*, that controls the trait was mapped to a 4.5 Mb region containing 65 genes in the latter chromosome. Furthermore, the SNP array was used to evaluate the genetic diversity of a panel of 399 *C. annuum* elite and landrace pepper lines originated from China. The relative low genetic diversity level found within this panel indicates the need to broaden the genetic variation of the germplasm used for breeding.

Genomewide association study (GWAS) was conducted in a diverse collection of 94 *C. annuum* accessions to identify significant genomic regions affecting capsaicinoids content and fruit weight (Nimmakayala et al. 2016). SNPs' discovery and genotyping were done by the genotyping-by-sequencing (GBS) method (Elshire et al. 2011). A total of 66,960 SNPs were identified among the accessions and mapped to the reference genome of CM334 (Kim et al. 2014), of which a set of 7331 SNPs was used for the QTL study. For both traits, multiple genomic regions with relatively small effects were found to contain significant SNPs. Several significant SNPs were found in candidate genes that have related biological function in other species. For capsaicinoid content, 30 and 56 SNPs were found to be associated with capsaicin and dihydrocapsaicin, respectively; 14 SNPs were common for both traits. Both capsaicinoid content and fruit weight are important traits for pepper domestication. In accordance, many significant SNPs for these traits were located within regions in the genome that exhibits selective sweep signatures.

6.3 Genetic Mapping in *C. baccatum*

Most mapping populations in pepper have been constructed in *C. annuum* intraspecific crosses or in interspecific crosses between *C. annuum* and *Capsicum chinense* or *C. frutescens*. Few genetic studies have been performed in the other cultivated species *Capsicum baccatum* and *Capsicum pubescens*. *C. baccatum* consists of both cultivated and wild subspecies and possesses high variability of fruit-related traits and sources for resistance to diseases such as anthracnose and powdery mildew. Therefore, mapping efforts have been performed in the latter species to aid in mapping and introgression of the resistance genes and other traits for use in breeding (Kim et al. 2010; Lee et al. 2010; Eggink et al. 2014;

Mahasuk et al. 2016). To determine the genome structure of *C. baccatum* and its comparison to *C. annuum*, an intraspecific *C. baccatum* cross was used to map 395 SNPs identified by resequencing the two mapping parents (Lee et al. 2016). Comparison of the map to the *C. annuum* reference genome of CM334 revealed translocations between chromosomes 1 and 8 as previously shown in interspecific crosses in *Capsicum*. Furthermore, additional reciprocal translocations were detected between chromosomes 3 and 5 and between chromosomes 3 and 9. These translocations may act as genetic barriers between *C. baccatum* and *C. annuum* and explain the difficulties in crossing these species.

To study population diversity in *C. baccatum*, a panel of 283 and 94 accessions of *C. baccatum* and *C. annuum*, respectively, was genotyped by genotyping by sequencing (GBS; Elshire et al. 2011) and assessed for population structure, linkage disequilibrium (LD) and QTL mapping by GWAS analysis. Approximately 13,000 SNPs were detected in the *C. baccatum* panel (Nimmakayala et al. 2016). The population was phenotyped for peduncle length that differentiates cultivated and wild accessions. Significant associations were detected in 10 out of the 12 chromosomes, cumulatively explaining 21% of the variation for the trait.

The potential of using *C. baccatum* for the improvement of fruit quality was tested by introgressing multiple chromosome segments into *C. annuum* backgrounds. Multi-parent backcrossing coupled with embryo rescue allowed the construction of BC_2S_1 population from an interspecific cross of *C. annuum* and *C. baccatum* which was evaluated for attributes of fruit quality and subjected for QTL mapping (Eggink et al. 2014). Fruit phenotyping included volatile profiling, chemical composition, morphology, and sensory attributes. Subsequently, near-isogenic lines were developed to confirm QTLs detected in the BC_2S_1 population. The QTL with the strongest effect (LOD = 40.1) was detected for immature fruit color. This QTL likely corresponds to *GOLDEN2-like* (*CaGLK2*) that was identified as controlling chlorophyll content in the immature fruit (Brand et al. 2014). Sensory and metabolomic analyses allowed the identification of a QTL allele originated from *C. baccatum* that confers a strong effect on volatile content and flavor in chromosome 3. Since this QTL is mapped to a small introgression without apparent linkage drag, it is an important candidate for use in breeding for improved flavor. Additional potential sources for improved flavor are a QTL for increased content of sugars in chromosome 3 that does not coincide with reduced fruit size. Furthermore, QTLs for increased content of terpenoids were detected in chromosomes 1 and 10; their phenotypic effect on plant adaptation is yet to be determined.

6.4 QTLs for Economically Important Traits

High yield, early flowering, biotic and abiotic stress tolerance, enriched metabolite content, desired fruit size and shape and reduced postharvest water loss have been major targets for pepper improvement mostly by classical breeding efforts. More recently, molecular breeding techniques such as QTL mapping and introgression, identification of causative genes, and molecular marker development have been utilized for breeding enhancement. A compilation of QTL data for various economical traits associated with plant and fruit growth is presented in Table 6.1, and the major results are summarized in the following paragraphs.

6.4.1 Plant Growth

Shoot architectural components such as the length of the primary stem, internode length, leaf size, degree of lateral branching, and timing of flowering initiation determine the overall plant growth. QTLs for plant development in a cross of Yolo Wonder × Criollo de Morelos 334 RIL population were identified by Barchi et al. (2009). Colocalization of QTLs affecting flowering time, primary axis (stem) length, internode length, axis growth speed, and internode growth time was observed in chromosomes P2, P4, P9,

Table 6.1 List of QTLs for plant growth and fruit traits in pepper

Trait	Population	No. of QTLs	Major effect QTLs[a]	References
Plant architecture				
Axis growth speed	YW (*C. annuum*) × CM334 (*C. annuum*)	4		Barchi et al. (2009)
Internode growth time	YW (*C. annuum*) × CM334 (*C. annuum*)	2		Barchi et al. (2009)
Internode length	YW (*C. annuum*) × CM334 (*C. annuum*)	3		Barchi et al. (2009)
Internode length	Perennial (*C. annuum*) × Dempsey (*C. annuum*)	5		Han et al. (2016)
Primary axis length	YW (*C. annuum*) × CM334 (*C. annuum*)	5		Barchi et al. (2009)
Main stem length	Perennial (*C. annuum*) × Dempsey (*C. annuum*)	6		Han et al. (2016)
Plant height	CW (*C. annuum*) × LCA235 (*C. annuum*)	1	*Qpht.iivr.5.1*	Dwivedi et al. (2015)
Plant height	Perennial (*C. annuum*) × Dempsey (*C. annuum*)	6		Han et al. (2016)
Plant width	Perennial (*C.annuum*) × Dempsey (*C. annuum*)	2		Han et al. (2016)
Branching	2814-6 (*C. frutescens*) × NuMexRNAKY (*C. annuum*)	6		Yarnes et al. (2012)
Lateral branch number	Perennial (*C. annuum*) × Dempsey (*C. annuum*)	2	*LBN-2.1, LBN-2.2*	Han et al. (2016)
Trichome density	CM334 (*C. annuum*) × Chilsungcho (*C. annuum*)	11	*Ptel1, Ptel2, Ptel9,*	Kim et al. (2011)
Flowering time	YW (*C. annuum*) × CM334 (*C. annuum*)	5		Barchi et al. (2009)
Flowering time	2814-6 (C. frutescens) × NuMexRNAKY (C. annuum)	8	*2.6, 2.8*	Yarnes et al. (2012)
Flowering time	PI 527325 (*C. annuum*) × PI 511887 (*C. annuum)*	1	*Flw2.1*	Borovsky et al. (2015)
Number of leaves	CW (*C. annuum*) × LS2341 (*C. annuum*)	2	*Nle1.1, Nle12.1*	Mimura et al. (2010)
Number of leaves	YW (*C. annuum*) × CM334 (*C. annuum*)	2		Alimi et al. (2013)
Number of leaves	BA3 (*C. annuum*) × YNXML (*C. frutescens*)	6	*Nle2.2*	Tan et al. (2015)
Yield				
Number of fruits/plant	CW (*C. annuum*) × LCA235 (*C. annuum*)	1		Dwivedi et al. (2015)
Ten fruits weight	CW (*C. annuum*) × LCA235 (*C. annuum*)	1	*Qtfw.iivr-2.1*	Dwivedi et al. (2015)
Total fruit weight	CW (*C. annuum*) × LCA235 (*C. annuum*)	3	*Qtofw.iivr-1.1*	Dwivedi et al. (2015)
Fruit size/weight				
Fruit size	Maor (*C. annuum*) × Perennial (*C. annuum*)	5		Ben Chaim et al. (2001)
Fruit size	Maor(*C. annuum*) × BG 2816 (*C. frutescens*)	8		Rao et al. (2003)
Fruit weight	PI 152225 (*C. chinense*) × 100/63 (*C. annuum*)	3	*fw2.1, fw4.1, fw4.2*	Zygier et al. (2005)
Fruit weight	NuMex Rnaky (*C. annuum*) × BG 2814-6 (*C. frutescens*)	2	*fw2.1, fw3.1*	Ben Chaim et al. (2006)
Fruit weight	YW (*C. annuum*) × CM334 (*C. annuum*)	7		Barchi et al. (2009)
Fruit weight	PEN45 (*C. baccatum)* × SM,GNM (*C. annuum*)	3	*LG1_8*	Eggink et al. (2014)
Fruit weight	CW (*C. annuum*) × LCA235 (*C. annuum*)	1	*Qfw.iivr-2.1*	Dwivedi et al. (2015)
Fruit weight	94 accessions of *C. annuum*	16		Nimmakayala et al. (2016)
Fruit weight	Perennial (*C. annuum*) × Dempsey (*C.annuum*)	6		Han et al. (2016)

(continued)

Table 6.1 (continued)

Trait	Population	No. of QTLs	Major effect QTLs[a]	References
Pericarp				
Pericarp thickness	YW (*C. annuum*) × CM334 (*C. annuum*)	8		Barchi et al. (2009)
Pericarp thickness	2814-6 (*C. frutescens*) × NuMexRNAKY (*C. annuum*)	10	*4.4*	Yarnes et al. (2012)
Pericarp thickness	Early Jalapeno (*C. annuum*) × CM334 (*C. annuum*)	1	*3.1*	Naegele et al. (2014)
Pericarp thickness	CW (*C. annuum*) × LCA235 (*C. annuum*)	1	*Qpt.iivr-2.1*	Dwivedi et al. (2015)
Pericarp area	2814-6 (*C. frutescens*) × NuMexRNAKY (*C. annuum*)	8		Yarnes et al. (2012)
Fruit shape				
Fruit diameter	YW (*C. annuum*) × CM334 (*C. annuum*)	12		Barchi et al. (2009)
Fruit diameter	Perennial (*C. annuum*) × Dempsey (*C. annuum*)	5	*FD-1, FD-3.2*	Han et al. (2016)
Fruit width	PEN45 (*C. baccatum*) × SM,GNM (*C. annuum*)	7	*LG1_8, LG9*	Eggink et al. (2014)
Fruit length	YW (*C. annuum*) × CM334 (*C. annuum*)	4	*Frl4.1*	Barchi et al. (2009)
Fruit length	TF68 (*C. annuum*) × Habanero (*C. chinense*)	5	*3.1*	Lee et al. 2011
Fruit length	PEN45 (*C. baccatum*)X SM,GNM (*C. annuum*)	4	*LG10.1*	Eggink et al. (2014)
Fruit length	CW (*C. annuum*) × LCA235 (*C. annuum*)	2	*Qfl.iivr.3.2*	Dwivedi et al. (2015)
Fruit length	Perennial (*C. annuum*) × Dempsey (*C. annuum*)	6	*FL-3.1, FL-3.2, FL-3.3*	Han et al. (2016)
Fruit shape index	Maor (*C. annuum*) × Perennial (*C. annuum*)	3	*fs3.1*	Ben Chaim et al. (2001)
Fruit shape index	Maor (*C. annuum*) × BG 2816 (*C. frutescens*)	5	*fs3.1*	Rao et al. (2003)
Fruit shape index	5226 (*C. annuum*) × PI 159234 (*C. chinense*)	1	*fs10.1*	Ben Chaim et al. (2003a)
Fruit shape index	PI 152225 (*C. chinense*) × 100/63 (*C. annuum*)	3	*fs4.2*	Zygier et al. (2005)
Fruit shape index	YW (*C. annuum*) × CM334 (*C. annuum*)	7		Barchi et al. (2009)
Fruit shape index	1154 (*C. annuum*) × PI 152225 (*C. chinense*)	3	*fs1.1, fs10.1*	Borovsky and Paran (2011)
Fruit shape index	2814-6 (*C. frutescens*) × NuMexRNAKY (*C. annuum*)	51	*2.5, 2.6,2.8, 2.10, 4.4, 11.4*	Yarnes et al. (2012)
Fruit shape index	Early Jalapeno (*C. annuum*) × CM334 (*C. annuum*)	5	*4.1*	Naegele et al. (2014)
Fruit shape index	PEN45 (*C. baccatum*) × SM,GNM (*C. annuum*)	5	*LG1_8*	Eggink et al. (2014)
Fruit shape index	Perennial (*C. annuum*) × Dempsey (*C. annuum*)	4	*FS-3.1, FS-3.2*	Han et al. (2016)
Fruit color				
Color ripe	PEN45 (*C. baccatum*) × SM,GNM (*C. annuum*)	3		Eggink et al. (2014)
Color unripe	PEN45 (*C. baccatum*) × SM,GNM (*C. annuum*)	3	*LG10.1*	Eggink et al. (2014)
Immature fruit color	Perennial (*C. annuum*) × Dempsey (*C. annuum*)	6	*IFC-10.2*	Han et al. (2016)
Metabolites				
Biochemical composition	PEN45 (*C. baccatum*) × SM,GNM (*C. annuum*)	8	*LG1_8, LG10.1*	Eggink et al. (2014)
Capsaicin	*NB1 (C. annuum)* × Bhut Jolokia (*C. chinense*)	3	*Qcap3.1, qcap6.1*	Lee et al. (2016)
Capsaicin	NuMex Rnaky (*C. annuum*) × BG 2814-6 (*C. frutescens*)	5	*cap7.2*	Ben Chaim et al. (2006)

(continued)

Table 6.1 (continued)

Trait	Population	No. of QTLs	Major effect QTLs[a]	References
Capsaicinoid	2814-6 (*C. frutescens*) × NuMexRNAKY (*C. annuum*)	12	*4.2, 4.14, 4.15*	Yarnes et al. (2012)
Capsaicinoid	Maor (*C. annuum*) × BG 2816 (*C. frutescens*)	1	*cap*	Blum et al. (2003)
Capsaicinoids	94 accessions of *C. annuum*	14		Nimmakayala et al. (2016)
Dihydrocapsaicin	*C. annuum* 'NB1' × *C. chinense* 'Bhut Jolokia'	2	*Qhdc2.1, qdhc2.2*	Lee et al. (2016)
Dihydrocapsaicin	NuMex Rnaky (*C. annuum*) × BG 2814-6 (*C. frutescens*)	4	*dhc4.1*	Ben Chaim et al. (2006)
Nordihydrocapsaicin	NuMex Rnaky (*C. annuum*) × BG 2814-6 (*C. frutescens*)	1	*ndhc7a.1*	Ben Chaim et al. (2006)
Total capsaicinoids	NuMex Rnaky (*C. annuum*) × BG 2814-6 (*C. frutescens*)	5	*total7.2*	Ben Chaim et al. (2006)
Chlorophyll content	1154 (*C. annuum*) × PI 152225 (*C. chinense*)	2	*pc8.1*	Brand et al. (2012)
Metabolites	AC1979 (*C. annuum*) × No. 4661 (*C. chinense*)	279		Wahyuni et al. (2014)
Other fruit traits				
Number of locules	YW (*C. annuum*) × CM334 (*C. annuum*)	9	*NloLG25.1*	Barchi et al. (2009)
Pedicel length	YW (*C. annuum*) × CM334 (*C. annuum*)	6		Barchi et al. (2009)
Fruit firmness	Early Jalapeno (*C. annuum*) × CM334 (*C. annuum*)	1	*12.1*	Naegele et al. (2014)
Fruit position	Perennial (*C. annuum*) × Dempsey (*C. annuum*)	6	*FP-12.1,FP-12.2, FP-12.3*	Han et al. (2016)
Fruit orientation	BA3 (*C. annuum*) × YNXML (*C. frutescens*)	1	*Up12.1*	Cheng et al. (2016)
Postharvest water loss	1154 (*C. annuum*) × PI 593611 (*C. chinense*)	3	*PWL10.1, PWL10.2*	Popovsky-Sarid et al. (2017)

[a]QTL with effect larger than 20% explained phenotypic variation

and LG47 which may indicate pleiotropic effects of these QTLs. Twelve QTLs for several plant growth traits were detected in an intraspecific *C. annuum* doubled haploid (DH) population, explaining 14–34% of the phenotypic variability for various plant architectural traits (Mimura et al. 2010). The strongest QTL effect was detected for flowering time in *LG8* that was not conclusively assigned to a specific chromosome. QTLs for pepper growth traits such as leaf size, plant height, flowering time, days to breaker fruit stage, and branching were also mapped in an RIL population from an interspecific cross of 2814-6 (*C. frutescens*) × NuMex RNaky (*C. annuum*). A total of 23, 15 and 17 QTLs mostly with minor effects were identified for leaf traits, floral traits, and whole plant morphology, respectively (Yarnes et al. 2012). Additionally, a major QTL for plant height was reported in *LG5* (Dwivedi et al. 2015), and two major QTLs for lateral branch number were reported in chromosome 2 (Han et al. 2016).

Several studies have focused specifically on mapping flowering time QTLs. A major QTL for flowering time was identified in chromosome 2 in a cross of *C. annuum* blocky-fruited type accession (early flowering) and *C. annuum* var. *glabriusculum* wild accession (late flowering) (Fig. 6.2, Borovsky et al. 2015). This QTL was co-localized with the flowering suppressor gene *CaAPETALA2* that is disrupted in the EMS (ethylmethane sulfonate)-induced early flowering mutant. The same genomic region has been detected as a major flowering time QTL in an independent study in a *C. annuum* × *C. frutescens* cross (Tan et al. 2015). Additional five minor QTLs were detected in the latter study. In addition to these QTLs, several flowering

promoter and suppressor genes that also control shoot architecture were identified by using EMS-induced mutants including *CaJOINTLESS* (Cohen et al. 2012), *CaBLIND* (Jeifetz et al. 2011), *CaS* (Cohen et al. 2014) and *CaFASCICULATE* (Elitzur et al. 2009).

Trichome density on the plant stem and flower calyx varies among accessions and are often present in wild accessions and landraces. QTL mapping for the trait was performed in intraspecific *C. annuum* populations involving CM334 as a trichome-rich parent. A major QTL, *Ptel1*, controlling trichome density in the main stem and in the calyx was detected in LG24 corresponding to Chromosome 10 (Kim et al. 2011). Additional 10 minor QTLs were detected in other chromosomes. Recently, an RIL population from a cross of CM334 × Maor was analyzed for several fruit and growth traits including the degree of stem pubescence that was scored by a visual scan (Chunthawodtiporn et al. 2018). Three QTLs were detected in chromosomes 2, 10, and 11, the QTL in chromosome 10 having the largest effect on the trait which likely corresponds to *Ptel1*. Two candidate genes, *TRICHOME BIREFRINGENCE-LIKE 5* and a *C2H2* zinc-finger transcription factor which are putatively involved in the formation of trichomes based on Arabidopsis studies, were located in the vicinity of the QTL.

6.4.2 Fruit Traits

6.4.2.1 Fruit Size and Yield

Fruit size/weight QTLs have been identified in multiple studies (Table 6.1 and summarized by Paran and van der Knaap 2007; Hill et al. 2017). Several QTL studies in mapping populations consisting of crosses of a common blocky-fruited cv. Maor (*C. annuum*) with small-fruited *C. annuum*, *C. frutescens,* and *C. chinense* accessions have been carried out. Two major QTLs for fruit weight, *fw2.1* and *fw4.2,* are conserved in the three *Capsicum* species. *fw2.1* had the most significant effect in multiple populations (Ben Chaim et al. 2001; Rao et al. 2003; Zygier et al. 2005). A putative tomato orthologous fruit weight QTL, *fw2.2*, is located in a syntenic region in chromosome 2 (Frary et al. 2000). The gene that underlies *fw2.2* in tomato is *CELL NUMBER REGULATOR* (*CNR*). However, the syntenic region in pepper consists of multiple genes associated with organ size regulation including the ortholog of *OVATE*, a fruit shape gene in tomato (Hill et al. 2017). Therefore, high-resolution mapping will be required to precisely map this QTL and identify the underlying gene.

Another possible orthologous fruit weight QTL in pepper and tomato is *fw3.2* that is associated with the gene *KLUH*, a P450 coding enzyme in both species (Chakrabarti et al. 2013). A cluster of minor QTLs for fruit weight, fruit shape, fruit diameter, and pericarp thickness is located in *P11* and *P12* (Barchi et al. 2009). 16 significant SNPs associated with fruit weight were identified in a GWAS study of 94 accessions (Nimmakayala et al. 2016). Except for chromosome 7, all other chromosomes had at least one significant SNP. Out of the 16 SNPs, seven were located in known genes that control organ size such as *STYLOSA*, *FASCIATED*, *WUSCHEL*, and *CLAVATA1*.

The yield of pepper is affected by parameters such as number of fruits per plant, fruits weight, and total fruit yield. QTL mapping for these traits was performed in a *C. annuum* intraspecific RIL population (Dwivedi et al. 2015). A total of 10 QTLs were detected for yield-related traits. Colocalization of five QTLs in chromosome 2 (*Qtofw.iivr-2.1*, *Qtfw.iivr-2.1*, *Qfw.iivr-2.1*, *Qnfp.iivr-2.1,* and *Qpt.iivr-2.1*) with significant additive effects was identified which might be due to the linkage of different QTLs or pleiotropic effects of the same genes. Other QTL studies for fruit-related traits in pepper reported clustering of QTLs for fruit traits in the same region of chromosome 2 (Ben Chaim et al. 2001; Rao et al. 2003; Zygier et al. 2005; Barchi et al. 2009; Chunthawodtiporn et al. 2018).

Pericarp thickness is positively correlated with fruit weight (Ben Chaim et al. 2001), and therefore, QTLs for both traits are often located in common genomic positions. Two major QTLs for pericarp thickness were identified in different

intraspecific populations of *C. annuum* in chromosomes 2 and 3 (*Qpt.iivr-2.1, 3.1*). *Qpt.iivr-2.1* is located in the same genomic region in chromosome 2 that contains a QTL for fruit weight (Dwivedi et al. 2015; Naegele et al. 2014). Several linked QTLs for pericarp thickness were identified in chromosome 4 in a cross of *C. annuum* and *C. frutescens* (Yarnes et al. 2012). Several minor QTLs for pericarp thickness were also identified by Barchi et al. (2009).

An additional factor that may be associated with fruit size is locule number. The locule number locus, *lcn2.1,* in tomato affects fruit size via changing carpal numbers (Lippman and Tanksley 2001). Only limited data is available for mapping this trait in pepper. Low positive correlation between fruit weight and locule number was reported by Barchi et al. (2009). Few minor QTLs for locule number were identified in this cross; the strongest QTL being *NloLG25.1*.

6.4.2.2 Fruit Shape

Despite the large variation in fruit shape that exists in pepper, only few shape attributes were studied that include fruit width, fruit length, and fruit shape index (length/width). The two most significant QTLs for fruit shape index were *fs3.1* and *fs10.1* that have been identified as controlling fruit elongation (Ben Chaim et al. 2001, 2003a, b; Borovsky and Paran 2011). Both QTLs control most of the trait variation in *Capsicum*, explaining 67 and 44% of the phenotypic variation for *fs3.1* and *fs10.1*, respectively (Ben Chaim et al. 2003a, b). Major QTLs regulating fruit shape variation within *C. annuum* have been identified in multiple populations in chromosomes 1, 3, and 4 (Ben Chaim et al. 2001; Barchi et al. 2009; Naegele et al. 2014; Dwivedi et al. 2015; Han et al. 2016). Fruit shape QTLs were also identified in interspecific crosses between *C. annuum* and *C. chinense* in chromosomes 1, 3, 4, and 10 (Ben Chaim et al. 2003a; Zygier et al. 2005; Borovsky and Paran 2011). Four major QTLs in chromosome 4 and one QTL in each of the chromosomes 3, 4, and 11 were identified in a cross of *C. annuum* and *C. frutescens* (Yarnes et al. 2012). Fruit shape QTLs in chromosomes 10 and *LG1_8* have been identified in a cross of *C. annuum* and *C. baccatum* (Eggink et al. 2014). To-date, none of the genes underlying natural fruit shape variation in pepper has been identified. A pepper homolog of *OVATE,* a tomato fruit shape QTL (Liu et al. 2002), was shown to be associated with fruit shape variation by down regulation using virus-induced gene silencing (VIGS; Tsaballa et al. 2011), indicating the possibility that the pepper gene may regulate fruit shape variation in natural populations.

6.4.2.3 Fruit Color and Chlorophyll Content

Since the fruit pigments are associated with fruit color, nutrition, and flavor, pepper fruit color is important to the breeder and to the consumer. The color of the ripe fruits is determined primarily by carotenoids and that of the immature fruits by anthocyanins and chlorophyll. Mutations in genes in the carotenoid biosynthesis pathway result in change of color from red to yellow or to orange (Popovsky and Paran 2000; Thorup et al. 2000; Huh et al. 2001; Borovsky et al. 2013). Additional pepper fruit color variation is associated with chlorophyll catabolism, anthocyanin accumulation and chloroplast compartment size (Borovsky et al. 2004; Borovsky and Paran 2008; Pan et al. 2013).

Both qualitative and quantitative variation in fruit pigment content exists, however, its genetic control in largely unknown. QTL analysis for chlorophyll content in the immature fruit in a cross between a dark green-fruited *C. annuum* inbred 1154 and a light green-fruited *C. chinense* accession PI 152225 revealed the presence of two major QTLs, *pc8.1* and *pc10.1*, that control the trait (Fig. 6.3, Brand et al. 2012). One of the QTLs, *pc10*, was found to correspond to the pepper homolog of *GOLDEN2-like* transcription factor (*GLK2*) that controls chloroplast compartment size in the immature fruit (Brand et al. 2014). Major QTLs that regulates unripe fruit color were identified in chromosome 10 in an interspecific cross of *C. annuum* and *C. baccatum* and in an intraspecific *C. annuum* cross which likely correspond to *pc10.1* (Eggink et al. 2014; Han et al. 2016).

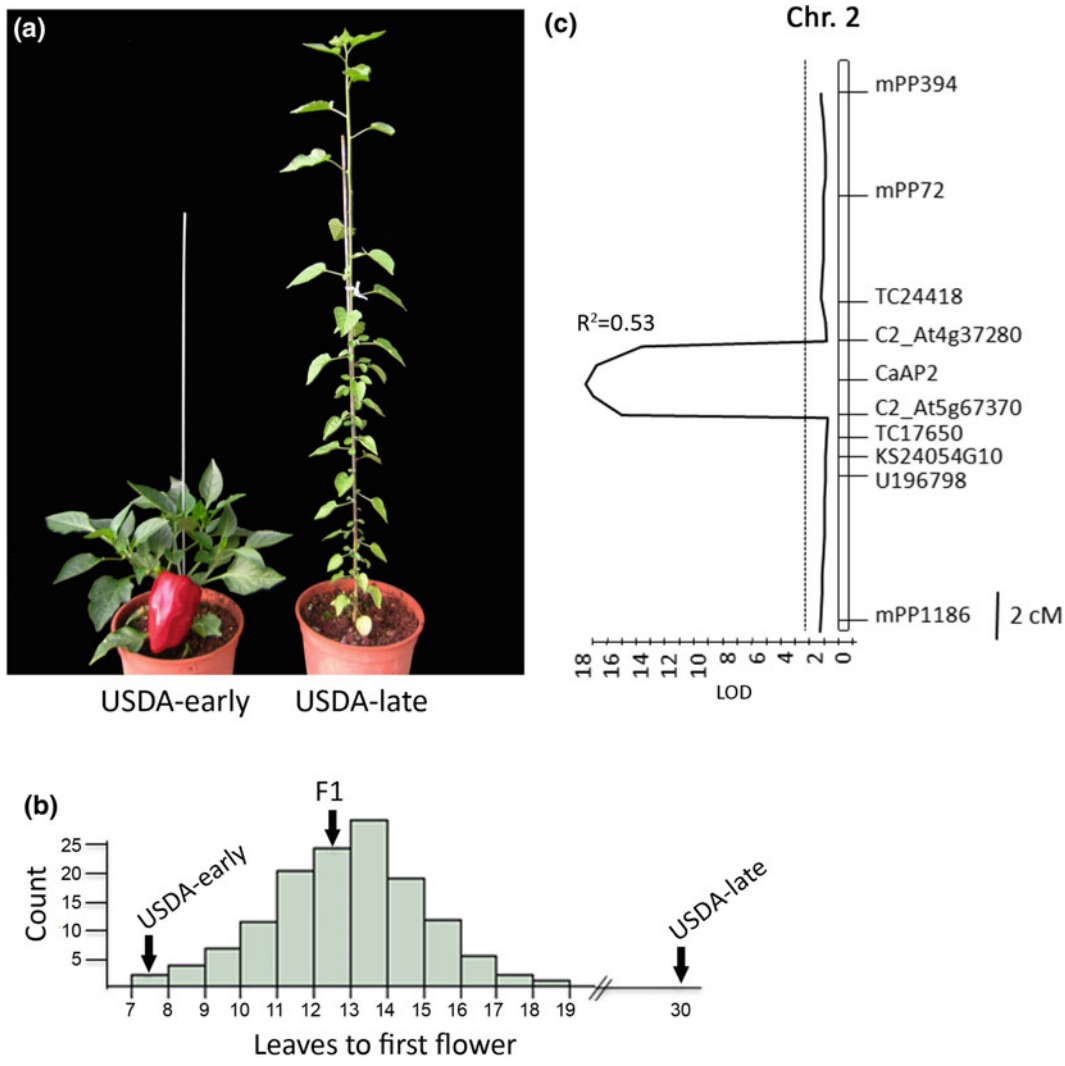


Fig. 6.2 QTL mapping of flowering time in a cross of *C. annuum* accession PI 527325 (USDA early) and *C. annuum* var. *glabriusculum* wild accession PI 511887 (USDA late). **a** Pictures of the early- and late-flowering parents used for QTL mapping. **b** Distribution of flowering time in the F_2 population. **c** Interval QTL mapping of flowering time in a region containing *CaAP2* in chromosome 2. Reprinted from Borovsky et al. (2015) by permission

6.4.2.4 Metabolites Content

A QTL study for metabolites content associated with flavor was performed in an interspecific cross between *C. annuum* and *C. baccatum* (Eggink et al. 2014). A strong effect on flavor was found in a small introgression of chromosome 3. This QTL explained 38.7% of the variation for odor and was associated with an intense odor of *C. baccatum.* NILs for this QTL showed an increase in intensity of the compound 6-methyl-4-oxo-5-heptenal and decrease of the compound (Z)-butanoic acid 3-hexenyl ester and 2-isobutyl-3-methoxypyrazine. Additional minor QTLs associated with sensory attributes were detected in different chromosomes. Two major QTLs that control variation in biochemical composition were identified in *LG1_8* and *LG10.1*. A significant QTL was found for brix,

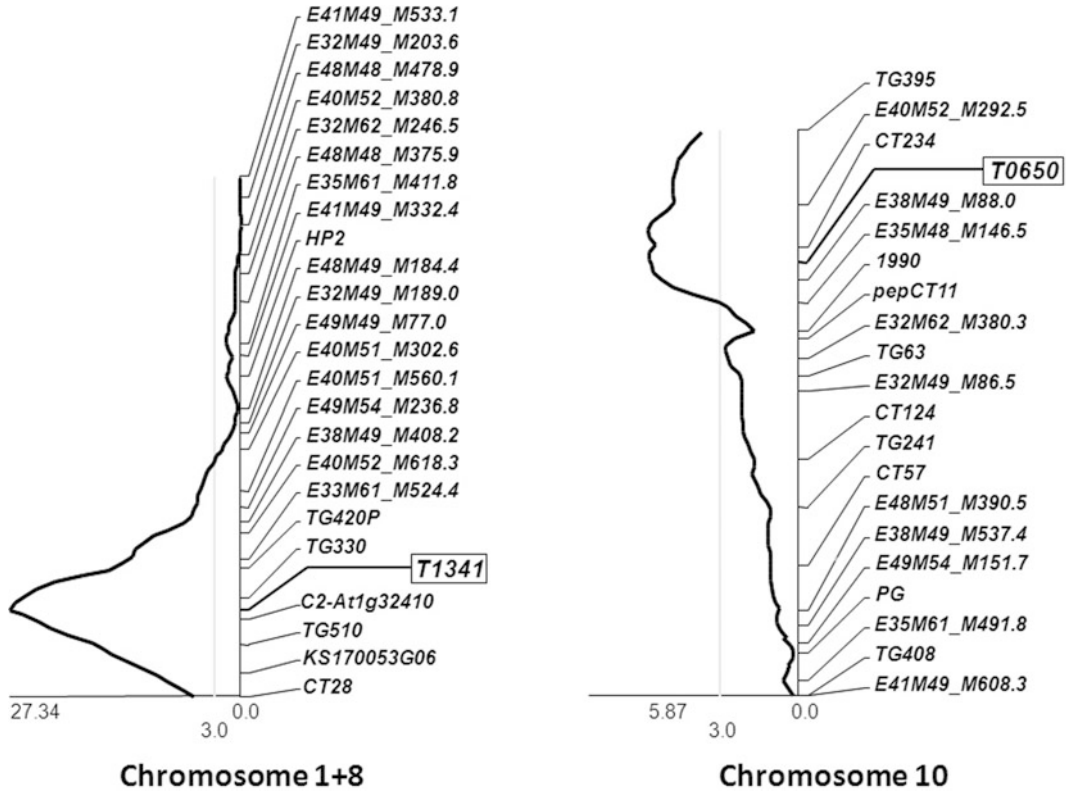


Fig. 6.3 Position of QTLs for pigment content in the F_2 of 1154 × PI 152225. The most significant markers at the two QTLs are boxed. Numbers on the horizontal axis represent LOD values. Reprinted from Brand et al. (2012) by permission

glucose, fructose, malate, and citrate in *LG1_8* that coincided with a major fruit size QTL. Because the *C. baccatum* allele was associated with an increased metabolite content and decreased fruit size, it was concluded that the increase in the metabolites concentration was a result of smaller fruits and not an effect of increased metabolism. In contrast, a QTL for increased BRIX value was detected in LG3 which was unaffected by fruit size. 46% of the variation in the metabolic content between the two parents was due to a group of 15 terpenoids controlled by QTLs in *LG10.1* and *LG1*.

The major QTL, *pc8.1*, affecting chlorophyll content in the immature fruit was also associated with an increase of other metabolites accumulated in the chloroplast such as tocopherols and carotenoids (Brand et al. 2012). This association is likely due to the effect of the QTL in modulating chloroplast compartment size. A QTL study was performed to dissect the molecular basis for variation in flavonoid content in a cross between *C. annuum* and *C. chinense* (Wahyuni et al. 2014). LCMS metabolic profiling of semi-polar metabolites allowed the identification of 52 annotated metabolites. A total of 279 mQTL were detected; however, most QTLs were clustered in few chromosomal regions creating QTL hotspots in chromosome 9. Furthermore, genes controlling flavonoids biosynthesis were mapped and some exhibited colocalization with mQTLs in chromosomes 1, 6, and 9.

6.4.2.5 Pungency

Pungency in pepper fruit is due to the unique accumulation of alkaloid compounds termed capsaicinoids. A single dominant gene at the *Pun1* locus in chromosome 2 is required for the

production of capsaicinoids (Stewart et al. 2005). In addition for the qualitative difference in the presence or absence of pungency, large variation in the capsaicinoid content exists which result in cultivars with varying degree of pungency. Several studies on QTL mapping for capsaicinoid content have been performed in diverse genetic backgrounds. Twelve QTLs were identified in six chromosomes by Yarnes et al (2012). Several QTLs were detected in chromosome 4, similar to Ben Chaim et al. (2006). A large effect QTL detected by bulked segregant analysis (BSA) was mapped in chromosome 7 (Blum et al. 2003). 14 significant SNPs scattered throughout the genome were associated with capsaicin and dihydrocapsaicin content in a GWAS study of 94 accessions (Nimmakayala et al. 2016). A QTL mapping study was also conducted in a cross of one of the hottest chili peppers 'Bhut Jolokia' (Lee et al. 2016). Two QTLs for capsaicin content were detected in chromosomes 3 and 6, while two different QTLs for dihydrocapsaicin content were detected in chromosome 2. A study conducted using a diversity panel of 40 lines consisting of 21 pungent and 19 non-pungent lines revealed several fixed regions for non-pungency (NP). Out of the 17 fixed regions for NP, 14 are overlapped with QTLs for fruit size or shape. The most significant fixed regions were located in chromosome 2, spanning *PUN1* and *CaOVATE*. In addition to *PUN1,* six genes regulating capsaicin biosynthesis were located in NP regions in chromosomes 1, 3, and 6, implicating their importance in breeding of non-pungent cultivars (Hill et al. 2017). The large variation in QTLs positions in different genetic backgrounds implicates the complexity of this trait and that markers used for selection in breeding programs will have to be developed in a genotype-specific manner.

6.4.2.6 Fruit Postharvest Water Loss

Fruit postharvest water loss (PWL) results in reduction in the overall fruit quality and thus affects the marketing of peppers. Based on screening of a wide germplasm for variation in the trait, two parents that exhibited large difference in PWL were selected for QTL mapping in an interspecific cross between *C. annuum* and *C. chinense*. Two linked QTLs, *PWL10.1* and *PWL10.2*, were identified for fruit PWL in chromosome 10 in multiple generations (Popovsky-Sarid et al. 2017). Several genes associated with cuticle biosynthesis, cell wall metabolism, and fruit ripening were identified as QTL candidates using transcriptome analysis of near-isogenic line (NILs) that differ for the QTL.

6.5 Concluding Remarks and Future Prospects

In recent years, numerous QTL studied have been conducted for economically important traits in pepper which can be exploited for introgression of beneficial QTL alleles into elite lines. The deciphering of the pepper genome sequence allowed searching for candidate genes that co-localize with the QTLs; however, only in few cases the causative genes underlying the QTLs were unequivocally identified. High-resolution mapping, expression studies, and functional assays for candidate genes will be required to expedite the discovery of such genes. One obstacle that hinders functional genomic studies in pepper is the lack of an efficient transformation system. The recent report on successful application of pollen-mediated transformation in pepper (Zhao et al. 2017) may open the way for a large-scale use of genome editing techniques in this species.

References

Alimi NA, Bink M, Dieleman JA, Magán JJ, Wubs AM et al (2013) Multi-trait and multi-environment QTL analyses of yield and a set of physiological traits in pepper. Theor Appl Genet 126:2597–2625

Ashrafi H, Hill T, Stoffel K, Kozik A, Yao J et al (2012) De novo assembly of the pepper transcriptome (*Capsicum annuum*): a benchmark for *in silico* discovery of SNPs, SSRs and candidate genes. BMC Genom 13:571

Barchi L, Lefebvre V, Sage-Palloix A-M, Lanteri S, Palloix A (2009) QTL analysis of plant development and fruit traits in pepper and performance of selective phenotyping. Theor Appl Genet 118:1157–1171

Ben Chaim A, Paran I, Grube R, Jahn M, van Wijk R, Peleman J (2001) QTL mapping of fruit related traits in pepper (*Capsicum annuum*). Theor Appl Genet 102:1016–1028

Ben Chaim A, Borovsky E, De Jong W, Paran I (2003a) Linkage of the *A* locus for the presence of anthocyanin and *fs10.1*, a major fruit-shape QTL in pepper. Theor Appl Genet 106:889–894

Ben Chaim A, Borovsky E, Rao GU, Tanyolac B, Paran I (2003b) *fs3.1*: a major fruit shape QTL conserved in *Capsicum*. Genome 46:1–9

Ben Chaim A, Borovsky Y, Falise M, Mazourek M, Kang BC et al (2006) QTL analysis for capsaicinoid content in *Capsicum*. Theor Appl Genet 113:1481–1490

Blum E, Mazourek M, O'Connell M, Curry J, Thorup T et al (2003) Molecular mapping of capsaicinoid biosynthesis genes and QTL analysis for capsaicinoid content in *Capsicum*. Theor Appl Genet 108:79–86

Borovsky Y, Paran I (2008) Chlorophyll breakdown during pepper fruit ripening in the *chlorophyll retainer* mutation is impaired at the homolog of the senescence-inducible stay-green gene. Theor Appl Genet 117:235–240

Borovsky Y, Paran I (2011) Characterization of *fs10.1*, a major QTL controlling fruit elongation in *Capsicum*. Theor Appl Genet 123:657–665

Borovsky Y, Oren Shamir M, Ovadia R, De Jong W, Paran I (2004) The *A* locus that controls anthocyanin accumulation in pepper encodes a MYB transcription factor homologous to *Anthocyanin2* of Petunia. Theor Appl Genet 109:23–29

Borovsky Y, Tadmor Y, Bar E, Meir A, Lewinsohn E, Paran I (2013) Induced mutation in *BETA-carotene hydroxylase* results in accumulation of beta-carotene and conversion of red to orange color in pepper fruit. Theor Appl Genet 126:557–565

Borovsky Y, Sharma VK, Verbakel H, Paran I (2015) *CaAP2* transcription factor is a candidate gene for a flowering repressor and a candidate for controlling natural variation of flowering time in *Capsicum annuum*. Theor Appl Genet 128:1073–1082

Brand A, Borovsky Y, Meir S, Rogachev I, Aharoni A, Paran I (2012) *pc8.1*, a major QTL for pigment content in pepper fruit, is associated with variation in plastid compartment size. Planta 235:579–588

Brand A, Borovsky Y, Hill T, Rahman KA, Bellalou A, Van Deynze A, Paran I (2014) *CaGLK2* regulates natural variation of chlorophyll content and fruit color in pepper fruit. Theor Appl Genet 127:2139–2148

Chakrabarti M, Zhang N, Sauvage C, Munos S et al (2013) A cytochrome 450 regulates a domestication trait in cultivated tomato. Proc Natl Acad Sci USA 110:17125–17130

Cheng J, Qin C, Tang X, Zhou H, Hu Y et al (2016) Development of a SNP array and its application to genetic mapping and diversity assessment in pepper (*Capsicum* spp.). Sci Rep 13:33293

Chunthawodtiporn J, Hill T, Stoffel K, Van Deynze A (2018) Quantitative trait loci controlling fruit size and other horticultural traits in bell pepper (*Capsicum annuum*). Plant Genome 11. https://doi.org/10.3835/plantgenome2016.12.0125

Cohen O, Borovsky Y, David-Schwartz R, Paran I (2012) *CaJOINTLESS* is a MADS-box gene involved in suppression of vegetative growth in all shoot meristems in pepper. J Exp Bot 63:4947–4957

Cohen O, Borovsky Y, David-Schwartz R, Paran I (2014) *Capsicum annuum* S (*Ca*S) promotes reproductive transition and is required for flower formation in pepper (*Capsicum annuum*). New Phytol 202:1014–1023

Dwivedi N, Kumar R, Paliwal R, Kumar U, Kumar S et al (2015) QTL mapping for important horticultural traits in pepper (*Capsicum annuum* L.). J Plant Biochem Biotechnol 24:154–160

Eggink PM, Tikunov Y, Maliepaard C, Haanstra JP, De Rooij H et al (2014) Capturing flavors from *Capsicum baccatum* by introgression in sweet pepper. Theor Appl Genet 127:373–390

Elitzur T, Nahum H, Borovsky Y, Pekker I, Eshed Y, Paran I (2009) Co-ordinated regulation of flowering time, plant architecture and growth by *FASCICULATE*: the pepper orthologue of *SELF PRUNING*. J Exp Bot 60:869–880

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K et al (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE 6(5):e19379

Frary A, Nesbitt TC, Frary A, Grandillo S, van der Knapp E et al (2000) *fw2.2*: a quantitative trait locus key to the evolution of tomato fruit size. Science 289:85–88

Han K, Jeong HJ, Yang HB, Kang SM, Kwon JK (2016) An ultra-high-density bin map facilitates high-throughput QTL mapping of horticultural traits in pepper (*Capsicum annuum*). DNA Res 6:81–91

Hill T, Ashrafi H, Chin-Wo SR, Stoffel K, Truco MJ (2015) Ultra-high density, transcript-based genetic maps of pepper define recombination in the genome and synteny among related species. Gene Genet Genom 5:2341–2355

Hill TA, Chunthawodtiporn J, Ashrafi H, Stoffel K, Weir A, Van Deynze A (2017) Regions underlying population structure and the genomics of organ size determination in *Capsicum annuum*. Plant Genome 10:3. https://doi.org/10.3835/plantgenome2017.03.0026

Huh JH, Kang BC, Nahm SH (2001) A candidate gene approach identified phytoene synthase as the locus for mature fruit color in red pepper (*Capsicum* spp.). Theor Appl Genet 102:524–530

Hulse-Kemp AM, Ashrafi H, Plieske J, Lemm J, Stoffel K et al (2016) A HapMap leads to a *Capsicum annuum* SNP infinium array: a new tool for pepper breeding. Hort Res 3:16036

Jeifetz D, David-Schwartz R, Borovsky Y, Paran I (2011) *CaBLIND* regulates axillary meristem initiation and transition to flowering in pepper. Planta 234:1227–1236

Kim S, Kim KT, Kim DH, Yang EY, Cho MC et al (2010) Identification of quantitative trait loci associated with anthracnose resistance in chili pepper (*Capsicum* spp.). Korean J Hort Sci Technol 28:1014–1024

Kim HJ, Han JH, Kim S, Lee HR, Shin JS et al (2011) Trichome density of main stem is tightly linked to PepMoV resistance in chili pepper (*Capsicum annuum* L.). Theor Appl Genet 122:1051–1058

Kim S, Park M, Yeom SI, Kim YM, Lee JM, Lee HA et al (2014) Genome sequence of the hot pepper provides insights into the evolution of pungency in *Capsicum* species. Nat Genet 46:270–278

Lee J, Hong JH, Do JW, Yoon JB (2010) Identification of QTLs for resistance to anthracnose to two *Colletotrichum* species in pepper. J Crop Sci Biotechnol 13:227–233

Lee HR, Kim KT, Kim HJ, Han JH, Kim JH (2011) QTL analysis of fruit length using rRAMP, WRKY, and AFLP markers in chili pepper. Hort Environ Biotechnol 52:602–613

Lee J, Park SJ, Hong SC, Han JH, Doil C, Yoon JB (2016) QTL mapping for capsaicin and dihydrocapsaicin content in a population of *Capsicum annuum* 'NB1' × *Capsicum chinense* 'Bhut Jolokia'. Plant Breed 135:376–383

Lefebvre V, Palloix A, Caranta C, Pochard E (1995) Construction of an intraspecifc integrated linkage map of pepper using molecular markers and doubled-haploid progenies. Genome 38:112–121

Lippman Z, Tanksley S-D (2001) Dissecting the genetic pathway to extreme fruit size in tomato using a cross between the small fruited wild species *L. pimpinellifolium* and *L. esculentum* var. Giant Heirloom. Genetics 158:413–422

Liu J, Van Eck J, 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

Mahasuk P, Struss D, Mongkolporn O (2016) QTLs for resistance to anthracnose identified in two *Capsicum* sources. Mol Breed 36:10. https://doi.org/10.1007/s11032-016-0435-5

Mimura Y, Minamiyama Y, Sano H (2010) Mapping for axillary shooting, flowering date, primary axis length, and number of leaves in pepper (*Capsicum annuum*). J Jpn Soc Hort Sci 79:56–63

Naegele RP, Ashraffi H, Hill TA, Chin-Wo SR, Van Deynze A, Hausbeck MK (2014) QTL mapping of fruit rot resistance to the plant pathogen *Phytophthora capsici* in a recombinant inbred line *Capsicum annuum* population. Phytopathology 104:479–483

Nimmakayala P, Abburi VL, Saminathan T, Alparthi SB, Almeida A, Davenport B et al (2016) Genome-wide diversity and association mapping for capsaicinoids and fruit weight in *Capsicum annuum* L. Sci Rep 6:38081. https://doi.org/10.1038/srep3808

Pan Y, Bradley G, Pyke K, Ball G, Lu C et al (2013) Network inference analysis identifies an *APRR2-Like* gene linked to pigment accumulation in tomato and pepper fruits. Plant Physiol 161:1476–1485

Paran I (2013) Molecular linkage maps of *Capsicum*. In: Kang BC, Kole C (eds) Genetics, genomics and breeding of peppers and eggplants. CRC Press, Boca Raton, FL, USA, pp 40–55

Paran I, van der Knaap E (2007) Genetic and molecular regulation of fruit and plant domestication traits in tomato and pepper. J Exp Bot 58:3841–3852

Paran I, Ben Chaim A, Kang BC, Jahn M (2006) Capsicum. In: Kole C (ed) Genome mapping and molecular breeding, vol 5. Vegetables. Springer, Heidelberg, Berlin, New York, Tokyo, pp 209–226

Popovsky S, Paran I (2000) Molecular analysis of the *Y* locus in pepper: its relation to capsanthin-capsorubin synthase and to fruit color. Theor Appl Genet 101:86–89

Popovsky-Sarid S, Borovsky Y, Faigenboim A, Parsons EP, Lohrey GT et al (2017) Genetic and biochemical analysis reveals linked QTLs determining natural variation for fruit post-harvest water loss in pepper (*Capsicum*). Theor Appl Genet 130:445–459

Qin C, Yu C, Shen Y, Fang X, Chen L et al (2014) Whole-genome sequencing of cultivated and wild peppers provides insights into *Capsicum* domestication and specialization. Proc Natl Acad Sci USA 111:5135–5140

Ramchiary N, Kehie M, Brahma V, Kumaria S, Tandon P (2014) Application of genetics and genomics towards Capsicm translational research. Plant Biotechnol Rep 8:101–123

Rao GU, Ben Chaim A, Borovsky E, Paran I (2003) Mapping of yield related QTLs in pepper in an inter-specific cross of *Capsicum annuum* and *C. frutescens*. Theor Appl Genet 106:1457–1466

Stewart CJ, Kang BC, Liu K, Mazourek M, Moore SL et al (2005) The *Pun1* gene for pungency in pepper encodes a putative acyltransferase. Plant J 42:675–688

Tan S, Cheng JW, Zhang L, Qin C, Nong DG et al (2015) Construction of an interspecific genetic map based on InDel and SSR for mapping the QTLs affecting the initiation of flower primordia in pepper (*Capsicum* spp.). PLoS ONE 10:e0119389. https://doi.org/10.1371/journal.pone.0119389

Thorup TA, Tanyolac B, Livingstone KD, Popovsky S, Paran I, Jahn M (2000) Candidate gene analysis of organ pigmentation loci in the Solanaceae. Proc Natl Acad Sci USA 97:11192–11197

Tsaballa A, Pasentsis K, Darzentas N, Tsaftaris AS (2011) Multiple evidence for the role of an Ovate-like gene in determining fruit shape in pepper. BMC Plant Biol 11:46

Wahyuni Y, Stahl-Hermes V, Ballester AR, de Vos RC, Voorrips RE (2014) Genetic mapping of semi-polar metabolites in pepper fruits (*Capsicum* sp.): towards unravelling the molecular regulation of flavonoid quantitative trait loci. Mol Breed 33:503–518

Yarnes SC, Ashrafi H, Reyes-Chin-Wo S, Hill TA, Stoffel KM, Van Deynze A (2012) Identification of QTLs for capsaicinoids, fruit quality, and plant architecture-related traits in an interspecific *Capsicum* RIL population. Genome 56:61–74

Zhao X, Meng Z, Wang Y, Chen W, Sun C et al (2017) Pollen magnetofection for genetic modification with magnetic nanoparticles as gene carriers. Nat Plants 3:956–964

Zygier S, Chaim AB, Efrati A, Kaluzky G, Borovsky Y, Paran I (2005) QTLs mapping for fruit size and shape in chromosomes 2 and 4 in pepper and a comparison of the pepper QTL map with that of tomato. Theor Appl Genet 111:437–445

7 Genes/Quantitative Trait Loci and Associated Molecular Mechanisms Identified in *Capsicum* Genome for Tolerance to Abiotic and Biotic Stresses

Vandana Jaiswal, Vijay Gahlaut, Meenakshi Dubey and Nirala Ramchiary

Abstract

Capsicum is one of the most important vegetable crops of the family Solanaceae and is widely used as spice due to its pungent nature. Besides, *Capsicum* fruit rich in metabolites and vitamins; and also has anti-cancerous property, which further increases the importance of this crop. However, *Capsicum* crop is highly affected by abiotic/biotic stresses such as drought, heat, cold, salinity, and pathogens. To overcome these stresses, plants adapted several mechanisms such as the production of osmoprotectant, proline, galactinol and raffinose, and the reduction of reactive oxygen species. Autophagy also plays an important role to provide tolerance against stresses through degradation of toxins. Among the others, transcription factors and plasma membrane intrinsic proteins, and plant endophytes are found to be involved in regulating stress tolerance mechanism. Furthermore, in *Capsicum* genome, a number of genes and quantitative trait loci (QTLs) involved in stress tolerance mechanism have been identified. In this chapter, a detail compilation of important molecular mechanisms and associated genes/QTLs involved toward imparting abiotic and biotic stress tolerance in *Capsicum* genome is made.

7.1 Introduction

Environmental stresses including both abiotic and biotic stresses have major effects on different developmental processes in plants. To overcome these stresses, plants adopted different mechanisms including production/accumulation of osmoprotectants, chaperones, and increasing superoxide radical scavengers. Among the major abiotic stresses, drought, cold, heat, salinity, and cold stresses are the most common in *Capsicum* crop. Beside abiotic stresses, several pathogens also damage *Capsicum* crop by causing several diseases. For example, *Phytophthora capsici* causes rot disease on various plant parts such as root, shoot, leaf, and fruits. Several other diseases including leaf spot (caused by *Xanthomonas campestris*), viral disease (caused by tobacco mosaic virus TMV, cucumber mosaic virus CMV, tomato spotted wilt virus TSWV, and potyvirus) also damage *Capsicum* plants severely. These diseases cause retorted growth and development and ultimately reduced yield and quality of fruits.

V. Jaiswal · M. Dubey · N. Ramchiary (✉)
Translational and Evolutionary Genomics Lab, School of Life Sciences, Jawaharlal Nehru University, New Delhi 110067, India
e-mail: nramchiary@jnu.ac.in

V. Gahlaut
Department of Plant Molecular Biology, University of Delhi, South Campus, New Delhi, India

M. Dubey · N. Ramchiary
Department of Biotechnology, Delhi Technological University, Delhi 110042, India

N. Ramchiary and C. Kole (eds.), *The Capsicum Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-319-97217-6_7

In winter crops including *Capsicum*, low temperature in root zone is most deleterious and cause turgor loss due to osmotic stress (Islam et al. 2014). At molecular level, low temperature in root zone may lead to several other imbalances like protein denaturation, membrane disorganization and damage, increased production of reactive oxygen species (ROS), change in cytoplasm viscosity, and enzyme activity (Janska et al. 2010; Krasensky and Jonak 2012). These abnormalities further affect different plant growth and developmental processes and may cause premature senescence, reduced fertility, wilting, chlorosis, reduced leaf expansion, necrosis, and ultimately plant death (Mahajan and Tuteja 2005). In *Capsicum*, significant economic loss may occur due to poor fruit set and quality due to biotic/abiotic stresses (Sanghera et al. 2011). To acclimate under stress, plant produces increased level of compatible solutes, such as proline, raffinose, and glycine betaine which stabilizes different cellular structures, and removes excess ROS and maintain redox balance.

Similarly, abscisic acid pathway is widely known to provide tolerance against abiotic stresses mainly in drought and osmotic stress. Upon availability of endogenous ABA, ABA-responsive element (ABRE) and MYC/MYB systems become functional which is required for expression of *rd22* gene (Abe et al. 1997). Sequences encoding MYC and MYB genes are essential for the ABA- and drought-responsive expression of *rd22*. Furthermore, NAC transcription factors (containing *AP2* domain to increase tolerance) are also induced under drought stress and in the presence of endogenous ABA.

Different parameters are used to measure stress tolerance in *Capsicum* including gas exchange, plant height, shoot dry mass, root morphology (like length, projected area, root tips' number, and dry mass), pattern of central as well as secondary metabolites in different tissues like leave, root shoot, and carbon remobilization. For example, a cold stress tolerant variety exhibits more carbon/nitrogen ratio in leaves than roots and shows a higher level of γ-aminobutyric acid (GABA), proline, galactinol, and raffinose (stress related) in roots (Aidoo et al. 2017).

In the present chapter, an attempt has been made to compile several reported research findings in identification and characterization of functional role of important genes involved in abiotic and biotic stress tolerance in *Capsicum*, and subsequently, the crosstalk between abiotic and biotic stress signaling pathways is also discussed.

7.2 Genes and Associated Molecular Mechanism Identified for Abiotic Stress Tolerance in *Capsicum* Genome

Abiotic stress tolerance involves a complex mechanism. Sometimes, more than one stresses act in combination and affect the plant growth. In *Capsicum*, several genes involved in important pathway of tolerance against abiotic stresses have been characterized (Table 7.1). For example, Sheong and Wang (2008) identified a protein encoded by *CaAbsi1* gene which has a putative zinc finger protein in its C-terminus and is upregulated in early stage of salt stress (high concentrations of NaCl or mannitol), and after six hours under cold stress. Besides, up-regulation in response to oxidative stress, methyl viologen, hydrogen peroxide, and abscisic acid suggested that *CaAbsi1* plays an important role in multiple abiotic stresses tolerance mechanism.

During abiotic stresses, highly toxic ROS (single oxygen, superoxides, peroxide, and hydroxyl radicles) are produced in mitochondria, chloroplast, and peroxisomes and may damage to cellular components including DNA, RNA, protein, lipid, therefore, need immediate detoxification by certain enzymatic or non-enzymatic scavenging systems (Apel and Hirt 2004). One of the important scavenging systems involves methionine sulfoxide reductases (MSR) which convert methionine sulfoxide back to methionine. Capsicum MSR-B2 (CaMSRB2) has been shown to provide tolerance against drought stress in rice. Transgenic rice (CaMSRB2) showed less

Table 7.1 List of genes reported to be involved in abiotic stresses in *Capsicum*

Gene/family	Description	Stress	Reference/s
BI1	BAX inhibitor 1 involved in program cell death	Cold, salinity, drought, flood, and heavy metal	Isbat et al. (2009)
CaAbsi1	Putative zinc finger protein in its C-terminus	Salt and cold	Seong and Wang (2008)
F-box genes	Sub-unit of E3, involved in ubiquitination activity	Cold, salt, and osmotic	Chen et al. (2014)
CaMSRs	Methionine sulfoxide reductases, reduces level of ROS	Drought	Kim et al. (2014a, b)
CaDHNs	Dehydrin and lysine-rich hydrophilic protein	Drought and cold	Szabala et al. (2014), Jing et al. (2016)
CaPAL1	Phenylalanine ammonia-lyase; major gene of phenylpropanoid metabolism	Pathogen defence	Kim and Hwang (2014)
CaATGs	Autophagy-related gene	Heat	Zhai et al. (2016)
CaPUB1	Pepper U-box E3 Ubiquitin Ligase	Drought	Min et al. (2016)
CaWRKYs	Contain WRKY domain	Heat, salinity, and drought	Oh et al. (2006), Cheng et al. (2016)
CaZFP1	Cys2/His2-type zinc finger transcription factor	Drought tolerance	Kim et al. (2004)
CaNACs	NAM, ATAF, and CUC transcription factors	Cold, salt, and drought	Guo et al. (2015), Diao et al. (2018)
CaKR1	Ankyrin-repeat domain C(3)H(1) zinc finger protein	Cold	Seong et al. (2007)
CaBiPs	Binding protein	Heat, drought, osmotic, and salinity	Wang et al. (2017)
CaChis	Chitin-binding proteins	Pathogen defence, cold, and salinity	Ali et al. (2018)
CaPIPs	Plasma membrane intrinsic proteins, aquaporins	Chilling, salt	Yin et al. (2015)
CaPGIPs	Polygalacturonase-inhibiting proteins	Cold treatment	Wang et al. (2013)
CaXTHs	Xyloglucan endotransglucosylase/hydrolase	Drought, high salinity, and cold	Cho et al. (2006b), Choi et al. (2011)
CaRma1H1	RING E3 Ub ligase	Drought and salt	Lee et al. (2009), Seo et al. (2012)
CabZIPs	Basic leucine zipper	Drought	Lee et al. (2006), Moon et al. (2015)
CaRAVs	Related to ABI3/VP1, transcription factor	Drought, salt, and pathogen	Sohn et al. (2006)
CaGLIP1	GDSL-type lipase	Salicylic acid, ethylene, and methyl jasmonate	Hong et al. (2008)
CaMLO2	Mildew resistance locus O	Abscisic acid and drought	Kim and Hwang (2012), Lim and Lee (2014)
CaRING1	Ring-type protein	Drought	Lim et al. (2015a, b)
AGO/DCL/RDR	Argonaut protein, Dicer-like protein, and RNA-dependent RNA polymerase	Cold, drought, and salinity	Qin et al. (2018)

(continued)

Table 7.1 (continued)

Gene/family	Description	Stress	Reference/s
CaARFs	Auxin-responsive factors	Salinity, cold, and heat stresses	Yu et al. (2017)
CaDRT1	*Capsicum annuum* DRought Tolerance 1	Drought	Baek et al. (2016)
CaWDP1	WPP Domain protein, involved in ABA signaling	Drought and NaCl treatments	Park et al. (2017)

oxidative stress, increased level of yield, and survival rate (Kim et al. 2014a, b). Further, it has also been suggested that *CaMSRB2* may target porphobilinogen deaminase (*PBGD*), which is involved in chlorophyll synthesis.

Dehydrins are hydrophilic proteins produced in response to abiotic stress to provide tolerance to plant. Dehydrin contains highly conserved lysine-rich amino acid sequence (EKKGIMDKIKEKLPG, also called K segment) at C-terminus, and serine residues (S-segment), and a consensus sequence (Y-segment) at N-terminus. SKn are acidic dehydrins which are mostly accumulated in plant cell in response to freezing stress (Rorat 2006). *DHN24* (a SK_3 dehydrin) found upregulated in phloem cells under drought and cold stresses (Szabala et al. 2014) suggested that it might play a role in drought tolerance. Similarly, *DHN3* was found to be associated with cold and salt stresses (Jing et al. 2016).

Autophagy also plays a vital role in stress tolerance through the degradation of damaged and denatured protein and thus reduces toxic level. In *Capsicum*, 15 autophagy-related genes (ATG) called *CaATGs* have been identified which got upregulated during abiotic stresses like salt, drought, heat, and cold. During heat stress, *CaATG* genes have higher expression in heat-tolerant genotype than heat-sensitive genotype. It has also been found that *CaATGs* interact with heat shock proteins of HSP90 family (Zhai et al. 2016).

7.2.1 Role of Ubiquitin Genes During Abiotic Stress Tolerance

Ubiquitin is one of the key regulators of several cellular functions such as protein sorting, endocytosis, and hormone signaling and mostly function through protein degradation. It is a peptide having highly conserved 76 amino acids. Three main enzymes named E1, E2, and E3 are involved in protein degradation through ubiquitin. E1 activates ubiquitin, E2 forms complex with activated ubiquitin and attached to the target site, and E3 catalyzes the isopeptide bonds. On the basis of sub-units, E3 can be RING-type/U-box E3 class and SKP-type cullin/CDC53-F-box. Both the types of E3 sub-units have been well characterized and found to be involved in abiotic stress tolerance in *Capsicum* in separate studies. Cho et al. (2006a, b) isolated a peptide called putative U-box protein 1 (*CaPUB1*) with U-box motif (essential for E3 activity) from water-stressed hot pepper. *CaPUB1* is found to be induced under different abiotic stress conditions like drought, salinity, and cold stress. Overexpression of *CaPUB1* in transgenic *Arabidopsis* showed longer hypocotyls and root, higher plant growth rate, and early bolting than wild-type. However, under abiotic stress conditions such as drought and low temperature, transgenic *Arabidopsis* plants showed increased sensitivity than wild-type plants suggesting *CaPUB1* gene to be a negative regulator of abiotic stress tolerance. Similarly, in another study conducted in rice, overexpression of *CaPUB1* showed hypersensitivity under drought stress (Min et al. 2016); however, under cold stress, overexpression of *CaPUB1* provided tolerance in transgenic rice. Moreover, cold inducer marker genes including *DREBs* and cytochrome *P450* also showed higher expression in overexpressing *CaPUB1* rice lines compared to the wild-type plants suggesting *CaPUB1* to be a positive regulator of cold stress. On the other hand, F-box protein, a member of SCF (Skp–Cullin–F-box) protein complex (another subunit

type of E3) was also found to play an important role in gene regulation during stress response (Chen et al. 2014). In *Capsicum*, *CaF-box* gene has been found to be differentially expressed predominantly during salt stress along with cold stress, and also in response to abscisic acid (ABA) and salicylic acid (SA).

7.2.2 Role of Plant Aquaporin Genes During Abiotic Stresses

Plasma membrane intrinsic proteins (PIPs) are membrane-bound proteins that allow transmembrane transfer of water (Chaumont et al. 2001). It has been suggested that besides water transmission, PIPs also play important role in the transportation of solutes and CO_2, and other physiological processes like stomatal opening, cell elongation, seed germination, and ripening (Forrest and Bhave 2007). Moreover, plant aquaporin also takes part in providing tolerance against biotic/abiotic stresses. In *Capsicum*, upregulation of *PIP-1* (isolated from P70) under cold and salt stresses suggested that it may be involved in providing tolerance to these stresses and increased susceptibility against salt of silent *PIP1* further confirmed its involvement in stress tolerance (Yin et al. 2015).

7.3 Role of Transcription Factors in Abiotic Stress Tolerance

Transcription factors are key regulators of cell signaling both internally and externally. In plants, several transcription factors have been characterized to play important role in abiotic stress tolerance (Gahlaut et al. 2016). In *Capsicum* also, several transcription factors including BAX inhibitor 1, WRKY, NAC, CAZFP1, bZIP like, RAV, GRAS, Dof, ARF, and PF1 have been found to be involved in abiotic stress tolerance.

7.3.1 BAX Inhibitor 1

In response to environmental stresses, plants follow programmed cell death (PCD) to eliminate damaged cells. BCL2-associated x protein (BAX) is found to be important regulator of PCD and balanced by the activity of *BAX inhibitor-1* (*BI-1*). In *Capsicum*, *CaBI-1* has been cloned and found to be upregulated in response to different abiotic stresses like cold, salinity, drought, flood, and heavy metal stresses and provides tolerance to plants against these stresses (Isbat et al. 2009). Loss of function of *CaBI-1* enhances cell death and shows more susceptibility toward cold stress.

7.3.2 WRKY and NAC Transcription Factor Genes

WRKY is one of the largest transcription factor families in higher plants, which contains WRKY domain (WRKYGQK peptide and Cx4–5Cx22–23HxH or Cx7Cx23HxC zing-finger structure). *WRKY* transcription factors have been found to be involved in several biological and physiological processes including stress tolerance. Totally 61 *WRKY*s genes (called *CaWRKY*s) have been identified in *Capsicum* (Cheng et al. 2016). Constitutive expression of 16 *CaWRKY*s suggested an involvement of WRKYs in fundamental developmental processes in *Capsicum*. Most of the WRKY genes (60%) are expressed in fruit tissues. Differential expression of 26, 27, and 14 WRKY genes under heat, salinity, and drought stresses, respectively, suggested active involvement of these WRKY genes in fruit development under abiotic stresses (Cheng et al. 2016).

NAC is also a well-characterized transcription factor family involved in stress tolerance in plants; however, a limited study is available in *Capsicum* (Guo et al. 2015; Diao et al. 2018). Recently, *CaNAC2* has been isolated in *Capsicum* (Guo et al. 2015). *CaNAC2* has conserved NAC domain at N-terminus which encodes 410 amino acids' long polypeptide. Induced expression of *CaNAC2* after cold and salt stresses suggested that *NAC2* may be involved in stress mechanism. Loss of function mutants showed enhanced susceptibility against chilling stress and delayed the salt-induced leaf chlorophyll

degradation. Recently, *104 CaNAC* genes have been identified and found to be distributed on all the 12 chromosomes of *Capsicum* (Diao et al. 2018). Under abiotic stress condition, several NAC genes showed differential expression. For example, *CaNAC72* showed >600-fold increased expression upon salt stress treatment along with 10 other *CaNAC* genes showing average 10-fold higher expression. Similarly, upon heat stress, total 10 NAC genes (*CaNAC13, CaNAC20, CaNAC27, CaNAC29, CaNAC35, CaNAC37, NAC53, CaNAC61, CaNAC72, and CaNAC102*) are found to be significantly upregulated; however, *CaNAC41* and *CaNAC86* get downregulated under stress condition. Further, under drought stress, more than 70-fold increased expression of two NAC genes (*CaNAC72,* and *CaNAC79*) suggested the involvement of these NAC genes in drought stress (Diao et al. 2018).

7.3.3 bZIP Transcription Factor Genes

Basic leucine zipper (bZIP), a large TFs family, consists of a 40–80 amino acid containing DNA-binding domain and a leucine zipper dimerization domain. In *Arabidopsis* and rice, a total of 75 and 89 bZIP TFs, respectively, are known to be involved in multiple mechanisms of biotic and abiotic stresses, plant development, seed maturation, etc (Muszynski et al. 2006). Group A bZIP genes (ABFs/AREBs) are found to be involved mainly in drought and salinity stresses (Yoshida et al. 2010). In *Capsicum*, *CaBZ1* has been characterized to be involved in salt and abiotic stresses (Moon et al. 2015). Ectopic expression of *CaBZ1* in potato provides tolerance against drought (Moon et al. 2015). Similarly, *CAbZIP1* provides tolerance against abiotic stresses in *Arabidopsis* (Lee et al. 2006).

7.3.4 ERF/AP2-Type and RAV Transcription Factor Genes

In *Capsicum*, Yi et al. (2004) characterized an ERF transcription factor gene (called *CaPF1*) for cold tolerance. Like other ERF/AP2-type TFs, *CaPF1* binds to GCC and CRT/DRE cis-elements. Higher expression of *CaPF1* has been observed under different treatments including chilling stress in transgenic *Arabidopsis*.

RAV (related to ABI3/VP1) is a new group of DNA-binding proteins transcription factors and contains two different plant-specific DNA-binding domains—(i) AP2/ERF DNA-binding domain at N-terminal and (ii) B3 DNA-binding domain of VP1/ABI3 at C-terminal (Kim et al. 2005; Sohn et al. 2006). A number of AP2/ERF-domain-containing proteins (such as DREBs, Tsi1, and CBFs) and VP1/B3 DNA-binding proteins (*VP1, ABI3*, and *ARF1*) are widely known to be involved in plant responses to biotic and abiotic stress (Gutterson and Reuber 2004; Kasuga et al. 1999; Park et al. 2001; Kirsten et al. 1998). In *Capsicum*, it has been found that *CaRAV1* interacts with oxidoreductase protein (*CaOXR1*) and provides extreme tolerance against osmotic and salinity stresses to the overexpressed (*CaOXR1/CaRAV1*) lines in *Arabidopsis* (Lee et al. 2010).

7.3.5 Auxin-Responsive Factors (ARFs) and DNA-Binding One Zinc Finger (DoF) Transcription Factor Genes

In *Capsicum*, 22 *CaARF* genes have been identified (Yu et al. 2017). These genes are grouped into six clusters and distributed on all the 12 *Capsicum* chromosomes. Most of the above-mentioned *CaARFs* showed different expression under abiotic stresses like salinity, cold, and heat stresses. Under salinity stress, nine and ten *CaARFs* got up- and downregulated, respectively. Under cold stress, expression of *CaARFs* differs in different tissues, for example, expression of *CaARF11* got upregulated in shoot, and however, its expression goes down in root at the same time. Similarly, differential expression of 11 *CaARFs* under heat stress condition suggested these *CaARFs* may be involved in heat stress tolerance (Yu et al. 2017).

Similarly, 33 *CaDoFs* have been identified in *Capsicum* (Wu et al. 2016) and found to be

distributed across 11 *Capsicum* chromosomes (excluding Chromosome 7). Several *CaDoFs* showed significant differential expression under two stresses including heat and salinity (Wu et al. 2016).

7.4 Genes Involved in Biotic Stress Tolerance

In *Capsicum*, several genes have been characterized to play a vital role in providing tolerance against biotic stresses like bacteria, virus, and nematodes (Table 7.2). Bacteria called *X. campestris* causes leaf blight disease in *Capsicum*. Choi et al. (2007) identified a *CaPO2* gene to provide tolerance against this disease. A knockdown mutant of *CaPO2* showed increased susceptibility against *Xanthomonas*. Similarly, *CaMLO2* also reported to show resistance against *Xanthomonas* and silencing of this also show increased susceptibility toward disease represented by cell death and increased ROS (Kim and Hwang 2012; Zheng et al. 2013). *CaMLO2* interacts with a calmodulin-related gene, involved in cell death (*CaCaM1*), and

Table 7.2 List of genes reported to be involved in different biotic stress tolerance in *Capsicum*

Biotic stress	Genes	Description	Reference/s
Xanthomonas campestris resistance	*Bs Genes*	Bacterial spot	Romer et al. (2010), Vallejos et al. (2010)
	CaPO2	Peroxidase	Choi et al. (2012)
	CaMLO2	Mildew resistance locus O; associated with powdery mildew	Kim and Hwang (2012), Zheng et al. (2013)
	CaCaM1	Calmodulin 1; involved in hypersensitive cell death	Kim et al. (2014a, b)
Pseudomonas syringae resistance	*CaLOX1*	Lipoxigenase	Hwang and Hwang (2010), Lim et al. (2015a, b)
Phytophthora capsici resistance	*CaMsrB2*	Methionine sulfoxide reductase B2	Hong Truong et al. (2013), Oh et al. (2010)
	CaRGA2	Resistance gene analogs	Zhang et al. (2013)
	Ipcr	Inhibitor of *P. capsici* resistance	Reeves et al. (2013), Wang et al. (2015)
Ralstonia solanacearum resistance	*CaHDZ27* Related genes	Homeodomain–Leucine Zipper I	Mou et al. (2017)
Cucumber mosaic virus resistance	*cmv11.1*	Cucumber mosaic virus	Ben-Chaim et al. (2001), Yao et al. (2013)
ToMV resistance	*Cmr1*	The gene showed synteny with ToMV-resistance locus	Kang et al. (2010)
Potato virus resistance	*pvr's (4E (eIF4E))*	Potato virus Y (PVY) resistance	Ruffel et al. (2006), Hwang et al. (2009)
	PVY, PepMoV, and *PMMoV*	Potato virusY (PVY) resistance	Banerjee et al. (2014), Rubio et al. 2008)
	PVMV-HN	Potato virus Y (PVY) resistance	Gao et al. (2014)
Nematode resistant	*RKN*	Root-knot nematodes	Djian-Caporalino et al. (1999, 2001)
	Me4, *Mech1*, and *Mech2*	Meloidogyne species or its populations	Djian-Caporalino et al. (2001, 2007)
	CaMi	Nematode-resistant gene	Chen et al. (2007), Fazari et al. (2012)

regulated disease tolerance mechanism (Kim et al. 2014a, b). Likewise, a lipoxygenase-related gene (*CaLOX1*) was found to provide tolerance against *Pseudomonas syringae* (Hwang and Hwang 2010). *P. capsici* is one of the most harmful bacteria for *Capsicum* causing rot disease. In *Capsicum*, a gene related to reactive oxygen species (ROS) production called *CaMsrB2* has been characterized to provide resistance against rot disease (Oh et al. 2010). Similarly, *CaRGA2* and *Ipcr* (disease resistance inhibitor) also provide resistance against *P. capsici* (Zhang et al. 2013; Reeves et al. 2013).

Similarly, several genes are also reported to be involved in tolerance against viral disease like cucumber mosaic virus (cmv11.1), tobacco mosaic virus (TMV), potato virus Y (PVY), and potyviruses [including veinal mottle virus (PVMV), tobacco etch virus (TEV), chili veinal mottle virus (ChiVMV), PVY, and PepMoV], and has been reviewed recently by us (for details, see Chhapekar et al. 2018).

7.5 Crosstalk Between Abiotic and Biotic Stress Responses

Signaling pathways that are involved in plant defense against abiotic and biotic stresses share some common modules like involvement of transcriptional factors, ROS, signaling pathways (calcium signaling, ABA signaling, jasmonic acid signaling, and mitogen-activated protein cascades (Moller et al. 2007; Wong and Shimamoto 2009; Ton et al. 2009; Fonseca et al. 2009; Pitzschke et al. 2009; Walley and Dehesh 2010; Galon et al. 2010). These convergent nodes help plants to swiftly adapt to a changed environment involving abiotic/biotic stresses via these signaling crosstalks (Fujita et al. 2006; Atkinson and Urwin 2012, Fig. 7.1). Here, we describe a few examples of above-mentioned abiotic and biotic crosstalks in plants including *Capsicum*.

In plant, several hormones are involved in defense pathways including abiotic (ABA) and

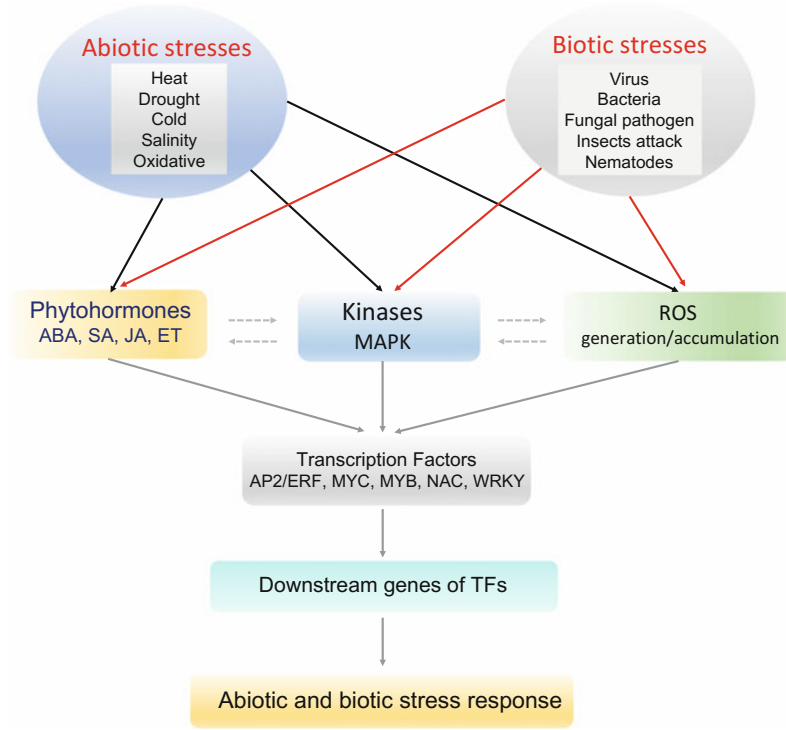


Fig. 7.1 Schematic diagram showing crosstalks between phytohormones, kinases (MPKK), reactive oxygen species (ROS), and transcription factor genes in plants during abiotic and biotic stresses

biotic (SA, JA, and ethylene) stresses. Further, ABA is also found to be involved in tolerance against biotic stresses, thus making this hormone enabled to create a crosstalk among different signaling pathways (Fujita et al. 2006; Yasuda et al. 2008; Lim and Lee 2014). In different crops like rice and tomato, as well as model plant *Arabidopsis*, ABA is found to be involved in resistance mechanism of different biotic stresses caused by *P. syringae*, *B. cinereal,* and *Magnaporthe grisea* (Audenaert et al. 2002; Mohr and Cahill 2003; Koga et al. 2004; Asselbergh et al. 2008; Jiang et al. 2010) through repression of the systemic acquired resistance (SAR) pathway and also through reduction of defense-related compounds like lignins and phenylpropanoids (Yasuda et al. 2008; Kusajima et al. 2010). On the other hand, examples are also available where ABA acts as positive regulator in biotic stress resistance (Asselbergh et al. 2008; Ton et al. 2009). In *Capsicum, CaMLO2* is transcriptionally induced under stress conditions (Kim and Hwang 2012) and is found to be upregulated under ABA treatment and drought. The overexpression of the *CaMLO2* gene in *Arabidopsis* reduces sensitivity toward ABA in germination and seedling growth stages. These results suggest that ABA signaling via *CaMLO2* may regulate drought stress (Lim and Lee 2014).

Mitogen-activated protein (MAP) kinase (MAPK/MPK) cascades are responsible for phosphorylation and dephosphorylation of proteins which significantly affect the regulation of physiological, morphological, and cellular processes and are also involved in defense mechanism involving hormone signaling and ROS (Jonak et al. 2002; Xiong and Yang 2003; Nakagami et al. 2005; Rodriguez et al. 2010; Atkinson and Urwin 2012). MEKK1/MKK2/MPK4/MPK6 cascades are found to be involved in signaling under biotic and abiotic stress conditions (Teige et al. 2004; Rodriguez et al. 2010) and play a crucial role of crosstalk signaling between abiotic and biotic stress mechanisms (Fujita et al. 2006; Zhang et al. 2006; Takahashi et al. 2011; Atkinson and Urwin 2012). MKK2–MPK4/MPK6 cascades are involved in cold and salt stress signaling (Ichimura et al. 2000; Teige et al. 2004); however, MEKK1–MKK4/MKK5–MPK3/MPK6 cascades are involved in pathogen defense response pathway via the regulation of expression of *WRKY 22*/*WRKY 29* genes (Nuhse et al. 2000). Recently in *Capsicum*, the involvement of MAPK cascade in response to abiotic (*CaMPK1* and *CaMPK3*; salinity and heat) and biotic stress (*CaMPK4; Ralstonia solanacearum* infection) has also been reported (Liu et al. 2015).

Another important player that is involved in crosstalk signaling during abiotic and biotic stresses is ROS (Fujita et al. 2006; Ton et al. 2009; Atkinson and Urwin 2012; Baxter et al. 2014). The ROS' signaling network is vastly conserved among plants and regulates various biological processes such as plant growth, development, and responses to biotic and/or abiotic stresses (Mittler et al. 2011; Baxter et al. 2014). During different types of stresses, ROS function differently. In general, ROS concentration induces during various abiotic stress conditions, i.e., drought, heat, and salinity stresses) and pathogen infection to minimize cell injury (Apel and Hirt 2004; Mittler and Blumwald 2010). Several, research findings in plants revealed that biotic and abiotic stress responses are mediated by a temporal–spatial synchronization between ROS and some other signals that rely on the production of several stress-specific compounds, chemicals, and hormones in plants (Baxter et al. 2014). Further, certain TFs integrate ROS-scavenging mechanisms in response to various types of stresses. *Arabidopsis* zinc-finger TF, ZAT12 regulates H_2O_2 levels in plants and its transcripts were induced by wounding, abiotic and biotic stresses. It also induces the expression of its downstream gene ascorbate peroxidase (*APX1*) and when overexpressed conferred tolerance to oxidative stress, freezing, and high light (Davletova et al. 2005; Vogel et al. 2005; Fujita et al. 2006). In *Capsicum*, a gene related to ROS production known as *CaMsrB2* has been characterized to provide resistance against rot disease (Oh et al. 2010). Another gene in *C. annuum,* i.e., receptor-like protein kinase 1 (*CaRLK1*), is also induced by pathogen infection and application of

exogenous H_2O_2 (Yi et al. 2010). These findings suggest that the ROS signaling might mediate crosstalk between biotic and abiotic stress-responsive gene expression.

TFs are another convergent node that play a crucial role in signal crosstalk under abiotic and biotic stress. For example, *MYC2* induced by ABA (key regulator of biotic/abiotic stress signaling pathway) suggested its involvement in crosstalks (Abe et al. 2003; Anderson et al. 2004; Asselbergh et al. 2008; Pieterse et al. 2009). Beside ABA, *MYC2* TF also acts as a positive regulator of JA-induced defense genes, however, negatively regulates combined JA/ethylene induced genes (Anderson et al. 2004; Pieterse et al. 2009). Another TF family MYB has also been found to be involved in the regulation of both biotic and abiotic stress regulation in plants (Dubos et al. 2010). For example, *MYB96* is upregulated under drought stress and also promotes ABA-dependent stress tolerance (Seo et al. 2011); and also, under biotic stress, *MYB96* regulates pathogenesis-related (PR) gene expression via ABA-dependent SA biosynthesis, thus acting as a node for crosstalk among stress responses (Seo and Park 2010). Other MYB TFs, i.e., *OsMYB4*, *AtBOS1,* and *TaPIMP1,* were involved in the regulation of broad-spectrum of different stresses including drought, salt, and pathogens (Mengiste et al. 2003; Vannini et al. 2006, 2007; Liu et al. 2011). In addition to MYB/MYC TFs, NAC and AP2/ERF TFs are also widely known to be involved in stress signaling (Xu et al. 2011). RD26, an NAC TF in *Arabidopsis*, is upregulated by JA, ABA, drought, salinity, and pathogen via regulation of ROS detoxification genes (Fujita et al. 2004; Atkinson and Urwin 2012). Similarly, in rice, *OsNAC6* was reported to be involved in tolerance against drought, salinity, and rice blast (Nakashima et al. 2007). Further, in wheat *TaNAC4* is upregulated in response to salinity, cold stress, and rust stripe fungus (Xia et al. 2010). Above-mentioned studies suggested that the NAC TFs also regulate cross-signaling between stress response pathways. Recently, in *Capsicum*, it is reported that the expression of NAC TFs (*CaNAC2, CaNAC72, CaNAC102*) was induced in response to cold, heat, and salt stress (Guo et al. 2015; Diao et al. 2018). AP2/ERF TF gene (*TSI1*) from tobacco is also involved in the regulation of both abiotic stress and pathogen response pathways. *TSI1* induces the expression of PR genes and can confer resistance to bacterial pathogen and salinity (Park et al. 2001). In *Capsicum*, a AP2/ERF TF gene *RFP1* was found to be involved in osmotic stress and pathogen defense (Hong et al. 2007; Asselbergh et al. 2008). Moreover, WRKY and DREB TFs also act as a key player in defense against biotic and abiotic stresses in many plant species including *Capsicum* (Qiu and Yu 2009; Tsutsui et al. 2009; Peng et al. 2011; Cheng et al. 2016). Altogether, these studies suggest that the TFs might mediate crosstalk between biotic and abiotic stress-responsive gene-expression networks. A list of TFs in *Capsicum* that may be crucial in controlling the response to biotic and abiotic stresses is given in Tables 7.1 and 7.2.

7.6 QTL Mapping for Abiotic/Biotic Stresses in *Capsicum* Genome

QTL mapping is a widely known approach to identify genomic loci associated with quantitative traits particularly complex traits. In agricultural crops, such as wheat, rice, maize, and tomato, a number of QTLs have been identified for abiotic stress tolerance including heat, drought, and cold. However, in *Capsicum*, no QTL mapping study is available for abiotic stress tolerance and majority of the QTL mapping studies are focused on pungency, fruit traits like color, shape, and other important agronomic traits (Chhapekar et al. 2018). Dozens of studies were also conducted to identify QTLs for biotic stress tolerance in *Capsicum* (Table 7.3). Using different marker systems starting from RAPD to SNPs, several QTLs have been reported for many biotic stresses caused by virus, fungus, bacteria, and nematodes. For *Phytophthora* resistance, a number of QTLs have been identified using different mapping populations including F_2, backcross, recombinant inbred lines, and doubled haploids. Interestingly, in most of the studies,

Table 7.3 List of QTLs identified to be associated with different biotic stress tolerance mechanism in *Capsicum*

Trait	Population (parents)	Marker type	Method	Reference
Phytophthora resistance	$F_{2:3}$ (CM334/Chilsungcho)	RFLP, SSR, and gene based	CIM	Kim et al. (2008)
	RILs (YCM334/Tean)	SNP and SPP	BSA and CIM	Liu et al. (2014)
	DHs (H3/Van; Per/YW) and F_2 (YW/CM334)	RFLP, RAPD, and AFLP	IM and CIM	Thabuis et al. (2003)
	F_2 (CM334/JEP)	RAPD, SCAR, and AFPL	CIM	Quirin et al. (2005)
	BC (Yolo Wonder/CM334)	AFLP, SCAR, and CAPS	CIM	Thabuis et al. (2004)
	RILs (CM334/Early Jalapeno)	SPP	IM	Naegele et al. (2014)
Cucumber mosaic virus resistance	$F_{2:3}$ (BJ0747/XJ0630)	SLAF	IM and MQM	Li et al. (2018)
	F_2 and backcross (BJ0747/XJ0630)	SSR and ISSR	CIM	Yao et al. (2013)
	DH (H3/Vania)	RAPD, RFLP, and AFLP	IM and CIM	Caranta et al. (2002)
	DH (Yolo wonder/Perennial)	–	MQM	Tamisier et al. (2017)
	F_2, BC_1 and $F_{2:3}$ (PBC688/G29)	SLAF	MQM	Guo et al. (2017)
Root-knot nematodes	$F_{2:3}$ (Yolo Wonder/Doux Long des Landes)	SCAR, SSR, and SNP	regression, SIM, CIM, and nonparametric interval mapping	Barbary et al. (2016)
Anthracnose resistance	F_2 (Jatilaba/PRI95030)	AFLP, SSR, and gene based	MQM	Voorrips et al. (2004)
	Three way popu. (PBC932C/PBC80C)	SCAR-Indel and SSR-HpmsE032	Regression	Suwor et al. (2017)
	BC (17013/PBC932)	SSR, InDel, and CAPS	ICIM	Sun et al. (2015)
Thripe resistances	F_2 (AC 1979/4661)	AFLP, SSR, and SNP	IM and MQM	Maharijaya et al. (2015)
Powdery mildew resistance	DH (H3/Vania)	Gene based	IM and CIM	Lefebvre et al. (2003)
Potato Virus	DH (Yolo wonder/Perennial)	–	MQM	Tamisier et al. (2017)
PepMoVirus resistance	F_2 (CM334/Chilsungcho)	RAMP, RFLP, SSR, CAPS, AFLP, and BAC-end sequences	CIM	Kim et al. (2011)

RILs recombinant inbred lines; *DH* double haploid; *BC* backcross; *RFLP* restriction fragment length polymorphism; *SSR* simple sequence repeats; *SNP* single nucleotide; *RAPD* random amplified polymorphic DNA; *AFLP* amplified fragment length polymorphism; *SCAR* sequence characterized amplified region; *CAPS* cleaved amplified polymorphic sequence; *SPP* single position polymorphism; *SLAF* specific length amplified fragment; *ISSR* inter-simple sequence repeats

CIM composite interval mapping; *BSA* bulk segregant analysis; *IM* interval mapping; *MQM* multiple QTL mapping; *SIM* simple interval mapping

CM334 (potentially resistance against *Phytophthora*) was used as one of the parents along with a susceptible parent (like Chilsungcho, Tean, Yolo Wonder, and Early Jalapeno). Kim et al. (2008) identified four major QTLs (cumulative PVE ~66%) for rot resistance. Similarly, Liu et al. (2014) identified a major gene for *Phytophthora* resistance on chromosome 5. QTL mapping studies conducted for biotic stresses in *Capsicum* is summarized in Table 7.3.

7.7 Role of Plant Endophytes in Abiotic/Biotic Stress Tolerance

Plant endophytes mainly consist of bacteria present in plant tissues symptomatically and do not cause any visible infection. These endophytes are mainly present in intercellular spaces as well as vascular tissues. A number of bacterial species have been isolated from different plant organs like root, stem, leaves, and seed. Under stress conditions (abiotic/biotic), endophytes are found to provide tolerance to host against stresses. For example, some bacteria are found to provide better nutrition through nitrogen fixation under stress condition (Vessey 2003). Further, through the production of indoleacetic acid and cytokinin, endophytes provide better growth even under abiotic/biotic stress condition (Beyeler et al. 1999; Timmusk et al. 1999).

Ethylene is an important signaling molecule under abiotic/biotic stresses, and high level of ethylene may found harmful for plant growth except fruit ripening (Czarny et al. 2006). Basically, methionine acts as precursor for ethylene production and are converted via methionine–*S* adenosyl L methionine–1-aminocyclopropane 1–carboxylic acid (ACC)–ethylene. In *Capsicum*, some rhizosphere bacteria produce enzymes with deaminase activity which cleave ACC molecule and ultimately control product of ethylene under stress condition (Mayak et al. 2004). In another study (Sziderics et al. 2007), out of five bacterial strains isolated from *Capsicum*, four were found to produce indoleacetic acid and thus provide better growth under osmotic stress condition. Beside better growth, these strains were also found to be involved in the regulation of osmotic pressure and proline content. Two strains of azotobacter (EZB4 and EZB8) reduced the expression of two stress-inducible genes *CaACCO* and *CaLTPI* under abiotic stresses (Sziderics et al. 2007).

Acknowledgements This work was supported by UGC Resource Networking Grant, UGC-SAP, and DST FIST program to School of Life Sciences, Jawaharlal Nehru University, New Delhi, India.

References

Abe H, Yamaguchi-Shinozaki K, Urao T, Tiwasaki T, Hosokawa D, Shinozaki K (1997) Role of Arabidopsis *MYC* and *MYB* homologs in drought- and abscisic acid regulated gene expression. Plant Cell 9:1859–1868

Abe H, Urao T, Ito T, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2003) *Arabidopsis AtMYC2* (bHLH) and *AtMYB2* (*MYB*) function as transcriptional activators in abscisic acid signaling. Plant Cell 15:63–78

Aidoo MK, Sherman T, Lazarovitch N, Fait A, Rachmilevitch S (2017) A bell pepper cultivar tolerant to chilling enhanced nitrogen allocation and stress-related metabolite accumulation in the roots in response to low root-zone temperature. Physiol Plant 161:196–210

Ali M, Luo D-X, Khan A, Haq S ul, Gai W-X, Zhang H-X, Cheng G-X, Muhammad I, Gong Z-H (2018) Classification and genome-wide analysis of chitin-binding proteins gene family in pepper (*Capsicum annuum* L.) and transcriptional regulation to *Phytophthora capsici*, abiotic stresses and hormonal applications. Intl J Mol Sci 19. https://doi.org/10.3390/ijms19082216

Anderson JP, Badruzsaufari E, Schenk PM, Manners JM, Desmond OJ, Ehlert C, Maclean DJ, Ebert PR, Kazan K (2004) Antagonistic interaction between abscisic acid and jasmonate-ethylene signaling pathways modulates defense gene expression and disease resistance in *Arabidopsis*. Plant Cell 16:3460–3479

Apel K, Hirt H (2004) Reactive oxygen species: metabolism, oxidative stress, and signal transduction. Annu Rev Plant Biol 55:373–399

Asselbergh B, Achuo AE, Hofte M, Van Gijsegem F (2008) Abscisic acid deficiency leads to rapid activation of tomato defence responses upon infection with *Erwinia chrysanthemi*. Mol Plant Pathol 9:11–24

Atkinson NJ, Urwin PE (2012) The interaction of plant biotic and abiotic stresses: from genes to the field. J Exp Bot 63:3523–3544

Audenaert K, De Meyer GB, Hofte MM (2002) Abscisic acid determines basal susceptibility of tomato to *Botrytis cinerea* and suppresses salicylic acid-dependent signaling mechanisms. Plant Physiol 128:491–501

Baek W, Lim S, Lee SC (2016) Identification and functional characterization of the pepper *CaDRT1* gene involved in the ABA-mediated drought stress response. Plant Mol Biol 91:149–160

Banerjee A, Dutta R, Roy S, Ngachan SV (2014) First report of chilli veinal mottle virus in Naga chilli (*Capsicum chinense*) in Meghalaya, India. Virus Dis 25:142–143

Barbary A, Djian-Caporalino C, Marteu N, Fazari A, Caromel B, Castagnone-Sereno P, Palloix A (2016) Plant genetic background increasing the efficiency and durability of major resistance genes to root-knot nematodes can be resolved into a few resistance QTLs. Front Plant Sci 7:632

Baxter A, Mittler R, Suzuki N (2014) ROS as key players in plant stress signalling. J Exp Bot 65:1229–1240

Ben-Chaim A, Grube RC, Lapidot M, Jahn M, Paran I (2001) Identification of quantitative trait loci associated with resistance to cucumber mosaic virus in *Capsicum annuum*. Theor Appl Genet 102:1213–1220

Beyeler M, Keel C, Michaux P, Haas D (1999) Enhanced production of indole-3-acetic acid by a genetically modified strain of *Pseudomonas fluorescens* CHA0 affects root growth of cucumber, but does not improve protection of the plant against *Pythium* root rot. FEMS Microbiol Ecol 28:225–233

Caranta C, Pflieger S, Lefebvre V, Daubèze AM, Thabuis A, Palloix A (2002) QTLs involved in the restriction of cucumber mosaic virus (CMV) long-distance movement in pepper. Theor Appl Genet 104:586–591

Chaumont F, Barrieu F, Wojcik E, Chrispeels MJ, Jung R (2001) Aquaporins constitute a large and highly divergent protein family in maize. Plant Physiol 125:1206–1215

Chen RG, Li HX, Zhang LY, Zhang JH, Xiao JH, Ye ZB (2007) CaMi, a root-knot nematode resistance gene from hot pepper (*Capsicum annuum* L.) confers nematode resistance in tomato. Plant Cell Rep 26:895–905

Chen R, Guo W, Yin Y, Gong Z-H (2014) A novel F-Box protein CaF-box is involved in responses to plant hormones and abiotic stress in pepper (*Capsicum annuum* L.). Intl J Mol Sci 15:2413–2430

Cheng Y, Ahammed GJ, Yu J, Yao Z, Ruan M, Ye Q, Li Z, Wang R, Feng K, Zhou G, Yang Y, Diao W, Wan H (2016) Putative *WRKYs* associated with regulation of fruit ripening revealed by detailed expression analysis of the *WRKY* gene family in pepper. Sci Rep 6:39000

Chhapekar SS, Jaiswal V, Ahmad I, Gaur R, Ramchiary N (2018) Progress and prospects in capsicum breeding for biotic and abiotic stresses. In: Vats S (ed) Biotic and abiotic stress tolerance in plants. Springer Nature, Singapore, pp 279–322

Cho SK, Chung HS, Ryu MY, Park MJ, Lee MM, Bahk Y-Y, Kim J, Pai HS, Kim WT (2006a) Heterologous expression and molecular and cellular characterization of *CaPUB1* encoding a hot pepper U-Box E3 ubiquitin ligase homolog. Plant Physiol 142:1664–1682

Cho SK, Kim JE, Park J-A, Eom TJ, Kim WT (2006b) Constitutive expression of abiotic stress-inducible hot pepper *CaXTH3*, which encodes a xyloglucan endotransglucosylase/hydrolase homolog, improves drought and salt tolerance in transgenic *Arabidopsis* plants. FEBS Lett 580:3136–3144

Choi HW, Hwang BK (2012) The pepper extracellular peroxidase *CaPO2* is required for salt, drought and oxidative stress tolerance as well as resistance to fungal pathogens. Planta 235:1369–1382

Choi HW, Kim YJ, Lee SC, Hong JK, Hwang BK (2007) Hydrogen peroxide generation by the pepper extracellular peroxidase *CaPO2* activates local and systemic cell death and defence response to bacterial pathogens. Plant Physiol 145:890–904

Choi JY, Seo YS, Kim SJ, Kim WT, Shin JS (2011) Constitutive expression of *CaXTH3*, a hot pepper xyloglucan endotransglucosylase/hydrolase, enhanced tolerance to salt and drought stresses without phenotypic defects in tomato plants (*Solanum lycopersicum* cv. Dotaerang). Plant Cell Rep 30:867–877

Czarny JC, Grichko VP, Glick BR (2006) Genetic modulation of ethylene biosynthesis and signaling in plants. Biotechnol Adv 24:410–419

Davletova S, Rizhsky L, Liang H, Shengqiang Z, Oliver DJ, Coutu J, Shulaev V, Schlauch K, Mittler R (2005) *CYTOSOLIC ASCORBATE PEROXIDASE 1* is a central component of the reactive oxygen gene network of *Arabidopsis*. Plant Cell 17:268–281

Diao W, Snyder JC, Wang S, Liu J, Pan B, Guo G, Ge W, Dawood MHSA (2018) Genome-wide analyses of the NAC transcription factor gene family in pepper (*Capsicum annuum* L.): chromosome location, phylogeny, structure, expression patterns, cis-elements in the promoter, and interaction network. Intl J Mol Sci 19. https://doi.org/10.3390/ijms19041028

Djian-Caporalino C, Pijarowski L, Januel A, Lefebvre V, Daubeze A, Palloix A, Dalmasso A, Abad P (1999) Spectrum of resistance to root-knot nematodes and inheritance of heat stable resistance in pepper (*Capsicum annuum* L.) Theor Appl Genet 99:496–502

Djian-Caporalino C, Pijarowski L, Fazari A et al (2001) High-resolution genetic mapping of the pepper (*Capsicum annuum* L.) resistance loci Me3 and Me4 conferring heat-stable resistance to root-knot nematodes (*Meloidogyne* Spp.) Theor Appl Genet 103:592–600

Djian-Caporalino C, Fazari A, Arguel MJ et al (2007) Root-knot nematode (*Meloidogyne* spp.) Me resistance genes in pepper (*Capsicum annuum* L.) are clustered on the P9 chromosome. Theor Appl Genet 114:473–486

Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L (2010) MYB transcription factors in *Arabidopsis*. Trend Plant Sci 15:573–581

Fazari A, PalloixA WL, Hua YM, Sage-Palloix AM, Zhang BX, Djian-Caporalino C (2012) The root-knot nematode resistance N-gene co-localizes in the Me-genes cluster on the pepper (*Capsicum annuum* L.) P9 chromosome. Plant Breed 131:665–673

Fonseca S, Chico JM, Solano R (2009) The jasmonate pathway: the ligand, the receptor and the core signalling module. Curr Opin Plant Biol 12:539–547

Forrest KL, Bhave M (2007) Major intrinsic proteins (MIPs) in plants: a complex gene family with major impacts on plant phenotype. Funct Integr Genom 7:263

Fujita M, Fujita Y, Maruyama K, Seki M, Hiratsu K, Ohme-Takagi M, Tran LSP, Yamaguchi-Shinozaki K, Shinozaki K (2004) A dehydration-induced NAC protein, RD26, is involved in a novel ABA-dependent stress-signaling pathway. Plant J 39:863–876

Fujita M, Futija Y, Noutoshi Y, Takahashi F, Narusaka Y, Yamaguchi-Shinozaki K, Shinozaki K (2006) Cross-talk between abiotic and biotic stress responses: a current view from the points of convergence in the stress signaling networks. Curr Opin Plant Biol 9: 436–442

Gahlaut V, Jaiswal V, Kumar A, Gupta PK (2016) Transcription factors involved in drought tolerance and their possible role in developing drought tolerant cultivars with emphasis on wheat (*Triticum aestivum* L.). Theor Appl Genet 129:2019–2042

Galon Y, Finkler A, Fromm H (2010) Calcium-regulated transcription in plants. Mol Plant 3:653–669

Gao F, Chang F, Shen J, Shi F, Xie L, Zhan J (2014) Complete genome analysis of a novel recombinant isolate of potato virus Y from China. Arch Virol 159:3439–3442

Guo W-L, Wang S-B, Chen R-G, Chen B-H, Du X-H, Yin Y-X, Gong Z-H, Zhang Y-Y (2015) Characterization and expression profile of *CaNAC2* pepper gene. Front Plant Sci 6:755

Guo G, Wang S, Liu J, Pan B, Diao W, Ge W, Gao C, Snyder JC (2017) Rapid identification of QTLs underlying resistance to cucumber mosaic virus in pepper (*Capsicum frutescens*). Theor Appl Genet 130:41–52

Gutterson N, Reuber TL (2004) Regulation of disease resistance pathways by AP2/ERF transcription factors. Curr Opin Plant Biol 7:1–7

Hong Truong HT, Kim JH, Cho MC, Chae SY, Lee HE (2013) Identification and development of molecular markers linked to Phytophthora root rot resistance in pepper (*Capsicum annuum* L.) Eur J Plant Pathol 135:289–297

Hong JK, Choi HW, Hwang IS, Hwang BK (2007) Role of a novel pathogen-induced pepper C3-H-C4 type RING-finger protein gene, *CaRFP*1, in disease susceptibility and osmotic stress tolerance. Plant Mol Biol 63:571–588

Hong JK, Choi HW, Hwang IS, Kim DS, Kim NH, Choi DS, Kim YJ, Hwang BK (2008) Function of a novel GDSL-type pepper lipase gene, *CaGLIP1*, in disease susceptibility and abiotic stress tolerance. Planta 227:539–558

Hwang IS, Hwang BK (2010) The pepper 9-Lipoxygenase gene *CaLOX1* functions in defense and cell death responses to microbial pathogens. Plant Physiol 152:948–967

Hwang JN, Li J, Liu WY, An SJ, Cho H, Her NH, Yeam I, Kim D, Kang B (2009) Double mutations in eIF4E and eIFiso4E confer recessive resistance to Chilli Veinal mottle virus in pepper. Mol Cells 27:329–336

Ichimura K, Mizoguchi T, Yoshida R, Yuasa T, Shinozaki K (2000) Various abiotic stresses rapidly activate *Arabidopsis* MAP kinases *ATMPK4* and *ATMPK6*. Plant J 24:655–665

Isbat M, Zeba N, Kim SR, Hong CB (2009) A BAX inhibitor-1 gene in *Capsicum annuum* is induced under various abiotic stresses and endows multi-tolerance in transgenic tobacco. J Plant Physiol 166:1685–1693

Islam MM, Haque MS, Hossain MK, Hasan MM (2014) Diverse antioxidative effects in Pui vegetable (*Basella alba*) induced by high temperature stress. Intl J Agron Agri Res 5:135–147

Janska A, Marsik P, Zelenkova S, Ovesna J (2010) Cold stress and acclimation: what is important for metabolic adjustment? Plant Biol 12:395–405

Jiang CJ, Shimono M, Sugano S, Kojima M, Yazawa K, Yoshida R, Inoue H, Hayashi N, Sakakibara H, Takatsuji H (2010) Abscisic acid interacts antagonistically with salicylic acid signalling pathway in rice–*Magnaporthe grisea* interaction. Mol Plant-Micr Interact 23:791–798

Jing H, Li C, Ma F, Ma J-H, Khan A, Wang X, Zhao L-Y, Gong Z-H, Chen R-G (2016) Genome-wide identification, expression diversification of dehydrin gene family and characterization of *CaDHN3* in pepper (*Capsicum annuum* L.). PLoS ONE 11:e0161073

Jonak C, Okresz L, Bogre L, Hirt H (2002) Complexity, cross talk and integration of plant MAP kinase signaling. Curr Opin Plant Biol 5:415–424

Kang WH, Hoang NH, Yang HB et al (2010) Molecular mapping and characterization of a single dominant gene controlling CMV resistance in peppers (*Capsicum annuum* L.) Theor Appl Genet 120:1587–1596

Kasuga M, Liu Q, Miura S, Yamaguchi-shinozaki K, Shinozaki K (1999) Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. Nat Biotechnol 17:287–291

Kim DS, Hwang BK (2012) The pepper MLO gene, *CaMLO2*, is involved in the susceptibility cell-death response and bacterial and oomycete proliferation. Plant J 72:843–855

Kim DS, Hwang BK (2014) An important role of the pepper phenylalanine ammonia-lyase gene (*PAL1*) in salicylic acid-dependent signalling of the defence response to microbial pathogens. J Exp Bot 65:2295–2306

Kim SH, Hong JK, Lee SC, Sohn KH, Jung HW, Hwang BK (2004) *CAZFP1*, Cys2/His2-type zinc-finger transcription factor gene functions as a pathogen-induced early-defense gene in *Capsicum annuum*. Plant Mol Biol 55:883–904

Kim S-Y, Kim Y-C, Lee J-H, Oh S-K, Chung E, Lee S, Lee Y-H, Choi D, Park JM (2005) Identification of a

CaRAV1 possessing an AP2/ERF and B3 DNA-binding domain from pepper leaves infected with Xanthomonas axonopodis pv. glycines 8ra by differential display. Biochim Biophys Acta 1729: 141–146

Kim H-J, Nahm S-H, Lee H-R, Yoon G-B, Kim K-T, Kang B-C, Choi D, Kweon OY, Cho M-C, Kwon J-K, Han J-H, Kim J-H, Park M, Ahn JH, Choi SH, Her NH, Sung J-H, Kim B-D (2008) BAC-derived markers converted from RFLP linked to *Phytophthora capsici* resistance in pepper (*Capsicum annuum* L.). Theor Appl Genet 118:15–27

Kim HJ, Han J-H, Kim S, Lee HR, Shin J-S, Kim J-H, Cho J, Kim YH, Lee HJ, Kim B-D, Choi D (2011) Trichome density of main stem is tightly linked to *PepMoV* resistance in chili pepper (*Capsicum annuum L.*). Theor Appl Genet 122:1051–1058

Kim DS, Choi HW, Hwang BK (2014a) Pepper mildew resistance locus O interacts with pepper calmodulin and suppresses Xanthomonas AvrBsT-triggered cell death and defense responses. Planta 240:827–839

Kim JS, Park H-M, Chae S, Lee T-H, Hwang D-J, Oh S-D, Park J-S, Song D-G, Pan C-H, Choi D, Kim Y-H, Nahm BH, Kim Y-K (2014b) A pepper *MSRB2* gene confers drought tolerance in rice through the protection of chloroplast-targeted genes. PLoS ONE 9: e90588

Kirsten R, Jaglo-ottosen SJ, Gilmour DG, Zarka OS, Thomashow MF (1998) *Arabidopsis CBF1* overexpression induces COR genes and enhances freezing tolerance. Science 280:104–106

Koga H, Dohi K, Mori M (2004) Abscisic acid and low temperatures suppress the whole plant-specific resistance reaction of rice plants to the infection of *Magnaporthe grisea*. Physiol Mol Plant Pathol 65:3–9

Krasensky J, Jonak C (2012) Drought, salt, and temperature stress-induced metabolic rearrangements and regulatory networks. J Exp Bot 63:1593–1608

Kusajima M, Yasuda M, Kawashima A, Nojiri H, Yamane H, Nakajima M, Akutsu K, Nakashita H (2010) Suppressive effect of abscisic acid on systemic acquired resistance in tobacco plants. J Gen Plant Pathol 76:161–167

Lee SC, Choi HW, Hwang IS, Choi DS, Hwang BK (2006) Functional roles of the pepper pathogen-induced bZIP transcription factor, *CAbZIP1*, in enhanced resistance to pathogen infection and environmental stresses. Planta 224:1209–1225

Lee HK, Cho SK, Son O, Xu Z, Hwang I, Kim WT (2009) Drought stress-induced *Rma1H1*, a RING membrane-anchor E3 ubiquitin ligase homolog, regulates aquaporin levels via ubiquitination in transgenic *Arabidopsis* plants. Plant Cell 21:622–641

Lee SC, Choi DS, Hwang IS, Hwang BK (2010) The pepper oxidoreductase *CaOXR1* interacts with the transcription factor *CaRAV1* and is required for salt and osmotic stress tolerance. Plant Mol Biol 73:409–424

Lefebvre V, Daubèze A-M, Rouppe van der Voort J, Peleman J, Bardin M, Palloix A (2003) QTLs for resistance to powdery mildew in pepper under natural and artificial infections. Theor Appl Genet 107:661–666

Li N, Yin Y, Wang F, Yao M (2018) Construction of a high-density genetic map and identification of QTLs for cucumber mosaic virus resistance in pepper (*Capsicum annuum* L.) using specific length amplified fragment sequencing (SLAF-seq). Breed Sci 68:233–241

Lim CW, Lee SC (2014) Functional roles of the pepper MLO protein gene, *CaMLO2*, in abscisic acid signaling and drought sensitivity. Plant Mol Biol 85:1–10

Lim CW, Han SW, Hwang IS, Kim DS, Hwang BK, Lee SC (2015a) The pepper lipoxygenase *CaLOX1* plays a role in osmotic, drought and high salinity stress response. Plant Cell Physiol 56:930–942

Lim CW, Hwang BK, Lee SC (2015b) Functional roles of the pepper RING finger protein gene, *CaRING1*, in abscisic acid signaling and dehydration tolerance. Plant Mol Biol 89:143–156

Liu HX, Zhou XY, Dong N, Liu X, Zhang HY, Zhang ZY (2011) Expression of a wheat MYB gene in transgenic tobacco enhances resistance to *Ralstonia solanacearum*, and to drought and salt stresses. Funct Integr Genom 11:431–443

Liu W-Y, Kang J-H, Jeong H-S, Choi H-J, Yang H-B, Kim K-T, Choi D, Choi GJ, Jahn M, Kang B-C (2014) Combined use of bulked segregant analysis and microarrays reveals SNP markers pinpointing a major QTL for resistance to *Phytophthora capsici* in pepper. Theor Appl Genet 127:2503–2513

Liu Z, Shi L, Liu Y, Tang Q, Shen L, Yang S, Cai J, Yu H, Wang R, Wen J, Lin Y, Hu J, Liu C, Zhang Y, Mou S, He S (2015) Genome-wide identification and transcriptional expression analysis of mitogen-activated protein kinase and mitogen-activated protein kinase kinase genes in Capsicum annuum. Front Plant Sci 6:780

Mahajan S, Tuteja N 92005) Cold, salinity and drought stresses: an overview. Arch Biochem Biophys 444: 139–58

Maharijaya A, Vosman B, Steenhuis-Broers G, Pelgrom K, Purwito A, Visser RGF, Voorrips RE (2015) QTL mapping of thrips resistance in pepper. Theor Appl Genet 128:1945–1956

Mayak S, Tirosh T, Glick BR (2004) Plant growth-promoting bacteria that confer resistance to water stress in tomatoes and peppers. Plant Sci 166:525–530

Mengiste T, Chen X, Salmeron J, Dietrich R (2003) The *BOTRYTIS SUSCEPTIBLE1* gene encodes an R2R3MYB transcription factor protein that is required for biotic and abiotic stress responses in *Arabidopsis*. Plant Cell 15:2551–2565

Min HJ, Jung YJ, Kang BG, Kim WT (2016) *CaPUB1*, a hot pepper U-box E3 ubiquitin ligase, confers enhanced cold stress tolerance and decreased drought stress tolerance in transgenic rice (*Oryza sativa* L.). Mol Cells 39:250–257

Mittler R, Blumwald E (2010) Genetic engineering for modern agriculture: challenges and perspectives. Annu Rev Plant Biol 61:443–462

Mittler R, Vanderauwera S, Suzuki N, Miller G, Tognetti VB, Vandepoele K, Gollery M, Shulaev V, Van Breusegem F (2011) ROS signaling: the new wave? Trends Plant Sci 16:300–309

Mohr PG, Cahill DM (2003) Abscisic acid influences the susceptibility of *Arabidopsis thaliana* to *Pseudomonas syringae* pv. tomato and *Peronospora parasitica*. Funct Plant Biol 30:461–469

Moller IM, Jensen PE, Hansson A (2007) Oxidative modifications to cellular components in plants. Annu Rev Plant Biol 58:459–481

Moon S-J, Han S-Y, Kim D-Y, Yoon IS, Shin D, Byun M-O, Kwon H-B, Kim B-G (2015) Ectopic expression of a hot pepper bZIP-like transcription factor in potato enhances drought tolerance without decreasing tuber yield. Plant Mol Biol 89:421–431

Mou S, Liu Z, Gao F, Yang S, Su M, Shen L, Wu Y, He S (2017) *CaHDZ27*, a homeodomain leucine zipper i protein, positively regulates the resistance to *Ralstonia solanacearum* infection in pepper. Mol Plant-Micr Interact https://doi.org/10.1094/MPMI-06-17-0130-R

Muszynski MG, Dam T, Li B, Shirbroun DM, Hou Z, Bruggemann E, Archibald R, Ananiev EV, Danilevskaya ON (2006) *Delayed flowering1* encodes a basic leucine zipper protein that mediates floral inductive signals at the shoot apex in maize. Plant Physiol 142:1523–1536

Naegele RP, Ashrafi H, Hill TA, Chin-Wo SR, Van Deynze AE, Hausbeck MK (2014) QTL mapping of fruit rot resistance to the plant pathogen *Phytophthora capsici* in a recombinant inbred line *Capsicum annuum* population. Phytopathology 104:479–483

Nakagami H, Pitzschke A, Hirt H (2005) Emerging MAP kinase pathways in plant stress signaling. Trends Plant Sci 10:339–346

Nakashima K, Tran LSP, Van Nguyen D, Fujita M, Maruyama K, Todaka D, Ito Y, Hayashi N, Shinozaki K, Yamaguchi-Shinozaki K (2007) Functional analysis of a NAC-type transcription factor OsNAC6 involved in abiotic and biotic stress-responsive gene expression in rice. Plant J 51:617–630

Nuhse TS, Peck SC, Hirt H, Boller T (2000) Microbial elicitors induce activation and dual phosphorylation of the *Arabidopsis thaliana* MAPK 6. J Biol Chem 275:7521–7526

Oh S-K, Yi SY, Yu SH, Moon JS, Park JM, Choi D (2006) *CaWRKY2*, a chili pepper transcription factor, is rapidly induced by incompatible plant pathogens. Mol Cells 22:58–64

Oh SK, Baek KH, Seong ES et al (2010) *CaMsrB2*, pepper methionine sulfoxide reductase B2, is a novel defense regulator against oxidative stress and pathogen attack. Plant Physiol 154:245–261

Park JM, Park CJ, Lee SB, Ham BK, Shin R, Paek KH (2001) Overexpression of the tobacco *Tsi1* gene encoding an EREBP/AP2-type transcription factor enhances resistance against pathogen attack and osmotic stress in tobacco. Plant Cell 13:1035–1046

Park C, Lim CW, Baek W, Kim J-H, Lim S, Kim SH, Kim K-N, Lee SC (2017) The pepper WPP domain Protein, *CaWDP1*, acts as a novel negative regulator of drought stress via ABA signaling. Plant Cell Physiol 58:779–788

Peng XX, Tang XK, Zhou PL, Hu YJ, Deng XB, He Y, Wang HH (2011) Isolation and expression patterns of rice WRKY82 transcription factor gene responsive to both biotic and abiotic stresses. Agri Sci China 10:893–901

Pieterse CMJ, Leon-Reyes A, Van der Ent S, Van Wees SCM (2009) Networking by small-molecule hormones in plant immunity. Nat Chem Biol 5:308–316

Pitzschke A, Schikora A, Hirt H (2009) MAPK cascade signalling networks in plant defence. Curr Opin Plant Biol 12:421–426

Qin L, Mo N, Muhammad T, Liang Y (2018) Genome-Wide Analysis of *DCL, AGO*, and *RDR* Gene families in pepper (*Capsicum annuum* L.). Intl J Mol Sci 19. https://doi.org/10.3390/ijms19041038

Qiu YP, Yu DQ (2009) Over-expression of the stress-induced *OsWRKY45* enhances disease resistance and drought tolerance in *Arabidopsis*. Environ Exp Bot 65:35–47

Quirin EA, Ogundiwin EA, Prince JP, Mazourek M, Briggs MO, Chlanda TS, Kim K-T, Falise M, Kang B-C, Jahn MM (2005) Development of sequence characterized amplified region (SCAR) primers for the detection of Phyto. 5.2, a major QTL for resistance to Phytophthora capsici Leon. in pepper. Theor Appl Genet 110:605–612

Reeves G, Monroy-Barbosa A, Bosland PW (2013) A novel *Capsicum* gene inhibits host-specific disease resistance to *Phytophthora capsici*. Phytopathology 103:472–478

Rodriguez MCS, Petersen M, Mundy J (2010) Mitogen-activated protein kinase signaling in plants. AnnuRev Plant Biol 61:621–649

Romer P, Jordan T, Lahaye T (2010) Identification and application of a DNA-based marker that is diagnostic for the pepper (*Capsicum annuum*) bacterial spot resistance gene Bs3. Plant Breed 129:737–740

Rorat T (2006) Plant dehydrins—tissue location, structure and function. Cell Mol Biol Lett 11:536–556

Rubio M, Caranta C, Palloix A (2008) Functional markers for selection of potyvirus resistance alleles at the pvr2-eIF4E locus in pepper using tetra-primer ARMS-PCR. Genome 51:767–771

Ruffel S, Gallois JL, Moury B, Robaglia C, Palloix A, Caranta C (2006) Simultaneous mutations in translation initiation factors eIF4E and eIF(iso)4E are required to prevent pepper veinal mottle virus infection of pepper. J Gen Virol 87:2089–2098

Sanghera GS, Wani SH, Hussain W, Singh NB (2011) Engineering cold stress tolerance in crop plants. Curr Genom 12:30–43

Seo PJ, Park CM (2010) MYB96-mediated abscisic acid signals induce pathogen resistance response by promoting salicylic acid biosynthesis in *Arabidopsis*. New Phytol 186:471–483

Seo PJ, Lee SB, Suh MC, Park MJ, Go YS, Park CM (2011) The MYB96 transcription factor regulates

cuticular wax biosynthesis under drought conditions in *Arabidopsis*. Plant Cell 23:1138–1152

Seo YS, Choi JY, Kim SJ, Kim EY, Shin JS, Kim WT (2012) Constitutive expression of *CaRma1H1*, a hot pepper ER-localized RING E3 ubiquitin ligase, increases tolerance to drought and salt stresses in transgenic tomato plants. Plant Cell Rep 31:1659–1665

Seong ES, Wang M-H (2008) A novel *CaAbsi1* gene induced by early-abiotic stresses in pepper. BMB Rep 41:86–91

Seong ES, Choi D, Cho HS, Lim CK, Cho HJ, Wang M-H (2007) Characterization of a stress-responsive ankyrin repeat-containing zinc finger protein of *Capsicum annuum* (*CaKR1*). J Biochem Mol Biol 40:952–958

Sohn KH, Lee SC, Jung HW, Hong JK, Hwang BK (2006) Expression and functional roles of the pepper pathogen-induced transcription factor RAV1 in bacterial disease resistance, and drought and salt stress tolerance. Plant Mol Biol 61:897–915

Sun C, Mao SL, Zhang ZH, Palloix A, Wang LH, Zhang BX (2015) Resistances to anthracnose (*Colletotrichum acutatum*) of *Capsicum* mature green and ripe fruit are controlled by a major dominant cluster of QTLs on chromosome P5. Sci Hort 181:81–88

Suwor P, Sanitchon J, Thummabenjaone P, Kumar S, Hawongstien ST (2017) Inheritance analysis of anthracnose resistance and marker-assisted selection in introgression populations of chili (*Capsicum annuum* L.). Sci Hort 220:20–26

Szabala BM, Fudali S, Rorat T (2014) Accumulation of acidic SK_3 dehydrins in phloem cells of cold- and drought-stressed plants of the *Solanaceae*. Planta 239:847–863

Sziderics AH, Rasche F, Trognitz F, Sessitsch A, Wilhelm E (2007) Bacterial endophytes contribute to abiotic stress adaptation in pepper plants (*Capsicum annuum* L.). Can J Microbiol 53:1195–1202

Takahashi F, Mizoguchi T, Yoshida R, Ichimura K, Shinozaki K (2011) Calmodulin-dependent activation of MAP kinase for ROS homeostasis in *Arabidopsis*. Molr Cell 41:649–660

Tamisier L, Rousseau E, Barraillé S, Nemouchi G, Szadkowski M, Mailleret L, Grognard F, Fabre F, Moury B, Palloix A (2017) Quantitative trait loci in pepper control the effective population size of two RNA viruses at inoculation. J Gen Virol 98:1923–1931

Teige M, Scheikl E, Eulgem T, Doczi R, Ichimura K, Shinozaki K, Dangl JL, Hirt H (2004) The MKK2 pathway mediates cold and salt stress signaling in *Arabidopsis*. Mol Cell 15:141–152

Thabuis A, Palloix A, Pflieger S, Daubèze A-M, Caranta C, Lefebvre V (2003) Comparative mapping of *Phytophthora* resistance loci in pepper germplasm: evidence for conserved resistance loci across *Solanaceae* and for a large genetic diversity. Theor Appl Genet 106:1473–1485

Thabuis A, Lefebvre V, Bernard G, Daubèze AM, Phaly T, Pochard E, Palloix A (2004) Phenotypic and molecular evaluation of a recurrent selection program for a polygenic resistance to *Phytophthora capsici* in pepper. Theor Appl Genet 109:342–351

Timmusk S, Nicander B, Granhall U, Tillberg E (1999) Cytokinin production by *Paenibacillus polymyxa*. Soil Biol Biochem 31:1847–1852

Ton J, Flors V, Mauch-Mani B (2009) The multifaceted role of ABA in disease resistance. Trends Plant Sci 14:310–317

Tsutsui T, Kato W, Asada Y et al (2009) DEAR1, a transcriptional repressor of DREB protein that mediates plant defense and freezing stress responses in *Arabidopsis*. J Plant Res 122:633–643

Vallejos CE, Jones V, Stall RE et al (2010) Characterization of two recessive genes controlling resistance to all races of bacterial spot in peppers. Theor Appl Genet 121:37–46

Vannini C, Iriti M, Bracale M, Locatelli F, Faoro F, Croce P, Pirona R, Di Maro A, Coraggio I, Genga A (2006) The ectopic expression of the rice *Osmyb4* gene in *Arabidopsis* increases tolerance to abiotic, environmental and biotic stresses. Physiol Mol Plant Pathol 69:26–42

Vannini C, Campa M, Iriti M, Genga A, Faoro F, Carravieri S, Rotino GL, Rossoni M, Spinardi A, Bracale M (2007) Evaluation of transgenic tomato plants ectopically expressing the rice *Osmyb4* gene. Plant Sci 173:231–239

Vessey JK (2003) Plant growth promoting rhizobacteria as biofertilizers. Plant Soil 255:571–586

Vogel JT, Zarka DG, Van Buskirk HA, Fowler SG, Thomashow MF (2005) Roles of the CBF2 and ZAT12 transcription factors in configuring the low temperature transcriptome of *Arabidopsis*. Plant J 41:195–211

Voorrips RE, Finkers R, Sanjaya L, Groenwold R (2004) QTL mapping of anthracnose (*Colletotrichum spp*.) resistance in a cross between *Capsicum annuum* and *C. chinense*. Theor Appl Genet 109:1275–1282

Walley JW, Dehesh K (2010) Molecular mechanisms regulating rapid stress signaling networks in *Arabidopsis*. J Integr Plant Biol 52:354–359

Wang X, Zhu X, Tooley P, Zhang X (2013) Cloning and functional analysis of three genes encoding polygalacturonase-inhibiting proteins from *Capsicum annuum* and transgenic *CaPGIP1* in tobacco in relation to increased resistance to two fungal pathogens. Plant Mol Biol 81:379–400

Wang P, Liu X, Guo J, Liu C, Fu N, Shen H (2015) Identification and expression analysis of candidate genes associated with defense responses to *Phytophthora capsici* in pepper line "PI 201234". Intl J Mol Sci 16:11417–11438

Wang H, Niu H, Zhai Y, Lu M (2017) Characterization of *BiP* genes from pepper (*Capsicum annuum* L.) and the role of *CaBiP1* in response to endoplasmic reticulum and multiple abiotic stresses. Front Plant Sci 8:1122

Wong HL, Shimamoto K (2009) Sending ROS on a bullet train. Sci Signal 2: pe60

Wu Z, Cheng J, Cui J, Xu X, Liang G, Luo X, Chen X, Tang X, Hu K, Qin C (2016) Genome-wide identification and expression profile of dof transcription factor gene family in pepper (*Capsicum annuum* L.). Front Plant Sci 7:574

Xia N, Zhang G, Liu XY, Deng L, Cai GL, Zhang Y, Wang XJ, Zhao J, Huang LL, Kang ZS (2010) Characterization of a novel wheat NAC transcription factor gene involved in defense response against stripe rust pathogen infection and abiotic stresses. Mol Biol Rep 37:3703–3712

Xiong L, Yang Y (2003) Disease resistance and abiotic stress tolerance in rice are inversely modulated by an abscisic acid-inducible mitogen-activated protein kinase. Plant Cell 15:745–759

Xu ZS, Chen M, Li LC, Ma YZ (2011) Functions and application of the AP2/ERF transcription factor family in crop improvement. J Integr Plant Biol 53:570–585

Yao M, Li N, Wang F, Ye Z (2013) Genetic analysis and identification of QTLs for resistance to cucumber mosaic virus in chili pepper (*Capsicum annuum* L.). Euphytica 193:135–145

Yasuda M, Ishikawa A, Jikumaru Y et al (2008) Antagonistic interaction between systemic acquired resistance and the abscisic acid-mediated abiotic stress response in *Arabidopsis*. Plant Cell 20:1678–1692

Yi SY, Kim J-H, Joung Y-H, Lee S, Kim W-T, Yu SH, Choi D (2004) The pepper transcription factor *CaPF1* confers pathogen and freezing tolerance in *Arabidopsis*. Plant Physiol 136:2862–2874

Yi SY, Lee DJ, Yeom SI, Yoon J, Kim YH, Kwon SY, Choi D (2010) A novel pepper (*Capsicum annuum*) receptor-like kinase functions as a negative regulator of plant cell death via accumulation of superoxide anions. New Phytol 185:701–715

Yin Y-X, Wang S-B, Zhang H-X, Xiao H-J, Jin J-H, Ji J-J, Jing H, Chen R-G, Arisha MH, Gong Z-H (2015) Cloning and expression analysis of *CaPIP1-1* gene in pepper (*Capsicum annuum* L.). Gene 563:87–93

Yoshida T, Fujita Y, Sayama H, Kidokoro S, Maruyama K, Mizoi J, Shinozaki K, Yamaguchi-Shinozaki K (2010) AREB1, AREB2, and ABF3 are master transcription factors that cooperatively regulate ABRE-dependent ABA signaling involved in drought stress tolerance and require ABA for full activation. Plant J 61:672–685

Yu C, Zhan Y, Feng X, Huang Z-A, Sun C (2017) Identification and expression profiling of the auxin response factors in *Capsicum annuum* L. under abiotic stress and hormone treatments. Intl J Mol Sci 18: https://doi.org/10.3390/ijms18122719

Zhai Y, Guo M, Wang H, Lu J, Liu J, Zhang C, Gong Z, Lu M (2016) Autophagy, a conserved mechanism for protein degradation, responds to heat, and other abiotic stresses in *Capsicum annuum* L. Front Plant Sci 7:131

Zhang T, Liu Y, Yang T, Zhang L, Xu S, Xue L, An L (2006) Diverse signals converge at MAPK cascades in plant. Plant Physiol Biochem 44:274–283

Zhang YL, Jia QL, Li DW, Wang JE, Yin YX, Gong ZH (2013) Characteristic of the pepper *CaRGA2* gene in defense responses against *Phytophthora capsici* Leonian. Intl J Mol Sci 14:8985–9004

Zheng Z, Nonomura T, Appiano M et al (2013) Loss of function in *Mlo* orthologs reduces susceptibility of pepper and tomato to powdery mildew disease caused by *Leveillula taurica*. PLoS ONE 8(7):e70723

8 Genome Sequencing of *Capsicum* Species: Strategies, Assembly, and Annotation of Genes

Pasquale Tripodi, Alberto Acquadro, Sergio Lanteri and Nunzio D'Agostino

Abstract

Pepper (*Capsicum* spp.) belongs to the *Solanaceae*, which is an economically important family of flowering plants consisting of about 102 genera and over 2500 species. The *Solanaceae* family includes crops of agronomic importance for which the efforts in genome sequencing are ongoing by almost 10 years (https://www.solgenomics.net/organism/sol100/view). Since the beginning of 2014, various consortia have released the genome sequences of domesticated and wild *Capsicum* species. The first effort was focused on the whole-genome sequencing of *Capsicum annuum* CM334 and of *Capsicum chinense* PI159236, which were widely used as founders of mapping populations and carry important disease resistance traits. Just a couple of months later, the genome sequences of *C. annuum* Zunla-1 and of the wild species Chiltepin (C. *annuum* var. *glabriusculum*) were published. Both studies reported a pepper genome size of ~3–3.5 Gb, rich in repetitive elements (over 80%) with about 35 thousand genes. The improved version of the reference genome CM334 as well as of *C. chinense* PI159236 together with the sequencing of the domesticated *Capsicum baccatum* revealed evolutionary relationships and estimated lineage divergence times occurring in *Capsicum*. Recently, the linked-read sequencing technology has been applied for the sequencing of a *C. annuum* accession that was an F_1 cross hybrid of CM334 and a non-pungent pepper breeding line. Furthermore, genome resequencing studies have been performed with the aim to analyze *loci* of interest related to biotic/abiotic stresses and to qualitative features. In this chapter, we provide an overview of the genome sequencing and annotation strategies and describe the main results disclosed by all the whole and targeted genome sequencing projects in *Capsicum*.

8.1 International Initiatives in Pepper Genome Sequencing

Since the release of the first whole-genome sequence of a plant species (Arabidopsis Genome Initiative 2000), various national and international initiatives have led to the sequencing and assembling of genomes of both crop and non-crop plants from different clades. During the last five years, transnational research consortia released the genome sequences of domesticated and wild peppers. The main information arising from *Capsicum* spp. genome sequencing has been obtained by two international groups: the

P. Tripodi (✉) · N. D'Agostino
CREA Research Centre for Vegetable and Ornamental Crops, Pontecagnano Faiano, Italy
e-mail: pasquale.tripodi@crea.gov.it

A. Acquadro · S. Lanteri
DISAFA, Plant Genetics and Breeding, University of Torino, Turin, Italy

N. Ramchiary and C. Kole (eds.), *The Capsicum Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-319-97217-6_8

former composed by scientists of 27 institutes from Korea, Israel, and USA (Kim et al. 2014), the latter by researchers of 13 institutes from China, Mexico, France, and USA (Qin et al. 2014). Since then, two other complete genome sequences have been released into the public domain (Kim et al. 2017; Hulse-Kemp et al. 2018). Descriptions of the genome sequencing and annotation strategies as well as the main results disclosed are reported below.

8.1.1 The Genome Sequence of Hot Pepper

In January 2014, Kim et al. (2014) reported the whole-genome sequencing and assembly of the Mexican landrace *C. annuum* cv. Criollo de Morelos 334 (hereafter CM334) and the *Capsicum chinense* accession PI159236 as the foundation for interspecies comparative analysis (Table 8.1). Both accessions were selected for being resistant to diseases including *Phytophthora* spp., Nematodes, *Tobacco Mosaic Virus* (TMV), *Potato Virus Y* (PVY), *Tomato Spotted Wilt Virus* (TSWV), *Pepper Mottle Virus* (PepMoV), *Tobacco Etch Potyvirus* (TEV). The authors also accomplished the resequencing of two cultivated peppers: *C. annuum* cv. 'Perennial' and *C. annuum* cv. 'Dempsey', which were the parents of a 120 recombinant inbred F_8 line (RIL) population used for the development of high-density genetic and physical maps. Paired-end (PE) and mate-pair (MP) libraries were sequenced on Illumina platforms (GAII and HiSeq 2000). As for CM334, a total of 650.2 Gb of genome sequence (coverage 186.6×) was generated from genomic libraries with insert sizes ranging from 180 bp to 20 Kb and read lengths of 36, 76, or 101 bp. In case of *C. chinense*, a total of 289.6 Gb of genomic sequence data was generated. Prior to the genome assembly, an *in-house* preprocessing pipeline was adopted to filter out low-quality sequences for short reads, by eliminating contamination from publicly available bacterial genome sequences (identity > 98%, coverage > 50%), duplicated reads, low-quality sequences as well as correcting substitution sequencing errors. A 19-mer analysis was performed to determine the genome size of CM334 and PI159236, which were estimated to be 3.48 and 3.14 Gb, respectively. To generate initial contigs, all the reads from both CM334 and *C. chinense* libraries were first merged to single reads ignoring pair information and then assembled into 37,989 scaffolds with N_{50} (the average length value of fragments for the 50% of the genome) of 2.47 Mb. The assembled CM334 genome sequence was validated with 27 bacterial artificial chromosomes (BACs) from euchromatic/heterochromatic regions with insert size larger than 70 Kb. All scaffolds matched the complete BAC sequences with more than 99.9% identity and 26 BACs were covered by single scaffolds. The quality of PI159236 genome assembly was assessed on the basis of about 600 *C. chinense* expressed sequence tags (ESTs) and mRNAs and additional 35,000 annotated CM334 genes, of which 97% matched with the PI159236 assembly. The validation confirmed about 23,000 genes in the *C. chinense* genome and made it possible the identification of a core gene set shared by the two accessions (Kim et al. 2014). To support the scaffolding process and construct pseudomolecules, a high-density genetic map was generated through low-depth (1×) whole-genome sequencing of 120 intraspecific *C. annuum* RILs ('Perennial' × 'Dempsey'). Over 3 million single nucleotide polymorphism (SNP) markers were identified among CM334 and the parents of the RILs, and a set of 21,121 markers were selected for map construction (Kim et al. 2014). The final map consisted of 12 linkage groups and 6281 markers covering 3796 cM. A subset of 4562 markers (73%) allowed to anchor 86% of scaffolds (2.63 Gb; 1357 scaffolds) to 12 pseudo-molecules. Accuracy of the linkage map was validated using the conserved ortholog set II (COSII) maps previously developed (Wu et al. 2009). Resequencing data revealed the presence of 10.9 and 11.9 million divergent SNPs in 'Perennial' and 'Dempsey', respectively, when compared with the CM334 reference sequence. Sequencing of *C. chinense* highlighted 56.6 million SNPs compared to CM334. As a result,

94.5, 94.3, and 89.6% of the CM334 genome was covered by 'Perennial', 'Dempsey' and *C. chinense* sequences, respectively. Transposable elements (TEs), which have played a key role in shaping the DNA landscape of genomes during evolution and led to the conversion of euchromatin into heterochromatin, were found to represent a preponderant portion on the whole-genome in respect to tomato as well as other sequenced *Solanaceae* genomes, being the 76.4% (i.e., 2.34 Gb) in CM334 and 79.6% (i.e., 2.35 Gb) in *C. chinense*. TEs were widely dispersed throughout the pepper genome and their distribution was inversely correlated with gene density. The most frequent TEs were long terminal repeat (LTR) elements, representing more than 70% of the identified TEs in the two genomes. The composition of these repetitive DNA sequence motifs widely differs from the one detected in other crops, as the *Gypsy* elements were 12-fold more than the *Copia* elements and their proliferation caused the expansion of the pepper genome. On the other hand, the accumulation of *Copia* and *Tat* elements (a subgroup of the *Gypsy* clade) was responsible for the expansion of hot pepper euchromatin. A consensus annotation of 34,903 protein-coding genes was generated for CM334 (Pepper Genome Annotation v. 1.5). Over 93% of the predicted protein-coding sequences was supported by ~20 Gb Illumina RNA-seq data from four tissues/organs (flower, root, leaf, and fruit) at different stages of plant development. Furthermore, 177 microRNAs, corresponding to 37 microRNA families, were identified in CM334. The number of annotated genes was similar to the one previously identified in tomato and potato. Overall, the authors reported 23,245 hot pepper genes distributed in 16,345 families. By comparing pepper and tomato genomes, it was possible to identify 17,397 orthologous genes, whose expression was investigated through RNA-seq from tissues at the same developmental stage of the plants. In both crop species, it was possible to identify a high number of differentially expressed genes (DEGs) in pericarp and placenta (34 and 24.7% on average, respectively) while in root and leaf the number of DEGs was relatively low (15.1 and 8.8%, respectively). The distribution of orthologous gene families of six crops (hot pepper, tomato, potato, Arabidopsis, grape, and rice) allowed to identify 7826 shared gene families and 756 unique families to hot pepper. Furthermore, variations in family size were found in many hot pepper gene families, such as those involved in disease resistance and

Table 8.1 Comparison of the main features of the *Capsicum* sequenced genomes

	C. annuum CM334 (v1.55)[a]	*C. chinense* PI159236[a]	Zunla-1[b]	Chiltepin[b]	*C. baccatum* PBC81[c]	*C. chinense* PI159236[c]	CM334 F_1[d]
Total sequence length (Gb)	650.2	289.6	477.37	295.85	526.7	425.7	104.7
Sequencing depth (X)	186.6	83.2	146.43	96.37	136.1	132.2	56
Genome size (Gb)	3.48	3.14	3.26	3.07	3.9	3.2	3.2
Scaffold no	37.989	239.495	28,149*	30,293*	2.083	1.557	83.391
TE elements %	76.4	79.6	80.9	81.4	85	85	*na*
Genes number	34.903	33.788	35.336	34.476	35.874	35.009	*na*

[a]Kim et al. (2014)
[b]Qin et al. (2014)
[c]Kim et al. (2017)
[d]Hulse-Kemp et al. (2018)
* >2 K bp
na not available (annotation not provided)

cellular functions (i.e., *cytochrome P450* and heat shock protein genes).

8.1.2 The Genome Sequence of Cultivated and Wild Peppers

Few months after the release of the CM334 genome, Qin et al. (2014) published the reference genome sequences of the cultivated pepper Zunla-1 (*C. annuum* L.) and its wild progenitor Chiltepin (*C. annuum* var. *glabriusculum,* also termed *C. annuum* var. *aviculare*) (Table 8.1). Zunla-1 is an F_9 inbred line derived from a cross between two *C. annuum* cultivars grown by small farmers in China, while Chiltepin is a landrace collected in the north-central Mexico.

Eleven (6 PE and 5 MP) and nine (5 PE and 4 MP) Illumina libraries with different insert sizes were prepared for Zunla-1 and Chiltepin, respectively. These libraries were sequenced using the Illumina Genome Analyzer II device and generated 477.37 Gb (146.43× coverage) of raw sequencing data for Zunla-1 and 295.85 Gb (96.37× coverage) for Chiltepin.

After filtering out low-quality and duplicate reads, a total of 325.29 Gb (99.78× coverage) of high-quality sequence data for Zunla-1 and of 204.86 Gb (66.73× coverage) for Chiltepin was retained. For Zunla-1, after filling the gaps, the total scaffolds size was ~3.35 Gb (N_{50} = 1.22 Mb) and the total contig size was ~3.21 Gb (N_{50} = 55.43 Kb). For Chiltepin, after gap filling, the total scaffold size was ~3.48 Gb (N_{50} = 444.59 Kb), while the total contig size was ~3.3 Gb (N_{50} = 52.23 Kb). Based on an intraspecific *C. annuum* F_2 population, a high-resolution genetic map with 7657 SNP markers was generated and used to anchor and orient 4956 scaffolds from Zunla-1 to the 12 chromosome pseudo-molecules. Overall, 78.95% of the assembly (~2.64 Gb; 1822 scaffolds) was successfully anchored to the 12 pseudo-chromosomes. The unplaced 3134 scaffolds (705 Mb in total) were assigned to a pseudo-chromosome designated as '00'. By comparing the genetic and the physical distances, similar patterns of recombination were detected, which were markedly reduced in broad pericentromeric regions and consistent at chromosome ends. LASTZ (Large-Scale Genome Alignment Tool) (Harris 2007) was used to align the assembly of Chiltepin chromosomes to the Zunla-1 reference genome. The completeness and quality of the assemblies were evaluated by aligning pepper ESTs available at dbEST (https://www.ncbi.nlm.nih.gov/dbEST) as well as Illumina reads generated from short insert size libraries onto Zunla-1 and Chiltepin genomes, respectively. Similarly to what previously reported for the CM334 genome (Kim et al. 2014), more than 81% (~2.7 Gb) of the Zunla-1 and Chiltepin genomes is composed by transposable elements (TEs), most of which are LTR retrotransposons of the *Gypsy* clade (54.5%) followed by *Copia* (8.6%). Divergence analysis allowed to date the insertion time of LTRs ~0.3 million years ago (Mya), suggesting that the expansion of the pepper genome was quite recent during the evolution of the *Solanaceae* family.

In total, 35,336 and 34,476 protein-coding genes were predicted with high confidence in Zunla-1 and Chiltepin, respectively. Furthermore, over 90% of predicted genes were supported by different items of evidence (ESTs; RNA-seq data; homologous proteins). Gene discovery and annotation benefited from the generation of 30 RNA-seq libraries (over 90 Gb of sequence data) from various tissues/organs at different developmental stages. RNA-seq expression profiles highlighted constitutively expressed (over 31%) as well as tissue-specific genes. Discovery and annotation of long non-coding RNAs (lncRNAs) as well as of short interference RNAs (siRNAs) and microRNAs (miRNAs) in Zunla-1 was also performed by the RNA-sequencing of a flower bud library and five small RNA libraries from different tissues. Over 6500 lncRNAs, 5500 phased siRNAs and 176 miRNAs were identified. A set of 141 miRNAs were in common with other *Solanaceae*, while 35 miRNAs were classified as pepper-specific. Over 1100 target genes were identified, mostly coding for transcription factors, of which 78% have putative functions.

8.1.3 The Genome Sequence of *C. baccatum* and *C. chinense*

Recently, researchers of the consortium which previously released the CM334 genome, performed the sequencing and assembly of the genome of *C. baccatum* PBC81 (hereafter, *Baccatum*) and provided an improved version of the reference genome of both CM334 and *C. chinense* PI159236 (hereafter, *Chinense*) (Kim et al. 2017; Table 8.1).

The Illumina HiSeq 2500 platform was used for the sequencing of libraries with insert sizes in the range of 200 bp–10 Kb. In total, 526.7 Gb (136.1× coverage) and 425.7 Gb (132.2× coverage) of the *Baccatum* and *Chinense* genomes were generated. On the basis of 19-mer analysis, the estimated genome sizes were 3.9 and 3.2 Gb, respectively. For scaffold anchoring, high-genetic density maps were developed following genotype by sequencing of an F_2 *C. baccatum* intraspecific population (obtained by crossing lines 'Golden-aji' and 'PI594137') as well as segregating interspecific populations obtained by crossing *C. annuum* and *C. chinense*. The assembled genomes of *Baccatum* and *Chinense* were organized into 12 chromosomes-scale pseudo-molecules, being 3.2 and 3.0 Gb in size with scaffold N_{50} of 2.0 and 3.3 Mb, respectively. The total length of successfully anchored scaffolds were 2.8 Gb in (2083 scaffolds) for *Baccatum* and 2.8 Gb (1557 scaffolds) for *Chinense*, accounting for the 87 and 89% of the pepper genome, respectively.

As expected, repeated sequences represented the 85% of the entire genome and, in each species, over half was made up of LTR retrotransposons of the *Gypsy* clade. In the *Baccatum* genome, *Athila* elements were found to be more abundant (>two fold) and contributed to species-specific genome expansion in the *C. baccatum* lineage. On average, about 35,000 genes were annotated in both the *Baccatum* and *Chinense* genomes. In addition, a comparison between the updated and previous protein-coding gene annotation of CM334 revealed differences in ~10,000 gene models, most of which were associated with TEs in the previous genome annotation.

The phylogenetic analysis on *Baccatum*, *Chinense* and CM334 revealed a first lineage divergence between *Baccatum* and a progenitor of the other two peppers at about 1.7 Mya, followed by divergence between CM334 and *Chinense* at 1.1 Mya. It is noteworthy that comparison between *Baccatum*, *Chinense* and CM334 disclosed important dynamic genome rearrangements involving translocations among chromosomes 3, 5, and 9 differentiating *C. baccatum* from the other two species.

8.1.4 Linked-Read Sequencing of Reference Genome

In 2018, the pioneering linked-read sequencing technology has been applied in *C. annuum* (Hulse-Kemp et al. 2018) and generated a highly ordered and more contiguous sequence assembly in respect to the available *C. annuum* reference genomes (Table 8.1). This technology was used to sequence a F_1 heterozygous individual from a cross between CM334 and a non-pungent blocky accession. The authors used Illumina HiSeq × Ten sequencer (10× Chromium technology) to produce 2 × 150 paired-end sequences (56× coverage). The Supernova Assembler (Weisenfeld et al. 2017) was used to resolve complex repeats and separate chromosomes based on haplotype information. It produced locally phased haplotype blocks, or pseudohaps, as output. In particular, two individual haplotypes were generated. With the aim of generating a reference assembly (hereafter UCD10X), a single pseudohap was utilized. Indeed, the pseudohap1 assembly was made up of 83,391 scaffold sequences for a total size of 3.21 Gb. Over 83% of the assembled sequence (~2.67 Gb) was anchored to the 12 chromosomes along with 541 Mb of unplaced sequence. The N_{50} was 123 Kb, 3.69 Mb and 227.2 Mb for contigs, scaffolds, and pseudo-molecules, respectively. The quality of the assembly was assessed by comparing the order of contigs with four high-density pepper genetic maps (three

transcriptome-based and one genomic-based) and highlighted a concordant marker order. Furthermore, physical location of markers was also compared with the CM334 genome V1.55 pointing out that marker positioning in pericentromeric regions is more reliable in case of UCD10X. This assembly was, in the end, compared with the other publicly available pepper genomes (i.e. CM334V. 1.55, Zunla-1V. 2.0, Chiltepin V 2.0) in terms of length of scaffold sequences and overall size of the assembly. All the genome assemblies were comparable even if the quality within pseudo-chromosomes was variable, especially in heterochromatic regions. On the whole, although some regions were not accurately assembled with the linked-read library technology, the latter demonstrated to provide a valuable tool also for the de novo assembly of complex, highly repetitive, and heterozygous plant genomes.

8.1.5 Insight into Genome Expansion

Similar to what was observed in tomato (Tomato Genome Consortium 2012) and petunia (Bombarely et al. 2016), pepper is a paleohexaploid as its genome is the results of ancient triplication event. Since its speciation within the *Solanaceae* family, the pepper genome experienced no additional whole-genome duplication; however, its size is approximately four times than the one of tomato and threefold larger than the one of potato. Tomato and pepper genomes share syntenic blocks highly conserved and a high representation of LTR retrotransposons. The genome released by Kim et al. (2014) evidenced that pepper chromosomes highly expanded in both euchromatic and heterochromatic regions with respect to other *Solanaceae*. Most regions of the pepper genome are very rich in constitutive heterochromatin, which consists mostly of repetitive sequences and transposable elements. Comparison with the tomato genome suggested that the gene-rich regions near heterochromatin in tomato became heterochromatic regions in the pepper by accumulating repetitive sequences. In both species, the distribution of LTR retrotransposons was investigated highlighting a major representation of Del (Gypsy superfamily) and the existence of specific elements in pepper (Pseudovirus, Sire, CMR) and tomato (CoDi-D). Moreover, all repetitive elements in tomato were found in pericentromeric heterochromatin regions while in pepper, their distribution was observed in both heterochromatic and euchromatic regions.

Both Qin et al. (2014) and Kim et al. (2014) evidenced how the genes responsible of pungency synthesis underwent to duplication events, highlighting the existence of independent duplications in 13 gene families compared with *Arabidopis*, tomato and potato. Recently, retroduplication events (RTE) were described in NLR genes which are the major contributors of resistances in plants (Kim et al. 2017). The authors confirmed how RTE are common phenomena involved in the evolution of plants.

8.1.6 Gene Families

In pepper, 16,956 gene families were reported accounting for 22,885 genes identified among the predicted 34,447 protein-coding sequences (Kim et al. 2014). A similar number was found in the genome released by Qin et al. (2014), in which 16,770 families, including 26,444 genes out of 35,336 protein-coding sequences, were detected. Over two thousands transcription factors and transcriptional regulators (6.25% of predicted genes), which cluster in 80 families, were identified. The number was comparable with that of other plant species, although ABI3VP1 and RWP-RK families were most represented. Overall, 85% (1829) of TF genes were anchored to the pseudo-chromosomes with higher and lower concentration on chromosome 3 (257) and 10 (102), respectively. Among transcription factors, 73 families were represented by WRKY genes and 106 by NAC, both involved in plant development and defense mechanisms. Expression profiles of NAC genes evinced that about 30% of the genes were highly expressed during fruit developmental stages and 13% with the highest abundance in developing fruits.

One hundred and twenty-three genes belonging to the AP2/ERF superfamily (cellular responses, growth, and development) were identified. This family was mainly represented by different subfamilies including ERF (80%), AP2 (17%), RAV (1.6%).

The cytochrome P450s family was represented by 447 genes distributed in 9 of the 11 groups identified in land plants all involved in different metabolic tasks.

Other families of genes involved in developmental mechanism include the flowering truss gene family which was represented by 16 members responsible for flowering regulation and shoot architecture and the cuticle biosynthesis genes, which play a key role in preserving plants from various abiotic and biotic stresses regulating water and gas exchanges.

Phospatase families included serine/threonine classified into two groups: PPP (Ser/Thr-specific phosphoprotein phosphatase) and PPM/PP2C (magnesium dependent protein phosphatise).

Members of other gene families involved in growth, defense and physiological activities and including RNA-binding proteins, auxin Response Factors (ARFs), receptor-like kinases (867 genes), nucleotide binding site (684) and glycoside hydrolase gene families have been also identified within the pepper genomes.

8.2 Generalized Workflow for Genome Assembly, Structural and Functional Annotation

In Fig. 8.1, it is reported a generalized flowchart of the pepper genome assembly pipeline. The genome annotation pipeline included two phases: 'structural annotation' and 'functional annotation'. The former refers to the identification of DNA elements (e.g., repetitive elements, protein-coding genes, etc.) embedded in the genome, while the latter allows attaching biological information to these elements. Even if genome annotation pipelines differ in details, they share a core set of features and best practices (Fig. 8.1).

Prior to gene prediction, a thorough annotation of repetitive sequences in newly sequenced genomes is of utmost importance (Maumus and Quesneville 2016). To this end, a combination of de novo and similarity-based approaches was used for the identification and classification of repetitive DNA sequence motifs in the genome (Maumus and Quesneville 2016) (Fig. 8.2a). Protein-coding genes were predicted using ab initio gene finder tools in combination with comparative methods. The former is based on the identification of regions with coding potential and on the detection of signals within the DNA known as typical of gene structures; the latter relies on the use of homologous sequences (ESTs, mRNAs, RNA-Seq tags, proteins) to deduce gene structure. Among all forms of evidence, RNA-Seq tags have the greatest potential to improve the accuracy of gene annotations (Yandell and Ence 2012).

In case of ab initio gene prediction, an array of different gene finders was independently run to predict coding genes (Fig. 8.2b). Since ab initio gene finders need to be trained on a set of known genes, the first step was the construction of a training dataset (D'Agostino et al. 2007). To accomplish this task, available full-length cDNA and assembled RNA sequences were splice-aligned versus genome sequences.

As for comparative methods, protein-to-genome alignments as well as EST/RNA-seq tag-to-genome alignments laid the foundations to identify evidence-based gene *loci* and define gene structure. Previously identified transcripts and full-length proteins from pepper as well as available sequences from model or phylogenetically related species were used.

In the final step, ab initio gene predictions and diverse similarity-based evidence types were combined into consensus gene structures. This was performed using a 'combiner' algorithm in conjunction with manual curation of miss-annotated genes (Lewis et al. 2002) (Fig. 8.2b). The final set of gene annotations were further filtered to remove weakly supported genes.

Biological description (i.e., gene functions) was assigned to protein-coding genes based on BLAST similarity searches against UniProt (The

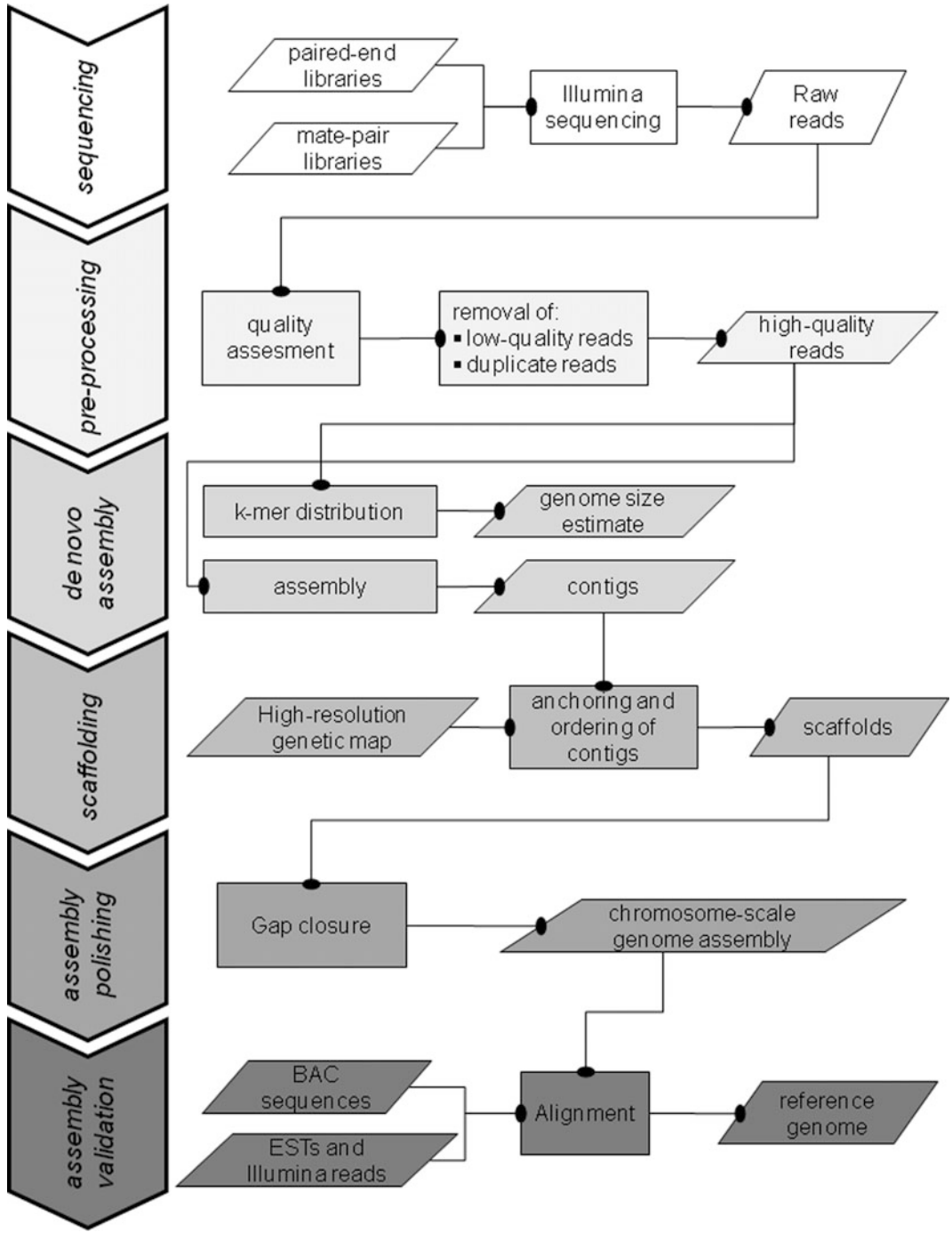


Fig. 8.1 Generalized flowchart of the pepper genome assembly pipeline. It can be divided into a core set of steps from the sequencing of Illumina paired-end and mate-pair libraries to the validation of the chromosome-scale genome assembly

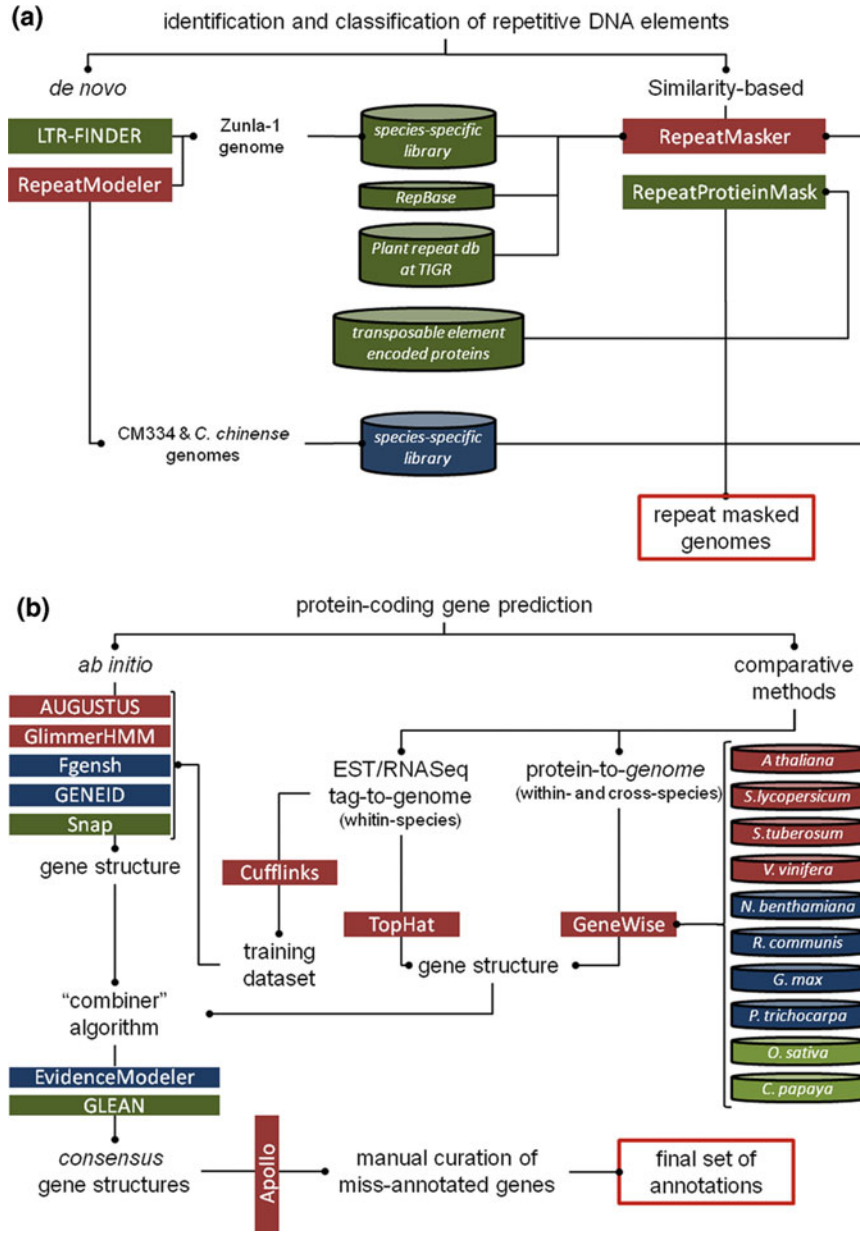


Fig. 8.2 Generalized flowchart of the structural genome annotation. Modules and corresponding software/tools are shown. In red are tools and databases common to both pipelines; in green those Zunla-1/Chiltepin specific; in blue those CM334 specific. **a** Identification and classification of repetitive DNA sequence motifs. **b** Prediction of protein-coding genes

UniProt Consortium 2015) and TAIR (Leonore et al. 2017) databases. In addition, motifs and domains within predicted protein sequences were identified using InterProScan (Jones et al. 2014). Sequence function description was integrated with gene ontologies (GO) (The Gene Ontology Consortium 2017) and the Enzyme Commission identifiers (EC; https://web.archive.org/web/20060218084611/; http://www.expasy.org/enzyme/) for a standard classification of gene products.

Where possible, the KEGG Orthology (KO) system was used to map proteins onto KEGG metabolic and signaling pathways (Kanehisa et al. 2017).

8.3 The Resequencing of *Capsicum* Genotypes

Thanks to the availability of pepper whole-genome sequences, NGS (next-generation sequencing) has been applied for the resequencing of additional pepper accessions in order to gather information on the structure, function, and evolution of genomes as well as to spot in detail allelic/structural variants.

To date, the resequencing of target genomic regions as well as of whole genomes of pepper accessions has been carried out with the goal to: (i) study how artificial selection traces embedded in the pepper genome correlates with breeding history, (ii) identify and fine mapping genomic *loci* conferring resistance against biotic stressors, (iii) reconstruct and structurally/functionally annotate the genomes of local pepper landraces.

8.3.1 Resequencing and Identification of Genes Involved in the Process of Pepper Domestication

Plant domestication is a complex evolutionary process in which human use of plant species led to morphological and physiological changes that distinguish domesticated taxa from their wild ancestors (Purugganan and Fuller 2009). The use of domestication as a model for the evolutionary process stems from an understanding of events associated with the origins of crop species and knowledge of the selective pressures experienced by domesticated taxa. The paper published by Qin et al. (2014) reported also the resequencing, at a coverage ranging from 20 to 30-fold depth, of 18 cultivated accessions representative the major varieties of *C. annuum* and two semi-wild/wild peppers, with the goal to provide insights into the identification of genes involved in the process of pepper domestication.

The alignment to the reference led to the identification of more than 9 M SNPs and 200 K small InDels, and both neighbor-joining tree and population structure highlighted that the wild and domesticated peppers are genetically distinguishable. To identify genomic footprints of artificial selection, a genetic bottleneck approach was used (Li et al. 2013). Genetic diversity was estimated by calculating θ_π (average pairwise divergence within a population) and θ_w Watterson's estimator, Watterson (1975). The regions showing significantly lower θ_π and θ_w in cultivars relative to the semi-wild/wild accessions were considered as potentially subjected to artificial selection. Only 2.6% of the genome, e.g., 115 regions containing 511 genes, appeared to be strongly affected by artificial selection in the cultivated peppers. The 511 spotted genes were mainly related to transcription regulation, stress and/or defense response, protein–DNA complex assembly, growth and fruit development, and ripening-associated biological processes. Among them 34 transcription factors (TFs), including activating protein (AP2), ethylene-responsive element-binding factor (ERF), and basic helix-loop-helix (bHLH) families as well as 10 disease resistance protein containing the NB-ARC domain were spotted. Those results show how resequencing approaches may contribute to the identification of genes related to morphological and physiological differences between cultivated and wild peppers and underlying pepper domestication and genetic improvement.

8.3.2 Resequencing Approaches for Genetic Analyses of Biotic Stress Resistance

The resequencing of *Capsicum* genomic regions and whole genomes has made it possible to identify molecular markers tightly linked to genes affecting resistance to biotic stresses and exploitable for marker-assisted breeding (MAS).

Devran and co-authors (2015) employed NGS technology in combination with bulk segregant analysis (BSA) for the identification of new molecular markers tightly linked to the Pvr4 *locus*, located on chromosome 10 and conferring dominant resistance to three pathotypes of Potyvirus (PVY, *Potato Virus Y*) as well as to *Pepper mottle virus* (PepMoV; Caranta et al. 1996, 1999). The susceptible *C. annuum* cv. 'SR-231' was crossed with the resistant accession 'Criollo de Morelos 334' (CM334) and a F_2 segregating progeny was generated from a single F_1 plant. The DNA of 15 resistant and 15 susceptible F_2 plants was at first pooled in two bulks, which were Illumina sequenced together with the parents. Due to the high synteny of tomato and pepper chromosome 10 (Wu et al. 2009), the sequence of tomato chromosome 10 was used as a reference for the alignment of Illumina reads of pepper parental lines, while reads from the two bulks were used to confirm the recognized polymorphisms. Some of the identified SNPs were then converted into CAPS (Cleaved Amplified Polymorphism sequence) marker, and the Pvr4 *locus* was mapped between the CAPS markers MY262 and MY69.

Thanks to the subsequent availability of the *Capsicum* genome sequences, pepper and tomato genomic regions including the Pvr4 *locus* were compared. A high degree of synteny was detected although the pepper chromosome 10 resulted inverted compared to tomato, and a tomato DNA region of approximately 1 Mb aligned against the corresponding pepper region that is three times wider. More of 5000 polymorphic sites (InDels and SNPs) were further spotted and markers developed. This allowed the fine mapping of *Pvr4* between two flanking markers (MY1176 and MY5009), with only one estimated recombination event on either side (less than 1 cM genetic distance away from the *locus*). The identification of two tightly linked flanking markers represents a highly reliable tool for easily transferring the Pvr4 *locus* to pepper breeding lines via marker-assisted backcrossing (MAB) selection.

Kang et al. (2016) performed the resequencing of the pepper cultivars 'YCM344' and 'Taean'. The former is highly resistant against Bacterial wilt (BW) which is caused by the soil-borne bacterium *Ralstonia solanacearum,* a pathogen distributed from tropical to temperate areas and which affects a broad range of dicot and monocot hosts, being particularly harmful for solanaceous crops. The two cultivars were Illumina resequenced at a coverage of 10× and the reads showed mapping rates higher than 93% to the CM334 reference genome (Kim et al. 2014). Approximately 7 K SNPs were detected in both accessions with frequencies ranging from 1.95 SNPs/Kb in 'Taean' to 2.01 SNPs/Kb in 'YCM334'.

The resequenced genomes were compared to each other with the goal to identify the most informative alleles related to BW resistance. More than 5, 6 M polymorphic SNPs and 149 K InDels were identified. This dataset allowed to identify genetic markers able to distinguish both these cultivars from CM334 and the two cultivars from each other. More than 100 K of the polymorphic SNPs were within gene regions, while ~36 K were in coding sequences (CDS), of which 23,396 showed non-synonymous (non-Syn) protein changes in 9102 genes.

Among the ten most polymorphic genes between 'YCM344' and 'Taean', two encodes for a 'Putative disease resistance protein' (CA10g15480 and CA12g20430) and were assigned to the 'Late blight resistance protein R1' gene family (IPR021929). This result suggested that the detected polymorphisms could be responsible for the different response to disease of the two cultivars. Other highly polymorphic genes included polyproteins, LRR like receptor kinases, N-like proteins, CC-NBS-LRR proteins, and putative phosphatidylinositol 4-kinase. A comparative analysis of SNPs located in genomic regions showing high similarity with known resistance genes in the tomato genomes was also performed. Among them, a total of seven genes showed non-Syn changes between 'YCM334' and 'Taean', which may be related to functional differences between the cultivars and represent strong candidate *loci* that contribute to BW disease in the cultivar 'YCM334'.

Recently Ahn et al. (2018), with the goal to discover SNPs associated with Powdery Mildew

(PM) resistance, resequenced via Illumina (~11× coverage) the resistant *C. baccatum* line'PRH1' and the susceptible *C. annuum* line 'Saengryeg' (~10× coverage). The agent of PM is *Leveillula taurica*, which is spread in a wide range of environments and represents a devastating fungal disease in pepper. The level of resistance to PM was assessed in both the lines, as well as in 45 individuals of their RIL F_4 population, through co-cultivation with powdery mildew and by using a scale ranging from one (resistant) to five (susceptible).

About 6 M SNPs, whose majority was classified as homozygous, were detected in both lines and found differentially distributed among the chromosomes. About 4.8 M SNPs were polymorphic between the two lines and were used for the design of 306,871 high-resolution melting (HRM) marker primer sets. The highest number of heterozygous SNPs was detected on chromosome 1 of PRH1 (i.e.: 23,932) and on chromosome 12 (i.e.,: 15,942) in 'Saengryeg', while the lowest one on chromosome 8 in both pepper lines (11,915 in 'PRH1' and 7229 'Saengryeg'). Based on their position within the genome sequence, the SNPs were then classified into intergenic or genic, and these latter sub-classified as intron SNPs, which were more frequent, and coding SNPs.

With the goal to gain deeper insight into SNPs associated with genes involved in disease resistance and stress tolerance processes, a chromosome wide functional annotation of the polymorphic variants among the two lines was performed. In introns and coding regions up to 6281 SNPs, associated with 46 RGA (Resistance Genes Analogues) carrying nucleotide binding site-leucine-rich repeat (NBS-LRR) motifs, were identified, found predominantly distributed on chromosome 4. NBS-LRR represents a large family of proteins that are encoded by RGA and are involved in pathogen recognition, including powdery mildew (Meyers et al. 2003; Coleman et al. 2009). Since the highest number of NB-LRR-linked SNPs was present in the PM resistant line 'PRH1' compared to the susceptible line 'Saengryeg', the authors assumed that NB-LRR resistance genes might play a key role in PM resistance. A subset of the identified SNPs was validated through HRM assay and, among the 36 primers applied, 19 significantly distinguished both parental lines and the resistant and susceptible plants in the F_4 progeny.

8.3.3 Resequencing of Pepper Landraces

Farmers' selection and adaptation to local climate and low-input agricultural practices has resulted in a plethora of pepper landraces that differ in growth habit, fruit shape and size, and organoleptic properties and that frequently carry resistance genes that are effective against abiotic and biotic stress. In the Piedmont region (northwest Italy) valuable and morphologically distinguishable landraces are grown, which are the result of a long selection process for adaptation to specific ecological niches. Thanks to the recent availability of *Capsicum spp.* genome sequences (Kim et al. 2014; Qin et al. 2014), Barchi and colleagues (2017) performed the genome resequencing of inbred lines of the four main landraces grown in the Piedmont region, namely: 'Cuneo' and 'Quadrato' (blocky types), 'Corno' (long type) and 'Tumaticot' (with small, sub-spherical fruits). The sequencing of the four genotypes was performed through Illumina technology, at coverage of ~35X, and each genomic sequence was assembled into 12 chromosome-scale pseudo-molecules. Approximately 35 k genes were identified of which about 75% contained at least one IPR domain. The protein complements of the four reconstructed genomes, together with that of the reference (CM334), were analyzed to identify orthologs and orthogroups. More than 170 K sequences were clustered into 34,664 gene families (excluding singletons) of which 26,270 resulted to be shared among the five accessions, while only 152 gene families were in common between the two blocky types ('Quadrato' and 'Cuneo').

By aligning reads of the resequenced genotypes to the CM334 genome using standard pipelines, a set of about 19 M SNP/InDel was detected, ranging from 16.33 M ('Tumaticot') to

18.07 M ('Corno'). As expected for a selfing crop, the heterozygosity was rather low and ranged from ~0.2% in 'Corno' to ~0.1% in 'Tumaticot'.

A survey of the SNPs within genes that generally affect fruit size and shape in the *Solanaceae* (Chunthawodtiporn et al. 2018), were performed and mutations in the coding sequences of fw2.2, WUSCHEL (WUS) and fw3.2 were found to be common to the 4 genotypes while single deleterious mutation in sun-like ortholog gene was predicted. Differently, regulatory regions were rich in mutations with the exception of those related to the fw2.2 and fw3.2 *loci*. The large allelic diversity identified in the four resequenced accessions suggests that the pepper landraces under investigation can be considered as highly valuable pre-breeding resources.

References

Ahn YK, Manivannanbinaya A, Sandeep K, Jun TH, Yang EY, Choi S, Kim JH, Kim DS, Lee E-S (2018) Whole genome resequencing of *Capsicum baccatum* and *Capsicum annuum* to discover single nucleotide polymorphism related to powdery mildew resistance. Sci Rep 8:5188. https://doi.org/10.1038/s41598-018-23279-5

Barchi L, Acquadro A, Portis E, Comino C, Nourdine M, Borras D, Bustos Lopez M, Giordano R, Monge S, Carli C, Lanteri S (2017) Genome re-sequencing of piedmontese pepper ecotypes. In: The XIV *Solanaceae* and III cucurbitaceae genomics joint conference, 3–6 Sept, Valencia, Spain

Bombarely A, Moser M, Amrad A, Bao M, Bapaume L et al (2016) Insight into the evolution of the *Solanaceae* from the parental genomes of *Petunia hybrida*. Nat Plants 2(6):16074. https://doi.org/10.1038/nplants.2016.74

Caranta C, Palloix A, Gebre-Selassie G, Lefebvre V, Moury B, Daubeze AM (1996) A complementation of two genes originating from susceptible *Capsicum annuum* lines confers a new and complete resistance to pepper veinal mottle virus. Phytopathology 86:739–743. https://doi.org/10.1094/Phyto-86-739

Caranta C, Thabuis A, Palloix A (1999) Development of a CAPS marker for the Pvr4 locus: a tool for pyramiding potyvirus resistance genes in pepper. Genome 42:1111–1116. https://doi.org/10.1139/gen-42-6-1111

Chunthawodtiporn J, Hill T, Stoffel K, Van Deynze A (2018) Quantitative trait *Loci* controlling fruit size and other horticultural traits in bell pepper (*Capsicum annuum*). Plant Genome 11(1)

Coleman C, Copetti D, Cipriani G, Hoffmann S, Kozma P, Kovács L, Morgante M, Testolin R, Di Gaspero G (2009) The powdery mildew resistance gene REN1 co-segregates with an NBS-LRR gene cluster in two Central Asian grapevines. BMC Genet 10:89

D'Agostino N, Traini A, Frusciante L, Chiusano ML (2007) Gene models from ESTs (GeneModelEST): an application on the *Solanum lycopersicum* genome. BMC Bioinf 8(Suppl 1):S9–S9. https://doi.org/10.1186/1471-2105-8-S1-S9

Devran Z, Kahveci E, Özkaynak E, Studholme DJ, Tör M (2015) Development of molecular markers tightly linked to *Pvr*4 gene in pepper using next-generation sequencing. Mol Breed 35(4):101

Harris RS (2007) Improved pairwise alignment of genomic DNA. Ph.D. thesis, The Pennsylvania State University, University Park, USA

Hulse-Kemp AM, Maheshwari S, Stoffel K, Hill TA, Jaffe D, Williams S et al (2018) Reference quality assembly of the 3.5-Gb genome of *Capsicum annuum* from a single linked-read library. Hort Res. https://doi.org/10.1038/s41438-017-0011-0

Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C et al (2014) InterProScan 5: genome-scale protein function classification. Bioinformatics 30(9):1236–1240. https://doi.org/10.1093/bioinformatics/btu031

Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K (2017) KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucl Acids Res 45(Database issue), D353–D361. https://doi.org/10.1093/nar/gkw1092

Kang YJ, Ahn YK, Kim KT, Jun TH (2016) Resequencing of *Capsicum annuum* parental lines (YCM334 and Taean) for the genetic analysis of bacterial wilt resistance. BMC Plant Biol 16:235

Kim S, Park M, Yeom SI, Kim YM, Lee JM, Lee HA et al (2014) Genome sequence of the hot pepper provides insights into the evolution of pungency in *Capsicum* species. Nat Genet 46:270–278

Kim S, Park J, Yeom SI, Kim YM, Seo E, Kim KT et al (2017) New reference genome sequences of hot pepper reveal the massive evolution of plant disease-resistance genes by retroduplication. Genome Biol 18:210. https://doi.org/10.1186/s13059-017-1341-9

Leonore R, Shabari S, Donghui L, Eva H (2017) Using the arabidopsis information resource (TAIR) to find information about arabidopsis genes. Curr Prot Bioinformat 60(1):11111–111145. https://doi.org/10.1002/cpbi36

Lewis SE, Searle SMJ, Harris N, Gibson M, Iyer V, Richter J et al (2002) Apollo: a sequence annotation editor. Genome Biol 3(12). https://doi.org/10.1186/gb-2002-3-12-research0082

Li YH, Zhao SC, Ma JX, Li D, Yan L, Li J et al (2013) Molecular footprints of domestication and improvement in soybean revealed by whole genome re-sequencing. BMC Genom 14:579. https://doi.org/10.1186/1471-2164-14-579

Maumus F, Quesneville H (2016) Impact and insights from ancient repetitive elements in plant genomes. Curr Opin Plant Biol 30:41–46. https://doi.org/10.1016/jpbi201601003

Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW (2003) Genome-wide analysis of NBS-LRR–encoding genes in Arabidopsis. Plant Cell 15:809–834

Purugganan MD, Fuller DQ (2009) The nature of selection during plant domestication. Nature 457 (7231):843–848

Qin C, Yu C, Shen Y, Fang X, Chen L, Min J et al (2014) Whole-genome sequencing of cultivated and wild peppers provides insights into *Capsicum* domestication and specialization. Proc Natl Acad Sci USA 111:5135–5140

The Gene Ontology Consortium (2017) Expansion of the gene ontology knowledgebase and resources. Nucl Acids Res 45(Database issue):D331–D338. https://doi.org/10.1093/nar/gkw1108

The UniProt Consortium (2015) UniProt: a hub for protein information. Nucl Acids Res 43(Database issue):D204–D212. https://doi.org/10.1093/nar/gku989

Tomato Genome Consortium (2012) The tomato genome sequence provides insights into fleshy fruit evolution. Nature 30,485(7400):635–641. https://doi.org/10.1038/nature11119

Watterson GA (1975) On the number of segregating sites in genetical models without recombination. Theor Popul Biol 7:256–276. https://doi.org/10.1016/0040-5809(75)90020-9

Weisenfeld NI, Kumar V, Shah P, Church DM, Jaffe DB (2017) Direct determination of diploid genome sequences. Genome Res 27:757–767

Wu F, Eanetta NT, Xu Y, Durrett R, Mazourek M, Jahn MM, Tanksley SD (2009) A COSII genetic map of the pepper genome provides a detailed picture synteny with tomato and new insights into recent chromosome evolution in the genus *Capsicum*. Theor Appl Genet 118:1279–1293. https://doi.org/10.1007/s00122-009-0980-y

Yandell M, Ence D (2012) A beginners guide to eukaryotic genome annotation. Nat Rev Gen 13:329. https://doi.org/10.1038/nrg3174

Sequencing of *Capsicum* Organellar Genomes

9

Nitin Kumar, Khushbu Islam and Nirala Ramchiary

Abstract

In early 1980s, DNA sequencing was heralded as a game changer in the way we look at biology. However, the cost of operation and scale of handling large contiguous reads were the biggest bottleneck. Therefore, research and experimentation were coerced to relatively simpler genomes of prokaryotes and organelle genomes of eukaryotes such as mitochondria and a diverse range of plastid genomes. Mitochondria and chloroplast are two important cell organelles which have their own genetic material called as mitogenome and plastome, respectively. They are believed to be an endosymbiotic relic of a free-living organism in the past, but now tightly integrated into their host cells. One exclusive property about organellar genome is maternal inheritance, which ensures that genome remains highly compact with almost null genome expansion by crossing over and external evolutionary forces. They show very low genome diversity yet high adaptability to host environment, which has intrigued researchers to understand plant kingdom through organelles including Solanaceae plants. *Capsicum* is one of the most important vegetable crops of Solanaceae family. In this chapter, an effort was made to enumerate the sequencing and their derived information of chloroplast and mitochondria genomes of different *Capsicum* species. Till date, a total of 13 accessions of whole chloroplast genome sequences (the size of which ranges from 156,807 to 157,145 bp) have been reported from *Capsicum annuum*, *C. chinense*, *C. frutescens*, *C. tovarii*, *C. chacoense*, *C. baccatum*, *C. galapagoense, C. eximium* and *C. lycianthoids*, and only one mitochondrial genome sequence of *C. annuum* has been reported.

9.1 The Capsicum Chloroplast Genome

A plant cell is distinguished from an animal cell mainly due to the presence of chloroplasts (typically 20–40 per cell) which was identified as the site of photosynthesis by T. Engelmann 1881 (Drews 2005). Chloroplasts are double enveloped structures, the internal membrane is called *grana* that is formed by stacking of thylakoids, and outer one is called stroma which harbors the chloroplast (cp) DNA. The average size of

N. Kumar
Department of Bioengineering and Technology, Institute of Science and Technology, Gauhati University, Gopinath Bordoloi Nagar, Guwahati 781014, Assam, India

N. Kumar · K. Islam · N. Ramchiary (✉)
Translational and Evolutionary Genomics Lab, School of Life Sciences, Jawaharlal Nehru University, New Delhi 110067, India
e-mail: nramchiary@mail.jnu.ac.in

N. Ramchiary
Department of Biotechnology, Delhi Technological University, Delhi 110042, India

N. Ramchiary and C. Kole (eds.), *The Capsicum Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-319-97217-6_9

chloroplast genome in higher plants ranges from 150 to 220 kb. Cp DNA has a simple circular arrangement and lacks histone association (Campbell et al. 2009). Apart from photosynthesis, chloroplast carries out other biological functions also, such as synthesis of fatty acids, amino acids, nucleotides, vitamins and phytohormones, and nitrogen and sulfur assimilation. Cp genome is a compact circular quadripartite, typically coding for 100–200 genes vital for photosynthetic pathways and chloroplast's internal functions. Chloroplasts are result of endosymbiosis of cyanobacteria (Gibbs 1981). In *Capsicum* species, the size of chloroplast genome ranges from 156,807 bp (Park et al. 2016; GenBank: KU041709) in *Capsicum Chinense* to 157,145 bp (GenBank: KR078314) for *Capsicum baccatum* var. *Baccatum*. The cp genome has a large single copy (LSC) region and a small single copy (SSC) region separated by a pair of inverted repeats (IRs, IRA, and IRB). IRs are believed to be good examples of concerted evolution and serve a very important role in stabilizing the overall chloroplast structure (Palmer and Thompson 1982; Palmer 1985). Many earlier comparative studies spanning over cp genome of unrelated plants such as tobacco and liverwort show significant homology between the two (Ohyama et al. 1986; Shinozaki et al. 1986), and most of the variations were confined to the IR region. *tRNA* clusters are the highly divergent hotspot in LSC region of the cp genome. Another remarkable aspect of cp genome is its ability to undergo successive loss of functionality to acquire specialized roles such as chromoplast. *Capsicum* chloroplast is reported to undergo a drastic change by the total loss of ribosomes after transforming into chromoplast, hence adopting various specialized roles (Carde et al. 1988).

9.1.1 Overview of Capsicum Chloroplast Genome Sequencing Projects

The Solanaceae family has long been a model system for comparative and evolutionary genomics. The first chloroplast genome to be sequenced was that of *Nicotiana tabacum* (Tobacco; Shinozaki et al. 1986). As of May 2018, we have 2583 listed accessions for organellar genomes from land plants which include a total of 1971 chloroplast and 291 mitochondrial genomes at National Center for Biotechnology Information database (NCBI; https://www.ncbi.nlm.nih.gov/genome/browse#!/organelles/). The Cp genome of *Capsicum* was not sequenced until 2010 even when sequencing of chloroplasts genomes of seven other members of the Solanaceae family including tobacco (155,844 bp), tomato (155,461 bp) and potato (155,312 bp) were already available (Shinozaki et al. 1986; Chung et al. 2006; Kahlau et al. 2006). As of now, a total of 13 unique accessions of whole chloroplast genome sequences belonging to genus *Capsicum* are available for the exploration in NCBI. These are from *Capsicum annuum* (GenBank: JX270811/NC_018552), *C. annuum* var. *annuum* (GenBank: KR078313), *C. annuum* var. *glabriusculum* (GenBank: KJ619462; KR078311), *C. baccatum* var. *baccatum* (GenBank: KR078314), *C. chacoense* (GenBank: KX913218/NC_033525), *C. chinense* (GenBank: KU041709/NC_030543; KX913217), *C. eximium* (GenBank: KX913220/NC_033527), *C. frutescens* (GenBank: KR078312/NC_028007), *C. galapagoense* (GenBank: KX913216/NC_033524), *C. lycianthoids* (GenBank: KP274856/NC_026551), and *C. tovarii* (GenBank: KX913219). Full genome alignment clearly demonstrates regions of homology and divergence among various accessions. Quadripartite demarcations are also clearly visible in Fig. 9.1. The total GC content for genus *Capsicum* is found to be approximately 37.7%, having most contribution from IR region, owing to GC rich rRNA genes.

The *C. annuum* var. *glabriusculum* (Raveendar et al. 2015b; GenBank: KR078311) cp genome is 156,817 bp in size. The inverted repeats (IR) constitute 43.05% of the total GC content where LSC region comprises 35.74% and SSC region with 32.01%. It has 87,380 bp long LSC and 17,853 bp long SSC regions, separated by a pair of IRs (25,792 bp). A total of 132 genes are predicted, of which 87 are protein-coding, 8 are rRNA, and 37 are tRNA genes. Eight

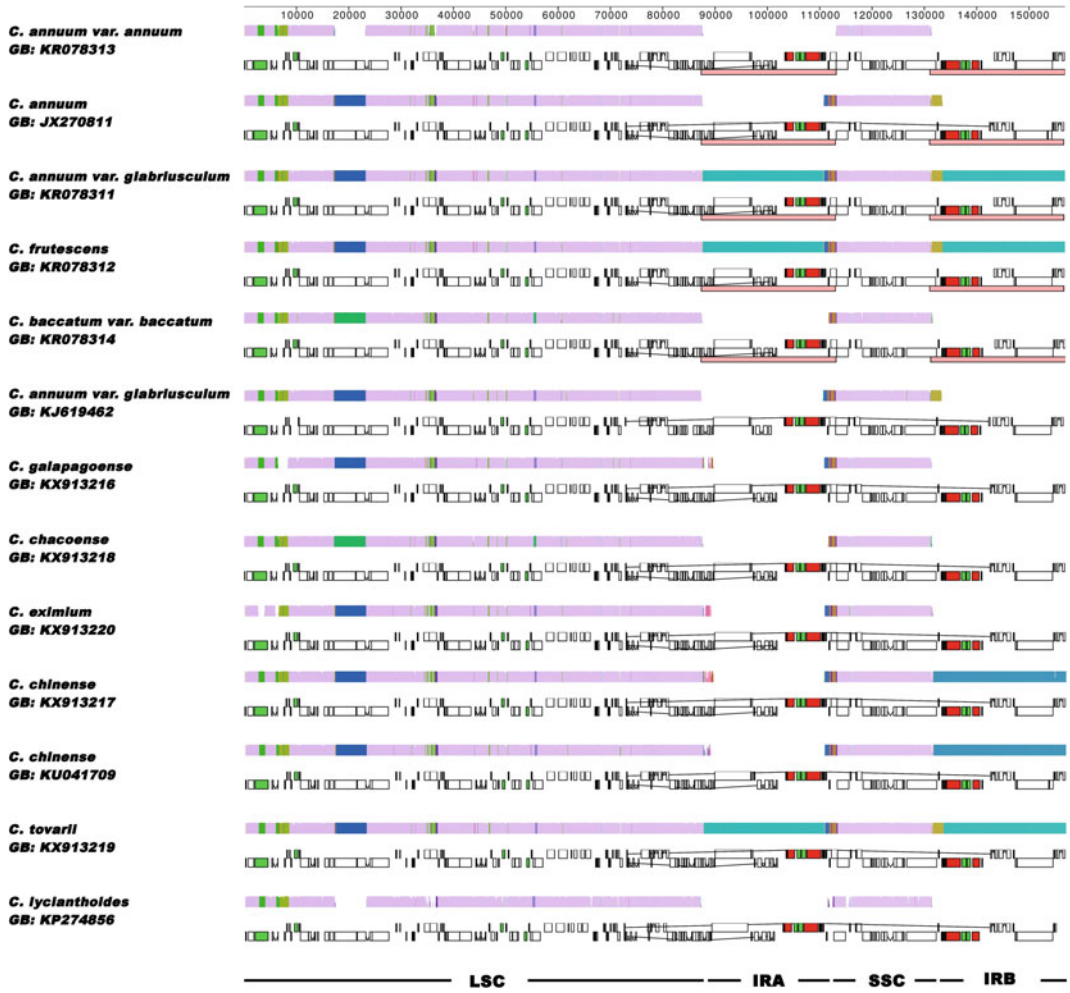


Fig. 9.1 Collinear blocks drawn by Mauve Aligner (Darling et al 2004) for 13 *Capsicum* Cp genomes. Two major tracks are (1) backbone similarity and (2) annotation features. Two layers of boxes represent major features, upper layer indicates forward strand transcribed features, while reverse strand features are depicted as inverted boxes. CDS regions are depicted in plain white color, rRNAs as red, tRNA as green, while kinked connecting lines highlight introns and genes. Overall quadripartite regions and major conserved syntenic blocks are also clearly recognizable

protein-coding genes (*ndhB, rpl2, rpl23, rps7, rps12, ycf1, ycf2, ycf15*), four rRNA genes, and seven *tRNA* genes have duplicate copies in the IR region (Table 9.1). Nine protein-coding genes [*rpoC1, petB, petD, Rps16, atpF, rpl16, rpl2*(IR), *ndhA, ndhB*(IR)] and six tRNA genes have single intron, whereas three other protein-coding genes (*clpP, rps12 and ycf3*) contain double introns. The size of cp genome of the Korean landrace 'Subicho' *C. annuum* var. *annuum* is 156,878 in length (Raveendar et al. 2015a). Two IR regions of 25,801 bp separate LSC (87,347 bp) and SSC (17,929 bp) regions. The genome harbors 132 unique genes whose content and order are identical to *C. annuum* var. *glabriusculum* (Genbank: KR078311). The sequencing of *C. chinense* cp genome with 156,936 bp length was reported by Park et al. (2016; Genbank: KU041709) and Raveendar et al. (2017; Genbank: KX913217). This genome has 87,330 bp LSC region and 17,912 bp SSC region which are interrupted by IRA and IRB regions of 25,847 bp length. Out of

Table 9.1 Details of genes reported in different *Capsicum* plastome sequencing studies

	C. frutescens (Shim et al. 2016)	*C. annuum* var. *glabriusculum* (Raveendar et al. 2015b)	*C. baccatum* (Kim et al. 2016)	*C. chinense* (Raveendar et al. 2017)	*C. annuum* var. *annuum* (Raveendar et al. 2015a)
Photosystem I	*psaA, B, C, I, J, ycf3*[*2], ycf4*	*psaA, B, C, I, J, ycf3*[*2], ycf4*	*psaA, B, C, I, J, ycf3*[*2], ycf4*	*psaA, B, C, I, J, ycf3*[*2], ycf4*	*psaA, B, C, I, J, ycf3*[*2], ycf4*
Photosystem II	*psbA, B, C, D, E, F, H, I, J, K, L, M, N, T, Z*	*psbA, B, C, D, E, F, H, I, J, K, L, M, N, T, Z*	*psbA, B, C, D, E, F, H, I, J, K, L, M, N, T, Z*	*STpsbA, B, C, D, E, F, H, I, J, K, L, M, PpsbN, T, Z*	*psbA, B, C, D, E, F, H, I, J, K, L, M, N, T, Z*
Cytochrome b_6f	*petA, B*[*1], D*[*1], G, L, N*	*petA, B*[*1], D*[*1], G, L, N*	*petA, B*[*1], D*[*1], G, L, N*	*petA, B*[*1], D*[*1], G, L, N*	*petA, B*[*1], D*[*1], G, L, N*
ATP synthase	*atpA, B, E, F*[*1], H, I*	*atpA, B, E, F*[*1], H, I*	*atpA, B, E, F*[*1], H, I*	*atpA, B, E, F*[*1], H, I*	*atpA, B, E, F*[*1], H, I*
RuBisCO	*rbcL*	*rbcL*	*rbcL*	*rbcL*	*rbcL*
NADH oxidoreductase	*ndhA*[*1], B*[*1*3], C, D, E, F, G, H, I, J, K*	*ndhA*[*1], B*[*1*3], C, D, E, F, G, H, I, J, K*	*ndhA*[*1], B*[*1*3], C, D, E, F, G, H, I, J, K*	*ndhA*[*1], B*[*1*3], C, D, E, F, G, H, I, J, K*	*ndhA*[*1], B*[*1*3], C, D, E, F, G, H, I, J, K*
Large subunit ribosomal proteins	*Rpl2*[*1*3], 14, 16*[*1], 20, 22, 23*[*3], 32, 33, 36*	*Rpl2*[*1*3], 14, 16*[*1], 20, 22, 23*[*3], 32, 33, 36*	*Rpl2*[*1*3], 14, 16*[*1], 20, 22, 23*[*3], 32, 33, 36*	*Rpl2*[*1*3], 14, 16*[*1], 20, 22, 23*[*3], 32, 33, 36*	*Rpl2*[*1*3], 14, 16*[*1], 20, 22, 23*[*3], 32, 33, 36*
Small subunit ribosomal proteins	*Rps2, 3, 4, 7*[*3], 8, 11, 12*[*2*3*4], 14, 15, 16*[*1], 18, 19*	*Rps2, 3, 4, 7*[*3], 8, 11, 12*[*2*3*4], 14, 15, 16*[*1], 18, 19*	*rps2, 3, 4, 7c, 8, 11, 12*[*2*3*4], 14, 15, 16*[*1], 18, 19*	*Rps2, 3, 4, 7*[*3], 8, 11, 12*[*2*3*4], 14, 15*[*3], 16*[*1], 18, 19*[*3]	*Rps2, 3, 4, 7*[*3], 8, 11, 12*[*2*3*4], 14, 15, 16*[*1], 18, 19*
RNA polymerase	*rpoA, B, C1*[*1], C2*	*rpoA, B, C1*[*1], C2*	*rpoA, B, C1*[*1], C2*	*rpoA, B, C1*[*1], C2*	*rpoA, B, C1, C2*
Protein-coding genes function unknown	*Ycf1*[*3], ycf2*[*3], ycf15*[*3]*	*ycf1*[*3], 2*[*3], 15*[*3]*	*Ycf1*[*3], 2*3, 15*[*3]*	*Ycf1*[*3], ycf2*[*3], ycf15*[*3]*	*Ycf1*[*3], 2*[*3], 15*[*3]*
Other genes	*accD, ccsA, cemA, clpP*[*2], matK*	*accD, ccsA, cemA, clpP*[*2], matK*	*accD, ccsA, cemA, clpP*[*2], matK*	*ccsA, cemA, clpP1, clpP2, matK*	*accD, ccsA, cemA, clpP*[*2], matK*
Ribosomal RNAs	*Rrn16*[*3], 23*[*3], 4.5*[*3], 5*[*3]*	*Rrn16*[*3], 23*[*3], 4.5*[*3], 5*[*3]*	*Rrn16*[*3], 23*[*3], 4.5*[*3], 5*[*3]*	*Rrn16*[*3], 23*[*3], 4.5*[*3], 5*[*3]*	*rrn16*[*3], 23*[*3], 4.5*[*3], 5*[*3]*
Transfer RNAs	*trnA-UGC*[*1*3], C-GCA, D-GUC, E-UUC, F-GAA, G-UCC*[*1], G-GCC, H-GUG, I-CAU*[*3], I-GAU*[*1*3], K-UUU*[*1], L-UAA*[*1], L-UAG, L-CAA*[*3], fM-CAU, M-CAU, N-GUU*[*3], P-UGG, Q-UUG, R-ACG*[*3], R-UCU, S-GCU, S-GGA, S-UGA, T-*	*trnA-UGC*[*1*3], C-GCA, D-GUC, E-UUC, F-GAA, G-UCC*[*1], G-GCC, H-GUG, I-CAU*[*3], I-GAU*[*1*3], K-UUU*[*1], L-UAA*[*1], trnL-UAG, L-CAA*[*3], fM-CAU, M-CAU, N-GUU*[*3], P-UGG, Q-UUG, R-ACG*[*3], R-UCU, S-GCU, S-GGA, S-UGA, T-*	*trnA-UGC*[*1*3], C-GCA, D-GUC, E-UUC, F-GAA, G-UCC*[*1], G-GCC, H-GUG, I-CAU*[*3], I-GAU*[*1*3], K-UUU*[*1], L-UAA*[*1], L-UAG, L-CAA*[*3], fM-CAU, M-CAU, N-GUU*[*3], P-UGG, Q-UUG, R-ACG*[*3], R-UCU, S-GCU, S-GGA, S-UGA, T-*	*trnA-UGC*[*1*3], C-GCA, D-GUC, E-UUC, F-GAA, G-UCC*[*1], G-GCC, H-GUG, I-CAU*[*3], I-GAU*[*1*3], K-UUU*[*1], L-UAA*[*1], L-UAG, L-CAA*[*3], M-CAU, fM-CAU, N-GUU*[*3], P-UGG, Q-UUG, R-ACG*[*3], R-UCU, S-GCU, S-GGA, S-UGA, T-*	*trnA-UGC*[*1*3], C-GCA, D-GUC, E-UUC, F-GAA, G-UCC*[*1], G-GCC, H-GUG, I-CAU*[*3], I-GAU*[*1*3], K-UUU*[*1], L-UAA*[*1], L-UAG, L-CAA*[*3], fM-CAU, M-CAU, N-GUU, P-UGG, Q-UUG, R-ACG*[*3], R-UCU, S-GCU, S-GGA, S-UGA, T-GGU,*

(continued)

Table 9.1 (continued)

	C. frutescens (Shim et al. 2016)	*C. annuum* var. *glabriusculum* (Raveendar et al. 2015b)	*C. baccatum* (Kim et al. 2016)	*C. chinense* (Raveendar et al. 2017)	*C. annuum* var. *annuum* (Raveendar et al. 2015a)
	GGU, T-UGU, V-UAC[*1], *V-GAC*[*3], *W-CCA, Y-GUA*	*GGU, T-UGU, V-UAC*[*1], *V-GAC*[*3], *W-CCA, Y-GUA*	*GGU, T-UGU, V-UAC*[*1], *V-GAC*[*3], *W-CCA, Y-GUA*	*GGU, T-UGU, V-UAC*[*1], *V-GAC*[*3], *W-CCA, Y-GUA*	*T-UGU, V-UAC*[*1], *V-GAC*[*3], *W-CCA, Y-GUA*

[*1]Genes containing a single intron, [*2]genes containing two introns, [*3]two gene copies in IRs, [*4]Trans-splicing genes
A certain nomenclature style has been followed to compile huge data sets in compact view. Alphabets such as A, B, and C after full gene names depict its various domains. Table has been adapted from various sources, properly cited in discussion

113 unique genes of *C. chinense* cp genome reported, 79 code for proteins, 30 for *tRNA* and four for *rRNA* genes, among them, 21 are found to be duplicated in the IR region. Single introns are found in 15 protein-coding and six *tRNA* genes, while two genes (*rps12* and *ycf3*) have double introns (Raveendar et al. 2017). The size of *C. frutescens* cp genome is 156,817 bp (Shim et al. 2016). It is 36 bp longer than the reported *C. annuum* (GenBank: NC_018552) cp genome and 205 bp longer than that of *C. annuum* var. *glabriusculum* (GenBank: KJ619462). The LSC is 14 bp shorter and 167 bp longer than the above-mentioned chloroplast genomes, respectively. The SSC and IR show only slight differences in size. The number and content of genes are identical to reported *C. annuum* cp genome. The *C. baccatum* cp genome comprises 25,910 bp IRs, separated by 87,351 bp LSC and 17,974 bp SSC regions (Kim et al. 2016). The number of predicted genes is similar to that of *C. annuum* var. *glabriusculum*. It contains nine single intron genes [*atpF, petB, petD, Rps16, rpoC1, rpl16, rpl2*(IR), *ndhA* and *ndhB*(IR)], while three genes, namely *rps12, ycf3, and clpP,* have two introns, and *Rps12* is a trans-splicing gene. The sequencing of chloroplast genome of *C. tovarii* was reported by Shin et al. (2017). The total of 156,816 bp cp genomes consists of LSC of 87,379 bp, SSC of 17,853 bp, and a pair of IRs of 25,792 bp length. The reported number of unique genes was lesser than the chloroplast genomes of other *Capsicum* species, thereby suggesting gene loss, which might be due to the transfer of chloroplast genes to nucleus or mitochondria during the course of evolution (Shin et al. 2017).

High degree of conservation exists in chloroplast genomes of different plant species including *Capsicum* as they regulate many vital processes in plant cell. Nearly 46 genes in *Capsicum* species code for important genes of photosynthesis such as *NADH oxidoreductase*, *photosystem I and II*, *cytochrome b6f complex*, *ATP synthase* and *ribulose 1,5-bisphosphate carboxylase/oxygenase,* etc. (RuBisCo; Raveendar et al. 2015a). RuBisCo is the most abundant protein on the Earth's surface (Ellis 1979). The functional RuBisCo enzyme has eight large and eight small subunits. In algae, both the large and small subunits are encoded by chloroplast genome (Samiee and Kohnehrouz 2015). However, alternative mechanisms exist in higher plants, where only large subunits are encoded by chloroplast genome and synthesized on its own ribosomes, while small subunits are contributed by nucleus. Apart from major photosynthetic functions, 21 genes encode ribosomal subunits (12 *rps* genes for small subunit and nine *rpl* genes for large subunit), while four genes (*rpoA, rpoB, rpoC1,* and *rpoC2*) code for DNA-directed RNA polymerases.

9.1.2 Application of Capsicum Chloroplast Genome Sequence

9.1.2.1 Molecular Marker Development

Most direct application of cp genome is in marker development to study DNA polymorphism and phylogeny. Single nucleotide polymorphism (SNP) and simple sequence repeat (CpSSR) markers have been predicted in chloroplast genomes of *C. baccatum* var. *baccatum*, *C. annuum* var. *annuum*, *C. annuum* var. *glabriusculum*, *C. chinense* Jacq., and *C. frutescens*. For SSR motif prediction, Sputnik has been a widely used software in most of the published genome sequencing projects (Cardle et al. 2000). While MicroSAtellite identification tool (MISA; Thiel et al. 2003), which is a collection of perl scripts (http://pgrc.ipk-gatersleben.de/misa/), have also been used in *C. chinense* jacq (Raveendar et al. 2017) and *C. frutescens* (Shim et al. 2016). The *C. annuum* var. *glabriusculum* (Raveendar et al. 2015b) plastome renders about 125 potential SSR motifs with a frequency of 1 per 1250 bp, located mostly in the noncoding regions. The tetra- and trinucleotide repeats constituted 50 and 26% of the total SSRs, respectively. *C. annuum* var. *annuum* (Subicho) could reveal 144 potential SSR motifs (Table 9.2). The tetranucleotide (50%) and pentanucleotide (21.5%) repeats constituted the majority of the SSRs predicted. Cp genome of *C. frutescens* has exactly same number of SSRs as *C. annuum* var. *glabriusculum*. A total of 117 potential SSR motifs were reported in *C. chinense* Cp genome. Most of them are trinucleotide repeats (58.11%) and dinucleotides repeats (36.75%) unlike *C. annuum* and *C. frutescens*. *C. baccatum* cp genome was reported to have 34 SSR motifs with nearly 86.5% of them located in the intergenic region. The observed frequency of SSRs is approximately 1 per 1170 bp, which is greater than most of the *Capsicum* species and has 43.3% tetranucleotides and 27.5% trinucleotide SSR repeats. In *C. tovarii*, a total of 144 SSR markers were reported.

SSR markers play direct role in species characterization and diversity analysis. Forty-three accessions of ten *Capsicum* species were investigated for examining microsatellite polymorphism in cpDNA, and a total of 33 allelic variants were identified with just six cpDNA microsatellite loci. Vast majority of accessions from this study had a unique haplotype for each species. While 27 *C. annuum* accessions were either monomorphic or dimorphic, no clear polymorphism was reported for most of the accessions, which indicates a very low plastome variation. Reported low cp genome variation in *C. annuum* accessions was consistent with nuclear genome polymorphism data which reveals a high degree of conservation within the species (Ryzhova and Kochieva 2004).

To identify single nucleotide polymorphisms (SNPs) and insertions/deletions (indels) variants in *Capsicum* species widely used workflow employs Burrows–Wheeler aligner (BWA; Li and Durbin 2009) and SAMtools software (Li et al. 2009a). Variant detection is enabled by mapping various *Capsicum* plastomes against *C. annuum* cp genome (GenBank: JX270811). The comparative analysis of *C. annuum* var. *glabriusculum* (GenBank: KR078311) with *C. annuum* cp genome (GenBank: JX270811) could identify a total of 15 SNPs and 33 indels, collectively a total of 48 mutations, of which 43 were in LSC and five were in SSC regions, respectively. Other *Capsicum* cp genomes were also compared with *C. annuum* reference genome (GenBank: JX270811) for detection of SNPs and indels. For *C. annuum* var. *annuum*, 96 mutations (45 SNPs and 51 indels) were identified (Table 9.3). Among these, 78 mutations are located in the LSC, 17 in SSC, and one in IR region. *C. frutescens* has a total of 34 mutations (18 SNPs and 16 indels), where 29 mutations lie in the LSC region and 5 in the SSC region. For *C. chinense* Jacq., there were a total of 174 mutations (82 SNPs and 92 indels) in the detected variants. In case of *C. baccatum* var. *baccatum,* 282 mutations (218 SNPs and 64 indels) were identified and found to be distributed as 216 in LSC, 62 in SSC, and 4 in IRA regions, respectively. Other *Capsicum* species like *C. tovarii* have revealed a total of 96 variants.

Table 9.2 Number, types, distribution, and abundance of simple sequence repeats in chloroplast genome of different *Capsicum* species

	C. annuum var. *annuum* (Raveendar et al. 2015a)	*C. chinense* Jacq. (Raveendar et al. 2017)	*C. annuum* var. *glabriusculum* (Raveendar et al. 2015b)	*C. frutescens* (Shim et al. 2016)	*C. baccatum* (Kim et al. 2016)
Localization	Intergenic	Intergenic	Intergenic	Intergenic	Mostly intergenic
Total SSRs	144	117	125	125	134
High prevalence	50% Tetra	58.11% Tri	50% Tetra	50% Tetra	43.3% Tetra
Low prevalence	28.50% Di + Tri 21.50% Penta	5.12% Tetra	26% Tri 25% Di + Penta	26% Tri 25% Di + Penta	27.5% Tri
Tetra nucleotide motifs	ATAA/TAAA/AAAT, AAAT/AATA/ATAA TTTG/TTGT/TGTT, TCTT/CTTT/TTTC and AATT/ATTA/TTAA	–na–	AAAT/AATA/ATAA, ATAA/TAAA/AAAT, TTTG/TTGT/TGTT, TCTT/CTTT/TTTC AATT/ATTA/TTAA	AAAT/AATA/ATAA, ATAA/TAAA/AAAT, TTTG/TTGT/TGTT, TCTT/CTTT/TTTC AATT/ATTA/TTAA	AAAT/AATA/ATAA, ATAA/TAAA/AAAT, TTTG/TTGT/TGTT, TCTT/CTTT/TTTC AATT/ATTA/TTAA
Penta nucleotide motifs	ATATT/TATTA/ATTA, TTTTA/TTTAT/TTATT and TTATT/TATTT/ATTTT	–na–	TTTTA/TTTAT/TTATT TTATT/TATTT/ATTTT	TTTTA/TTTAT/TTATT TTATT/TATTT/ATTTT	TTTTA/TTTAT/TTATT TTATT/TATTT/ATTTT
Di- and trinucleotide motifs	TTC/TCT/CTT and TTA/TAT/ATT TA/AT	–na–	–na–	–na–	–na–

Table 9.3 Single nucleotide polymorphisms (SNPs) and insertions/deletions (indel) identified in chloroplast genomes of different *Capsicum* species

	C. annuum var. *annuum* (Raveendar et al. 2015a)	*C. chinense* Jacq. (Raveendar et al. 2017)	*C. baccatum* (Kim et al. 2016)	*C. annuum* var. *glabriusculum* (Raveendar et al. 2015b)	*C. frutescens* (Shim et al. 2016)
Total variants	96	174	282	48	34
Total SNPs	45	82	218	15	18
SNPs (coding region)	13	35	86	5	6
SNPs (LSC region)	78	–na–	262	43	29
SNPs (SSC region)	17	–na–	62	5	5
SNPs (IR region)	1	–na–	4	0	0
Sites of >1nt variant	46	69	60	32	15
Total Indels	51	92	64	33	16
Indels (coding region)	3	7	8	3	2

9.1.2.2 Chloroplast as a Tool for Phylogenetic Analysis

Olmstead and Palmer (1994) proposed guidelines for plant systematics using chloroplast DNA sequences and investigated the use of rapidly evolving sites such as introns, intergenic spacers, and noncoding portions over conventional approach of using slowly evolving gene sequences (those having low substitution rates) such as *rbcL* (coding the large subunit of Rubisco) and rRNA genes for phylogenetic studies. Several markers from *matK, ndhF*, *psaB,* and *trnL-trnF* regions have been widely used for inferring phylogeny in plants including *Capsicum* (Oxelman et al. 1999; Chiang and Schaal 2000; Soltis et al. 2000; Miz et al. 2008). The genus *Solanum* has been a quick reference point for classifying other Solanaceae members such as *Capsicum* and *Lycianthes*. Earlier reports have identified *Jaltomata* as the sister group of *Solanum* (Olmstead and Palmer 1992, 1997). Contrastingly, Bohs and Olmstead (1997) identified *Capsicum* and *Lycianthes* as the sister group to *Solanum* using *ndhF* region. The *ndhF* gene (2220 bp) codes for a subunit of *NADH dehydrogenase* (Sugiura 1989, 1992), which have been frequently used in various phylogenetic studies in plants (Olmstead and Sweere 1994; Clark et al. 1995; Kim and Jansen 1995; Olmstead and Reeves 1995; Scotland et al. 1995; Neyland and Urbatsch 1996; Olmstead et al. 2008). Walsh and Hoot (2001) investigated a long noncoding region of 800 bp between two highly conserved genes—*atpB* and *rbcL*—along with a single copy nuclear gene *waxy* (3 kb gene, 12 introns). *Waxy* encodes for an enzyme in granule-bound starch synthase (GBSS) pathway. In this study, 11 unique *Capsicum* species and 7 outgroups from Solanaceae were investigated to test monophyly of *Capsicum* genus. All the 11

Capsicum members segregated in a distinct clade from sister group *Lycianthes,* despite re-rooting the trees by different outgroups from Solanaceae family, affirming its strong relatedness. Almost similar patterns have also been observed in our own interpretations using multiple Solanaceae chloroplast genomes (Fig. 9.2).

To analyze the utility of a small noncoding intergenic spacer *trnE-trnT*, phylogenetic trees were constructed in a total of 23 species including a total of 16 Solanaceae members with *C. annuum* and *C. frutescens.* Based on sequence alignment and interpretation of one of the 12 most parsimonious trees, three groups were identified. *Physalis* segregated at 53% bootstrap value with *C. annuum* and *C. frutescens*, while *Petunia* and *Solanum* genus maintained their distinct clades (Passarin et al. 2008). To achieve greater phylogenetic resolution, Olmstead et al. (2008) used *ndhF* and *trnLF* together, to establish the relations among 89 genera including 190 species in total. Observations affirmed monophyletic nature of subfamily Solanoideae and Jaltomata being sister to *Solanum* and identified a clade comprising tribes Capsiceae (*Capsicum* and *Lycianthes*) and Physaleae. Garcia et al. (2016) analyzed the phylogenetic relationships and diversification of 34 *Capsicum* species and nine additional species using two chloroplast DNA markers *matK* and *psbA-trnH* and a nuclear gene *waxy*. They conducted maximum parsimony (MP) and Bayesian inference (BI) analysis using the individual markers and also for the combined data set. Two most parsimonious trees (length of 1051 steps) were produced using combined data set. However, individual markers *matK*, *psbA-trnH,* and *waxy* produced 5601, 9770 and 18 most parsimonious trees, respectively. The strict consensus showed low levels of resolution with the two plastid markers in *Capsicum*, while the nuclear gene *waxy* demonstrated the highest resolution at all steps. Thus, the MP consensus tree and BI phylogram resolved *Capsicum* species as monophyletic. *Lycianthes* was recognized as sister group to the genus *Capsicum*, which was also supported by the morphological data (Bohs and Olmstead 1997; Carrizo García et al. 2016). Solanaceae wide phylogram using 75 accessions

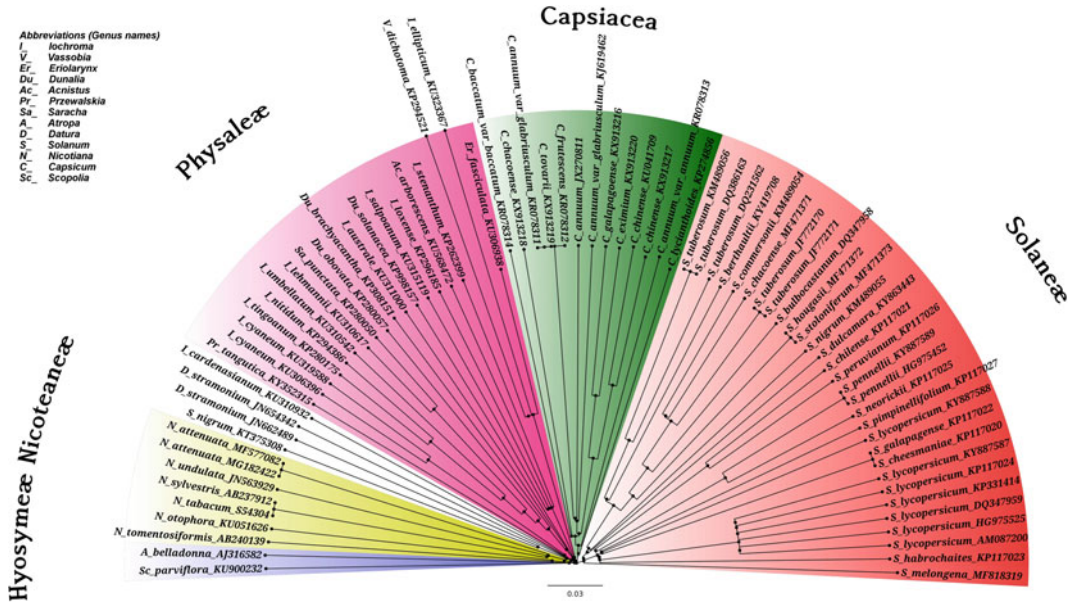


Fig. 9.2 Total 75 cp genome accessions representing 13 genera such as *Capsicum*, *Solanum*, *Nicotiana*, *Datura*, *Scopolia*, *Atropa*, *Saracha*, *Prezewalskia*, *Acnistus*, *Dunalia*, *Eriolarynx*, *Vassobia,* and *Iochroma* showing clear segregation at tribe level. As observed in many previous studies, Capsiacea tribe is closely related to Solaneae and Physaleae

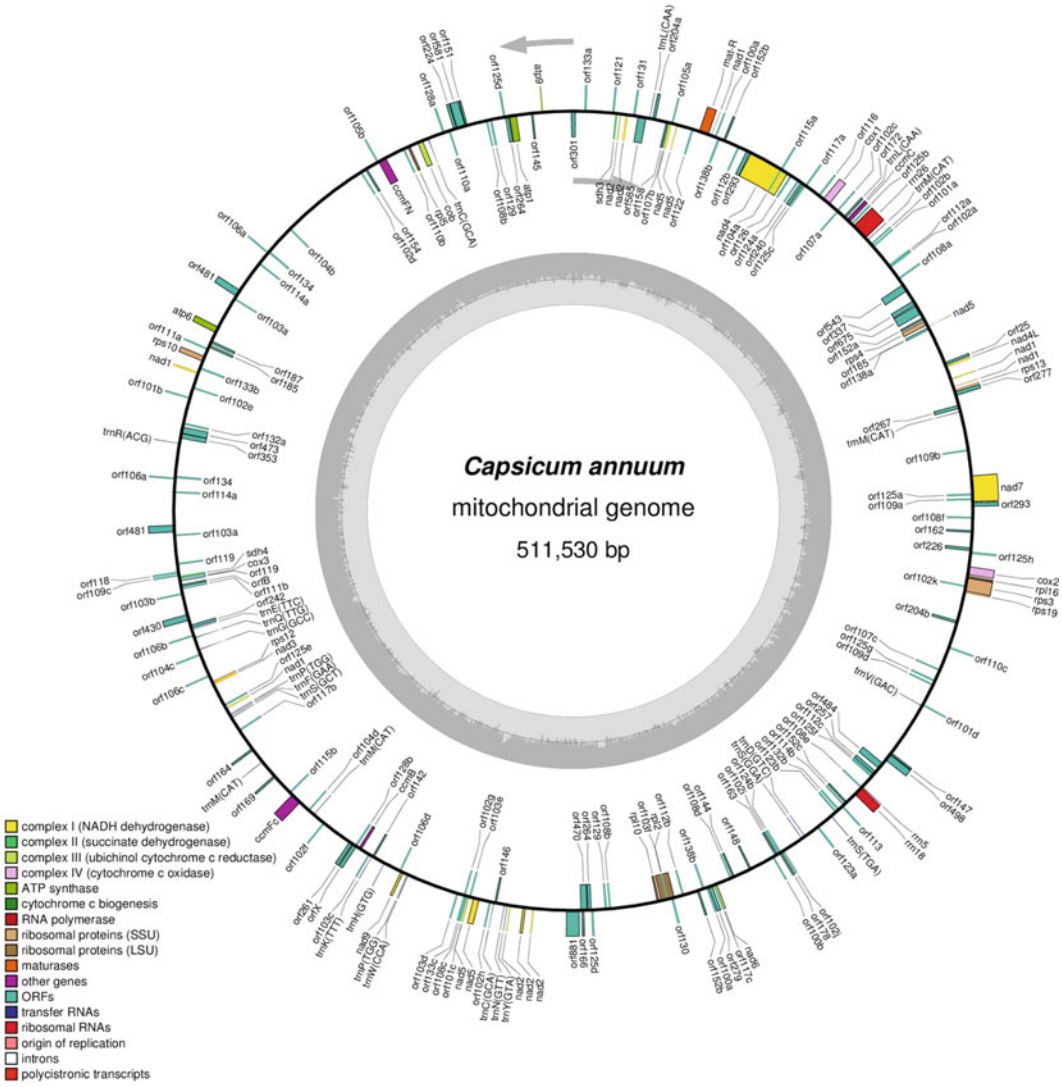


Fig. 9.3 Organization of mitochondrial genome of *C. annuum* cultivar Jeju (GenBank: KJ865410.1) drawn by using OGDRAW (Lohse et al. 2007). All major features including important genes, CDS, ORFs, tRNAs, rRNAs are rendered on circular tracks representing 511,530 bp; faint gray arrows indicate direction of replication. Forward transcribed sequence is represented on outer track, while inner circle indicates reverse strand genes

from 13 genera, obtained from NCBI, was constructed (Fig. 9.2) to demonstrate genus and tribe level relations among members.

9.1.2.3 Comparative Analysis of Cp Genomes of Capsicum and Other Plant Species

An in silico study compared the genome size, gene content, and homology of plastomes of ten Solanaceae members including *C. annuum* to understand chloroplast genome evolution (Kaur et al. 2014). *Capsicum* holds the largest LSC region, while on contrary, it has the smallest SSC region, despite that overall cp genome size was the largest among ten Solanaceae plants. Comparative analysis showed that a duplicated copy of *trnH* gene was observed only in *Capsicum* species. In all other members, *trnH* was a part of *ycf2* gene as

annotated by BLAST search. *Rps19* pseudogene was reported in all species including *Capsicum* except *Nicotiana sylvestris*, *N. tabacum,* and *N. undulata.* Identical orthologous sequence to *sprA* gene was reported in *C. annuum* and five other members. Originally, *sprA* was annotated in *N. sylvestris*, *N. tomentosiformis*, *Solanum lycopersicum*, and *S. tuberosum*. Multiple sequence alignments of many important genes were performed, and *3′* region of *ycf1* (fastest evolving gene) was found to hold species-specific indels, while maximum number of indels were reported in *Capsicum* (Kaur et al. 2014).

Pairwise alignment studies among cp genomes of *Capsicum* and seven other Solanaceae members (*Nicotiana tabacum*, *N. sylvestris*, *N. tomentosiformis*, *Solanum tuberosum*, *S. lycopersicum*, *S. bulbocastanum,* and *Atropa belladonna*) have revealed high level of variations in the protein-coding genes *accD* and *rpl20* and unveiled some unique features of *Capsicum* cp genome. A large insertion of 144 bp in the *accD* gene 674 bp downstream to its start codon was observed in the pepper cp genome. Subsequently, a repeat finder program (Benson 1999) predicted a 18 bp motif repeated seven times in the inserted sequence in *accD* gene; a pair of 15 bp long direct repeats was found along the flanking region of this insert. A stop codon in 3′ region of *rpl20* gene was a common observation in all Solanaceae members in the study, while this seems disrupted in *C. annuum* due to an insertion and a new stop codon is found 80 bp downstream to the usual position in *rpl20* gene, thus increasing the coding region length. Phylogenetic analysis using ClustalW (Thompson et al. 2003) highlighted conservation status of four ORFs (expected to be conserved across all), two ORFs (*orf79* and *orf71B*) being highly conserved across all eight members, and two ORFs (*orf70B* and *orf131*) were shortened in length in pepper plastome by creation of stop codons due to 4 bp insertions (Jo et al. 2011).

9.1.2.4 DNA Barcoding

Correct taxonomic classification of organisms has been a burning debate even before people started to sequence organisms. Early drives of sequencing and annotation of few compact genomes also enabled homology based-comparisons feasible. As available sequences grew in number, it was well established that DNA-based identification techniques can be sufficiently discriminating between species (Blaxter 2003). Initially, the idea was conceived for microscopic organisms like bacteria and viruses, where miniscule anatomy and complex morphology limited the scale of classification by traditional taxonomy (Hebert et al. 2003). Nowadays, it is a widely used approach for any new addition to phylogenetic tree. The regions of low mutation are appropriate for phylogenetic studies, and regions of high mutation rate are important for distinguishing between closely related species. Mitochondrial DNA has relatively smaller number of genes with certain genes having high degree of conservation along with regions of heavy diversity, additionally they are prevalent in almost every living organism. Mitochondrial *cytochrome oxidase 1* (*CO1*) has already been proposed for the purpose. The *CO1* gene is a widely used DNA barcode for animals (Hebert et al. 2003), while two chloroplast genes *rbcL* and *matK* were also investigated to provide a two-locus DNA barcode for land plants (Kress and Erickson 2007). The requirement for an effective barcode is: (i) significant species diversity at species level, (ii) short sequence length, and (iii) presence of conserved flanking sites to develop universal primers. A major setback of *rbcL* and *matK* is that primer universality and discriminating efficiency with these primers are compromised (Hollingsworth et al. 2009). However, in plants, chloroplast (cp) genome sequence is at the core of DNA barcoding. The relatively low evolutionary rate of chloroplast genome makes it a suitable choice among few others, for species identification, and phylogenetic analysis in higher plants. However, the evolutionary rate of chloroplast is higher in comparison with mitochondria (Yi et al. 2012). Chloroplast genome is haploid with very rare or no genetic recombination and follows maternal inheritance. It offers ease of amplification in PCR and sequencing. The selection of target sequences to be amplified in PCR is crucial for the

desired analysis. Furthermore, universal primers can easily amplify any target gene. A number of gene loci including *trnS-trnfM*, *trnL-trnT*, *trnH-psbA*, *trnF-trnL*, *trnD-trnT*, *trnC-rpoB*, *rps16* and *matK* and nuclear *waxy* introns have been investigated, and their feasibilities have been evaluated to use as DNA barcodes for easy differentiation and identification of seven *Capsicum* species (*C. annuum*, *C. baccatum*, *C. chinense*, *C. frutescens*, *C. pubescens*, *C. chacoense* and *C. rhomboideum*). However, none of cp DNA introns were able to differentiate individual members, except *trnL-trnT*, *trnF-trnL* and *trnH-psbA*, which could distinguish *C. annuum*, *C. chinense* and *C. frutescens* from remaining four species (Jarret 2008). Jeong et al. (2010) have also investigated intergenic spacer sequences between *trnL* and *trnF* genes from plastid genome for species identification in *Capsicum* in another study. The use of chloroplast genome will largely replace other markers for genotyping and identification of plant species.

9.1.2.5 Chloroplasts and Genome Ploidy

Doubled haploid plants are very useful in plant breeding. Chromosomal doubling of anthers is a common technique to develop doubled haploids. The variation of ploidy level from anther cultured pepper plants (*C. annuum* L.) was reported (Vaulx et al. 1981). Therefore, alternative methods were adopted to accurately determine the ploidy level. One such method is the chromosome count in the mitotic root tip cells or meiotic flower bud cells, but is impractical when dealing with a large number of plants. Alternatively, chloroplast number per guard cell pair is used as a reliable, fast, and convenient method for early determination of ploidy level in androgenic plants. The ploidy level predicted with chloroplast count is essentially consistent with chromosome count in a way that haploid plants have lesser number of chloroplasts per guard cell pair than diploid plants. Therefore, in vitro chloroplast counting method can be used to separate haploid and diploid individuals from the pool of androgenic plants easily. Similar methods for early determination of ploidy level of androgenic plants are also reported for *S. Lycopersicum* (Jacobs and Yoder 1989), *S. tuberosum* (Singsit and Veilleux 1991), and *Arachis hypogaea* (Singsit and Akins 1992).

9.2 Capsicum Mitochondrial Genome

Although mitochondria are well-studied organelle in animals, yet they are scantily reported in plants, especially in *Capsicum*. Mitochondria are believed to have been evolved by the endosymbiosis of proteobacterium in a single event (Davidov et al. 2006). Mitochondria are bean shaped organelle which serves as a primary source of ATP in a cell. The mitochondrial membrane consists of an outer double envelope and an inner membrane. The aqueous compartment of mitochondria known as matrix has several small and circular DNA molecules along with prokaryotic-like ribosomes. Mitochondria carry out important functions such as tricarboxylic acid cycle (TCA Cycle), respiratory electron transport chain (ETC), and ATP synthesis. Plant mitochondria have only 60 known genes for ribosomal proteins, ETC, tRNAs, and rRNAs; on the contrary, it has a proteome of more than 1000 proteins which are mostly imported from the nuclear genome. Genome packaging is in the form of mitochondrial nucleoprotein complexes called nucleoids which also act as heritable units of mtDNA. Each nucleoid package contains many mtDNA copies imparting functional isolation by having its own nucleoprotein complex. Unlike cp genome, mt genome is a multipartite structure, owing to significant genome size variations, explained by expansion of the intergenic regions, structural rearrangements, and intra- or intermolecular recombination events (Handa 2003) giving rise to several mitotypes. Notably, a high mobility group (HMG) box family of nucleoproteins is associated with mtDNA maintenance and packaging. Few others such as *Aco1* and *llv5* also have additional metabolic roles (Kucej and Butow 2007). Study of these nucleoid proteins also suggests a contrasting belief against mitochondrial endosymbiotic theory of origin (Gray 1989; Gray et al. 1999). For instance, tracing

back origins of HMG, they are revealed to be of eukaryotic ancestry, while mitochondria are believed to be prokaryotic. Notable application of mitochondrial gene formate dehydrogenase (*fdh1*; EC 1.2.1.2) is to confer basic innate immune response to plant, by regulating cell death and defense responses to bacterial pathogens (Choi et al. 2014). Recently, mitochondrial genome sequencing has been reported in *C. annuum* with nucleotide length of approximately 511,530 base pairs (Jo et al. 2011). Mitochondrial genome sequencing has been further discussed in details in the follow-up subsection.

9.2.1 Capsicum Mitochondrial Genome Sequencing

Several crop species including *Capsicum* have been subjected for complete mitochondrial(mt) genome sequencing which includes the sequence comparison analysis of normal (fertile) and mutant phenotypes such as cytoplasmic male sterile (CMS) line so as to identify genes responsible for CMS (Jo et al. 2014). By the year 2014, mitochondrial genomes of two Solanaceae members—tobacco and *Capsicum*—had been published (Sugiyama et al. 2005; Jo et al. 2014). The mitochondrial genome of *C. annuum* cultivar Jeju (GenBank: KJ865410.1) is 511,530 bases long, approximately threefold to fourfold in size, than the majority of the *Capsicum* chloroplast genomes (Fig. 9.3). The first complete mitochondrial genome sequence of *Capsicum* lines was reported by Jo et al. (2011). Two *C. annuum* lines—a CMS (FS4401) and a normal fertile line (Jeju)—were compared for sequence variation analysis among themselves and with the tobacco mt genome. Both the FS4401 and Jeju mt genomes were found to be significantly larger than tobacco mt genome (FS4401-507,452 bp and Jeju-511,530 bp). The two pepper mitochondrial genomes have 7.9 and 7.7% protein-coding genes, respectively, while tobacco had a slightly higher percentage of protein-coding genes due to the presence of repetitive genes (*nad2a*, *rrn26*, *sdh3,* and *trnM*). One notable difference between Jeju and FS4401 is the presence of an additional copy of *ATP synthase 6* (*atp6;* encodes subunit 6 in F_0 complex) gene in the latter. FS4401 also has an additional *tRNA* coding gene in comparison with Jeju. However, both the mt genomes have similar number of *rRNA* genes. The protein-coding genes and their functions in the pepper mt genome are summarized in Table 9.4.

Comparative analysis of protein-coding genes of the two pepper mt genomes—CMS line FS4401 and normal fertile Jeju—reveals sequence polymorphism in five genes namely, *atp4, atp8, rpl2, sdh3,* and *atp6*. The sequence polymorphism has led to changes in the protein sequence as well as structure. The nucleotide sequence variation in the above protein-coding genes among the two cultivars is speculated to be the reason behind CMS in FS4401. Total 45 and 30 unique open reading frames (ORFs), coding for proteins of more than 100 amino acids, other than those coding for known genes, were found to be present in FS4401 and Jeju, respectively.

Table 9.4 Protein-coding genes in mitochondrial genome of *Capsicum* species (Joet al. 2011)

Function	Gene names
Complex I proteins	*nad1, nad2, nad3, nad 4, nad4L, nad5, nad6, nad7, nad 9*
Complex II proteins	*Sdh3, sdh4*
Complex III proteins	*cob*
Complex IV proteins	*cox1, cox2, cox3*
ATP synthase units	*atp1, atp4, atp6, atp8, atp9*
Ribosomal proteins	*rpl2, rpl5, rpl10, rpl16, rps3, rps4, rps10, rps12, rps13, rps19*
Cytochrome C biogenesis	*ccmB, ccmC, ccmFc, ccmFN*
Maturase	*matR*
Protein translocation system subunit	*mttB*

Out of these, *orf507* in FS4401 was found to be a strong candidate for CMS. Another ORF, *ψatp6-2* has been reported by Kim and Kim (2006), Kim et al. (2007), as a possible candidate for CMS. Comparative studies of the structure of genomic regions around the *orf507* have revealed an interesting observation. The *orf507* gene in FS4401 is surrounded by *cox2* gene upstream and *ψatp6-2* gene 12 kb downstream to it. While in Jeju, *orf507* is not only absent, but also *cox2* and *ψatp6-2* are distantly located. Another repeated sequence (*R19*, *Ra*) is also located downstream to *cox2* in FS4401. A small number of nucleotides overlap with *Ra* and *orf407*, and sequences highly similar to these overlapping sequences are found in FS4401, Jeju, and tobacco in the 5′ upstream region of *nad9* gene. Jeju on the other hand has *CS1, CS2,* and *R21* sequence elements in the downstream region of *cox2*. These sequence elements are, however, present in different regions in FS4401. Similarly, gene sequences around *atp6 (ψatp6-2)* have shown DNA rearrangements too. The FS4401 line specifically has a repeated sequence *Rb* and *CS2-R21* sequence element in the downstream region to *atp6* gene. A small number of nucleotides overlapped the conserved regions of *atp6, Rb* and *CS2*. Conclusively, *R*21 in Jeju was duplicated in the downstream region of *atp6* in FS4401 which results in the generation of a repeated pair around *ψatp6-2* gene. Apart from this, multiple DNA rearrangements would have resulted due to the insertion of *orf407* and other sequence elements in question between *cox2* and *R21*. The complete mitochondrial genome sequencing of *Capsicum* species has enabled chloroplast genome-based molecular marker development, and an average of 45.50 SSR loci was identified in *C. annuum* (Cheng et al. 2016).

9.2.2 Application of Mitochondrial Genome Study

9.2.2.1 Understanding the Cytoplasmic Male Sterility

To force interbreeding, crossing is performed manually which is a labor-intensive operation of plant breeding. Therefore, male sterile lines have been developed to address this issue which enables easy crossing schemes with clear tracking of pedigree. Plant male sterility (Mayr 1986) is the inability to produce dehiscent anthers, functional pollen grains, and therefore viable male gametes. It could be cytoplasmic male sterility (CMS) or genic male sterility (GMS). CMS is caused by myriad of interactions between the mitochondrial and nuclear genes which together control male sterility and fertility restoration. It has applications in the commercial production of hybrid seeds harnessing heterosis or hybrid vigor.

The CMS-based hybrid seed technology uses three different genetic lines: (i) CMS line, (ii) maintainer line, and (iii) restorer line. The CMS line has a male sterile cytoplasm and is used as a female parent. It has a CMS causing gene (*CMS* gene) and lacks functional nuclear restorer of fertility (*Rf*) gene or genes. The maintainer line has the same nuclear genome as the CMS line but a normal fertile cytoplasm. This line serves as the male parent for the propagation of CMS line. The third line (restorer line) has the functional nuclear restorer of fertility (*Rf*) gene and serves as the male parent for CMS line to produce F_1 hybrid seeds. The F_1 plants are able to restore their fertility owing to the nuclear *Rf* gene from the restorer male parent. The combination of the nuclear genomes from the CMS line and the restorer line produces the fertile hybrid vigor. This is a classic example showing the mutual regulatory network between nuclear *Rf* gene and plastid *CMS* gene. *CMS/Rf* systems have been extensively studied to identify candidate cytoplasmic sterility genes in various crop plants. An *Rf*-linked region 10 kb upstream to a random amplified polymorphic DNA (RAPD) fragment has been sequenced for finding polymorphism. Apart from the dominant and the recessive alleles (98% homology reported among them), a third haplotype was identified and was found to cause unstable male sterility in exotic breeding lines (Min et al. 2008). The CMS genes are result of gene rearrangements in the mitochondrial genome. Most of the genes involved in the origination of CMS genes are essential genes

of electron transport chain (Chen and Liu 2014). No wonder that most of the CMS genes code for transmembrane proteins. Among the many CMS-associated mitochondrial genes in pepper, *orf456* is found at the 3′ end of *cox2* gene in *C. annuum* and encodes for a 17 kDa protein. The functional role of *orf456* was validated by transforming mitochondria-targeted gene construct in Arabidopsis by *Agrobacterium*-mediated gene transfer (Kim et al. 2007). This is also evident in pepper by the presence of *atp6* (an essential *ATP synthase*) sequences in *orf456* CMS-Peterson. The reduced activity of F_1F_oATP *synthase* in mitochondria has been reported to affect floral development and morphogenesis and therefore causes pollen grain abortion in *Capsicum* (Li et al. 2009b). Another report suggests the role of aberrant cytochrome c oxidase activity along with $F_1F_oATPase$ for causing anther abortion in the CMS line. The *ψatp6-2* gene controls $F_1F_oATPase$ activity, and the *orf507* controls *cytochrome c oxidase* activity and is responsible for their dysfunction, respectively (Ji et al. 2013). Apart from this, CMS is rather driven by rearrangements of mitochondrial DNA.

9.3 Overview of Capsicum Organellar Sequencing Protocols

This section has summarized practices followed in various *Capsicum* sequencing projects. Protocols may vary in some specific projects due to practicality or personal preference; more or less they are consistent across all projects.

9.3.1 DNA Extraction and Library Preparation

As a common practice, DNA is extracted from 40 days old *Capsicum* seedlings. After DNA extraction, chloroplast DNA libraries are created using library preparation kits from different sequencing companies. For *C. frutescens*, *C. chinense* (Raveendar et al. 2017), *C. annuum* var. *annuum*, *C. annuum* var. *glabriusculum* (GenBank: KR078311), and *C. baccatum* cp/mt, sequencing an Illumina paired-end cp DNA library (average insert size of 500 bp) was constructed using the Illumina TruSeq library preparation kit and library was sequenced with 2 × 300 bp on the MiSeq instrument. While for *C. chinense* (Park et al. 2016), Illumina paired-end (PE) genomic library of 300-bp insert was constructed and later sequenced using Illumina Hiseq 2000 platform. Illumina has been the platform of choice in case of *Capsicum* sequencing projects.

9.3.2 Sequence Assembly

De novo and reference assembly are preferred methods for assembling the short reads to full-length contigs. Many read quality filtration criteria are followed before going for assembly. Low-quality reads (Phred Q score <20) are initially filtered out. In majority of *Capsicum* cp genome sequencing projects, cp genome reads are filtered from a pool of whole genome reads in a downstream refinement cycle in genome assembly workflow. Starting with whole genome sequence data sets, BLASTN search against known cp sequences is used to select Cp contigs. These contigs are used for final assembly of cp genome. CLC genome assembler has been used for *C. baccatum*, *C. chinense* (Park et al. 2016), *C. chinense* (Raveendar et al. 2017), *C. frutescens,* and *C. annuum* var. *glabriusculum* (Raveendar et al. 2015b), while *C. annuum* var. *glabriusculum* (Zeng et al. 2016) Cp genome assembly was done with SOAPdenovo software (Li et al. 2010).

9.3.3 Annotation and Visualization

Annotation refers to characterizing genome features, which are used in downstream analysis and classification of various genomic features, and if coupled with co-ordinates, it can be used as a navigable genomic map. Unlike nuclear genomes, which show limited homology across species, relatively conserved nature of

mitogenome and plastome also boosts accuracy in the organellar genome annotation. There are dedicated pipelines to annotate chloroplast genomes. Dual Organellar GenoMe Annotator (DOGMA; Wyman et al. 2004) has been successfully used in majority of *Capsicum* sequence annotation projects including *C. baccatum*, *C. annuum* var. *annuum*, *C. chinense*, *C. annuum* var. *glabriusculum*, *C. frutescens*, while many Solanaceae members are also annotated using this program e.g. *S. dulcamara*, *S. bulbocastanum, and S. tuberosum*. Default parameters were used to predict various features including protein-coding, tRNA, and rRNA genes in all of the projects. DOGMA is a generic annotation tool for mitochondria as well as chloroplasts. FASTA formatted complete genomic sequences serve as input. It performs BLASTX searches against a custom database of published chloroplast genomes. CpGAVAS (Liu et al. 2012) is another notable tool which has been used in annotation of *C. baccatum*, *C. frutescens,* and *S. dulcamara*. Very often multiple tools are used in combination or independently, which gives researchers a new vantage point of described gene models and also for gap filling. All transfer RNA (tRNA) genes were amended with tRNAscan-SE (Lowe and Eddy 1997) in *C. baccatum* and *C. frutescens*. To validate CDS regions, gene positions, and identification of intron containing genes, simple BLASTN (Altschul et al. 1990) search has also been successfully used in case of *C. annuum* var. *glabriusculum* (Zeng et al. 2016),*C. chinense* (Raveendar et al. 2017),*C. annuum* var. *glabriusculum* (Raveendar et al. 2015b), and *C. annuum* var. *annuum*. Several RNA-seq data sets of corresponding species have also been used to predict start/stop codons, gene length, splice sites in *Solanaceae*. Verdant (McKain et al. 2017) and GeSeq (Tillich et al. 2017) are also good softwares for organellar genome annotation, and they have been used in related studies.

Plastome maps are visualized as circular tracks, unlike linear representations for most of the nuclear data sets. In real-world scenario, user workflow for data analysis may also involve some specific environment such as R (notable visualization packages such as circlize and circosR), perl (circos), python (chord diagramming in Bokeh library), or some Web-based scripting languages e.g., JavaScript (d3). They serve needs of most of the tech savvy users, but simple Web-based alternatives are also available, e.g., OGDRAW, CpGAVAS (quick draw module), GenomeVX, and CGView. Most of the mentioned programmes understand gff/genbank files as input and render beautiful circular diagrams. Majority of published cp genome maps in *Capsicum* and *Solanaceae* are constructed using the organellar genome DRAW (OGDRAW; Lohse et al. 2007) software.

9.4 Conclusion

The availability of complete information of whole genome sequences of chloroplast and mitochondria has opened up enormous opportunity to *Capsicum* researchers to study the details of gene number, their functions, and conservation and diversification with that of other plants. Furthermore, the structure and sequence variation in the chloroplast and mitochondrial genomes have helped in deriving evolutionary relationships within and between the phylogenetic clades. The thorough study to explore the biological functions of all the identified genes and their interactions with other genes from organelles or nuclear genome would shed more light in details of organellar genomes of *Capsicum* species.

Acknowledgements This work was supported by UGC Resource Networking Project, funded by University Grants Commission, New Delhi, to School of Life Sciences, Jawaharlal Nehru University, and DST PURSE grant from the Department of Science and Technology, Government of India, to the Jawaharlal Nehru University, Delhi, India.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360

Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. Nucl Acids Res 573–580

Blaxter M (2003) Molecular systematics: counting angels with DNA. Nature 421:122–124. https://doi.org/10.1038/421122a

Bohs L, Olmstead RG (1997) Phylogenetic relationships in *Solanum* (Solanaceae) based on *ndhF* sequences. Syst Bot 22:5–17. https://doi.org/10.2307/2419674

Campbell NA, Reece JB, Urry LA, Cain ML, Wasserman SA, Minorsky PV, Jackson RB (2009) Biology, 8th edn. Pearson, Benjamin Cummings, San Francisco, p 516

Carde JP, Camara B, Cheniclet C (1988) Absence of ribosomes in *Capsicum* chromoplasts. Planta 173:1–11. https://doi.org/10.1007/BF00394480

Cardle L, Ramsay L, Milbourne D, Macaulay M, Marshall D, Waugh R (2000) Computational and experimental characterization of physically clustered simple sequence repeats in plants. Genetics 156:847–854

Carrizo García C, Barfuss MHJ, Sehr EM, Barboza GE, Samuel R, Moscone EA, Ehrendorfer F (2016) Phylogenetic relationships, diversification and expansion of chili peppers (*Capsicum*, Solanaceae). Ann Bot 118:35–51. https://doi.org/10.1093/aob/mcw079

Chen L, Liu YG (2014) Male sterility and fertility restoration in crops. Ann Rev Plant Biol 65:579–606. https://doi.org/10.1146/annurev-arplant-050213-040119

Cheng J, Zhao Z, Li B, Qin C, Wu Z, Trejo-Saavedra DL, Luo X, Cui J, Rivera-Bustamante RF, Li S, Hu K (2016) A comprehensive characterization of simple sequence repeats in pepper genomes provides valuable resources for marker development in *Capsicum*. Sci Rep 6:18919

Chiang TY, Schaal BA (2000) Molecular evolution of the *atpB-rbcL* noncoding spacer of chloroplast DNA in moss family Hylocomiaceae. Bot Bull Acad Sin 41:85–92

Shin MJ, Lee GA, Lee JR, Cho YH, MA KH, Raveendar S (2017) The complete chloroplast genome sequence of *Capsicum tovarii*—a Peruvian wild pepper. J Plant Physiol Pathol Bangkok, Thailand 5:52

Choi DS, Kim NH, Hwang BK (2014) Pepper mitochondrial *Formate dehydrogenase1* regulates cell death and defense responses against bacterial pathogens. Plant Physiol 166:1298–1311. https://doi.org/10.1104/pp.114.246736

Chung HJ, Jung JD, Park HW, Kim JH, Cha HW, Min SR, Jeong WJ, Liu JR (2006) The complete chloroplast genome sequences of *Solanum tuberosum* and comparative analysis with Solanaceae species identified the presence of a 241-bp deletion in cultivated potato chloroplast DNA sequence. Plant Cell Rep 25:1369–1379. https://doi.org/10.1007/s00299-006-0196-4

Clark LG, Zhang W, Wendel JF (1995) A phylogeny of the grass family (Poaceae) based on *ndhF*sequence data. Syst Bot 20:436–460. https://doi.org/10.2307/2419803

Darling AC, Mau B, Blattner FR, Perna NT (2004) Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 7:1394–1403. https://doi.org/10.1101/gr.2289704

Davidov Y, Huchon D, Koval SF, Jurkevitch E (2006) A new α-proteobacterial clade of Bdellovibrio-like predators: implications for the mitochondrial endosymbiotic theory. Environ Microbiol 8:2179–2188. https://doi.org/10.1111/j.1462-2920.2006.01101.x

Drews G (2005) Contributions of Theodor Wilhelm Engelmann on phototaxis, chemotaxis, and photosynthesis. Photosynth Res 83:25–34. https://doi.org/10.1007/s11120-004-6313-8

Ellis RJ (1979) The most abundant protein in the world. Trends Biochem Sci 4:241–244. https://doi.org/10.1016/0968-0004(79)90212-3

Gibbs SP (1981) The chloroplasts of some algal groups may have evolved from endosymbiotic eukaryotic algae. Ann New York Acad Sci 361:193–208. https://doi.org/10.1111/j.1749-6632.1981.tb54365.x

Gray MW (1989) The evolutionary origins of organelles. Trends Genet 5:294–299. https://doi.org/10.1016/0168-9525(89)90111-X

Gray MW, Burger G, Lang BF (1999) Mitochondrial evolution. Science 283:1476–1481. https://doi.org/10.1126/science.283.5407.1476

Handa H (2003) The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (*Brassica napus* L.): comparative analysis of the mitochondrial genomes of rapeseed and *Arabidopsis thaliana*. Nucl Acids Res 31:5907–5916. https://doi.org/10.1093/nar/gkg795

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. ProcRoy Soc B BiolSci 270:313–321. https://doi.org/10.1098/rspb.2002.2218

Hollingsworth PM, Forrest LL, Spouge JL, Hajibabaei M, Ratnasingham S, Bank M van der, Chase MW, Cowan RS, Erickson DL, Fazekas AJ, Graham SW, James KE, Kim K-J, Kress WJ, Schneider H, AlphenStahl J van, Barrett SCH, Berg C van den, Bogarin D, Burgess KS, Cameron KM, Carine M, Chacón J, Clark A, Clarkson JJ, Conrad F, Devey DS, Ford CS, Hedderson TAJ, Hollingsworth ML, Husband BC, Kelly LJ, Kesanakurti PR, Kim JS, Kim Y-D, Lahaye R, Lee H-L, Long DG, Madriñán S, Maurin O, Meusnier I, Newmaster SG, Park C-W, Percy DM, Petersen G, Richardson JE, Salazar GA, Savolainen V, Seberg O, Wilkinson MJ, Yi D-K, Little DP (2009) A DNA barcode for land plants. Proc Natl Acad Sci USA106:12794–12797. https://doi.org/10.1073/pnas.0905845106

Jacobs JP, Yoder JI (1989) Ploidy levels in transgenic tomato plants determined by chloroplast number. Plant Cell Rep 7:662–664. https://doi.org/10.1007/BF00272055

Jarret RL (2008) DNA barcoding in a crop genebank: the *Capsicum annuum* species complex. Open Biol J 1:35–42. https://doi.org/10.2174/1874196700801010035

Jeong HJ, Jo YD, Park SW, Kang BC (2010) Identification of *Capsicum* species using SNP markers based on

high resolution melting analysis. Genome 53:1029–1040. https://doi.org/10.1139/G10-094

Ji J, Huang W, Yin C, Gong Z (2013) Mitochondrial *Cytochrome c Oxidase* and *F1Fo-ATPase*dysfunction in peppers (*Capsicum annuum* L.) with cytoplasmic male sterility and its association with *orf507* and *Ψatp6-2*genes. Intl JMol Sci 14:1050–1068. https://doi.org/10.3390/ijms14011050

Jo YD, Park J, Kim J, Song W, Hur CG, Lee YH, Kang BC (2011) Complete sequencing and comparative analyses of the pepper (*Capsicum annuum* L.) plastome revealed high frequency of tandem repeats and large insertion/deletions on pepper plastome. Plant Cell Rep 30:217–229. https://doi.org/10.1007/s00299-010-0929-2

Jo YD, Choi Y, Kim DH, Kim BD, Kang BC (2014) Extensive structural variations between mitochondrial genomes of CMS and normal peppers (*Capsicum annuum* L.) revealed by complete nucleotide sequencing. BMC Genomics 15:561. https://doi.org/10.1186/1471-2164-15-561

Kahlau S, Aspinall S, Gray JC, Bock R (2006) Sequence of the tomato chloroplast DNA and evolutionary comparison of solanaceous plastid genomes. J Mol Evol 63:194–207. https://doi.org/10.1007/s00239-005-0254-5

Kaur H, Singh BP, Singh H, Nagpal AK (2014) Comparative genomics of ten solanaceous plastomes. Adv Bioinformat 2014:13. https://doi.org/10.1155/2014/424873

Kim KJ, Jansen RK (1995) *ndhF* sequence evolution and the major clades in the sunflower family. Proc Natl Acad Sci USA 92:10379–10383. https://doi.org/10.1073/pnas.92.22.10379

Kim DH, Kim B-D (2006) The organization of mitochondrial *atp6* gene region in male fertile and CMS lines of pepper (*Capsicum annuum* L.). CurrGenet 49:59–67. https://doi.org/10.1007/s00294-005-0032-3

Kim DH, Kang JG, Kim BD (2007) Isolation and characterization of the cytoplasmic male sterility-associated *orf456* gene of chili pepper (*Capsicum annuum* L.). Plant Mol Biol 63:519–532. https://doi.org/10.1007/s11103-006-9106-y

Kim TS, Lee JR, Raveendar S, Lee GA, Jeon YA, Lee HS, Ma KH, Lee SY, Chung JW (2016) Complete chloroplast genome sequence of *Capsicum baccatum* var. baccatum. Mol Breed 36:110. https://doi.org/10.1007/s11032-016-0532-5

Kress WJ, Erickson DL (2007) A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. PLoS ONE 2:e508. https://doi.org/10.1371/journal.pone.0000508

Kucej M, Butow RA (2007) Evolutionary tinkering with mitochondrial nucleoids. Trends Cell Biol 17:586–592. https://doi.org/10.1016/j.tcb.2007.08.007

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R (2009a) The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li JJ, Jo YD, Kim YM, Kang BC (2009b) Cytoplasmic male sterility of *Capsicum* is caused by a reduced activity of *ATP synthase* in mitochondria. Proc Korean Soc Plant Biotechnol 2009:150

Li R, Zhu H, Ruan J, Qian W, Fang X, Shi Z, Li Y, Li S, Shan G, Kristiansen K, Li S, Yang H, Wang J, Wang J (2010) De novo assembly of human genomes with massively parallel short read sequencing. Genome Res 20:265–272. https://doi.org/10.1101/gr.097261.109

Liu C, Shi L, Zhu Y, Chen H, Zhang J, Lin X, Guan X (2012) CpGAVAS, an integrated web server for the annotation, visualization, analysis, and GenBank submission of completely sequenced chloroplast genome sequences. BMC Genom 13:715. https://doi.org/10.1186/1471-2164-13-715

Lohse M, Drechsel O, Bock R (2007) Organellar Genome DRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. Curr Genet 52:267–274. https://doi.org/10.1007/s00294-007-0161-y

Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucl Acids Res 25:955–964

Mayr E (1986) Joseph Gottlieb Kolreuter's contributions to biology. Osiris 2:135–176. https://doi.org/10.1086/368655

McKain MR, Hartsock RH, Wohl MM, Kellogg EA (2017) Verdant: automated annotation, alignment and phylogenetic analysis of whole chloroplast genomes. Bioinformatics 33:130–132. https://doi.org/10.1093/bioinformatics/btw583

Min W, Lim H, Lee YP, Sung SK, Kim BD, Kim S (2008) Identification of a third haplotype of the sequence linked to the *Restorer-of-fertility (Rf)* gene and its implications for male-sterility phenotypes in peppers (*Capsicum annuum* L.). Mol Cells 25:20–29

Miz RB, Mentz LA, Souza-Chies TT (2008) Overview of the phylogenetic relationships of some southern Brazilian species from section Torva and related sections of "spiny Solanum" (*Solanum* subgenus Leptostemonum, Solanaceae). Genetica 132:143–158. https://doi.org/10.1007/s10709-007-9156-3

Neyland R, Urbatsch LE (1996) Phylogeny of subfamily Epidendroideae (Orchidaceae) inferred from *ndhF* chloroplast gene sequences. Amer JBot 83:1195–1206. https://doi.org/10.1002/j.1537-2197.1996.tb13901.x

Ohyama K, Fukuzawa H, Kohchi T, Shirai H, Sano T, Sano S, Umesono K, Shiki Y, Takeuchi M, Chang Z, Aota S, Inokuchi H, Ozeki H (1986) Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. Nature 322:572–574. https://doi.org/10.1038/322572a0

Olmstead RG, Palmer JD (1992) A chloroplast DNA phylogeny of the *Solanaceae*: subfamilial

relationships and character evolution. Ann MO Bot Garden 79:346–360. https://doi.org/10.2307/2399773

Olmstead RG, Palmer JD (1994) Chloroplast DNA systematics: a review of methods and data analysis. Amer JBot 81:1205–1224. https://doi.org/10.1002/j.1537-2197.1994.tb15615.x

Olmstead RG, Palmer JD (1997) Implications for the phylogeny, classification, and biogeography of *Solanum* from cpDNA restriction site variation. Syst Bot 22:19–29. https://doi.org/10.2307/2419675

Olmstead RG, Reeves PA (1995) Evidence for the polyphyly of the Scrophulariaceae based on chloroplast *rbcL* and *ndhF* sequences. Ann Mo Bot Gard 82:176–193. https://doi.org/10.2307/2399876

Olmstead RG, Sweere JA (1994) Combining data in phylogenetic systematics: an empirical approach using three molecular data sets in the Solanaceae. Syst Biol 43:467–481. https://doi.org/10.1093/sysbio/43.4.467

Olmstead RG, Bohs L, Migid HA, Santiago VE, Garcia VF, Collier SM (2008) A molecular phylogeny of the Solanaceae. Taxon 57:1159–1181

Oxelman B, Backlund M, Bremer B (1999) Relationships of the Buddlejaceae s. 1. Investigated using Parsimony Jackknife and branch support analysis of chloroplast *ndhF* and *rbcL* sequence data. Syst Bot 24:164–182. https://doi.org/10.2307/2419547

Palmer JD (1985) Chloroplast DNA and molecular phylogeny. BioEssays 2:263–267. https://doi.org/10.1002/bies.950020607

Palmer JD, Thompson WF (1982) Chloroplast DNA rearrangements are more frequent when a large inverted repeat sequence is lost. Cell 29:537–550. https://doi.org/10.1016/0092-8674(82)90170-2

Park HS, Lee J, Lee SC, Yang TJ, Yoon JB (2016) The complete chloroplast genome sequence of *Capsicum chinense* Jacq. (Solanaceae). Mitochondrial DNA Part B 1:164–165. https://doi.org/10.1080/23802359.2016.1144113

Passarin DMM, Berger IJ, Dressano K, Martin VF, Oliveira GCX, Bock R, Carrer H (2008) Phylogenetic relationships in Solanaceae and related species based on cpDNA sequence from plastid *trnE-trnT* region. Crop Breed Appl Biotechnol 8:85–95. https://doi.org/10.12702/1984-7033.v08n01a12

Raveendar S, Jeon YA, Lee JR, Lee GA, Lee KJ, Cho GT, Ma KH, Lee SY, Chung JW (2015a) The complete chloroplast genome sequence of Korean landrace "Subicho" pepper (*Capsicum annuum* var. *annuum*). Plant Breed Biotechnol 3:88–94. https://doi.org/10.9787/PBB.2015.3.2.088

Raveendar S, Na YW, Lee JR, Shim D, Ma KH, Lee SY, Chung JW (2015b) The complete chloroplast genome of *Capsicum annuum* var. glabriusculum using Illumina sequencing. Molecules 20:13080–13088. https://doi.org/10.3390/molecules200713080

Raveendar S, Lee KJ, Shin MJ, Cho GT, Lee JR, Ma KH, Lee GA, Chung JW (2017) Complete chloroplast genome sequencing and genetic relationship analysis of *Capsicum chinense* Jacq. Plant Breed Biotechnol 5:261–268. https://doi.org/10.9787/PBB.2017.5.4.261

Ryzhova NN, Kochieva EZ (2004) Analysis of microsatellite loci of the chloroplast genome in the genus *Capsicum* (Pepper). Russ JGenet 40:892–896. https://doi.org/10.1023/B:RUGE.0000039723.90902.b0

Samiee M, Kohnehrouz BB (2015) Isolating *rbcL* gene and promoter of bell pepper (*Capsicum annuum* L.) and its sequence analysis using bioinformatic tools. Intl JBiosci 6:395–402. https://doi.org/10.12692/ijb/6.2.395-402

Scotland RW, Sweere JA, Reeves PA, Olmstead RG (1995) Higher-level systematics of Acanthaceae determined by chloroplast DNA sequences. Amer J Bot 82:266–275. https://doi.org/10.1002/j.1537-2197.1995.tb11494.x

Shim D, Raveendar S, Lee JR, Lee GA, Ro NY, Jeon YA, Cho GT, Lee HS, Ma KH, Chung JW (2016) The complete chloroplast genome of *Capsicum frutescens* (Solanaceae). Appl Plant Sci 4:1600002. https://doi.org/10.3732/apps.1600002

Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Yamaguchi-Shinozaki K, Ohto C, Torazawa K, Meng BY, Sugita M, Deno H, Kamogashira T, Yamada K, Kusuda J, Takaiwa F, Kato A, Tohdoh N, Shimada H, Sugiura M (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J 5:2043–2049. https://doi.org/10.1002/j.1460-2075.1986.tb04464.x

Singsit C, Akins OP (1992) Rapid estimation of ploidy levels in in-vitro-regenerated interspecific *Arachis* hybrids and fertile triploids. Euphytica 64:183–188. https://doi.org/10.1007/BF00046047

Singsit C, Veilleux RE (1991) Chloroplast density in guard cells of leaves of anther-derived potato plants grown in vitro and in vivo. HortScience 26:592–594

Soltis DE, Soltis PS, Chase MW, Mort ME, Albach DC, Zanis M, Savolainen V, Hahn WH, Hoot SB, Fay MF, Axtell M, Swensen SM, Prince LM, Kress WJ, Nixon KC, Farris JS (2000) Angiosperm phylogeny inferred from 18S rDNA, *rbcL*, and *atpB* sequences. Bot JLinn Soc 133:381–461. https://doi.org/10.1111/j.1095-8339.2000.tb01588.x

Sugiura M (1989) The chloroplast chromosomes in land plants. Annu Rev Cell Biol 5:51–70. https://doi.org/10.1146/annurev.cb.05.110189.000411

Sugiura M (1992) The chloroplast genome. Plant Mol Biol 19:149–168

Sugiyama Y, Watase Y, Nagase M, Makita N, Yagura S, Hirai A, Sugiura M (2005) The complete nucleotide sequence and multipartite organization of the tobacco mitochondrial genome: comparative analysis of mitochondrial genomes in higher plants. Mol Genet Genom 272:603–615. https://doi.org/10.1007/s00438-004-1075-8

Thiel T, Michalek W, Varshney RK, Graner A (2003) Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). Theor Appl Genet 106:411–422. https://doi.org/10.1007/s00122-002-1031-0

Thompson JD, Gibson TJ, Higgins DG (2003) Multiple sequence alignment using ClustalW and ClustalX. Curr Protocols Bioinformat 00:2.3.1–2.3.22. https://doi.org/10.1002/0471250953.bi0203s00

Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S (2017) GeSeq—versatile and accurate annotation of organelle genomes. Nucl Acids Res 45:W6–W11. https://doi.org/10.1093/nar/gkx391

Vaulx DR, Chambonnet D, Pochard E (1981) Culture in vitro d'anthères de piment (*Capsicum annuum* L.): amélioration des taux d'obtention de plantes chez différents génotypes par des traitements à + 35 °C. Agronomie 1:859–864

Walsh BM, Hoot SB (2001) Phylogenetic relationships of *Capsicum* (Solanaceae) using DNA sequences from two noncoding regions: the chloroplast atpB-rbcL spacer region and nuclear waxy Introns. Intl J Plant Sci 162:1409–1418. https://doi.org/10.1086/323273

Wyman SK, Jansen RK, Boore JL (2004) Automatic annotation of organellar genomes with DOGMA. Bioinformatics 20:3252–3255. https://doi.org/10.1093/bioinformatics/bth352

Yi DK, Lee HL, Sun BY, Chung MY, Kim KJ (2012) The complete chloroplast DNA sequence of *Eleutherococcus senticosus* (Araliaceae); comparative evolutionary analyses with other three Asterids. Mol Cells 33:497–508. https://doi.org/10.1007/s10059-012-2281-6

Zeng F, Gao C, Gao L (2016) The complete chloroplast genome sequence of American bird pepper (*Capsicum annuum* var. glabriusculum). Mitochondrial DNA Part A 27:724–726. https://doi.org/10.3109/19401736.2014.913160

Noncoding RNAs in *Capsicum* Genome

10

Ilyas Ahmad, Mukesh Nitin, Abdul Rawoof, Meenakshi Dubey and Nirala Ramchiary

Abstract

Noncoding RNAs were not only classified based on their structural variation but also on the basis of their regulatory roles in diverse cellular pathways. During the last two decades, several studies on noncoding RNAs have been reported and now it is well known that they regulate expression of genes involved in diverse biological functions. Therefore, in gene regulatory network, noncoding RNAs are considered as the important top-tier regulators. In Solanaceae plants, several studies on noncoding RNAs especially small noncoding RNAs are reported mostly in Tomato. However, in *Capsicum* (pepper), one of the most important vegetable crops, belonging to the same family Solanaceae as of Tomato, the identification and characterization of noncoding RNAs are still limited. Furthermore, recently the identification and characterization of long noncoding RNAs are being reported in plants including Solanaceae plants. Therefore, in this chapter, an attempt is being made to highlight the identification and characterization of noncoding RNAs in *Capsicum* species.

10.1 Introduction

The transcribed region of genome produces an array of RNA molecules with difference size, abundance, function, location, and protein-coding capability. Very small amount of these RNAs are translated into protein (Coding RNAs) and rest of these RNAs do not translate, which are collectively known as noncoding RNAs (ncRNA) and increasing evidences of several studies highlighted below shows that noncoding RNAs play important role in regulating genes involved in many biological processes. This largely neglected noncoding part of the genome is now gaining a lot of importance in the scientific community because of their involvement in many complex molecular mechanisms. They comprise the major class of epigenetic regulators which control the genes without influencing the DNA sequences (Delpu et al. 2016). These noncoding RNAs were considered as "transcriptional noise" earlier (Backofen et al. 2006; Herbig and Nieselt 2011). Noncoding RNAs are divided into two major groups, i.e., housekeeping ncRNAs and regulatory ncRNAs. Housekeeping ncRNA such as Ribosomal RNAs (rRNAs), transfer RNAs (tRNAs), small nuclear RNAs (snRNAs), and small nucleolar RNAs

I. Ahmad · M. Nitin · A. Rawoof · M. Dubey
N. Ramchiary (✉)
Translational and Evolutionary Genomics Lab, School of Life Sciences, Jawaharlal Nehru University, 110067 New Delhi, India
e-mail: nrudsc@gmail.com

M. Dubey · N. Ramchiary
Department of Biotechnology, Delhi Technological University, Delhi 110042, India

N. Ramchiary and C. Kole (eds.), *The Capsicum Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-319-97217-6_10

Table 10.1 Plant noncoding RNA databases

Database	Description	Web link	Reference
miRBase	The biggest online registry for miRNAs	http://www.mirbase.org/	Kozomara and Griffiths (2014)
Rfam	Contains collection of RNA families arrangement of miRNA is different from miRBase	rfam-help@ebi.ac.uk	Kalvari et al. (2017)
PmiRKB	Four major functional modules are provided for plant miRNAs	http://bis.zju.edu.cn/pmirkb/index.php	Meng et al. (2011)
PMRD	A plant-specific miRNA annotation database	http://bioinformatics.cau.edu.cn/PMRD	Zhang et al. (2010)
miRTarBase	Provides experimentally verified miRNA–target interactions	http://mirtarbase.mbc.nctu.edu.tw/index.html	Chou et al. (2018)
MicroPC	A comprehensive resource for predicting and comparing plant miRNAs	http://www3a.biotec.or.th/micropc/index.html	Mhuantong and Wichadakul (2009)

(snoRNAs) are expressed constitutively, whereas regulatory ncRNAs, such as microRNAs (mRNA), small interfering RNAs (siRNA), long noncoding RNAs (lncRNAs), and Piwi RNAs (piRNAs) express according to the situation, i.e., specific tissues and environments (Carthew and Sontheimer 2009). Apart from RNAs with a specific function, such as rRNAs and tRNAs which helps cell to translate the mRNA into protein, regulatory ncRNAs are classified into three major groups based on their size and origin. Small noncoding RNAs (sncRNAs) are of 18–30 nucleotides (nt) long, medium noncoding RNAs (mncRNAs) are of 31–200 nt and long noncoding RNAs (lncRNAs) are more than 200 nt long (Ponting et al. 2009). These ncRNAs regulate many fundamental processes such as chromosome segregation, transcription, chromatin structure, RNA processing, RNA stability, and translation. The discovery of small nuclear RNAs (snRNA) in the 1980s which has a role in intron excision has changed the notion about ncRNAs'. Among those, the most widely studied noncoding RNAs with the help of high-throughput sequencing technology are miRNAs. Majority of the miRNAs have shown evolutionarily conserved nature among the different families of plant kingdom; therefore, the identification of orthologous genes in other species is possible (Din et al. 2016). More or less miRNAs are involved in all the biological processes such as growth, cell signaling, and biotic and abiotic stresses, emphasizing their potentiality in crop genetic modification and improvement.

After the discovery of microRNAs (miRNAs), studies leading to the discovery and understanding of many new classes of ncRNAs took a new momentum in biological science. As a result, today, noncoding RNAs are well studied in many model organisms, both in animals as well as in plants such as *Caenorhabditis elegans and Arabidopsis thaliana*, but very less information is available for non-model organisms. It is believed that most of the conserved miRNAs regulate genes involved in basic developmental processes while turning to the non-conserved miRNAs, they regulate genes attached to specific developmental processes (Omidvar et al. 2015). Recent studies showed that ncRNAs modulate a wide range of genes and gene regulatory networks in plants which are involved in many important physiological functions such as seed maturation, floral development, pathogen resistance, and biotic and abiotic stresses' resistance (Shin and Shin 2016). With the advent of new sequencing technologies and bioinformatics approaches, the identification and characterization of ncRNAs are progressing rapidly (Zhu and Wang 2012). Although few crop genomes of Solanaceae family, such as tomato and potato are considerably explored for ncRNAs identification and characterization, only few similar studies are available in *Capsicum species (*Bokszczanin et al. 2015; Hou et al. 2017). The identified noncoding RNAs in different plant species are listed in Table 10.1. However, information regarding *Capsicum* lncRNAs is still not updated.

10.2 Identification of Small Noncoding RNAs in *Capsicum* Species

Among the well-characterized small RNAs in plants, miRNAs, and siRNAs, both synthesized from double-stranded precursors are well studied (Park et al. 2002; Din et al. 2016). Argonaut proteins which help in forming silencing complexes for target suppression are common for both miRNAs and siRNAs. MiRNAs, which are short, endogenous, and nonprotein coding RNAs of 18–26 nucleotides (nt) in length are largely studied in crop plants (Park et al. 2002; Din et al. 2016). MiRNAs are transcribed from miRNA gene(s) by *RNA polymerase* II in the plant as primary miRNA (Pri-miRNA), and further Dicer-like protein 1 (*DCL1*) modifies these pri-miRNAs into stem-loop structure known as precursor miRNA (pre-miRNA) (Bartel 2004). After the DCLs action, HUA ENHANCER 1 (*HEN1*) transfers the methyl group onto the 20-OH on two 30 termini of small RNA duplexes, which increase the stability of small RNAs (Shin and Shin 2016). These pre-miRNAs are further processed to form miRNA duplex, and mature miRNAs are then transported to the cytoplasm by *Hasty* gene in plants (Park et al. 2002). Mature miRNA binds with argonaute protein and executes RNA-induced silencing complex (RISC) (Chen 2004; Carthew and Sontheimer 2009; Voinnet 2009; Chellappan et al. 2010). Till date, despite being very important crop plant, there is no information on *Capsicum* miRNAs in MirBase database.

The identification of microRNAs can be done in several ways: (i) by genetic screening, (ii) using already identified microRNAs from other plant species which are available in public domain and doing BLAST against expressed sequences tags (ESTs) or transcriptome data of species of interest and structure prediction computationally, (iii) small RNA sequencing, and by (iv) degradome library sequencing (Zhang et al. 2017). In 2014, Qin et al. (2014) identified a total of 176 *Capsicum* miRNAs belonging to 64 families using plant EST database. Further, they found that out of 176 miRNAs, 141 were conserved in Solanaceae family and only 34 miRNAs (i.e., 19%) were specific to *Capsicum*. Using 118,572 *Capsicum* ESTs extracted from dbEST (database of EST) and by in silico structure prediction, Din et al. (2016) identified a total of 88 miRNAs belonging to 81 miRNA families and their 204 target genes. Manila et al. (2009) using similar approach identified a total of 85 miRNAs from 33,311 ESTs derived from dbEST. Using high-throughput sequencing technology, Hwang et al. (2013) processed 542,938,815 reads obtained from multiple tissues harvested (Mexican pepper landrace, Criollo de Morelos 334, CM334) at different fruit developmental stages. After removing low-quality reads, they obtained a total of 15,250,415 conserved and 1,322,036 novel reads. This suggests that the novel miRNAs are expressing less in general than conserved miRNAs. Totally, they identified 128 miRNAs belonging to 29 different miRNA families.

Fruit is the economical part in *Capsicum*, so to understand the regulation of different fruit developmental stages of *Capsicum*, Liu et al. (2017) sequenced two hot pepper varieties "Luosijiao (with more vitamin content)" and "06J19-1-1-1-2" (less vitamin content). They have taken fruits at different stages such as 30, 40 and 50 days after anthesis (DAA). Libraries were prepared from extracted RNA and then sequenced using Illumina HiSeq 2500 platform. After removal of the low-quality data, only 83,639,907 clean high-quality sequence reads were blasted against the miRNA database for the identification of new putative miRNAs and they finally identified a total of 59 known miRNAs and 310 novel miRNAs. Recently, Seo et al. (2018) did degradome sequencing using Illumina HiSeq 2500 platform of various tissues such as leaves, roots, stems, green and red fruits of *Capsicum annuum* (334) to identify microRNA and their target genes. By sequencing of parallel analysis of RNA ends (PARE) libraries, they got a total of 251,689,162 reads. After removal of adapter sequences, structural RNAs, repeats, and transposons, the remaining 151,815,030 reads were used for miRNA identification. Using

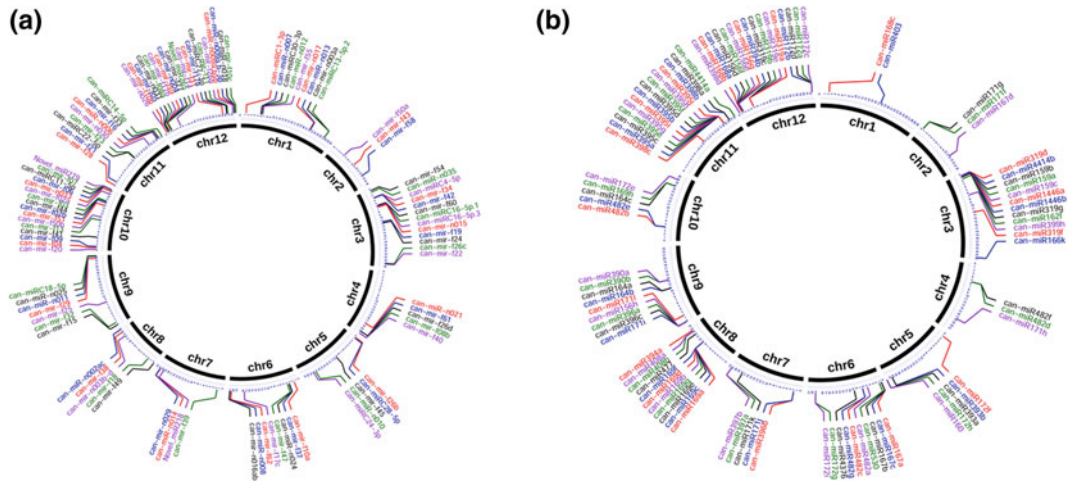


Fig. 10.1 Genomewide distribution of miRNAs in *Capsicum* genome, **a** distribution of novel miRNAs and **b** distribution of conserved miRNAs

CleaveLand4 pipeline (Brousse et al. 2014) 436 pairs of miRNAs/target genes were identified in *Capsicum* and out of these, 99 miRNAs were found to be targeting at least one target gene (Seo et al. 2018). Taller et al. (2018) sequenced small RNAs from seed, placenta, and fleshy fruit tissues using Illumina HiSeq 2000 platform and discovered a total of 193 non-redundant miRNA sequences of 38 conserved miRNA families and many miRNAs showed differential expression in those tissues of *C. annuum* suggesting specific role of miRNAs. They also noticed that the majority of miRNAs identified in their study were reported in miRBase and a fraction of these miRNAs were present in different plant of Solanaceae family, i.e., 143 miRNAs in *Solanum tuberosum*, 132 miRNAs in *Solanum lycopersicum*, and 110 miRNAs in *Nicotiana tabacum.* Distribution of conserved and novel miRNA in Capsicum shown in (Fig. 10.1).

10.3 Prediction of Target Genes for miRNAs Families in *Capsicum* Species

After identification of miRNAs, it is important to predict the miRNA: mRNA interaction. Many bioinformatics tools have been developed to predict miRNA to target genes. Riffo-Campos et al. (2016) listed most of the miRNA prediction tools such as, Diana tools, miRanda, and Pita. All these tools have their own advantage and disadvantage but one common problem with these tools is false-positive rate. Many available databases are being used for miRNA target prediction such as KOG (Clusters of Orthologous Groups of proteins), KEGG (Kyoto Encyclopaedia of Genes and Genomes), and GO (Gene Ontology). Din et al. (2016) used dual approaches for miRNA prediction; firstly, the newly identified miRNA was subjected to prediction tool psRNATarget (Dai and Zhao 2011). Secondary approach of "BLAST and RNA hybrid" was used for miRNAs that did not produce any potential targets. Using these tools, they annotated 409 target genes for *Capsicum* miRNAs and noticed that most of the targets were already identified in other plants. In this analysis, they observed that 31% of target genes were hypothetical proteins, 21% transcriptional factors, 24 target genes as transporters, 18 targets were involved in signal transduction, 17 were structural related proteins, 16 stress-related proteins {heat-shock protein, water stress-induced ER5 protein (ER5) etc.}, and 3% of newly identified miRNAs showed targeting the disease response genes such as NBS-LRR root-knot nematode resistance protein, blight resistance

protein *RGA1*, and TMV-induced protein 1–2. To explore miRNA targets by integrating miRNA sequencing, transcriptome sequencing, and degradome sequencing Zhang et al. (2017) applied a unique approach MiRTrans (Prediction of MiRNA targets by Trans-omics data). This approach comprised the following steps: the target transcript of miRNA can be predicted by scrutinizing their sequence which can be used as potential target pool and then the removal of false-positive targets. Degradome sequencing was utilized to capture miRNA targets. Qin et al. (2014) predicted a total of 1104 target genes for these miRNAs which were identified from ETS database and 78% of the target genes identified were found to have putative function. Additionally, they found that some of capsaicinoids biosynthesis pathway genes were potential targets of these *Capsicum* miRNAs, i.e., *dihydrolipoamide dehydrogenase* (Capana12g000245) was found targeted by can-miR5303 while α-CT (Capana09g001602) was found to be the potential target of few *Capsicum* miRNAs indicating probable role of these miRNAs in the regulation of capsaicinoids biosynthesis. By combining target prediction and GO enrichment analysis, Liu et al. (2017) predicted a total of 656 target genes for 310 novel miRNAs. Hwang et al. (2013) identified 334 potential target genes for 26 conserved miRNA families and most of the targets predicted for conserved miRNAs in pepper were transcription factors and they observed that unlike conserved miRNAs the targets of novel miRNAs were not enriched in transcription factors. Combined *in silco* study of ESTs from five different Solanaceae species including pepper (*C. annuum* L.), potato (*S. tuberosum* L.), tomato (*S. lycopersicum* L.), tobacco (*N. tabacum* L.), *Nicotiana benthamiana* helped out to identify 11 miRNAs and 54 target genes which belongs to at least nine different plant miRNA families: miRNA 156, 159, 166, 168, 172, 319, 395, 414, and 845 (Kim et al. 2011) .

10.4 Validation of miRNAs and Their Target Genes Expression in Capsicum Species

Various molecular techniques can be used for miRNA validation such as, northern blot analysis, polyacrylamide gel electrophoresis (PAGE), direct cloning of double-stranded RNA, fluorescence in situ hybridization with probes, quantitative polymerase chain reaction (qPCR) based methods such as stem-loop real-time PCR-based quantification of sncRNAs, and dual luciferase gene reporter assays to study miRNA function. Majority of miRNAs are found to be tissue specific or developmental stage specific, Hwang et al. (2013) used different tissues like leaf, stem, flower and fruit and root. Northern blot analysis, 5' RACE (rapid amplification of cDNA ends) and qRT-PCR were used to observe developmental stage-specific or tissue-specific expression patterns of miRNAs. In this study, a total of 19 miRNA families were tested on northern blot analysis, of which four families were not detected on blot. For further checking, Hwang et al. (2013) used carried out 5' RACE analysis for some targets like, *ARF*, *SBP*, *MYB*, and F-box. They got interesting results, such as *SBP* were targeted by two miRNAs, *can-miR156a-c* and *can-miR156d-g*, likewise in their study they validated other targets also *ARF*, *NAM*, *F-box*, *MYB*, *Argonaute*, *TCP*, and *sulfate transporter gene*. Liu et al. (2017) validated randomly selected four novel and four conserved miRNAs by qRT-PCR, they compared the data with fruit developmental genes and noticed that most of the miRNAs were negatively correlated to expression of their target genes, e.g., these three miRNAs (can-miR156a, canmiR160a, and can-miR396a-5p) were highly expressed in fruit at 50 DAA, while their target genes were downregulated. Similarly, nine randomly chosen miRNAs (can-mir-419,

can-mir-1512, can-mir-5032, can-mir-5503, can-mir-5504, can-mir-6024, can-mir-6434, can-mir-6443, and can-mir-8020a) were validated using RT-PCR in leaf tissue (Din et al. 2016). Taller et al. (2018) used DNA or LNA oligonucleotide probes for identification of some conserved miRNAs like miR171, miR172, miR396, miR159, and miR167 and they observed that the northern blot analysis result was similar to sequencing data with some differences. Conserved miRNAs of *Capsicum* species are listed in Table 10.2.

Table 10.2 Conserved miRNA and their target genes in *Capsicum spp*

MiRNA	Gene	Description	References
miRNA 156	Putative uncharacterized protein (A5BZT6); (A5BHV4), putative aconitase (B1Q486); Chr5 scaffold_2, whole-genome shotgun Sequence (A7NVZ2) SBP transcription factors	DNA binding Structural constituent of ribosome Regulate flowering time in plants	Kim et al. (2011) and Hwang et al. (2013)
miRNA 172	PHAP2A protein (Q9XHD4)	Transcription factor activity	Kim et al. (2011)
miRNA 414	Chr5 scaffold_72WGS (A7QAV5)	Structural constituent of ribosome	
can-miR160a	*ARF10, ARF17, ARF18*	Regulates auxin signal transduction during plant growth	Liu et al. (2017)
can-miR5079	Sorbitol transporter	Abiotic stressors responsive	Din et al. (2016)
can-miR1886	Nicotinate phosphoribosyl transferase-like protein		
can-miR838	*Phytoene synthase*	Transferase enzyme involved in the biosynthesis of carotenoids	
miR 395	Sulfate transporter, *ATP sulfurylases*	Regulate sulfur assimilation pathway	Hwang et al. (2013)
pap-miR15	*CUC3* gene	Multicellular organismal development, meristem initiation	Manila et al. (2009)
pap-miR25	*PTAC2* (*plastid transcriptionally active2*)	Positive regulation of transcription, transcription from plastid promoter and expressed in seed, embryo, cotyledon, hypocotyl, root, shoot	
pap-miR4	*CYP71B21*	Plasma membrane, vacuole, nucleoside transmembrane transporter activity, sugar: hydrogen symporter activity	
pap-miR8	PHB	Transcription regulation Determination of bilateral symmetry, adaxial/abaxial pattern formation	
pap-miR12	RPT5B	Glucose-mediated signaling and has ATPase activity	
miR159/319 families	MYB transcription factors, TCP transcriptional factor	Regulates meristem formation and seed development	Hwang et al. (2013)
Can-miR160 and can-miR167 families	ARF genes	Activators or repressors of auxin-responsive transcription (ARF) genes	
can-miR482 families	NBS-LRR disease-resistance proteins	Recognizing specific pathogen effectors and trigger resistance responses	
can-miR396	GRF, *DRM methyltransferase*	de novo methylation in all sequence contexts	
can-miR164a-b	NO APICAL MERISTEM		
miRNA-168	*AGO1* (Argonaute)	Small RNA biogenesis protein	
can-miR169a-g	Beta-mannosidase enzyme	Hydrolysis of terminal, non-reducing beta-D-mannose residues in beta-D-mannosides	
can-miR394	F-box	Interference. Vegetative and reproduction growth and development	
Can-miR319	TCP transcription factors	Cell proliferation, signaling pathways	
can-miR164	(NAM) No apical meristem	Regulates floral organ identity and lateral organ separation	

10.5 Identification of miRNAs Responsive to Biotic and Abiotic Stresses in *Capsicum* Species

Plants being sessile in nature experience a lot of biotic and abiotic stresses during their life cycle. Plants survive in these challenges by modulating their growth, development, and many physiological and biochemical changes, and several genes, proteins, and metabolites are shown to be involved in plant stress tolerance mechanism. Both biotic and abiotic stresses act as a primary limiting factor for plant growth and yield (Tuteja and Sopory 2008). MiRNAs act as the key regulators of many stress responsive genes, proteins, and transcriptional factor in *Capsicum* and in other plants. MiRNA acts upon stress responsive elements of genes thereby making either up or down-regulation of the target genes. Many of the miRNA families are considered to be conserved among the plant kingdom (Gong et al. 2013). A single miRNA could be involved in may stresses and many miRNAs in single stress. In environmental stresses such as high/low temperature, salinity, drought, and fertiliser deficiency miRNAs are shown to be involved in regulation of gene expressions (Wang et al. 2003; Zhang 2015).

Phosphorus (pi) is contemplated as the structural unit of nucleic acids, cellular membranes, and energy currency ATP (Guleria et al. 2012; Chou et al. 2018). MiR399 is responsible for the pi deprivation which has multiple target sites on the 5' untranslated regions of ubiquitin-conjugating E2 enzyme AtUBC24, this miRNA is also shown to be responsible for pi tolerance in *Capsicum* (Delhaize and Randall 1995; Schachtman et al. 1998; Sunkar and Zhu 2004). Sulfur is considered as a key component of many biochemical processes (Gonzalez-Ballester and Grossman 2009). Studies have shown that in sulfur-deficient condition in plants, miR395 was observed to accumulate and found to regulate the activities of low-affinity sulfate transporter (SULTR2;1) and ATP sulphurylases genes (*APS1*, *APS3* and *APS4*; Allen et al. 2004; Jones-Rhoades and Bartel 2004; Kawashima et al. 2009). In cold stress conditions miRNAs such as miR319 and miR167 are shown to upregulate to modulate cold responsive elements (Lv et al. 2010; Hwang et al. 2013). *In silico* identification and target prediction of miRNAs (Din et al. 2016) identified a total of 16 miRNAs which were potentially targeted the stress-related proteins such as, NBS/LRR resistance protein, acriflavin resistance protein, heat-shock protein 70 (HSP70-1), heat-shock protein 90, and water stress-induced ER5 protein (ER5). These all target proteins are extremely important for plants survival in both biotic and abiotic stress conditions. Some microRNA produces secondary small RNAs, such as phasiRNA which regulates large set of transcripts; most phasiRNAs are 22 nucleotides with some exception such as can-miR169a-g and can-miR171i, which are of 20 and 21 nucleotides in lengths, respectively. PhasiRNAs is likely to be associated with the basal defence response in pepper (Seo et al. 2018). These PHAS loci which are targeted by *Capsicum* specific miRNAs are associated with disease-resistance mechanism.

10.6 SiRNA-Mediated Gene Silencing

RNA interference (RNAi) was discovered in nematode *C. elegans (*Fire et al. 1998*)*. RNA interference (RNAi) involves the production of double-stranded RNA (dsRNA). Dicer enzymes convert/cleaves this dsRNAs into 21 nt long RNA known as small interfering RNAs (siRNAs, Mello and Conte 2004). These siRNAs then provide specificity to endonuclease-containing, RNA-induced silencing complex (RISC) which then targets homologous RNA for degradation. RNAi required Dicer-like protein 4 (*DCL4*) for siRNA production, some studies have shown that *DCL2* is also indirectly helps in RNAi mechanism. Although it is reported that siRNAs mainly involve in post-transcriptional gene modification/repression, it is also shown that it executes transcriptional gene silencing (Lippman and Martienssen 2004). RNAi is a powerful tool which is widely used for identification of gene

function by manipulating genes (Brodersen and Voinnet 2006). Abundantly available class of siRNAs (24 nt) are largely produced from transposons and repeat sequence elements (Kasschau et al. 2007; Taller et al. 2018). These siRNAs play an important role in developmental gene and transposon regulation, heterochromatin formation, genome stability and are also considered to have a crucial role in transgenerational epigenetic inheritance (Mosher et al. 2009; Castel and Martienssen 2013). After the accidental of discovery of loss of petunia flower pigmentation, RNAi mechanism became a novel strategy for gene silencing (Jorgensen 1990). There are many types of siRNA such as transacting miRNA (ta-si), repeat-associated siRNAs (ra-si), viral siRNAs (vi-si), natural antisense RNAs (nat-si) (Allen and Howell 2010). Furthermore, use of siRNA-mediated silencing was also reported in plants for generating resistance against viral disease. It was noted in *Capsicum* that TOM1 and TOM3 genes help in multiplication of Tomato Mosaic Virus (TMV) and the disease causes heavy loss (1–90%) to the Capsicum production (Chitra et al. 2002; Kumar et al. 2012). These genes were silenced in Capsicum using siRNA, which provided an efficient way to generate resistance in plants against virus using RNAi approach. Another instance of RNAi-based approach showed significantly reduced infection of chilli-infecting begomoviruses (CIBs) suggested more efficient use of small noncoding RNA molecules in providing resistance against disease.

10.7 Identification of Long Noncoding RNAs (LncRNAs) in *Capsicum* Species

Long noncoding RNAs (lncRNAs) are transcribed independently and show resemblance to mRNAs but do not encode for functional proteins. LncRNAs are categorized into intergenic, intronic, and antisense lncRNAs based on their positions with respect to coding genes (Ulitsky and Bartel 2013; Di et al. 2014) and are known to play important role biological functions, such as cellular differentiation, development, and biotic and abiotic stress (Shuai et al. 2014; Chekanova 2015; Cruz de Carvalho et al. 2015). To date, multiple studies were reported to characterize lncRNAs in animals but only few were reported for plants. Functioning of lncRNAs are subdivided into two major classes, i.e. either (1) cis or (2) trans based on the sequence complementarity or homology with respect to RNAs or DNAs which they control at the transcriptional and post-transcriptional level (Chekanova 2015). Studies show that endogeneous long noncoding RNA (lncRNAs) like Induced by Phosphate Starvation1 (IPS1) binds to miRNA *ath-miR399* (Franco-Zorrilla et al. 2007) and affects its cleavage functioning via blocking its potential target. Thus, lncRNAs (IPS1) serve as a potential decoy against miRNAs (ath-miR399) interfering its binding against other targets (Wu et al. 2013). So far a large number of lncRNAs have been identified in various plant species but their functioning and regulation in plants are still limited (Liu et al. 2012; Li et al. 2014; Wang et al. 2015; Kang and Liu 2015). Recently, LncNAT *COOLAIR* and intronic *COLDAIR* lncRNAs transcribed from the *Flowering locus C* in arabidopsis were found to play an important role in vernalization indicating its regulation to a specific developmental process in plants (Swiezewski et al. 2009; Sun et al. 2013; Wang et al. 2014). Furthermore, various lncRNA genes such as *ENOD40* in soybean and medicago (Crespi et al. 1994; Yang et al. 2018), *OsPI1* and *LDMAR* in rice (Wasaki et al. 2003; Ding et al. 2012), *TPS11* in tomato (Liu et al. 1997), *AtIPS1*, *COOLAIR,* and *COLDAIR* in arabidopsis (Franco-Zorrilla et al. 2007; Swiezewski et al. 2009; Heo and Sung 2011) were reported to play significant role in regulation of diverse biological processes (nodule formation, phosphate uptake, flowering time, and photo-sensitive male sterility).

Qin et al (2014) sequenced RNAs from flower buds of *C. annuum* and produced a total of 2,717,180 unique sequence tags from which they reported the identification of a total 6527 lncRNAs. Of these, 5976 were intergenic and 222 were intron spanning and the rest were

bidirectional lnRNAs. Similarly, lncRNAs in two *Capsicum* varieties, i.e., Luosijiao and 06J19-1-1-1-2 were also identified. Fruits at different developmental stages were used. The mature green fruits at 30 days after anthesis (DAA) were obtained from both the varieties while mature green fruit as well as ripened fruits at 40 and 50 DAA were collected only from Luosijiao variety for transcriptome sequencing. Around 1.90×10^8 clean reads were obtained. The clean reads aligned against *C. annuum* reference genome showed mapping of more than 76% of reads. The lncRNAs were identified using common lncRNAs identification pipeline utilizing various tools like Tophat (Trapnell et al. 2012), Cufflinks 2.0, cuffmerge, CPC-Coding Potential Calculator and CNCI—Coding Non-Coding Index (Kong et al. 2007; Sun et al. 2013). Finally, 2505 lncRNAs were identified out of which 1066 lncRNAs showed differential expression across different fruit developmental stages (Ou et al. 2017). Further these lnRNAs were classified into three groups based on fruit developmental stages in Capsicum. Moreover, they observed the differential expression pattern of lncRNAs specific to individual fruit development stage in these two phenotypically different varieties, thereby suggesting their association in fruit quality variations of hot pepper (Ou et al. 2017).

Recently, another study reported the identification lncRNAs expressing in fruit pericarp tissues of *Capsicum* green bell peppers (*C. annuum* L. cv. Jingtian) with different chilling temperature, i.e., 1 °C (treated) and 10 °C (used as control sample) by transcriptome sequencing (Zuo et al. 2018). The analysis with standard lncRNA identification pipeline could identify a total of 9848 lncRNAs. Of these, 84 lncRNAs were found to be similar to already reported lncRNAs in plants, and the rest 9764 were novel lncRNAs expressed in pericarp tissues of chilling treated samples of *Capsicum* fruits. Furthermore, they observed that 380 lncRNAs were significantly differentially expressed in chilling treated (1 °C) pericarp tissues compared to control (10 °C) samples (Zuo et al. 2018) suggesting potential role lncRNAs in abiotic stresses. From their study and other reported results earlier they have concluded that the identified *Capsicum* lncRNAs might play a vital role in stimulus against abiotic stress, signal transduction, hormone-mediated

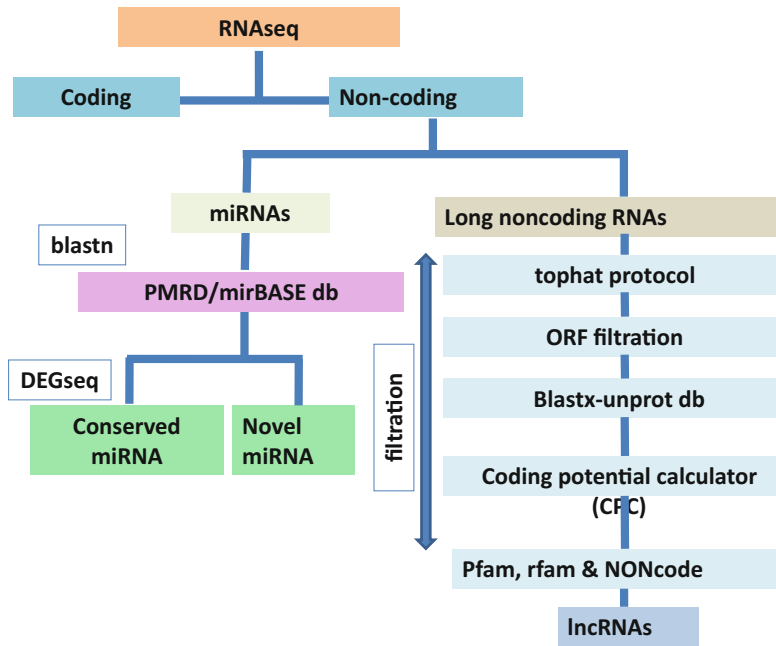


Fig. 10.2 Computational pipeline for noncoding RNAs prediction (miRNAs and LncRNAs) using different software, such as blastN, DEGSEQ-R statistical package.

signaling pathway, glyoxylate and dicarboxylate metabolism; and carbon metabolism (Zuo et al. 2018). Furthermore, using available transcriptome data from National Centre for Biotechnology Information (NCBI) gene expression omnibus database, in the present study, we identified lnCRNAs. We used raw reads with accession number GSE45037 containing transcriptomes from five different tissues, i.e., leaf, flower, early, breaker, and mature fruit of *C. annuum*. We have processed ~116 million raw reads using standard pipeline given in Fig. 10.2 for detection of lncRNAs. Finally, a total of 35,139 lncRNAs were predicted across all stages of *C. annuum*, of which 784 lncRNAs were found to be differentially expressed in five tissues. Continuous research on exploring various biological functions influenced directly or indirectly by lncRNAs in *Capsicum* is still undergoing.

10.8 Study Highlighting Solanaceae (Capsicum) LncRNAs

Currently, understanding the mode of action of lncRNAs and its underlying function in plant species is still at infancy stage limiting to only a handful of lncRNAs. The identification of lncRNAs in different plant species and comparative analysis opens a new challenge to study the rapid evolution of lncRNAs in different plant species. The recent development and availability of high-throughput next-generation sequencing technologies coupled with computational approaches have been enabling researchers to identify lncRNAs which are being deposited in public databases. Therefore, taking this advantage in this study we have taken lncRNAs of model plant arabidopsis (total 4371 lncRNAs), and other Solanaceae plant species like tomato (4716) and potato (5790) from CANTATAdb database (Szcześniak et al. 2016). We used only 784 lncRNAs of *C. annuum* which showed differential expression along with the above-mentioned lncRNAs from different species for homology search. The BLASTn was used with selective parameter (e-value < 0.001) to find sequence homology between lncRNAs across these species. We observed that only six lncRNAs were conserved between *Capsicum* and tomato and nine were conserved between *Capsicum* and potato, while only two lncRNAs were conserved among these three Solanaceae plants. This may be due to the taking of less number of lncRNAs from *C. annuum*, we got very less number of lncRNAs conservation among those species. Also, overall a total of 925 lncRNAs were found to be conserved between tomato and potato. Apart from this, a total of 769, 3777, 4849, and 7583 lncRNAs were reported as unique to *C. annuum,* tomato, potato, and Arabidopsis, respectively, as shown in Fig. 10.3.

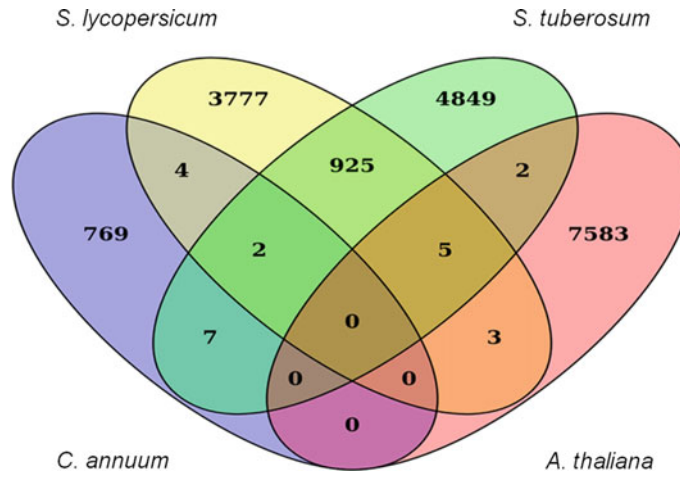


Fig. 10.3 Venn diagram showing comparative analysis of lncRNAs in four plant species: *Solanum lycopersicum, Solanum tuberosum, Capsicum annuum,* and *Arabidopsis thaliana*

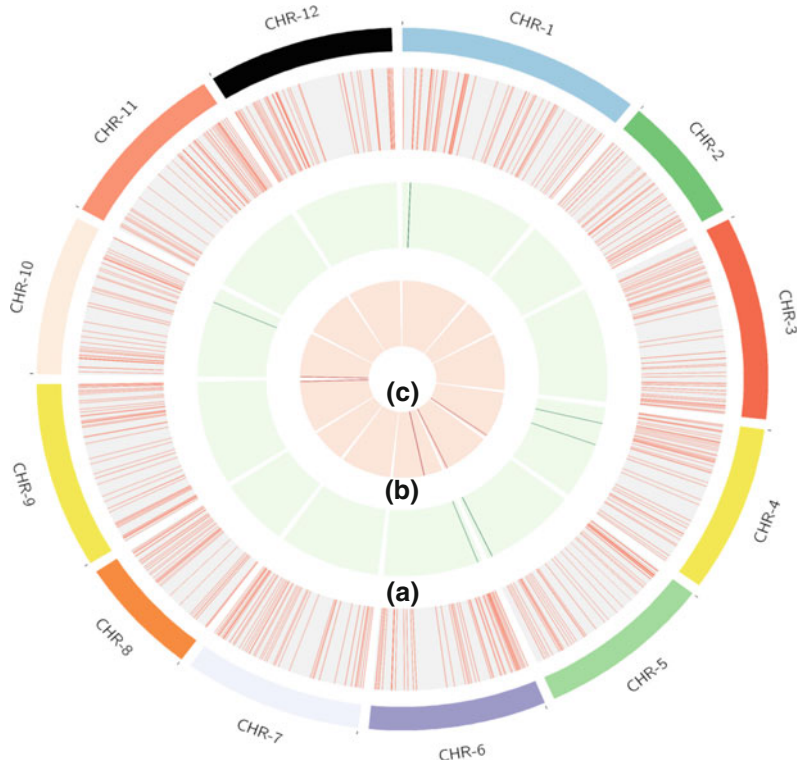


Fig. 10.4 Distribution of *Capsicum annuum* lncRNAs (784 lncRNAs, A) on 12 chromosomes of Capsicum along with homologous lncRNAs from *Solanum tuberosum* (9nine lncRNAs, B) and *Solanum lycopersicum* (six lncRNAs, C)

Further, conserved lncRNAs across these plant species (except Arabidopsis) were mapped against *C. annuum* genome and represented their distribution across 12 chromosomes as shown in Fig. 10.4.

10.9 Conclusion and Future Perspective

Till now, only a few studies identified limited number of regulatory noncoding RNAs in *Capsicum* mostly through high-throughput sequencing, and a substantial part of the noncoding RNAs are yet to be explored. Furthermore, the detail functional characterization of noncoding RNAs through overexpression and/or silencing is required. Further study will give insights into several novel ncRNAs with a new biological function that are yet unexplored in *Capsicum*. Once functional role of specific noncoding RNA is identified, researchers can manipulate economically important traits which will be immensely helpful in precision breeding of Capsicum.

Acknowledgements This work was supported under Special Assistance Programme (SAP) by the University Grant Commission (UGC), India, to Jawaharlal Nehru University, New Delhi, India.

References

Allen E, Howell MD (2010) miRNAs in the biogenesis of trans-acting siRNAs in higher plants. Semin Cell Dev Biol 21:798–804. https://doi.org/10.1016/j.semcdb.2010.03.008

Allen E, Xie Z, Gustafson AM et al (2004) Evolution of microRNA genes by inverted duplication of target gene sequences in *Arabidopsis thaliana*. Nat Genet 36:1282

Backofen R, Bernhart SH, Flamm C et al (2006) RNAs everywhere: genome-wide annotation of structured RNAs. J Exp Zool Part B Mol Dev Evol 308B:1–25. https://doi.org/10.1002/jez.b.21130

Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116(2):281–297

Bokszczanin KL, Krezdorn N, Fragkostefanakis S et al (2015) Identification of novel small ncRNAs in pollen of tomato. BMC Genom 16:714. https://doi.org/10.1186/s12864-015-1901-x

Brodersen P, Voinnet O (2006) The diversity of RNA silencing pathways in plants. Trends Genet 22:268–280. https://doi.org/10.1016/j.tig.2006.03.003

Brousse C, Liu Q, Beauclair L et al (2014) A non-canonical plant microRNA target site. Nucleic Acids Res 42:5270–5279

Carthew RW, Sontheimer EJ (2009) Origins and mechanisms of miRNAs and siRNAs. Cell 136:642–655. https://doi.org/10.1016/j.cell.2009.01.035

Castel SE, Martienssen RA (2013) RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. Nat Rev Genet 14:100

Chekanova JA (2015) Long non-coding RNAs and their functions in plants. Curr Opin Plant Biol 27:207–216. https://doi.org/10.1016/j.pbi.2015.08.003

Chellappan P, Xia J, Zhou X et al (2010) siRNAs from miRNA sites mediate DNA methylation of target genes. Nucleic Acids Res 38:6883–6894. https://doi.org/10.1093/nar/gkq590

Chen X (2004) A microRNA as a translational repressor of APETALA2 in Arabidopsis flower development. Science 303:2022–2025. https://doi.org/10.1126/science.1088060

Chitra TR, Prakash HS, Albrechtsen SE et al (2002) Indexing of leaf and seed samples of tomato and bell pepper for tobamoviruses. Indian Phytopathol 55:84–86

Chou C-H, Shrestha S, Yang C-D et al (2018) miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. Nucleic Acids Res 46:D296–D302. https://doi.org/10.1093/nar/gkx1067

Crespi MD, Jurkevitch E, Poiret M et al (1994) enod40, a gene expressed during nodule organogenesis, codes for a non-translatable RNA involved in plant growth. EMBO J 13:5099–5112

Cruz de Carvalho MH, Sun H-X, Bowler C, Chua N-H (2015) Noncoding and coding transcriptome responses of a marine diatom to phosphate fluctuations. New Phytol 210:497–510. https://doi.org/10.1111/nph.13787

Dai X, Zhao PX (2011) psRNATarget: a plant small RNA target analysis server. Nucleic Acids Res 39:W155–W159. https://doi.org/10.1093/nar/gkr319

Delhaize E, Randall PJ (1995) Characterization of a phosphate-accumulator mutant of *Arabidopsis thaliana*. Plant Physiol 107:207–213

Delpu Y, Larrieu D, Gayral M, et al (2016) Noncoding RNAs: clinical and therapeutic applications (Chapter 12). In: Egger G (ed) Arimondo PBT-DD in CE. Academic Press, Boston, pp 305–326

Di C, Yuan J, Wu Y et al (2014) Characterization of stress-responsive lncRNAs in *Arabidopsis thaliana* by integrating expression, epigenetic and structural features. Plant J 80:848–861. https://doi.org/10.1111/tpj.12679

Din M, Barozai MYK, Baloch IA (2016) Profiling and annotation of microRNAs and their putative target genes in chilli (*Capsicum annuum* L.) using ESTs. Gene Reports 5:62–69. https://doi.org/10.1016/j.genrep.2016.08.010

Ding J, Lu Q, Ouyang Y et al (2012) A long noncoding RNA regulates photoperiod-sensitive male sterility, an essential component of hybrid rice. Proc Natl Acad Sci USA 109:2654–2659. https://doi.org/10.1073/pnas.1121374109

Fire A, Xu S, Montgomery MK et al (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature 391:806

Franco-Zorrilla JM, Valli A, Todesco M et al (2007) Target mimicry provides a new mechanism for regulation of microRNA activity. Nat Genet 39:1033

Gong L, Kakrana A, Arikit S et al (2013) Composition and expression of conserved microRNA genes in diploid cotton (Gossypium) species. Genome Biol Evol 5:2449–2459. https://doi.org/10.1093/gbe/evt196

Gonzalez-Ballester D, Grossman AR (2009) Chapter 5—Sulfur: from acquisition to assimilation. In: Harris EH, Stern DB, Witman GBBT-TCS (eds) The chlamydomonas sourcebook, 2nd edn. Academic Press, London, pp 159–187

Guleria P, Goswami D, Mahajan M et al (2012) MicroRNAs and their role in plants during abiotic stresses BT. In: Ahmad P, Prasad MNV (eds) Environmental adaptations and stress tolerance of plants in the era of climate change. Springer, New York, pp 265–278

Heo JB, Sung S (2011) Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. Science 331(80):76–79

Herbig A, Nieselt K (2011) nocoRNAc: characterization of non-coding RNAs in prokaryotes. BMC Bioinform 12:40. https://doi.org/10.1186/1471-2105-12-40

Hou X, Du Y, Liu X et al (2017) Genome-wide analysis of long non-coding RNAs in potato and their potential role in tuber sprouting process. Int J Mol Sci 19:101. https://doi.org/10.3390/ijms19010101

Hwang D-G, Park JH, Lim JY et al (2013) The hot pepper (*Capsicum annuum*) microRNA transcriptome reveals novel and conserved targets: a foundation for understanding microRNA functional roles in hot pepper. PLoS ONE 8:e64238–e64238. https://doi.org/10.1371/journal.pone.0064238

Jones-Rhoades MW, Bartel DP (2004) Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. Mol Cell 14:787–799. https://doi.org/10.1016/j.molcel.2004.05.027

Jorgensen R (1990) Altered gene expression in plants due totrans interactions between homologous genes. Trends Biotechnol 8:340–344. https://doi.org/10.1016/0167-7799(90)90220-R

Kalvari I, Argasinska J, Olvera NQ, Nawrocki EP, Rivas E, Eddy ES, Bateman A, Finn RD, Petrov AI (2017) Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. Nucleic Acids Res. https://doi.org/10.1093/nar/gkx1038

Kang C, Liu Z (2015) Global identification and analysis of long non-coding RNAs in diploid strawberry *Fragaria vesca* during flower and fruit development. BMC Genom 16:815. https://doi.org/10.1186/s12864-015-2014-2

Kasschau KD, Fahlgren N, Chapman EJ et al (2007) Genome-wide profiling and analysis of Arabidopsis siRNAs. PLoS Biol 5:e57–e57. https://doi.org/10.1371/journal.pbio.0050057

Kawashima CG, Yoshimoto N, Maruyama-Nakashita A et al (2009) Sulphur starvation induces the expression of microRNA-395 and one of its target genes but in different cell types. Plant J 57:313–321. https://doi.org/10.1111/j.1365-313X.2008.03690.x

Kim H-J, Kwang-Hyun Baek, Bong-Woo Lee et al (2011) In silico identification and characterization of microRNAs and their putative target genes in Solanaceae plants. Genome 54:91–98. https://doi.org/10.1139/G10-104

Kong L, Zhang Y, Ye Z-Q et al (2007) CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. Nucleic Acids Res 35:W345–W349. https://doi.org/10.1093/nar/gkm391

Kozomara A, Griffiths-Jones S (2014) miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res 42(1):D68–D73

Kumar S, Dubey AK, Karmakar R et al (2012) Inhibition of TMV multiplication by siRNA constructs against TOM1 and TOM3 genes of Capsicum annuum. J Virol Methods 186:78–85. https://doi.org/10.1016/j.jviromet.2012.07.014

Li L, Eichten SR, Shimizu R et al (2014) Genome-wide discovery and characterization of maize long non-coding RNAs. Genome Biol 15:R40. https://doi.org/10.1186/gb-2014-15-2-r40

Lippman Z, Martienssen R (2004) The role of RNA interference in heterochromatic silencing. Nature 431:364

Liu C, Muchhal US, Raghothama KG (1997) Differential expression of TPS11, a phosphate starvation-induced gene in tomato. Plant Mol Biol 33:867–874. https://doi.org/10.1023/A:1005729309569

Liu J, Jung C, Xu J et al (2012) Genome-wide analysis uncovers regulation of long intergenic noncoding RNAs in Arabidopsis. Plant Cell 24:4333–4345. https://doi.org/10.1105/tpc.112.102855

Liu Z, Zhang Y, Ou L et al (2017) Identification and characterization of novel microRNAs for fruit development and quality in hot pepper (*Capsicum annuum* L.). Gene 608:66–72. https://doi.org/10.1016/j.gene.2017.01.020

Lv D-K, Bai X, Li Y et al (2010) Profiling of cold-stress-responsive miRNAs in rice by microarrays. Gene 459:39–47. https://doi.org/10.1016/j.gene.2010.03.011

Manila T M, Riju A, Lakshmi Priya Darshini, K, Chandrasekar A. Eapen SJ (2009) In silico microRNA identification from paprika (*Capsicum annuum*) ESTs. Available from Nature Precedings. http://hdl.handle.net/10101/npre.2009.3737.1

Mello CC, Conte D (2004) Revealing the world of RNA interference. Nature 31:338–342

Meng Y, Gou L, Chen D, Mao C, Jin Y, Wu P, Chen M (2011) PmiRKB: a plant microRNA knowledge base. Nucleic Acids Res. https://doi.org/10.1093/nar/gkq721

Mhuantong W, Wichadakul D (2009) MicroPC (μPC): comprehensive resource for predicting and comparing plant microRNAs, BMC Genom 10:366

Mosher RA, Melnyk CW, Kelly KA et al (2009) Uniparental expression of PolIV-dependent siRNAs in developing endosperm of Arabidopsis. Nature 460:283

Omidvar V, Mohorianu I, Dalmay T, Fellner M (2015) MicroRNA regulation of abiotic stress response in male-sterile tomato mutant. Plant Genome 8:13. https://doi.org/10.3835/plantgenome2015.02.0008

Ou L, Liu Z, Zhang Z et al (2017) Noncoding and coding transcriptome analysis reveals the regulation roles of long noncoding RNAs in fruit development of hot pepper (*Capsicum annuum* L.). Plant Growth Regul 83:141–156. https://doi.org/10.1007/s10725-017-0290-3

Park W, Li J, Song R et al (2002) Carpel factory, a dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*. Curr Biol 12:1484–1495

Ponting CP, Oliver PL, Reik W (2009) Evolution and functions of long noncoding RNAs. Cell 136:629–641. https://doi.org/10.1016/j.cell.2009.02.006

Qin C, Yu C, Shen Y, et al (2014) Whole-genome sequencing of cultivated and wild peppers provides insights into Capsicum domestication and specialization. Proc Natl Acad Sci 111:5135–5140

Riffo-Campos ÁL, Riquelme I, Brebi-Mieville P (2016) Tools for sequence-based miRNA target prediction: what to choose? Int J Mol Sci 17:1987. https://doi.org/10.3390/ijms17121987

Schachtman DP, Reid RJ, Ayling SM (1998) Phosphorus uptake by plants: from soil to cell. Plant Physiol 116:447–453

Seo E, Kim T, Park JH, et al (2018) Genome-wide comparative analysis in Solanaceous species reveals evolution of microRNAs targeting defense genes in Capsicum spp. DNA Res dsy025–dsy025

Shin S-Y, Shin C (2016) Regulatory non-coding RNAs in plants: potential gene resources for the improvement of agricultural traits. Plant Biotechnol Rep 10:35–47. https://doi.org/10.1007/s11816-016-0389-4

Shuai P, Liang D, Tang S et al (2014) Genome-wide identification and functional prediction of novel and drought-responsive lincRNAs in *Populus trichocarpa*. J Exp Bot 65:4975–4983. https://doi.org/10.1093/jxb/eru256

Sun L, Luo H, Bu D et al (2013) Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. Nucleic Acids Res 41:e166–e166. https://doi.org/10.1093/nar/gkt646

Sunkar R, Zhu J-K (2004) Novel and stress-regulated microRNAs and other small RNAs from Arabidopsis. Plant Cell 16:2001–2019

Swiezewski S, Liu F, Magusin A, Dean C (2009) Cold-induced silencing by long antisense transcripts of an Arabidopsis polycomb target. Nature 462:799

Szcześniak MW, Rosikiewicz W, Makałowska I (2016) CANTATAdb: a collection of plant long non-coding RNAs. Plant Cell Physiol 57:e8–e8. https://doi.org/10.1093/pcp/pcv201

Taller D, Bálint J, Gyula P et al (2018) Expansion of Capsicum annum fruit is linked to dynamic tissue-specific differential expression of miRNA and siRNA profiles. PLoS ONE 13:e0200207

Trapnell C, Roberts A, Goff L et al (2012) Differential gene and transcript expression analysis of RNA-seq experiments with tophat and cufflinks. Nat Protoc 7:562

Tuteja N, Sopory SK (2008) Chemical signaling under abiotic stress environment in plants. Plant Signal Behav 3:525–536

Ulitsky I, Bartel DP (2013) lincRNAs: genomics, evolution, and mechanisms. Cell 154:26–46. https://doi.org/10.1016/j.cell.2013.06.020

Voinnet O (2009) Origin, biogenesis, and activity of plant microRNAs. Cell 136:669–687. https://doi.org/10.1016/j.cell.2009.01.046

Wang M, Yuan D, Tu L et al (2015) Long noncoding RNAs and their proposed functions in fibre development of cotton (*Gossypium* spp.). New Phytol 207:1181–1197. https://doi.org/10.1111/nph.13429

Wang W, Vinocur B, Altman A (2003) Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. Planta 218:1–14. https://doi.org/10.1007/s00425-003-1105-5

Wang Z-W, Wu Z, Raitskin O et al (2014) Antisense-mediated FLC transcriptional repression requires the P-TEFb transcription elongation factor. Proc Natl Acad Sci USA 111:7468–7473. https://doi.org/10.1073/pnas.1406635111

Wasaki J, Yonetani R, Shinano T, et al (2003) Expression of the OsPI1 gene, cloned from rice roots using cDNA microarray, rapidly responds to phosphorus status. New Phytol 158:239–248. https://doi.org/10.1046/j.1469-8137.2003.00748.x

Wu H-J, Wang Z-M, Wang M, Wang X-J (2013) Widespread long noncoding RNAs as endogenous target mimics for microRNAs in plants. Plant Physiol 161:1875–1884. https://doi.org/10.1104/pp.113.215962

Yang W-C, Katinakis P, Hendriks P et al (2018) Characterization of GmENOD40, a gene showing novel patterns of cell-specific expression during soybean nodule development. Plant J 3:573–585. https://doi.org/10.1046/j.1365-313X.1993.03040573.x

Zhang B (2015) MicroRNA: a new target for improving plant tolerance to abiotic stress. J Exp Bot 66:1749–1761. https://doi.org/10.1093/jxb/erv013

Zhang L, Qin C, Mei J, et al (2017) Identification of microRNA targets of *Capsicum* spp. using mirtrans—a trans-omics approach. Front Plant Sci 8:495. https://doi.org/10.3389/fpls.2017.00495

Zhu Q-H, Wang M-B (2012) Molecular functions of long non-coding RNAs in plants. Genes (Basel) 3:176–190. https://doi.org/10.3390/genes3010176

Zuo J, Wang Y, Zhu B et al (2018) Analysis of the coding and non-coding RNA transcriptomes in response to bell pepper chilling. Int J Mol Sci 19:2001. https://doi.org/10.3390/ijms19072001

Zhang Z, Yu J, Li D, Zhang Z, Liu F, Zhou X, Wang T, Ling Y, Su Z (2010) PMRD: plant microRNA database. Nucleic Acids Res 38:D806–D813

Epigenome Landscape in *Capsicum* Genome

11

Abdul Rawoof, Ilyas Ahmad and Nirala Ramchiary

Abstract

At DNA level, methylation is one of the most studied epigenetic marks and plays an important role in plant growth and development via regulating gene expression, integrity, and mobility of genome as well as transposons. The epigenetic studies especially the DNA methylation have been investigated only in a few members of the Solanaceae family like tomato and potato. So far, cytosine methylation landscape in *Capsicum*, a diploid, self-pollinating crop of the Solanaceae family grown worldwide for fresh and processed products, is far less documented. In our research study in the laboratory, we found the overall high cytosine methylation in *Capsicum* fruit as compared to other plants. The *Capsicum* fruit shows at an average 89.1% of CG, 84.85% of CHG, and 24.9% CHH cytosine methylation globally. The variation in genome size reflects the variations in the global cytosine methylation across different species. The *Capsicum* genome which is 3–4-fold larger than that of tomato and potato is found to have ~1.2–2.7-fold higher cytosine methylation at all methylation contexts. The abundance of repetitive elements (REs) generally affects the variations in genome size across species and generally has dense cytosine methylation. The intraspecific variations in cytosine methylation as well as the miRNA-regulated methylation are unexplored in *Capsicum*, which could provide plausible evolutionary relationship between different species of Solanaceae family. DNA methylation is considered as one of the requisites for various developmental and transcriptional gene expression regulation, while it is also important for reprogramming of various biological processes and transcriptional gene regulation by trimming down their methylation profiles. Therefore, the collaborative role of methylation and demethylation phenomenon in DNA results in the global dynamic nature of cytosine methylation.

11.1 DNA Methylation: An Overview in Plants

Heritable alterations which are not due to variations in underlying DNA sequences include DNA methylation and histone modification resulting in alteration of chromatin structure and DNA accessibility, thereby ultimately modulating expression of several genes, and are termed as epigenetic

A. Rawoof · I. Ahmad · N. Ramchiary (✉)
Translational and Evolutionary Genomics Lab, School of Life Sciences, Jawaharlal Nehru University, New Delhi 110067, India
e-mail: nrudsc@gmail.com

N. Ramchiary
Department of Biotechnology, Delhi Technological University, Delhi 110042, India

N. Ramchiary and C. Kole (eds.), *The Capsicum Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-319-97217-6_11

modifications (Bonasio et al. 2010). DNA methylation, the most studied epigenetic mark, includes the formation of 5-methylcytosine via inclusion of methyl group to the 5th carbon of sugar residue of nucleotide bases and plays an vital role through regulation of different aspects of development in plants via influencing the gene expression, integrity, and versatility of genome and transposons, respectively (Richards and Elgin 2002). Cytosine methylation in animals generally occurs at CG context representing methylated cytosines (mCs) at both strand making it symmetrical, while plants have mCs at both symmetric (CG and CHG) and asymmetric (CHH) context, where H is considered as other than guanine (Law and Jacobsen 2010; Matzke and Mosher 2014). The cytosine methylation is mainly maintained by three well-distinguished DNA methyltransferases classes: (i) the *methyltransferase 1* (*MET1*), (ii) *chromomethylases* (*CMT*) *methyltransferases*, and (iii) domain rearranged methyltransferases (DRM) family containing *DRM1* and *DRM2*. The first two classes maintain symmetric methylation, whereas the last class (*DRM1* and *DRM2*) maintains asymmetric methylation context via siRNA-facilitated RNA-directed DNA methylation (RdDM) pathway. Moreover, RdDM pathway can also maintain cytosine methylation at all sequence contexts by de novo approach (Aufsatz et al. 2004; Law and Jacobsen 2010). These DNA methyltransferases ubiquitously showed a collaborative methylation activity at all methylation contexts in *RPS* locus of *Petunia hybrida* (Singh et al. 2008). The level and pattern of cytosine methylation in all contexts varies across plants due to size variation at genomic level. The overall cytosine methylation in *Arabidopsis thaliana* genome was observed in all contexts that is 22.3% for CG, 5.92% for CHG, and 1.51% for CHH, while in rice genome with 3-fold larger genome size has 59, 21 and 2.2% mCs of CG, CHG, and CHH, respectively (Feng et al. 2010). Furthermore, in *Zea mays* genome cytosine methylation was found to be 86% for CG, 74% for CHG, and 5.4% for CHH context, respectively (Gent et al. 2013). Likewise, our finding in *Capsicum annuum* suggests that cytosine at an average of 89.1% of CG, 84.85% of CHG, and 24.9% of CHH context was methylated. Further, levels of DNA methylation have been shown tissue-specific characteristic throughout the developmental phase (Gehring and Henikoff 2007). In *Solanum lycopersicum*, methylation level in mature tissues like leaf, fruit, and seed was higher than immature stem, leaves, protoplasts, and roots (Messeguer et al. 1991; Teyssier et al. 2008).

In general, methylation in promoter region is directly related to the gene silencing. For instance, a natural epigenetic mutation or epiallele with the hypermethylated promoter in the tomato *colorless non-ripening* (*Cnr*) gene was responsible for gene repression. This has encouraged researchers to study more on controlling fruit ripening via targeted DNA methylation (Manning et al. 2006; Zhong et al. 2013). Methylation is one of the requisite for regulation of various developmental stages; however, a process involving removal of mCs and replacing it back with original cytosine is also equally important for rescheduling of many biological processes, known as demethylation (Zhang and Zhu 2012). The demethylation of mCs can be induced by *DEMETER* (DME), a DNA glycosylase (Choi et al. 2002); *DEMETER-LIKE 2* (*DML2*); *DML3* (Penterman et al. 2007); and *REPRESSOR OF SILENCING 1* (*ROS1*) gene (Gong et al. 2002) which replace the mCs with non-methylated cytosine. In *Capsicum annuum* L., the cytosine methylation in germinating seed has been reported using methylation-sensitive amplified polymorphism (MSAP) marker and observed that demethylation of mCs is an important factor for transcriptional gene activation during seed germination (Portis et al. 2004). In tomato, active DNA demethylation was suggested to supervise fruit ripening via tomato *SlDML2*; thereby suggesting importance of demethylation in regulation of ripening-specific and ripening-restrained genes (Liu et al. 2015; Lang et al. 2017). In an another instance, in tomato *rin* (ripening inhibitor) mutant line, promoter hypermethylation was observed at *RIN*-binding sites in fruit-ripening genes like in *Polygalacturonase* (*PG*), suggesting the role of *SlDML2*-mediated demethylation in phase transition during fruit-ripening process (Zhong et al.

2013; Lang et al. 2017). Thus collaborative activities between replication, maintenance of methylation and demethylation instances at DNA level, results the global dynamic nature of methylation at different context.

11.2 Global Cytosine Methylation in *Capsicum*

Capsicum ($2n = 24$), a member of the Solanaceae family, is one of the most important crops grown for spices and vegetables worldwide. Plants like tomato, potato, tobacco, eggplant, and *Nicotiana tabacum* are closely related to *Capsicum*. *Capsicum* fruits are rich source of various alkaloids, pigments, vitamins and nutrients and most commonly used as spices (Aza-González et al. 2011). Wide variations are observed for fruit morphology (size, shape, and color) and biochemical contents in *Capsicum* fruits. The transcriptome study of *Capsicum* fruits observed differential gene expression and improved our understanding of capsaicin biosynthetic pathway (Liu et al. 2006; Kim et al. 2014). Importantly, it was reported that the epigenetic modification, including cytosine methylation, regulates the expression of diverse genes during fruit development (Gallusci et al. 2016). So far, epigenetic studies, especially DNA methylation in *Capsicum* species, are far less documented. Stable epigenetic marks are maintained and inherited by *MET1* and *CMT3* at both strands of daughter DNA making it symmetrical, while asymmetric methylation as name suggests only occurs at either strand of daughter DNA and maintained de novo throughout the each cell cycle (Zhang and Zhu 2012). Alteration in the gene expression or genomic instability could be correlated with dynamics of global cytosine methylation at genome level, and often plant genomes are found to be densely methylated. Furthermore, in *C. annuum* L., ~15–16% increased level of global cytosine methylation phenomenon was observed in water-deficit as well as in drought-affected plants treated with 200 mM H_2O_2 compared to normal plants (Rodríguez-Calzada et al. 2017).

Till now, variations in global cytosine methylation across different species are under-explored which could potentially reflect the evolutionary correlation between the species. To elucidate the global fruit methylome, we have performed whole-genome bisulfite sequencing (WGBS) in fruits of *C. annuum*. Fruit samples at different developmental stage (immature, breaker, and mature stages) were pooled together to get overall fruit methylome. Overall, 215,876,691 bisulfite sequencing reads were aligned to *C. annuum* reference genome (GCF). Almost 92.8% (200,317,410 reads) of total reads were aligned to reference genome out of that 88.7% (167,642,734 reads) of total aligned reads were uniquely mapped to the reference genome. Further, the status of methylated cytosine (mCs) was identified at CG, CHG, and CHH contexts. A sum of 5,143,414,121 cytosine was analyzed, of which 408,009,366, 624,328,406, 1,016,141,807 cytosines were found to be methylated at CG, CHG, and CHH contexts, respectively. In *C. annuum,* CHH (49.6%) methylation context has shown higher proportion of mCs followed by CHG (30.5%) and CG (19.9%) contexts which is approximately similar to tomato (Zhong et al. 2013). In *Capsicum*, the global cytosine methylation is the highest in all the methylation contexts (Fig. 11.1 and Table 11.1) compared to tomato (Zhong et al. 2013), potato (Wang et al. 2018), maize (Wang et al. 2018), rice, arabidopsis (Feng et al. 2010), soybean (An et al. 2017), and *Brassica rapa* (Chen et al. 2015). Further, it was found that plants show substantial variation at genome level due to the high abundance of REs which generally have highly dense regions of mCs regions in the genome (Rabinowicz et al. 2003; Fedoroff 2012); therefore, it was suggested that perhaps the genome size is related to global methylation level during the course of evolution (Alonso et al. 2015). This could be seen in view of *Capsicum* genome which has ~22-fold larger genome compared to *A. thaliana* and has shown significantly higher differences in global mCs at all contexts. As the genome size variation decreases, the variations in cytosine methylation

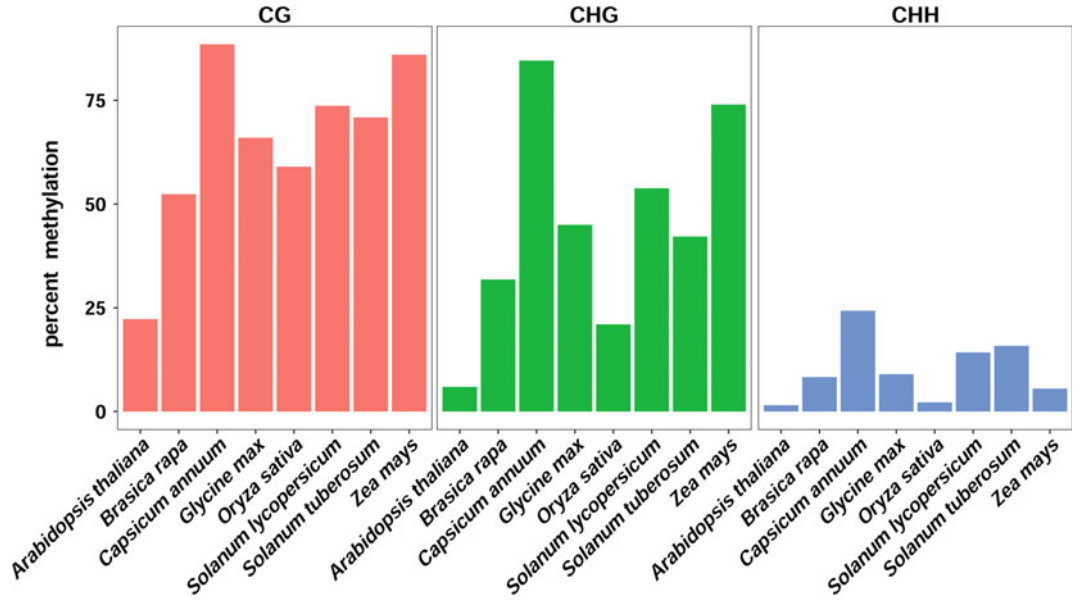


Fig. 11.1 Global cytosine methylation in *Capsicum annuum* compared to different plant species

Table 11.1 Global methylation level across different species along with their genome information at all cytosine methylation contexts

Species	Genome size	CG (%)	CHG (%)	CHH (%)
Capsicum annuum	2.8–3 gigabases	88.5	84.6	24.3
Arabidopsis thaliana	~135 Mbp	22.3	5.92	1.51
Brassica rapa	~485 Mbp	52.4	31.8	8.3
Oryza sativa	~500 Mbp	59	21	2.2
Solanum lycopersicum	~950 Mbp	73.7	53.82	14.26
Solanum tuberosum	~840 Mbp	70.9	42.19	15.84
Glycine max	1115 Mbp	66	45	9
Zea mays	2.4 gigabases	86	74	5.5

Table 11.2 Correlation of cytosine methylation variations at all methylation contexts compared to genome size variation with reference to *Capsicum annuum* genome

	Capsicum genome size (fold larger)	Increase in CG methylation in *Capsicum* (in fold)	Increase in CHG methylation in *Capsicum* (in fold)	Increase in CHH methylation in *Capsicum* (in fold)
Arabidopsis thaliana	~22	3.95	14.19	15.89
Brassica rapa, *Oryza sativa*	~6–7	1.5–1.7	2.7–4	2.8–10
Solanum lycopersicum, *Solanum tuberosum*, *Glycine max*	~3–4	1.2–1.4	1.6–1.8	1.6–2.7
Zea mays	~1.4	1.02	1.14	4.36

The genome size of *Capsicum annuum* is compared to genomes of different species, and increment in cytosine methylation (in fold) at all contexts is compared to methylation in different species

also decrease. *Capsicum* genome which is ~1.4-fold larger than the maize genome has shown less variation at cytosine methylation at CG and CHG contexts, while the cytosine methylation in CHH context has significantly higher variations (~4.36) compared to maize, suggesting its potential role in the *Capsicum* fruit development. Compared to *S. lycopersicum*, *S. tuberosum,* and *Glycine max*, the *Capsicum* genome is 3–4-fold larger and has 1.2–2.7-fold cytosine methylation variations across all contexts (Table 11.2).

11.3 Genebody Methylation Distribution

The term genebody methylation is self-evident, which indicates toward the enrichment of mCs within the transcribed regions of protein-coding genes itself and was first narrated in *A. Thaliana* (Tran et al. 2005). Cytosine methylation in intragenic region is mainly considered as genebody methylation, primarily occurring at euchromatic regions with high number of methylated CpG sites (Lister et al. 2008; Feng et al. 2010; Regulski et al. 2013). Additionally, context of cytosine methylation, methylation density, and methylation location within the gene could shed light on essential information on the enzymatic pathways and their functional consequences responsible for controlled regulation of methylation instances (Takuno and Gaut 2012, 2013). Methylation at genebody level is mainly characterized by enrichment of CG cytosine methylation (mCG) confined to the transcribed region along with reduction in cytosine methylation at transcriptional start site (TSS) and transcriptional termination site (TTS; Bewick and Schmitz 2017). The overall genebody methylation in *Capsicum* is higher as compared to genebody methylation in tomato and potato (Wang et al. 2018) of same family, and it was observed that gene body has shown similar pattern of mCs level at CG and CHH contexts to that in tomato and potato, but the overall average cytosine methylation was found to be highest in *Capsicum* than both tomato and potato. Another interesting observation is that in *Capsicum* genebody regions in mCHG context, after the TSS and before the TTS, were highly differentially methylated, while in tomato as well as in potato the regions to the genebody vicinity, i.e., prior to TSS and following to TTS, were mainly observed to have high level of differential mCs. This suggests that higher genebody methylation at CHG context is potentially responsible for the maintenance of genebody methylation in *Capsicum* species.

Furthermore, an overall high average cytosine methylation at genebody level across all contexts might be indicative of transcriptional repression of REs or activation in *Capsicum* (Fig. 11.2a–c). Further, it was noted that genebody methylation genes frequently come under the category of housekeeping genes with dense cytosine methylation and very often the pattern and type of cytosine methylation in gene body reflects their expression status. Further, genebody methylation is found to be common feature among interspecific transcriptionally active orthologous genes, indicating functional conservation (Feng et al. 2010; Takuno and Gaut 2013; Bewick et al. 2016). It was noted that the gene which is developmentally regulated or whose expression is regulated at specific developmental stage predominantly lacks the CpG genebody methylation (Coleman-Derr and Zilberman 2012). Further, chromosomal distribution of average methylation across all *Capsicum* chromosomes was mentioned in Fig. 11.2d where CG and CHG methylation is comparatively higher than CHH methylation at all chromosomes. Also, it was observed that chromosomes 1, 3–6, 8, and 12 preferentially have less methylation at their both ends compared to rest of the regions (Fig. 11.2d). Moreover, it was hypothesized that the genebody methylation has positive correlation with gene expression and may potentially regulate the alternative splicing by precisely improving the intron–exon definition (Maunakea et al. 2010; Regulski et al. 2013). Notwithstanding to their wider presence in the genome, the biological functions of genebody methylation largely remain unclear. It was also suggested that highly dense mCs at genebody region can silence the

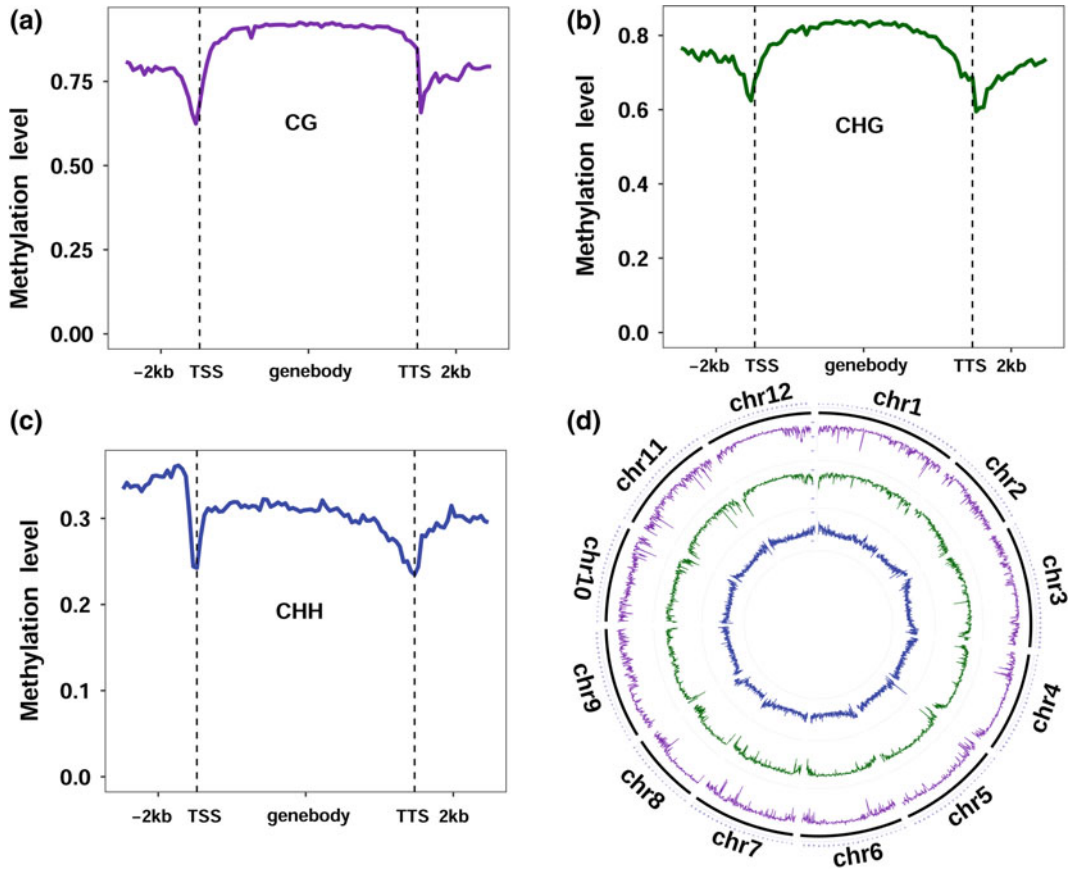


Fig. 11.2 Average cytosine methylation at CG, CHG, and CHH contexts in 2 kb upstream promoter region before TSS site, gene body region, and 2 kb downstream region after TTS site (**a**–**c**); circos representing average methylation distribution at all 12 *Capsicum* chromosomes. From outside to inside, the first lane represents chromosomes with their length information, 2–4th lane represents average methylation at CG, CHG, and CHH contexts, respectively (**d**)

repetitive DNA elements occurring within the gene body (Yoder et al. 1997). Moreover, methylation at genebody level is mainly maintained by the CHG methyltransferases of the chromomethylase gene family (Bewick and Schmitz 2017).

11.4 Promoter Methylation

In plants, cytosine methylation in promoter region is found to play a crucial role in managing the diverse developmental process by controlling the genes via repressing their expression. Unlike the mCs at gene body in CG context which shows positive regulation with transcribing genes, the promoter with dense mCs generally negatively controlled the expression of transcribing genes (Zemach et al. 2010; Wang et al. 2015). The CpG islands (CGIs) in promoter regions are generally unmethylated, thereby facilitating smooth binding between promoter region and proteins, and in arabidopsis, most of the endogenous genes were observed with less frequency of mCs in their promoter regions (Zhang 2008). Hitherto, several studies have been reported promoter methylation in its association with transcriptional gene silencing (Bewick et al. 2016; O'Malley et al. 2016; Lang et al. 2017). Till date, there is no direct study reported on dynamic of cytosine methylation in

Capsicum fruit and its development. We are working on categorizing the global cytosine methylation in *Capsicum* at different genomic levels. Our study suggested that compared to genebody methylation, the upstream 2 kb promoter region is less methylated in CG and CHH contexts, while in CHH methylation context, it is preferentially highly methylated than gene body, suggesting potential role of CHH cytosine methylation in regulation of gene expression as compared to CG and CHG contexts (Fig. 11.2a–c). Recently, it was found that promoter methylation is also responsible for regulation of different transition phases from early fruit to ripen fruit during its development and ripening. Manning et al. (2006) identified an epiallele in colorless non-ripening (*Cnr*) genes of tomato representing natural epigenetic mutation and whose hypermethylated promoter causes gene repression. In an another instance, promoter hypermethylation in *PcMYB10* gene which is responsible for anthocyanin accumulation in pear fruit skin drew attention toward plausible role of methylation in regulation of different aspects of development and ripening process in fruits (Wang et al. 2013).

11.5 Methylation of Transposable Elements

The mobile genetic elements aka transposable elements (TEs) are present in almost every genome and generally considered as 'parasitic' or 'selfish' elements. Mostly plant genome is enriched with long terminal repeat (LTR) retrotransposons and miniature inverted transposable elements (MITEs) among the diverse type of TEs (Casacuberta and Santiago 2003). TEs are integral part of constitutive heterochromatin and play a significant role in genome expansion and genome evolution (Slotkin and Martienssen 2007; Vicient and Casacuberta 2017). Due to larger in size, the LTR retrotransposons are predominant in almost all plant genomes, and in *Capsicum*, excess of single type of LTR retrotransposons could shed light on the genome expansion.

The *Capsicum* genome lacks whole-genome duplication, and ~81% of its genome comprises various transposable elements, while ~61% of tomato as well as potato (Park et al. 2012) genome were composed of TEs. Both LTR retrotransposons (70.3%) and DNA transposons (4.5%) were most abundant among the all plant TEs categories in *Capsicum* and LTR retrotransposons accorded more to genome expansion as compared to tomato (50.3%) and potato (47.2%; Park et al. 2012; Qin et al. 2014). The level of mCs acts as a key epigenetic signal which could repress the activation and transcription of TEs, controls the gene expression, and thereby can impact on the phenotypic variations (Zakrzewski et al. 2017). Moreover, the cytosine methylation pattern in transposable element is similar in *Capsicum*, tomato, and potato at both CG and CHG contexts which potentially provides mechanism of TE silencing and its inheritance. But the overall cytosine methylation at transposable element is highest than both tomato and potato, suggesting that the TEs were preferentially methylated in *Capsicum* genome. In contrast to CG and CHG methylation, CHH methylation at transposable element showed slight opposite pattern as those in tomato and potato and overall CHH methylation higher than both tomato and potato, suggesting potential maintenance of CHH methylation in transposable elements by de novo during the course of genome expansion (Fig. 11.3). Generally, active TEs target transcribing genes for potential insertion and can cause chromosomal breakage, genome rearrangement as well as illicit recombination. Like promoter methylation, cytosine methylation in TEs can suppress the expression of neighboring genes through posing as enhancers or promoters (Girard and Freeling 1999). Further, it was hypothesized that methylation of TEs is negatively correlated with diminishing expression of neighboring genes. Afterward, in *A. thaliana*, it was found that high amplitude of methylated TEs along with the abundance of TEs can reduce the expression of their neighboring genes which is independent of its chromosomal location (Hollister and Gaut 2009).

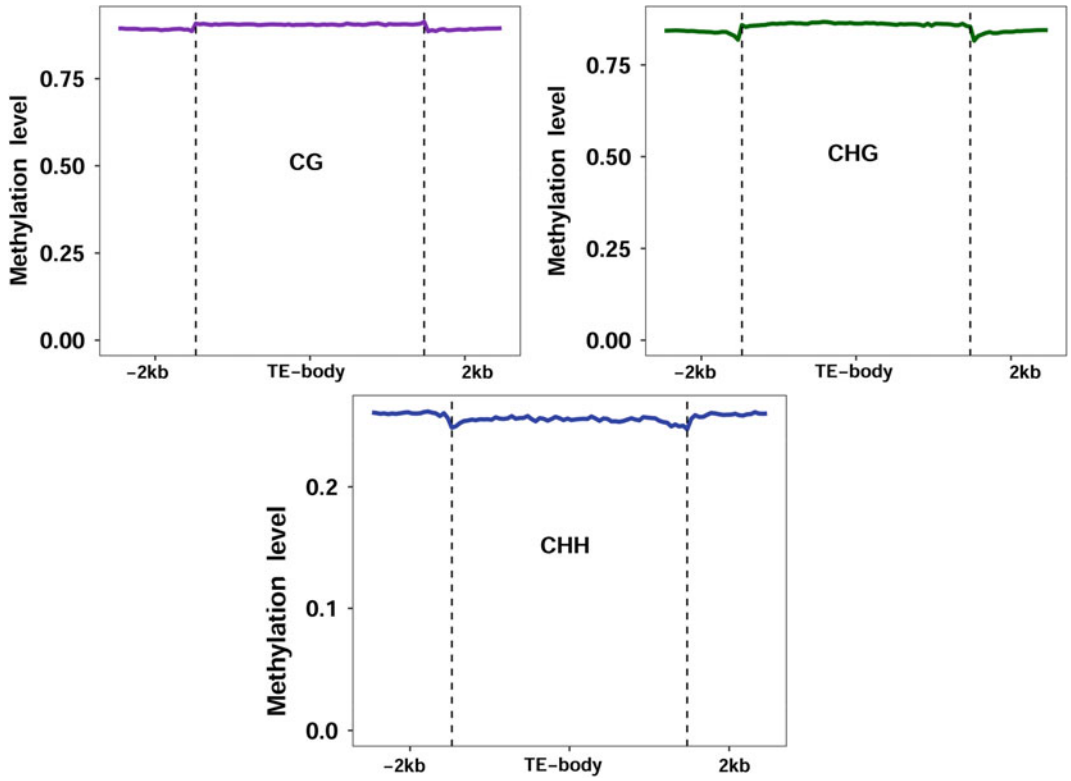


Fig. 11.3 DNA methylation patterns in *Capsicum annuum* across transposable elements (TEs). The metaplot shows cytosine methylation depicted in 2 kb vicinity of TEs at all three cytosine methylation contexts

11.6 Epigenetics of Cytosine Methylation in Hybrids

In plants, the status of cytosine methylation is easier to detect, which provided a wider scope in regulation of developmental process and tissue-specific expression of genes. Hybrid vigor or heterosis is an incident where hybrids from genetically different individuals show better traits in terms of enhanced phenotypic and functional features relative to their parents (Cheng et al. 2007). Till date, dynamics of mCs in hybrids and in their parent lines were reported substantially in *Capsicum* (Xu et al. 2015), arabidopsis (Kawanabe et al. 2016; Lauss et al. 2017), potato (Sanetomo and Hosaka 2011), *Zea mays* (Zhao et al. 2007; Lauria et al. 2014), rice (Xiong et al. 1999; Dong et al. 2006; Takamiya et al. 2008), and in sorghum (Yi et al. 2005; Zhang et al. 2007). These studies suggested notable variations in the level of mCs and their pattern in heterosis compared to parent plants (Zhao et al. 2008). The yield and quality of *Capsicum* has been improved through implementation of heterosis in *Capsicum* breeding, but the molecular and genetic bases of higher level of performance of heterosis relative to its parents remain elusive. Xu et al. (2015) analyzed the reciprocal hybrids with the help of MSAP from two genotypes of *hot pepper* having purple and green cotyledon and observed increased mCs level in hybrids. The overall observed DNA methylation in F1 hybrids of D85 × D34 and D34 × d85 was 67 and 64.36%, respectively, which was higher than mid-parent value (64.83%). Furthermore, in addition to the overall methylation, dynamic pattern of DNA methylation also varies during heterosis (Xu et al. 2015).

The term 'graft hybrid' defined as genetically distinguishable plants which are produced through asexual combination of different plant species. The grafting has been considerably used to improve the production of crops. Hitherto, studies based on grafting revealed the interchange of nucleic acid molecules across the plants used as grafting partners, thereby indicating toward the molecular basis of genetic variations facilitated by grafting (Wu et al. 2013). Furthermore, reports in *A. thaliana* concluded that graft hybrid shows epigenetic variations at mCs level induced due to grafting process compared to normal seed plants (Molnar et al. 2010), in tomato, eggplant, and pepper of Solanaceae family (Wu et al. 2013) and *Cucurbitaceae* family (Avramidou et al. 2015). Furthermore, after reciprocal interspecies grafting, considerable variations at mCs were detected in grafted Solanaceae species at genome wide level using MSAP, while significantly altered global mCs level in tomato, eggplant, and pepper scion was observed at both CG and CHG contexts. Moreover, self-pollinated progeny of graft hybrid was observed to inherit the variations of mCs, suggesting potentiality of grafting to introduce stable epigenetic variations transferable to the progeny (Wu et al. 2013). Further, in *C. annuum* grafting results in fruit shape and size variations, and graft hybrid indicated toward the potential role of siRNA-mediated epigenetic regulation of genes responsible for maintenance of fruit shape (Tsaballa et al. 2013).

11.7 MiRNA-Mediated Methylation in *Capsicum*

As a conserved epigenetic mechanism, DNA methylation mainly regulates gene expression by epigenetic silencing of transcription. Both small and long ncRNA (lncRNA) are involved in epigenetic regulation of cytosine methylation and maintenance using RdDM pathway. Most of the instances of DNA methylation occurring through RdDM pathway are triggered by siRNAs and are involved in de novo maintenance of mCs at different contexts. Though in plants several studies have documented the DNA methylation events directed by miRNAs in plants, their regulation mechanism is not yet fully elucidated (Jia et al. 2011). Small RNAs (sRNAs) behave as indispensable triggers to regulate cytosine methylation at all mCs context, thereby regulating the transcriptional gene networks in most of the eukaryotes (Zilberman et al. 2003; Onodera et al. 2005; Teotia et al. 2017). In plant, primary miRNAs are transcribed by pol-II enzyme and are further cleaved into pre-miRNAs by dicer-like 1 (DCL1). RdDM is a major methyltransferase enzyme involved in maintenance and regulation of methylation phenomenon in plants. Hwang et al. (2013) identified miRNA-directed cleavage of *Capsicum* DRM methyltransferase which regulates and maintains cytosine methylation through de novo. In *Capsicum*, microRNA *Ca-mir-396* family regulates transcriptional silencing of REs and is responsible for de novo maintenance of mCs through targeting methyltransferases (Hwang et al. 2013). In spite of several research progresses in miRNA-mediated methylation regulation, still there is a need to focus on their mechanism for thorough understanding in plants, and especially in *Capsicum*. A lot of insides are yet to be explored to determine the more specific role of miRNA and other non-coding RNA-mediated DNA methylation for *Capsicum*.

11.8 Conclusion and Future Prospective

The epigenetic variations are much overlooked in most of the plant-breeding program dependent on DNA-based molecular markers. With the emerging evidence, so far the epigenetic landscape in *Capsicum* is under-explored. This work of DNA methylation profiling of fruit development in *Capsicum* could provide some insight about the overall epigenetic modification during fruit transition from unripe to ripe in *Capsicum* species. However, many more studies using different developmental stages of fruits separately and from contrasting genotypes will shed more light. Furthermore, recently developed

high-throughput methylome or histone sequencing using high-throughput sequencing technologies will tremendously help to study epigenetics mechanism of fruit development in *Capsicum* species. Therefore, plant engineering equipped with epigenetic variations might be informative in developing improved crop varieties with economically important traits. Till now, the discoveries of various mutants to demonstrate the epigenetic regulation in fruit development, such as *cnr* mutant, *rin* mutant, and *sldml2* mutant, have been performed on tomato fleshy fruit model which could facilitate the better understanding of controlled fruit ripening in *Capsicum* as well. Such information could help in improving the fruit quality and fruit harvesting for longer period. In the case of fruit development from unripe to ripe and quality of fruit, the differentially methylated regions kindred with various genes responsible for fruit ripening and fruit repressed ripening could manifest the targets for analysis of epigenetic differences across the fruit varieties. Thus, the assessment of epigenetic variation at different fruit developmental stages may help in improving and expanding the selection strategies, thereby helping in improving fruits traits like shelf life and quality across the agronomically important crops.

Acknowledgements This work was supported by the financial grant from the University Grants Commission (UGC), India, in the form of University of Potential For Excellence II grant to Jawaharlal Nehru University, New Delhi, India.

References

Alonso C, Perez R, Bazaga P, Herrera CM (2015) Global DNA cytosine methylation as an evolving trait: phylogenetic signal and correlated evolution with genome size in angiosperms. Front Genet 5:1–9. https://doi.org/10.3389/fgene.2015.00004

An YC, Goettel W, Han Q et al (2017) Dynamic changes of genome-wide DNA methylation during soybean seed development. Sci Rep 7:12263. https://doi.org/10.1038/s41598-017-12510-4

Aufsatz W, Mette MF, Matzke AJM, Matzke M (2004) The role of MET1 in RNA-directed de novo and maintenance methylation of CG dinucleotides. Plant Mol Biol 54:793–804. https://doi.org/10.1007/s11103-004-0179-1

Avramidou E, Kapazoglou A, Aravanopoulos FA et al (2015) Global DNA methylation changes in *Cucurbitaceae* inter-species grafting. Crop Breed Appl Biotechnol 15:112–116. https://doi.org/10.1590/1984-70332015v15n2n20

Aza-González C, Núñez-Palenius HG, Ochoa-Alejo N (2011) Molecular biology of capsaicinoid biosynthesis in chili pepper (*Capsicum* spp.). Plant Cell Rep 30:695–706. https://doi.org/10.1007/s00299-010-0968-8

Bewick AJ, Schmitz RJ (2017) Gene body DNA methylation in plants. Curr Opin Plant Biol 36:103–110. https://doi.org/10.1016/j.pbi.2016.12.007

Bewick AJ, Ji L, Niederhuth CE et al (2016) On the origin and evolutionary consequences of gene body DNA methylation. Proc Natl Acad Sci USA 113:9111–9116. https://doi.org/10.1073/pnas.1604666113

Bonasio R, Tu S, Reinberg D (2010) Molecular signals of epigenetic states. Science 330:612–616

Casacuberta JM, Santiago N (2003) Plant LTR-retrotransposons and MITEs: control of transposition and impact on the evolution of plant genes and genomes. Gene 311:1–11. https://doi.org/10.1016/S0378-1119(03)00557-2

Chen X, Ge X, Wang J et al (2015) Genome-wide DNA methylation profiling by modified reduced representation bisulfite sequencing in *Brassica rapa* suggests that epigenetic modifications play a key role in polyploid genome evolution. Front Plant Sci 6:1–12. https://doi.org/10.3389/fpls.2015.00836

Cheng S-H, Zhuang J-Y, Fan Y-Y et al (2007) Progress in research and development on hybrid rice: a super-domesticate in China. Ann Bot 100:959–966. https://doi.org/10.1093/aob/mcm121

Choi Y, Gehring M, Johnson L et al (2002) DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in arabidopsis. Cell 110:33–42. https://doi.org/10.1016/S0092-8674(02)00807-3

Coleman-Derr D, Zilberman D (2012) Deposition of histone variant H2A.Z within gene bodies regulates responsive genes. PLoS Genet 8. https://doi.org/10.1371/journal.pgen.1002988

Dong ZY, Wang YM, Zhang ZJ et al (2006) Extent and pattern of DNA methylation alteration in rice lines derived from introgressive hybridization of rice and *Zizania latifolia* Griseb. Theor Appl Genet 113:196–205. https://doi.org/10.1007/s00122-006-0286-2

Fedoroff NV (2012) Transposable elements, epigenetics, and genome evolution. Science 338:758–767. Available at: http://science.sciencemag.org/content/338/6108/758.abstract

Feng S, Cokus SJ, Zhang X et al (2010) Conservation and divergence of methylation patterning in plants and animals. Proc Natl Acad Sci USA 107:8689–8694. https://doi.org/10.1073/PNAS.1002720107

Gallusci P, Hodgman C, Teyssier E, Seymour GB (2016) DNA methylation and chromatin regulation during

fleshy fruit development and ripening. Front Plant Sci 7:807. https://doi.org/10.3389/fpls.2016.00807
Gehring M, Henikoff S (2007) DNA methylation dynamics in plant genomes. Biochim Biophys Acta—Gene Struct Expr 1769:276–286. https://doi.org/10.1016/j.bbaexp.2007.01.009
Gent JI, Ellis NA, Guo L et al (2013) CHH islands: de novo DNA methylation in near-gene chromatin regulation in maize. Genome Res 23:628–637. https://doi.org/10.1101/gr.146985.112
Girard L, Freeling M (1999) Regulatory changes as a consequence of transposon insertion. Dev Genet 25:291–296. https://doi.org/10.1002/(sici)1520-6408(1999)25:4%3c291::aid-dvg2%3e3.0.co;2-5
Gong Z, Morales-Ruiz T, Ariza RR et al (2002) ROS1, a repressor of transcriptional gene silencing in arabidopsis, encodes a DNA glycosylase/lyase. Cell 111:803–814. https://doi.org/10.1016/S0092-8674(02)01133-9
Hollister JD, Gaut BS (2009) Epigenetic silencing of transposable elements: a trade-off between reduced transposition and deleterious effects on neighboring gene expression. Genome Res 19:1419–1428. https://doi.org/10.1101/gr.091678.109
Hwang DG, Park JH, Lim JY et al (2013) The hot pepper (*Capsicum annuum*) microRNA transcriptome reveals novel and conserved targets: a foundation for understanding microRNA functional roles in hot pepper. PLoS One 8. https://doi.org/10.1371/journal.pone.0064238
Jia X, Yan J, Tang G (2011) MicroRNA-mediated DNA methylation in plants. Front Biol (Beijing) 6:133–139. https://doi.org/10.1007/s11515-011-1136-4
Kawanabe T, Ishikura S, Miyaji N et al (2016) Role of DNA methylation in hybrid vigor in *Arabidopsis thaliana*. Proc Natl Acad Sci USA 113:E6704–E6711. https://doi.org/10.1073/pnas.1613372113
Kim S, Park M, Yeom S-I et al (2014) Genome sequence of the hot pepper provides insights into the evolution of pungency in *Capsicum* species. Nat Genet 46:270
Lang Z, Wang Y, Tang K et al (2017) Critical roles of DNA demethylation in the activation of ripening-induced genes and inhibition of ripening-repressed genes in tomato fruit. Proc Natl Acad Sci USA 114:E4511–E4519. https://doi.org/10.1073/pnas.1705233114
Lauria M, Piccinini S, Pirona R et al (2014) Epigenetic variation, inheritance, and parent-of-origin effects of cytosine methylation in maize (*Zea mays*). Genetics 196:653–666. https://doi.org/10.1534/genetics.113.160515
Lauss K, Wardenaar R, Oka R et al (2017) Parental DNA methylation states are associated with heterosis in epigenetic hybrids. Plant Physiol 176:1627–1645. https://doi.org/10.1104/pp.17.01054
Law JA, Jacobsen SE (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat Rev Genet 11:204–220
Lister R, O'Malley RC, Tonti-Filippini J et al (2008) Highly integrated single-base resolution maps of the epigenome in arabidopsis. Cell 133. https://doi.org/10.1016/j.cell.2008.03.029
Liu K, Jiang H, Moore SL et al (2006) Isolation and characterization of a lipid transfer protein expressed in ripening fruit of *Capsicum chinense*. Planta 223:672–683. https://doi.org/10.1007/s00425-005-0120-0
Liu R, How-Kit A, Stammitti L et al (2015) A DEMETER-like DNA demethylase governs tomato fruit ripening. Proc Natl Acad Sci USA 112:10804–10809. https://doi.org/10.1073/pnas.1503362112
Manning K, Tör M, Poole M et al (2006) A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. Nat Genet 38:948–952. https://doi.org/10.1038/ng1841
Matzke MA, Mosher RA (2014) RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. Nat Rev Genet 15:394
Maunakea AK, Nagarajan RP, Bilenky M et al (2010) Conserved role of intragenic DNA methylation in regulating alternative promoters. Nature 466:253–257. https://doi.org/10.1038/nature09165
Messeguer R, Ganal MW, Steffens JC, Tanksley SD (1991) Characterization of the level, target sites and inheritance of cytosine methylation in tomato nuclear DNA. Plant Mol Biol 16:753–770. https://doi.org/10.1007/BF00015069
Molnar A, Melnyk CW, Bassett A et al (2010) Small silencing RNAs in plants are mobile and direct epigenetic modification in recipient cells. Science 328:872–875
O'Malley RC, Huang SC, Song L et al (2016) Cistrome and epicistrome features shape the regulatory DNA landscape. Cell 165:1280–1292. https://doi.org/10.1016/J.CELL.2016.04.038
Onodera Y, Haag JR, Ream T et al (2005) Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. Cell 120:613–622. https://doi.org/10.1016/j.cell.2005.02.007
Park M, Park J, Kim S et al (2012) Evolution of the large genome in *Capsicum annuum* occurred through accumulation of single-type long terminal repeat retrotransposons and their derivatives. Plant J 69:1018–1029. https://doi.org/10.1111/j.1365-313X.2011.04851.x
Penterman J, Zilberman D, Huh JH et al (2007) DNA demethylation in the arabidopsis genome. Proc Natl Acad Sci USA 104:6752–6757. https://doi.org/10.1073/pnas.0701861104
Portis E, Acquadro A, Comino C, Lanteri S (2004) Analysis of DNA methylation during germination of pepper (*Capsicum annuum* L.) seeds using methylation-sensitive amplification polymorphism (MSAP). Plant Sci 166:169–178. https://doi.org/10.1016/j.plantsci.2003.09.004
Qin C, Yu C, Shen Y et al (2014) Whole-genome sequencing of cultivated and wild peppers provides insights into *Capsicum* domestication and specialization. Proc Natl Acad Sci USA 111:5135–5140. https://doi.org/10.1073/pnas.1400975111
Rabinowicz PD, Palmer LE, May BP et al (2003) Genes and transposons are differentially methylated in plants,

but not in mammals. Genome Res 13:2658–2664. https://doi.org/10.1101/gr.1784803

Regulski M, Lu Z, Kendall J et al (2013) The maize methylome influences mRNA splice sites and reveals widespread paramutation-like switches guided by small RNA. Genome Res 23:1651–1662. https://doi.org/10.1101/gr.153510.112

Richards EJ, Elgin SCR (2002) Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. Cell 108:489–500. https://doi.org/10.1016/S0092-8674(02)00644-X

Rodríguez-Calzada T, Torres-Pacheco I, Feregrino-Pérez AA et al (2017) Methylation profile and phenotypical changes in *Capsicum annuum* L. under water deficit and H_2O_2 application. In: 2017 13th international engineering congress CONIIN 0–3. https://doi.org/10.1109/coniin.2017.7968186

Sanetomo R, Hosaka K (2011) Reciprocal differences in DNA sequence and methylation status of the pollen DNA between F1 hybrids of *Solanum tuberosum* × *S. demissum*. Euphytica 182:219–229. https://doi.org/10.1007/s10681-011-0444-8

Singh A, Zubko E, Meyer P (2008) Cooperative activity of DNA methyltransferases for maintenance of symmetrical and non-symmetrical cytosine methylation in *Arabidopsis thaliana*. Plant J 56:814–823. https://doi.org/10.1111/j.1365-313X.2008.03640.x

Slotkin RK, Martienssen R (2007) Transposable elements and the epigenetic regulation of the genome. Nat Rev Genet 8:272–285. https://doi.org/10.1038/nrg2072

Takamiya T, Hosobuchi S, Noguchi T et al (2008) Inheritance and alteration of genome methylation in F1 hybrid rice. Electrophoresis 29:4088–4095. https://doi.org/10.1002/elps.200700784

Takuno S, Gaut BS (2012) Body-methylated genes in *Arabidopsis thaliana* are functionally important and evolve slowly. Mol Biol Evol 29:219–227. https://doi.org/10.1093/molbev/msr188

Takuno S, Gaut BS (2013) Gene body methylation is conserved between plant orthologs and is of evolutionary consequence. Proc Natl Acad Sci USA 110:1797–1802. https://doi.org/10.1073/pnas.1215380110

Teotia S, Singh D, Tang G (2017) DNA methylation in plants by microRNAs BT. In: Rajewsky N, Jurga S, Barciszewski J (eds) Plant epigenetics. Springer International Publishing, Cham, pp 247–262. https://doi.org/10.1007/978-3-319-55520-1_13

Teyssier E, Bernacchia G, Maury S et al (2008) Tissue dependent variations of DNA methylation and endoreduplication levels during tomato fruit development and ripening. Planta 228:391–399. https://doi.org/10.1007/s00425-008-0743-z

Tran RK, Henikoff JG, Zilberman D et al (2005) DNA methylation profiling identifies CG methylation clusters in arabidopsis genes. Curr Biol 15:154–159. https://doi.org/10.1016/J.CUB.2005.01.008

Tsaballa A, Athanasiadis C, Pasentsis K et al (2013) Molecular studies of inheritable grafting induced changes in pepper (*Capsicum annuum*) fruit shape. Sci Hort (Amsterdam) 149:2–8. https://doi.org/10.1016/j.scienta.2012.06.018

Vicient CM, Casacuberta JM (2017) Impact of transposable elements on polyploid plant genomes. Ann Bot 120:195–207. https://doi.org/10.1093/aob/mcx078

Wang Z, Meng D, Wang A et al (2013) The methylation of the PcMYB10 promoter is associated with green-skinned sport in Max Red Bartlett pear. Plant Physiol 162:885–896. https://doi.org/10.1104/pp.113.214700

Wang H, Beyene G, Zhai J et al (2015) CG gene body DNA methylation changes and evolution of duplicated genes in cassava. Proc Natl Acad Sci USA 112:13729–13734. https://doi.org/10.1073/pnas.1519067112

Wang L, Xie J, Hu J et al (2018) Comparative epigenomics reveals evolution of duplicated genes in potato and tomato. Plant J 93:460–471. https://doi.org/10.1111/tpj.13790

Wu R, Wang X, Lin Y et al (2013) Inter-species grafting caused extensive and heritable alterations of DNA methylation in *Solanaceae* plants. PLoS One 8. https://doi.org/10.1371/journal.pone.0061995

Xiong LZ, Xu CG, Saghai Maroof MA, Zhang Q (1999) Patterns of cytosine methylation in an elite rice hybrid and its parental lines, detected by a methylation-sensitive amplification polymorphism technique. Mol Gen Genet 261:439–446. https://doi.org/10.1007/s004380050986

Xu X, Li T, Li Y, Wang H (2015) Variation of DNA cytosine methylation patterns among parent lines and reciprocal hybrids in hot pepper. Chem Eng Trans 46:1345–1350. https://doi.org/10.3303/CET1546225

Yi Z, Niu T, Sun Y et al (2005) Patterns of DNA cytosine methylation between hybrids and their parents in sorghum genome. Acta Agron Sin 31

Yoder JA, Walsh CP, Bestor TH (1997) Cytosine methylation and the ecology of intragenomic parasites. Trends Genet 13:335–340. https://doi.org/10.1016/S0168-9525(97)01181-5

Zakrzewski F, Schmidt M, Van Lijsebettens M, Schmidt T (2017) DNA methylation of retrotransposons, DNA transposons and genes in sugar beet (*Beta vulgaris* L.). Plant J 90:1156–1175. https://doi.org/10.1111/tpj.13526

Zemach A, Kim MY, Silva P et al (2010) Local DNA hypomethylation activates genes in rice endosperm. Proc Natl Acad Sci USA 107:18729–18734. https://doi.org/10.1073/pnas.1009695107

Zhang X (2008) The epigenetic landscape of plants. Science 320:489–492

Zhang H, Zhu J-K (2012) Active DNA demethylation in plants and animals. Cold Spring Harb Symp Quant Biol 77:161–173. https://doi.org/10.1101/sqb.2012.77.014936

Zhang MS, Yan HY, Zhao N et al (2007) Endosperm-specific hypomethylation, and meiotic inheritance and variation of DNA methylation level and pattern in sorghum (*Sorghum bicolor* L.) inter-strain hybrids. Theor Appl Genet 115:195–207. https://doi.org/10.1007/s00122-007-0555-8

Zhao X, Chai Y, Liu B (2007) Epigenetic inheritance and variation of DNA methylation level and pattern in maize intra-specific hybrids. Plant Sci 172:930–938. https://doi.org/10.1016/j.plantsci.2007.01.002

Zhao Y, Yu S, Xing C et al (2008) Analysis of DNA methylation in cotton hybrids and their parents. Mol Biol 42:169–178. https://doi.org/10.1134/S0026893308020015

Zhong S, Fei Z, Chen YR et al (2013) Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. Nat Biotechnol 31:154–159. https://doi.org/10.1038/nbt.2462

Zilberman D, Cao X, Jacobsen SE (2003) ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. Science 299(5607):716–719. https://doi.org/10.1126/science.1079695

12 Revisiting Origin, Evolution, and Phylogenetics of *Capsicums* in the Genomics Era

Pasquale Tripodi

Abstract

The *Capsicum* genus comprises approximately thirty species with different ploidy and chromosome base number for which the understanding of the phylogenetic relationships has been a target of many researches since forty years. The earliest morphological, biochemical, and molecular studies allowed to identify three main groups (Annuum, Baccatum, and Pubescens) enclosing the most widely used species in terms of consumption and breeding. The advent of molecular markers gave new insight into the taxonomy of the genus better clarifying previous classifications. In addition, the progress made in the field of biology and the release of whole genome sequences have accelerated comparative mapping between pepper and the major solanaceous species such as tomato, potato, and eggplant. This chapter is a review of modern approaches that contributed to study the phylogenentic relationships and evolution of *Capsicum* species. Main comparative studies in pepper using genetic and genomic approaches are also discussed.

P. Tripodi (✉)
CREA Research Centre for Vegetable and Ornamental Crops, 84098 Pontecagnano Faiano, Italy
e-mail: pasquale.tripodi@crea.gov.it

12.1 Molecular Evolution and Phylogenetics of *Capsicums* Using Molecular Markers

Since half of the twentieth century, the study of the taxonomy of the *Capsicum* genus has been a main target of many researches. Botanical aspects, hybridization techniques, and biochemical approaches (McLeod et al. 1982; Hunziker 2001; Pickersgill 1988; Barboza and Bianchetti 2005) have been performed revealing the existence of three main complexes. These approaches of investigation, although allowed to obtain remarkable achievements, had as a drawback the impossibility to elucidate some inconsistencies observed in the relationships between species.

The progress made in the area of molecular biology and the development of genetic markers (able to detect polymorphisms in plant species caused by mutations in DNA regions) had a paramount role in the investigation of the phylogenetic relationships of pepper species (Table 12.1).

A first attempt has been performed to investigate variation in the coding regions and spacer of rRNA genes in five *Capsicum* species, including *C. annuum*, *C. chinense*, *C. frutescens*, *C. baccatum,* and *C. pubescens* (Park et al. 2000). By comparing gene sequences, the first three species formed one lineage, although *C. chinense* and *C. frutescens* were more related

N. Ramchiary and C. Kole (eds.), *The Capsicum Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-319-97217-6_12

Table 12.1 Phylogenetic studies in *Capsicum* spp. using molecular markers

Genotyping methodology*	No. of markers	No. of species	Species#	No. of accessions	Reference
5S rRNA	2 oligomer	5	CA, CB, CC, CF, CP	5	Park et al. (2000)
Chloroplast markers	*atpB-rbcL* spacer (4), *waxy* (2)	11	CA, $CB_{b/p}$, CC, CD, CE, CF, CG, CH, CL, CP, CT	17	Walsh and Hoot (2001)
cpDNA introns	*trn* (6), *rps* (1), *matK* (1), *waxy* (1)	7	CA, $CB_{b/p}$, CC, CF, CH, CP, CR	15 + 53	Jarret (2008)
Td-RAPD-PCR	2760	11	CA, CB, CC, CD, CE, CF, CG, CH, CM, CP, CT	24	Ince et al. (2010)
Chloroplast markers and CAPS via HRM	8 (6 CAPS, 1 *trn*, 1 *waxy*)	6	CA, CB, CC, CF, CH, CP	31	Jeong et al. (2010)
AFLP, SSR	4 combinations, 10	10	CA, $CB_{b/p}$, CC, CD, CE, CF, CH, CM, CP, CT	260	Ibiza et al. (2012)
SSR	28	11	CA, CA_g, $CB_{b/p}$, CC, CD, CE, CF, CG, CH, CM, CP	1352	Nicolai et al. (2013)
SSR	39	6	CA, CB, CC, CF, CH, CP	96	Gonzalez-Pérez et al. (2014)
SSR, CAPS	10, 10	8	CA, $CB_{b/p}$, CC, CF, CG, CH, CM, CP	59	Di Dato et al. (2015)
Chloroplast markers	*matK* (8), *psb*A-*trn*H (2), *waxy* (12)	34	All *Capsicum* species	34	Carrizo et al. (2016)
SNP—transcriptome-based	48	11	CA, $CB_{b/p}$, CC, CD, CE, CF, CG, CH, CM, CP, CT	4652	Lee et al. (2016) ^

#AFLP = amplified fragment length polymorphism; SSR = simple sequence repeats; CAPS = cleaved amplified polymorphic sequences; SNP = single-nucleotide polymorphism

*CA = *C. annuum*; CA_g = *C. annuum* var. *glabriusculum*; CB = *C. baccatum*; $CB_{b/p}$ = *C. baccatum* var. *baccatum* and var. *pendulum*; CC = *C. chinense*; CD = *C. cardenasii*; CE = *C. eximium*; CF = *C. frutescens*; CG = *C. galapagoense*; CH = *C. chacoense*; CL = *C. ciliatum*; CM = *C. praetermissum*; CP = *C. pubescens*; CR = *C. rhomboideum*; CT = *C. tovarii*

^no phylogenesis inferences

sharing 89.2% of the total bases. *C. baccatum* and *C. pubescens*, instead, grouped apart and suggested the existence of two different clusters, although the former was intermediate between the Annuum lineage and *C. pubescens*. A year later, phylogenesis in 11 *Capsicum* species was investigated using sequence data from both the chloroplast *atpB-rbcL* spacer region and a 1200-bp segment of the nuclear gene *waxy* which encoded for an essential enzyme in granule-bound starch synthesis (GBSS, Walsh and Hoot 2001). These types of markers had the advantage to contain a considerable number of nucleotide substitutions and were universally distributed across plant species.

Results revealed the existence of three main groups:

1. Annuum (*C. annuum*, *C. chinense*, *C. frutescens,* and *C. galapagoense*)
2. Baccatum (*C. baccatum* and *C. chacoense*)
3. Eximium (*C. eximium* and *C. cardenasii*).

The remaining species clustered separately (*C. ciliatum*) or were not assigned randomly to any group (*C. tovarii* and *C. pubescens*). The authors (Walsh and Hoot 2001) confirmed the existence of three main complexes reporting Eximium as the third group which, unlike what previously observed, did not include *C. pubescens*.

cpDNA markers and nuclear *Waxy* introns have also been used by Jarret (2008) in seven *Capsicum* species. The analysis revealed the wild *C. chacoense* as a member of the Annuum group, while *C. baccatum* and *C. pubescens* made two different clusters, although the former was much more closely related to the Annuum complex. Another wild relative was included in the study: *C. rhomboideum*, which resulted to be distant from the remaining species analyzed. A single bp insertion/deletion (In/Del) and a substitution within the introns separated *C. chinense* from *C. frutescens*, confirming what previously reported by Park et al. (2000) who suggested a very close relationship between the two species.

Similar observations were made by Jeong et al. (2010) using the high-resolution melting (HRM) system to detect single-nucleotide polymorphism (SNP) variation in *atpB-rbcL* and *waxy* genes, as well in the conserved ortholog set (COS) sequence previously mapped in pepper. The results evidenced the close evolutionary relationship of *C. annuum*, *C. frutescens,* and *C. chinense*, as well as between *C. baccatum* and *C. chacoense* as previously reported by Walsh and Hoot (2001).

An extensive study with different types of molecular markers was carried out by Ince and collaborators (2010), which investigated 11 *Capsicum* species using 2760 Td-RAPD-PCR (touch-down random amplified polymorphic DNA polymerase chain reactions). The authors confirmed the existence of three main complexes, although, beyond than Annuum group, some inconsistencies with earlier investigations were observed: a main clade combined *C. eximium* and *C. cardenasii* to *C. pubescens* (Pubescens group), while *C. baccatum* and *C. tovarii* made a new cluster with *C. praetermissum* (Baccatum group). The results of the study suggested different relationships between species, and consequently diverse regrouping within the complex. In fact, cytogenetical studies reported *C. eximium* as an intermediate between *C. baccatum* and *C. pubescens* (Moscone et al. 2003), while, on the basis of DNA sequencing data, an unclear relationship between *C. eximium* and *C. pubescens* was reported (Walsh and Hoot 2001). A fourth group, including only accessions of *C. chacoense,* was reported by the authors as intermediate between the Annuum and the Baccatum groups, although probably equally related also to Pubescens (Ince et al. 2010).

A combination of amplified fragment length polymorphism (AFLP) and microsatellites was used by Ibiza et al. (2012) in 260 accessions of nine species retrieved from the Andean region. The research confirmed what previously observed, plus reporting *C. eximium* and *C. cardenasii* as a unique species as well *C. chacoense* close to *C. baccatum*.

Microsatellites have also been used for pepper phylogeny, firstly by Nicolai et al. (2013) in a large collection (1300 individuals) of 11 wild and domesticated *Capsicum* species, and then by Gonzalez-Pérez et al. (2014) in 96 individuals of six *Capsicum* species. In both studies, the relationships observed were in agreement with the earlier studies, although accessions primarily clustered on the basis of the species. The two researches evidenced how *C. chacoense* was more related to *C. pubescens* and *C. baccatum* respect than *C. annuum*. Moreover, Nicolai et al. (2013) gave new insight on the affinity of *C. annuum* var. *glabriusculum* with the other species, demonstrating a closer relationship to *C. frutescens* and *C. annuum*. Gonzalez-Pérez et al. (2014), observed a very close clustering between *C. eximium* and *C. cardenasii* confirming what reported by Ibiza et al. (2012).

More recently, Di Dato et al. (2015) used a double approach to assess the genetic diversity in 59 accessions of nine different species based on both microsatellites as well markers linked to resistance genes and pungency. The authors reported how *C. chacoense* was intermediate between the Annuum and the Baccatum complexes, although much related to the former. The combination with markers able to select for resistances and pungency, allowed separating genotypes within species, without, however, giving any more development on the phylogenetic relationships.

The above-mentioned researches were circumstanced to wild and domesticated species belonging to the main pepper complex. A comprehensive study has been carried out recently in

34 *Capsicum* species using plastid markers (*mat*K and *psb*A-*trn*H) and the nuclear gene *waxy*, in combination with morphological, anatomical, and karyological features (Carrizo et al. 2016). The authors firstly evidenced a close link between *Capsicum* and *Lycianthes* better clarifying previous classifications proposing *Capsicum* closely related to other genera into the large subtribe Capsicinae of tribe Solaneae. Moreover, 11 different clades which reflected the geographical dispersion in the origin centers were observed. The principal complexes were confirmed as in previous findings; moreover, the authors reported the two wild species *C. galapagoense* and *C. chacoense* as part of the Annuum and the Baccatum clades, respectively. In the Baccatum clade, were also included *C. praetermissum*, a species considered by Hunziker (2001), a variety of *C. baccatum*, and *C. chacoense* which resulted to be strongly nested within this clade without any other closer affinities to other species. *C. chacoense* is considered to be an ancestor or an unmodified descendent of the ancestor of cultivated pepper and to date, its affinity to the main complex is still controversial. The authors report the need of more informative data to solve interspecific relationships within the Baccatum clade (Carrizo et al. 2016). The integration of data showed also how *C. cardenasii* and *C. eximium* were clustered with *C. eshbaughii* in a group other than *C. pubescens* and with whom shared the typical purple corolla giving the name of the clade. A better clarification on the relationship of *C. tovarii* was given evidencing how this species was also isolated in the Tovarii clade. Overall, the above-mentioned five clades (Annuum, Baccatum, Purple Corolla, Pubescens, and Tovarii) were reported to form a well-supported terminal superclade of *Capsicum*.

The recognition of the complexes and of the species included therein is crucial for breeding purposes. Interspecific crosses within the same complex allow to produce fertile or partially fertile hybrids. Crosses between species of different complexes, require instead aids such as embryo rescue, without, however ensure the success and the development of the related offspring. Therefore, most of breeding activities are limited to the Annuum complex, with few examples, adopting embryo rescue techniques or using bridge species (Yoon and Park 2005) to overcome the barriers between different gene pools such as *C. annuum* and *C. baccatum*.

12.2 Recent Highlights in *Capsicum* Comparative Mapping with Other Solanaceae Crops

Pepper was one of the first plant species on which comparative genetics studies have been carried out in order to reveal similarities and differences in gene content and gene order between genera (Tanksley et al. 1988). It has been known since early 70s that *C. annuum* comprises different chromosome structure with respect to its wild form and the other species of the Annuum complex (*C. chinense*, *C. frutescens*; Pickersgill 1971; Pochard 1970). Indeed, *C. annuum* comprises two acrocentric chromosomes and 10 meta-chromosomes, while the others have a single acrocentric chromosome and 11 meta-chromosomes (Pickersgill 1971; Lanteri and Pickersgill 1993). This change is due to a single reciprocal translocation occurring in the lineage leading to the development of the cultivated form (Wu et al. 2009).

Various genetic maps have been developed in the last 30 years, involving various types of molecular markers and mostly aimed to compare pepper and tomato genomes for various disease resistance genes and agronomic-related quantitative trait loci (QTLs) (Lefebvre et al. 1995; Thorup et al. 2004; Paran et al. 2004; Ben-Chaim et al. 2006; Minamiyama et al. 2006). In 2004, Thorup and collaborators identified homologous regions in tomato for genes involved in the biosynthetic pathway of carotenoid in pepper (Thorup et al. 2004). The same year, Thabuis and collaborators identified syntenic disease resistance genes in the two species (Thabuis et al. 2004). Since then, various studies aimed to identify tomato disease resistance genes on pepper chromosomes have been carried out for *Phythophthora capsici* (Lefebvre et al. 2002),

nematodes (Djian-Caporalino et al. 2001), and viral diseases (Lefebvre et al. 1995; Caranta et al. 1999; Moury et al. 2000). Other maps have been developed to identify syntenic regions for fruit shape traits between tomato and pepper (Zygier et al. 2005; Ben-Chaim et al. 2006). All these studies, although useful, did not report a complete and robust genetic map of the pepper genome, in which 12 linkage groups correspond to 12 pepper chromosomes.

Only in recent years, efforts toward this objective have been performed. In 2009, Wu and collaborators published a pepper genetic map consisting of 399 markers of which 263 represented the second generation of conserved orthologous (COSII) markers (Wu et al. 2009). COSII, being developed from a set of single-copy conserved orthologous genes in the Asterid species (Wu et al. 2006) represent a powerful tool for establishing a syntenic network between Solanaceae and Arabidopsis. The map allowed to identify 35 shared regions between the pepper and tomato genomes with an average range of about 30 cM for a total of 19 inversions and the six translocations involving all the chromosomes (Wu et al. 2009; Wu and Tanksley 2010).

The advent of genomics has helped to give better insight into the identification of comparative regions among pepper and other Solanaceae species. Hill and collaborators (2015) compared two high-density, transcript-based genetic maps of pepper to tomato and potato genome. The authors, in addition to the major translocations previously described between tomato and potato with pepper, identified two more translocations between the non-recombining region on chromosome 4 of pepper and chromosomes 11 and 12 of both tomato and potato. Rinaldi et al. (2016) reported the first comparative map based on coding DNA sequence (CDS) of pepper, tomato, and eggplant. Thanks to the genome dataset publicly available, the authors aligned 35,000 CDS of both pepper and tomato to the tomato ITAG2.5 and the pepper v1.55 genome sequence, respectively. Moreover, the COSII markers above described (Wu et al. 2009) were aligned to the pepper and tomato genomes. The syntenic analysis revealed the existence of 19,734 unique hits on the basis of the alignment of 34,727 tomato CDS to the pepper genome, as well as, 20,700 unique hits among the 34,899 pepper CDS aligned to the tomato genome. Moreover, the alignment of 347 COSII markers (Wu et al. 2009) allowed to identify 30,942 matches.

Results confirmed the 10 translocations and 14 inversions previously described (Wu et al. 2009; Hill et al. 2015), reporting furthermore the existence of 3 new translocations and 13 new inversions. However, probably due to errors in genome mapping and/or misalignment, the authors did not confirm 4 previously reported inversions on chromosomes 1, 4, and 11 and small translocations previously identified in centromere regions (Rinaldi et al. 2016).

The alignment of CDS regions reinforces the hypothesis that an illegitimate pairing and crossing over event occurred in relatively recent times between two non-homologous metacentric chromosomes in the ancestral genome of *C. annuum*. All these exchanges and rearrangements of the genetic material can explain the higher genome size of pepper with respect to tomato.

The alignment of pepper and eggplant CDS evidenced 14 translocations and two inversions in seven pepper chromosomes (Rinaldi et al. 2016). Furthermore, several orthologous QTLs for agronomic and fruit morphological traits were evidenced between eggplant and pepper. These results can be considered as the first direct syntenic analysis of these two species. All these comparative studies, integrated with the results on genome sequence, represent important findings for both basic research and applied crop improvement.

12.3 Novel Insight in the Evolution and Gene Orthology from Sequenced Genomes

Recent information released from the whole genome sequence of *Capsicum* gave new insight into the syntenic relationships between pepper and other major Solanaceae.

In the first sequenced genome (Kim et al. 2014), the analysis of orthologous gene sets (OGs) among pepper, tomato, and potato, revealed 16,524 orthologous genes between pepper and potato, and 17,397 between pepper and tomato, the majority of which differentially expressed in placenta and pericarp. The position of OGs in pepper and tomato demonstrated a well-conserved gene synteny in euchromatic regions and a total of 11 inversions, 6 translocations, and 8 inversion-associated translocations extensively distributed on all chromosomes (Kim et al. 2014).

The second genome sequenced by Qin and collaborators (2014) evidenced the existence of 430 (about 351 Mb) translocation blocks and 367 inverse blocks (about 361 Mb) between pepper and potato, while 612 (about 560 Mb) translocation blocks and 468 inverse blocks (about 590 Mb) were found between pepper and tomato. Overall, a synteny majorly conserved was found between tomato and potato in comparison to pepper due to the large segmental inversions and translocations of the latter.

In pepper, tomato and potato, a total of 108,205 sequences, were found to be clustered into 21,529 gene families, 63% of which shared among the three genomes. The comparative analysis of the three before mentioned Solanaceae, Arabidopsis, grape and rice revealed instead a total of 7826 shared families (36%), 1014 of which were unique to eudicotyledons. Although about 800 families contained over two thousand genes unique to pepper (Kim et al. 2014), each family shared several genes with the other species. A high degree of synteny between pepper, tomato, and potato was found for nucleotide-binding-site-encoding genes (NBS). The phylogenetic and evolutionary reconstruction indicated that NBS in pepper were over-represented after the speciation and different gene-duplication events had contributed to their expansion and to the lost of colinearity with tomato and potato. This mechanism did not occur instead between tomato and potato. These observations could explain the parallel evolution of the genus *Capsicum* through various types of gene duplication responsible for loss or gain of resistance genes. Gene duplication is known to be a main process which generates novel diversity, creating new genes during the evolution. Kim and collaborators (2014) gave insight into the evolution of genes involved in the biosynthetic pathway of capsaicinoids, revealing an extensive duplication in pepper which occurred also in tomato orthologues. The expression of these genes evidenced a minutely expression in the tomato placenta and in potato fruit tissues.

As reported in Chap. 8, there are several families within transcription factors (TF) and the overall number in peppers is comparable to those of other species. WKRY (the larger representative of TF) and NAC family were found to form specific orthologous clusters with the major sequenced Solanaceae (pepper, tomato, potato), whereas no representatives were found in the Arabidopsis, rice, and grape genomes. Within the AP2/ERF superfamily, the number of members found in pepper was lower than the other species excepting for grape (Kim et al. 2014). Phylogenetic analysis revealed the existence of 11 groups clusterizing differentially the Solanaceae and non-Solanaceae species evidencing a lower number of duplicated or triplicated genes in the AP2/ERF in comparison to Arabidopsis, rice, tomato, and potato. Among auxin response factors (ARFs), despite the structural diversity between pepper to other dicots, a conserved orthologous relationship and a tighter orthology with tomato was witnessed. Moreover, in both pepper and tomato, a uniform duplication was observed for this gene family although not massive as in Arabidopsis.

Comparing the number and the expression of the cytochrome P450s genes, large differences of the genes involved in diverse metabolic pathways during fruit maturation were found between pepper and tomato. Different subgroups in which P450s member clusterized were found to be differentially represented compared to tomato and potato, suggesting specific biological processes in pepper. Other gene families showed orthology between pepper and other species. Over a hundred of genes involved in cuticle biosynthesis as well three hundred of the receptor-like kinases (RLK) family were shared

between pepper, tomato, and Arabidopsis. On the contrary, the genes present in other species were not found in pepper such as the tomato single flower truss (SFT). Finally, the phylogenetic analysis of the serine/threonine phosphatase family (PPP/PPM) evidenced a strong conservation indicating that its members may have diverged before speciation. Moreover, although pepper was reported to contain a similar number of genes of the other clades, within this family a single additive gene was found with respect to Arabidopsis and tomato suggesting an event of duplication occurred after speciation.

More recently, a new version of the reference genome (*C. annuum* CM334) and the *novo* sequence of two domesticated species (*C. baccatum*, *C. chinense*) elucidated the relationship between pepper species in terms of evolution and chromosomal rearrangements (Kim et al. 2017) (see Chap. 8). Comparison among pepper, tomato, and potato revealed translocation between *C. baccatum* (chromosome 9) and the other two *Capsicum* species (chromosome 3) as well as between Solanum (chromosomes 4 and 6) and pepper (chromosomes 3 and 5). Moreover, it was demonstrated how the duplication events were predominant in the Baccatum lineage in particular for nucleotide-binding and leucine-rich-repeat proteins (NLR), a large group of functional disease resistance loci of plants, and for long-terminal repeat retrotransposons (LTR) which resulted in a major proportion of the members of *Athila* family in *C. baccatum*. These results gave new insight on the diversification of pepper species.

References

Barboza GE, Bianchetti LDB (2005) Three new species of *Capsicum* (Solanaceae) and a key to the wild species from Brazil. Syst Bot 30:863–871

Ben-Chaim A, Brodsky Y, Falise M, Mazourek M et al (2006) QTL analysis for capsaicinoid content in *Capsicum*. Theor Appl Genet 113:1481–1490

Caranta C, Thabuis A, Palloix A (1999) Development of a CAPS marker for the *Pvr4* locus: a tool for pyramiding potyvirus resistance genes in pepper. Genome 42:1111–1116

Carrizo GC, Barfuss HJ, Sehr EM, Barboza GE, Samuel R, Moscone EA, Ehrendorfer F (2016) Phylogenetic relationships, diversification and expansion of chili peppers (*Capsicum*, Solanaceae). Ann Bot 118:35–51

Di Dato F, Parisi M, Cardi T, Tripodi P (2015) Genetic diversity and assessment of markers linked to resistance and pungency genes in *Capsicum* germplasm. Euphytica 1:103–119

Djian-Caporalino C, Pijarowski L, Fazari A, Samson M, Gaveau L, O'Byrne C, Lefebvre V, Caranta C, Palloix A, Abad P (2001) High-resolution genetic mapping of the pepper (Capsicum annuum L.) resistance loci Me3 and Me4 conferring heat-stable resistance to root-knot nematodes (Meloidogyne spp.). Theor Appl Genet 103(4):592–600

Gonzalez-Pérez S, Garceés-Claver A, Mallor C, Sàenz de Miera LE, Fayos O, Pomar F, Merino F, Silvar C (2014) New insights into *Capsicum* spp. relatedness and the diversification process of *Capsicum annuum* in Spain. PLoS One 9(12):116276

Hill T, Ashrafi H, Chin-Wo SR, Stoffel K, Truco MJ, Kozik A, Michelmore R, Van Deynze A (2015) Ultra-high density, transcript-based genetic maps of pepper define recombination in the genome and synteny among related species. Genes Genomes Genet 5(11):2341–2355

Hunziker AT (2001) Genera Solanacearum: the genera of Solanaceae illustrated, arranged according to a new system. Gantner Verlag, Ruggell, Liechtenstein

Ibiza VP, Blanca J, Canizares J, Nuez F (2012) Taxonomy and genetic diversity of domesticated *Capsicum* species in the Andean region. Genet Resour Crop Evol 59:1077–1088

Ince AG, Karaca M, Onus AN (2010) Genetic relationships within and between *Capsicum* species. Biochem Genet 48:83–95

Jarret RL (2008) DNA barcoding in a crop genebank: the *Capsicum annuum* species complex. Opin Biol J 1:35–42

Jeong HJ, Jo YD, Park SW, Kang BC (2010) Identification of *Capsicum* species using SNP markers based on high resolution melting analysis. Genome 53:1029–1040

Kim S, Park M, Yeom SI, Kim YM, Lee JM, Lee HA et al (2014) Genome sequence of the hot pepper provides insights into the evolution of pungency in *Capsicum* species. Nat Genet 46:270–278

Kim S, Park J, Yeom SI, Kim YM, Seo E, Kim KT et al (2017) New reference genome sequences of hot pepper reveal the massive evolution of plant disease-resistance genes by retroduplication. Genome Biol 18:210

Lanteri S, Pickersgill B (1993) Chromosomal structural-changes in *Capsicum annuum* L. and *C. chinense* Jacq. Euphytica 67:155–160

Lee HY, Ro NY, Jeong HJ, Kwon JK, Jo J, Ha Y, Jung A, Han JW, Venkatesh J, Kang BC (2016) Genetic diversity and population structure analysis to construct

a core collection from a large *Capsicum* germplasm. BMC Genet 17:142

Lefebvre V, Palloix A, Caranta C, Pochard E (1995) Construction of an intra-specific integrated linkage map of pepper using molecular markers and double dhaploid progenies. Genome 38:112–121

Lefebvre V, Pflieger S, Thabuis A, Caranta C, Blattes A, Chauvet JC, Daubèze AM, Palloix A (2002) Towards the saturation of the pepper linkage map by alignment of three intraspecific maps including known-function genes. Genome 45:839–854

McLeod MJ, Guttman SI, Eshbaugh WH (1982) Early evolution of chili peppers (*Capsicum*). Econ Bot 36:361–368

Minamiyama Y, Tsuro M, Hirai M (2006) An SSR-based linkage map of *Capsicum annuum*. Mol Breed 18:157–169

Moscone EA, Baranyi M, Ebert I, Greilhuber J, Ehrendorfer F, Hunziker AT (2003) Analysis of nuclear DNA content in *Capsicum* (Solanaceae) by flow cytometry and Feulgen densitometry. Ann Bot 92:21–29

Moury B, Pflieger S, Blattes A, Lefebvre V, Palloix A (2000) A CAPS marker to assist selection of *Tomato spotted wilt virus* (TSWV) resistance in pepper. Genome 43:137–142

Nicolai M, Cantet M, Lefebvre V, Sage-Palloix AM, Palloix A (2013) Genotyping a large collection of pepper (*Capsicum* spp.) with SSR loci brings new evidence for the wild origin of cultivated *C. annuum* and the structuring of genetic diversity by human selection of cultivar types. Genet Resour Crop Evol 60:2375–2390

Paran I, Van der Voort JR, Lefebvre V, Jahn M et al (2004) An integrated genetic linkage map of pepper (*Capsicum* spp.). Mol Breed 13:251–261

Park Y-K, Park K-C, Park C-H, Kim N-S (2000) Chromosomal localization and sequence variation of 5Sr RNA gene in five *Capsicum* species. Mol Cells 10:18–24

Pickersgill B (1971) Relationships between weedy and cultivated forms in some species of chili peppers (genus *Capsicum*). Evolution 25:683–691

Pickersgill B (1988) The genus *Capsicum*: a multidisciplinary approach to the taxonomy of cultivated and wild plants. Biol Zentrabl 107:381–389

Pochard E (1970) Description of trisomic individuals of *Capsicum annuum* L. obtained in progeny of a haploid plant. Ann Amel Plantes 20:233–256

Qin C, Yu C, Shen Y, Fang X, Chen L, Min J et al (2014) Whole-genome sequencing of cultivated and wild peppers provides insights into *Capsicum* domestication and specialization. Proc Natl Acad Sci USA 111:5135–5140

Rinaldi R, Van Deynze A, Portis E, Rotino GL, Toppino L, Hill T, Ashrafi H, Barchi L, Lanteri S (2016) New insights on eggplant/tomato/pepper synteny and identification of eggplant and pepper orthologous QTL. Front Plant Sci 7:1031

Tanksley SD, Bernatzky R, Lapitan NL, Prince JP (1988) Conservation of gene repertoire but not gene order in pepper and tomato. Proc Natl Acad Sci USA 85:6419–6423

Thabuis A, Palloix A, Servin B, Daubèze AM, Signoret P, Hospital F, Lefebvre V (2004) Marker-assisted introgression of 4 *Phytophthora capsici* resistance QTL alleles into a bell pepper line: validation of additive and epistatic effects. Mol Breed 14:9–20

Thorup T, Tanyolac B, Livingstone K, Popovsky S, Paran I, Jahn M (2000) Candidate gene analysis of organ pigmentation loci in the Solanaceae. Proc Natl Acad Sci USA 97:11192–11197

Walsh BM, Hoot SB (2001) Phylogenetic relationships of *Capsicum* (Solanaceae) using DNA sequences from two non-coding regions: the chloroplast atpB-rbcL spacer region and nuclear waxy introns. Int J Plant Sci 162:1409–1418

Wu FN, Tanksley SD (2010) Chromosomal evolution in the plant family Solanaceae. BMC Genom 11:182

Wu FN, Mueller LA, Crouzillat D, Petiard V, Tanksley SD (2006) Combining bioinformatics and phylogenetics to identify large sets of single-copy orthologous genes (COSII) for comparative, evolutionary and systematic studies: a test case in the euasterid plant clade. Genetics 174:1407–1420

Wu FN, Eannetta NT, Xu YM, Durrett R, Mazourek M, Jahn M, Tanksley S (2009) A high density COSII map of the pepper genome provides a detailed picture of synteny with tomato and new insights into recent chromosome evolution in the genus. Theor Appl Genet 118:1279–1293

Yoon JB, Park HG (2005) Trispecies bridge crosses, (*Capsicum annuum* × *C. chinense*) × *C. baccatum*, as an alternative for introgression of anthracnose resistance from *C. baccatum* into *C. annuum*. J Kor Soc Hort Sci 46:5–9

Zygier S, Ben Chaim A, Efrati A, Kaluzky G et al (2005) QTL mapping for fruit size and shape in chromosomes 2 and 4 in pepper and a comparison of the pepper QTL map with that of tomato. Theor Appl Genet 111:437–445

Impact of Genomics on *Capsicum* Breeding

13

Kenta Shirasawa, Tomohiro Ban, Noriko Nagata and Toshiya Murakana

Abstract

The genus *Capsicum* includes several species of cultivated peppers, including *Capsicum annuum*, *C. frutescens*, *C. chinense*, *C. baccatum,* and *C. pubescens*. Information on plant genomics is essential not only for the efficient and accurate advancement of breeding programs, but also for the correct management of plant genetic resources. Genomes of six *Capsicum* lines which belong to *C. annuum*, *C. baccatum*, and *C. chinense* have been sequenced using next-generation sequencing (NGS) approaches. Consequently, high-throughput genomewide single-nucleotide polymorphism genotyping techniques based on NGS or DNA microarrays have been developed. Genotyping plant populations enable the evaluation of population structure and detection of genetic loci via quantitative trait locus (QTL) analysis and genomewide association study (GWAS). In this chapter, we introduce the current state of *Capsicum* genomics and its contribution to *Capsicum* breeding. Additionally, we report the results of our GWAS based on double-digest restriction-site associated DNA (ddRAD-Seq) analyses performed on 192 *Capsicum* lines stocked at The Kihara Institute for Biological Research, Yokohama City University, Japan. The genomics information summarized in this chapter will potentially be useful for the development of new and attractive pepper cultivars.

13.1 Introduction

Peppers (*Capsicum* spp.) are members of the family Solanaceae, along with other important crops, including potato (*Solanum tuberosum*), tomato (*S. lycopersicum*), tobacco (*Nicotiana tabacum*), and eggplant (*S. melongena*). Because these plant species are important agricultural crops, whole-genome sequences of these crops and related species have been done, released, and utilized in breeding programs (The Potato Genome Sequencing Consortium 2011; Bombarely et al. 2012; The Tomato Genome Consortium 2012; Liedl et al. 2013; Sierro et al. 2013, 2014; Bolger et al. 2014; Hirakawa et al. 2014; Razali et al. 2017). However, because of the large

K. Shirasawa (✉)
Kazusa DNA Research Institute, Kisarazu, Chiba 292-0818, Japan
e-mail: shirasaw@kazusa.or.jp

T. Ban
Kihara Institute for Biological Research, Yokohama City University, Yokohama, Kanagawa 244-0813, Japan

N. Nagata
Japan Women's University, Bunkyo, Tokyo 112-8681, Japan

T. Murakana
Osaka University, Suita, Osaka 565-0871, Japan

N. Ramchiary and C. Kole (eds.), *The Capsicum Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-319-97217-6_13

genome size of *Capsicum* spp. (approximately 3 Gb; Moscone et al. 2003), the investigation of *Capsicum* spp. has lagged behind that of other Solanaceae members.

The genus *Capsicum* includes several cultivated species, including *C. annuum* (bell pepper, Anaheim chile, New Mexico chile, ancho, jalapeño, and banana pepper), *C. baccatum* (Aji amarillo), *C. frutescens* (Tabasco pepper), *C. chinense* (habanero), and *C. pubescens* (rocoto pepper) (Paran et al. 2007). Because most of these species, except *C. pubescens*, are cross-compatible (Walsh and Hoot 2001), it is possible to transfer favorable traits from one species to another via interspecific hybridization. In breeding programs, *Capsicum* lines have usually been selected on the basis of agronomically important phenotypic traits, such as disease resistance, yield, and fruit morphology. Phenotypic evaluations have been also used for the management of genetic bioresources, in accordance with 12 standardized phenotypic index criteria of the International Plant Genetic Resource Institute, Asian Vegetable Research and Development Center, and Centro Agronomico Tropical de Investigación y Enseñanza of Costa Rica (IPGRI, AVRDC, and CATIE 1995). However, since phenotypes are often affected by environmental conditions, incorrect selection, and misidentification are a common occurrence in breeding programs and bioresource management, respectively. By contrast, DNA sequence information is stable and reliable, as it is unaffected by environmental conditions. Therefore, several DNA markers have been developed to facilitate the selection of favorable lines in breeding programs and accurate species identification for the management of genetic resources (Paran et al. 2007; Chhapekar et al. 2016). Whole-genome sequence information has also been used for the efficient breeding and management of crop plants (Varshney et al. 2005).

This chapter summarizes the current status of *Capsicum* genomics and introduces genomewide single-nucleotide polymorphism (SNP) genotyping techniques based on genome sequence information. In addition, we demonstrate genomewide SNP genotyping and genomewide association study (GWAS) for fruit orientation among a collection of 192 *Capsicum* lines stocked at The Kihara Institute for Biological Research, Yokohama City University, Japan.

13.2 Genome Sequences of Peppers

Advancements in next-generation sequencing (NGS) technologies have enabled the sequencing of *Capsicum* genomes. To the best of our knowledge, genomes of six *Capsicum* lines belonging to the species *C. annuum*, *C. baccatum*, and *C. chinense* have been published till date (Table 13.1). Two reference genomes of *C. annuum*, the main species of cultivated peppers, have been released in parallel, including that of a Mexican hot pepper landrace, Criollo de Morelos 334 (CM334) (Kim et al. 2014), and that of a cultivated pepper Zunla-1 (Qin et al. 2014). The genome of CM334 (3.06 Gb) has been sequenced to 186.6X coverage using Illumina short-read sequencing (San Diego, CA, USA), resulting in 37,989 scaffold sequences with an N50 length of 2.47 Mb (Table 13.1; Kim et al. 2014). Of the 3.06 Gb genome sequence of CM334, 2.63 Gb, comprising 1357 scaffold sequences, is genetically anchored to the 12 chromosomes. In the case of Zunla-1, 325 Gb of short-read sequences have been generated from various Illumina libraries (insert sizes ranging from 170 bp to 40 kb) (Qin et al. 2014). These sequences have been assembled into a 3.35 Gb genome sequence, comprising 967,017 scaffold sequences with an N50 length of 1.23 Mb (Table 13.1). These scaffold sequences are anchored to the chromosomes, in accordance with a genetic linkage map with 7657 SNP loci, to establish the 12 chromosome-level pseudomolecules consisting of 4956 scaffold sequences. Genomes of CM334 and Zunla-1 are predicted to harbor 34,903 and 35,336 protein-coding genes, respectively.

Recently, the genome of an F_1 hybrid derived from a cross between CM334 and a nonpungent blocky pepper breeding line was sequenced using Linked-Read technology supplied by 10X

Table 13.1 Statistics of genome assemblies in *Capsicum* spp.

Species	Line	Genome assembly size (Gb)	Number of scaffolds	N50 length (Mb)	Number of predicted genes	Reference
C. annuum	CM334	3.06	37,989	2.47	34,903	Kim et al. (2014)
	Zunla-1	3.35	967,017	1.23	35,336	Qin et al. (2014)
	F1 hybrid of an interspecific cross between CM334 and a nonpungent blocky pepper breeding line	3.21	83,391	3.69	NA[a]	Hulse-Kemp et al. (2018)
C. annuum var. *glabriusculum*	Chiltepin	3.48	1,973,483	0.45	34,476	Qin et al. (2014)
C. chinense	PI159236	3.01	51,917	3.30	35,009	Kim et al. (2017)
C. baccatum	PBC81	3.22	25,349	2.00	35,874	Kim et al. (2017)

[a]NA, data not available

Chromium library (10x Genomics, Pleasanton, CA, USA) (Hulse-Kemp et al. 2018). This resulted in the construction of two haploid genome sequences, where each haplotype phase is referred to as a 'pseudohap' assembly (Hulse-Kemp et al. 2018). One of the pseudohap assemblies, pseudohap1, consists of 83,391 scaffold sequences with a total size of 3.21 Gb and an N50 length of 3.69 Mb (Table 13.1). The pseudohap1 assembly includes 258,884 phase blocks with an N50 length of 1.72 Mb, which is equivalent to a total of 2.67 Gb of 1587 scaffold sequences anchored to the 12 chromosomes. In addition to the sequencing of *C. annuum* cultivars and landraces, the genome sequence of the wild progenitor of *Capsicum*, *C. annuum* var. *glabriusculum* (Chiltepin), has also been determined (Qin et al. 2014). The genome assembly of Chiltepin is 3.48 Gb in length, comprising 1,973,483 scaffold sequences with an N50 length of 445,585 bp, and is predicted to harbor 34,476 protein-coding genes (Table 13.1).

Genome sequences of *C. chinense* and *C. baccatum* have also been published. Although the first draft genome sequence of *C. chinense* accession PI159236 was highly fragmented, comprising 239,495 sequences with an N50 length of 60 kb and spanning a total length of 2.95 Gb (Kim et al. 2014), it has subsequently been updated by sequencing multiple Illumina libraries of different insert sizes (200 bp–10 kb) (Kim et al. 2017). The improved sequence of PI159236 is 3.01 Gb in length; it consists of 51,917 scaffold sequences with an N50 length of 3.30 Mb and is predicted to encode 35,009 protein-coding genes (Table 13.1). In the case of *C. baccatum*, the genome of PBC81 has been sequenced using multiple libraries with different insert sizes (200 bp–10 kb) (Kim et al. 2014). The genome of PBC81 is 3.22 Gb in length and consists of 25,349 scaffold sequences, with an N50 length of 2.00 Mb and a total of 35,874 predicted protein-coding genes (Table 13.1).

13.3 High-Throughput SNP Genotyping Techniques

There have been numerous efforts to develop DNA markers in *Capsicum* spp. Prior to the availability of *Capsicum* genome sequence, single-copy conserved orthologous markers were developed across the Solanaceae crops, including tomato, potato, eggplant, pepper, and tobacco

(Wu et al. 2009). Additionally, simple-sequence repeat (SSR) markers, or microsatellite markers, were randomly developed from expressed sequence tags (ESTs) and SSR-enriched genomic libraries (Minamiyama et al. 2006; Yi et al. 2006; Mimura et al. 2012; Sugita et al. 2013). Genome sequencing of tomato, a Solanaceae model crop, greatly facilitated the deduction of SSR markers in the *Capsicum* genome (Shirasawa et al. 2013). Latest developments in whole-genome sequencing of *Capsicum* lines have generated a large number of genomewide SNP markers.

13.3.1 Genomewide SNPs Based on DNA Array/Chip Technologies

Premade custom microarray-based SNP arrays/chips, such as Illumina GoldenGate and Infinium (San Diego, CA, USA) and Affymetrix GeneChip and Axiom (Santa Clara, CA, USA), have enabled high-throughput SNP genotyping in multiple species. ESTs from several genotypes have been used for SNP discovery (Hill et al. 2013). The construction of an Affymetrix GeneChip involves the assembly of sequences into unigenes to detect SNPs. In *Capsicum*, the Affymetrix GeneChip has been used to detect 33,401 SNPs in 13,323 unigenes of 43 diverse *Capsicum* lines, including 40 lines of *C. annuum* and one line each of *C. frutescens*, *C. chinense*, and *C. pubescens* (Table 13.2). By contrast, the number of SNP markers among the 40 *C. annuum* lines is only 6426. Additional types of SNP arrays have been designed from NGS data. Hulse-Kemp et al. (2016) have developed the Illumina Infinium Array, PepperSNP16K (Table 13.2), from whole-genome resequencing analysis (WGRS) of 22 chile and bell pepper lines of *C. annuum*. This array, comprising 14,877 informative SNPs, has been used to construct an interspecific genetic map with 5546 markers separated into 1361 bins across 12 linkage groups, representing a genetic distance of 1392.3 centimorgan (cM). Cheng et al. (2016) have developed an Illumina Infinium iSelect SNP array, *Cap*SNP15K (Table 13.2), with 8199 informative SNPs. This SNP array has been designed from WGRS data of *C. annuum* and *C. frutescens*; it has been used to perform quantitative trait locus (QTL) analysis based on a genetic linkage map with a total map length of 1628.83 cM and comprising 5569 SNPs distributed in 3826 bins.

13.3.2 Genomewide SNPs Based on NGS Technology

The speed and efficiency of genome sequencing have increased tremendously in recent years, whereas the cost of sequencing has continued to decrease. Consequently, SNP genotyping via NGS has gained popularity because of its flexibility and relatively low cost (Davey et al. 2011). The genotyping-by-sequencing (GBS) technique (Elshire et al. 2011) has been widely used in model and nonmodel plant species (Poland and Rife 2012). In GBS, genome DNA is subjected to digestion with restriction endonucleases to reduce genome complexity prior to sequencing. Since this method generates sequence reads from the same loci across multiple individuals, it is used to identify and genotype common SNPs within populations.

The GBS method has been used for the analysis of population structure among 222 *C. annuum* accessions (Taranto et al. 2016; Table 13.2). A total of 108,591 SNPs has been identified among the tested lines, with an average density of one SNP per 8.7 kb across the *12* chromosomes. Subsequent clustering analysis of these SNPs has revealed three major groups within the 222 accessions, according to their geographical origin and fruit morphology. Nimmakayala et al. (2016a) have applied GBS on 94 *C. annuum* accessions, thus identifying 66,960 SNPs (Table 13.2). In these *C. annuum* accessions, a total of 1189 haplotypes have been identified based on 3413 SNPs, and a linkage disequilibrium (LD) decay based on haplotype extensions has been estimated within an average physical distance of 139 kb (Nimmakayala et al. 2016a). Additionally, GWAS has been used to detect SNPs in candidate genes regulating the weight and capsaicinoid content of *Capsicum*

Table 13.2 Number of single-nucleotide polymorphisms (SNPs) detected among *Capsicum* species using different genotyping platforms

Technology[a]	Species (number of lines tested)	Number of SNPs detected	Reference
DNA array/chip			
Affymetrix GeneChip	*C. annuum* (40), *C. frutescens* (1), *C. chinense* (1) and *C. pubescens* (1)	33,401	Hill et al. (2013)
Affymetrix GeneChip	*C. annuum* (40)	6426	Hill et al. (2013)
Illumina Infinium Array (PepperSNP16K)	An interspecific mapping population (90)	5546	Hulse-Kemp et al. (2016)
Illumina Infinium Array (*Cap*SNP15K)	An interspecific mapping population (297)	5569	Cheng et al. (2016)
Next-generation sequencing			
GBS	*C. annuum* (222)	108,591	Taranto et al. (2016)
GBS	*C. annuum* (94)	66,960	Nimmakayala et al. (2016a)
GBS	*C. baccatum* (283) and *C. annuum* (94)	77,407	Nimmakayala et al. (2016b)
GBS	*C. baccatum* (283)	13,129	Nimmakayala et al. (2016b)
GBS	*C. annuum* (145), *C. chinense* (42) and *C. frutescens* (21)	109,610	Han et al. (2018)
GBS	An interspecific mapping population (120)	8587	Han et al. (2018)
ddRAD-Seq	*C. annuum* (30), *C. baccatum* (21), *C. chinense* (85), *C. frutescens* (25), *C. pubescens* (1), and unidentified species (24)	14,444	This study
WGRS	*C. baccatum* (1) and *C. annuum* (1)	4,887,031	Ahn et al. (2018)
WGRS	An intraspecific mapping population (120)	1,431,214	Han et al. (2016)

[a]GBS, genotyping-by-sequencing; ddRAD-Seq, double-digest restriction-site associated DNA sequencing; WGRS, whole-genome resequencing

fruits. The GBS method has been employed in 283 *C. baccatum* lines and 94 *C. annuum* accessions (Nimmakayala et al. 2016b; Table 13.2). While a total of 77,407 SNPs has been identified among the 377 *Capsicum* lines (283 *C. baccatum* and 94 *C. annuum*), 13,129 of these SNPs have been discovered in *C. baccatum* lines. Based on the 77,407 genomewide SNPs, 1742 haplotypes with 4420 SNPs and an LD block size of 99.1 kb have been identified. SNPs associated with peduncle length have been detected using GWAS. Han et al. (2018) have genotyped 208 *Capsicum* accessions, including 145 *C. annuum*, 42 *C. chinense*, and 21 *C. frutescens* accessions, using the GBS method (Table 13.2), and identified a total of 109,610 SNPs, which divide the accessions into three species-specific groups. Approximately 90% of the SNPs are located within 5513 LD blocks, with an average size of 409 kb. In addition, Han et al. (2018) have constructed a genetic map from a collection of 85 recombinant inbred lines (RILs) derived from a cross between *C. annuum* 'TF68' and *C. chinense* 'Habanero'. This map

consists of 12 linkage groups based on 8587 SNPs distributed in 1089 bins and covering a map distance of 1127.3 cM. QTLs controlling capsaicinoid content have been detected on this genetic map and validated via GWAS using genotypic data of 208 *Capsicum* lines.

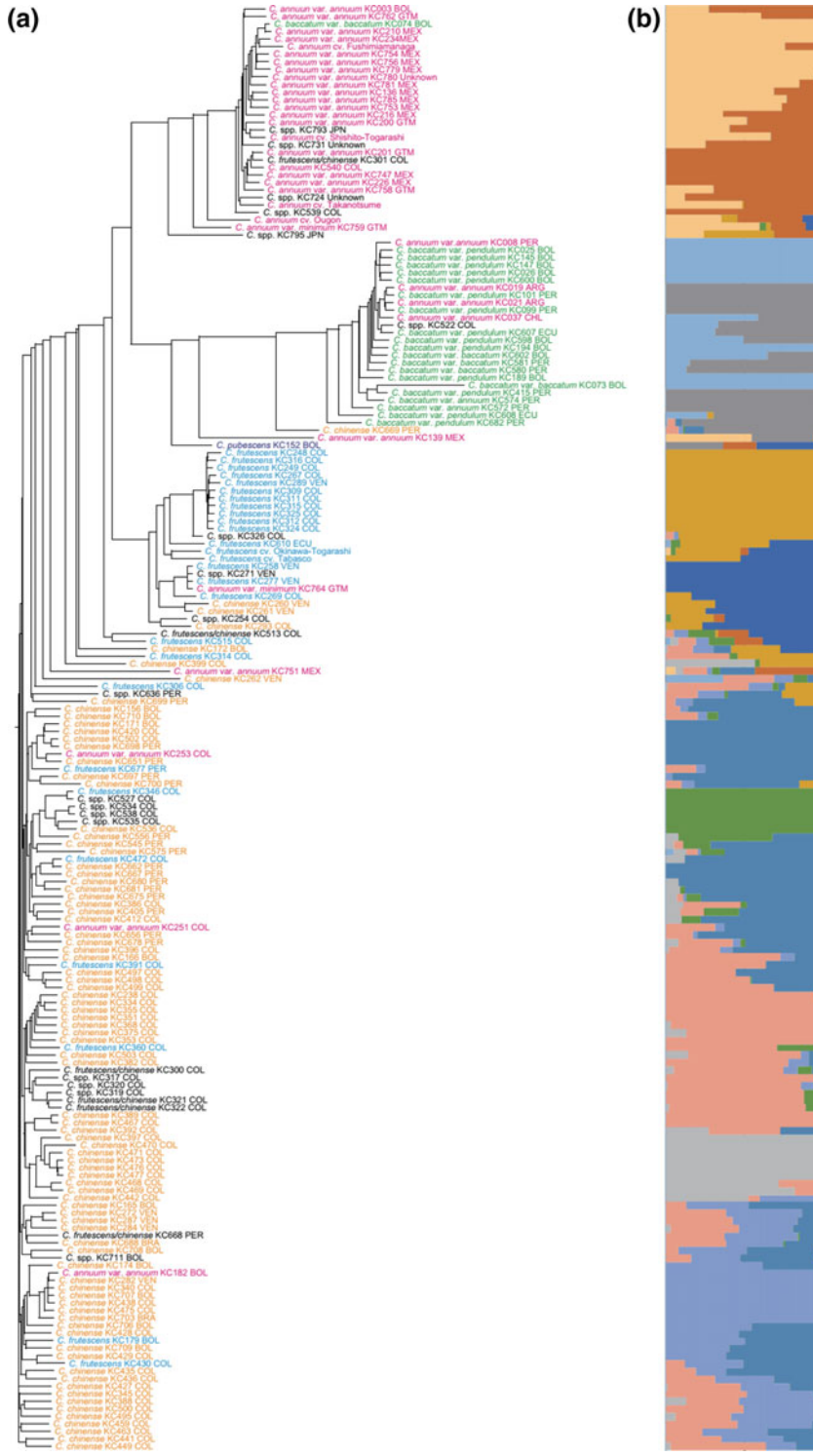
13.3.3 Whole-Genome SNPs Based on NGS

WGRS is a more powerful approach than GBS to detect genomewide SNPs, as restriction endonuclease sites do not limit the target DNA sequences in WGRS. Ahn et al. (2018) reported the WGRS of *C. baccatum* 'PRH1' and *C. annuum* 'Saengryeg' (Table 13.2). These *Capsicum* cultivars have been sequenced and covered approximately 5X of the *Capsicum* genome. The high-quality reads of *C. baccatum* and *C. annuum* have been aligned against the reference genome sequence at rates of 45 and 39%, respectively, spanning 39% of the genome in *C. battatum* and 91% in *C. annuum*. Totals of 6,213,009 and 6,804,889 genomewide SNPs have been identified in *C. baccatum* and *C. annuum*, respectively. Among these, 4,887,031 SNPs are polymorphic between the two species, of which only 150,932 and 39,955 SNPs are located within the coding sequences of genes in *C. baccatum* and *C. annuum*, respectively. Furthermore, WGRS has been used for genotyping a mapping population of 120 RILs (Han et al. 2016; Table 13.2). In this study, the parental lines, *C. annuum* 'Perennial' and *C. annuum* 'Dempsey', have been sequenced at 18X coverage, whereas RILs have been sequenced at only 1X coverage. Missing data due to low coverage have been imputed using a sliding window approach (Huang et al. 2009). This has resulted in the identification of 1,431,214 SNPs between the parental lines, which have been used to construct a genetic map of RILs spanning a distance of 1372 cM and comprising 2578 genetic bins, with an average interval of 0.53 cM between bins. This ultra-high-density bin map with short intervals between bins is a powerful tool for fine mapping QTLs for 17 horticultural traits.

13.4 Practical Workflow for GWAS

Previously, we have used 60 microsatellite markers derived from ESTs for clustering analysis of 192 *Capsicum* lines, including 34 *C. annuum*, 21 *C. baccatum*, 85 *C. chinense*, 27 *C. frutescens*, 1 *C. pubescens*, and 24 lines of unidentified *Capsicum* spp., stocked at The Kihara Institute for Biological Research, Yokohama City University, Japan (Shirasawa et al. 2013). Given the small number of markers available in 2013, no association studies could be conducted at that time. However, with significant advancements in NGS approaches, it has been possible to perform genomewide SNP analysis on these 192 *Capsicum* lines using double-digest restriction-site associated DNA sequencing (ddRAD-Seq), as an alternative method to GBS (Peterson et al. 2012).

Using the ddRAD-Seq workflow established in our group (Shirasawa et al. 2016), we have reanalyzed the genotypes of all 192 *Capsicum* lines (Table 13.2). Several steps were involved in this process. First, genomic libraries were prepared using two restriction enzymes, *Pst*I and *Msp*I, and sequenced on an Illumina HiSeq2000 to obtain 93-bp paired-end reads. The sequences generated were deposited at the DDBJ Sequence Read Archive (DRA) database under the accession numbers DRA006931 and DRA006932. Next, low-quality reads were trimmed, and adapter sequences were removed. The remaining high-quality reads, approximately 1.2 million reads per sample, were mapped onto the *C. annuum* reference genome sequence, Pepper.v.1.55 (Kim et al. 2014), using Bowtie2 with an average alignment rate of 81.5%. These sequences were also mapped onto the reference genomes of *C. baccatum* (Baccatum.v.1.2) and *C. chinense* (Chinense.v.1.2) with an alignment rate of 84.6 and 88.9%, respectively. The alignment of sequence reads against Pepper.v.1.55 yielded a total of 14,444 high-quality SNPs; these SNPs were called using BCFtools after eliminating low-quality SNPs using VCFtools with criteria: including only sites with a minor allele frequency of $\geq$ 0.05; including only sites with a number of alleles of 2; including only


(a)
(b)

◀ **Fig. 13.1** Investigation of population structure among 192 *Capsicum* lines based on 14,444 single-nucleotide polymorphisms (SNPs). **a** Dendrogram showing the genetic distances among *Capsicum* lines calculated using the neighbor-joining method in TASSEL. Names of lines are indicated with the countries of origin (ARG: Argentina; BOL: Bolivia; BRA: Brazil; CHL: Chile; COL: Colombia; ECU: Ecuador; GTM: Guatemala; JPN: Japan; MEX: México; PER: Peru; and VEN: Venezuela). **b** Population structure of *Capsicum* lines determined in ADMIXTURE. Each color represents a distinct group. Each horizontal bar indicates a different *Capsicum* line

genotypes supported by reads of $\geq$ 5; including only sites with quality value of $\geq$ 999; excluding sites with missing data of $\geq$ 75%; and excluding sites that contain an indel.

A dendrogram based on SNPs revealed the clustering of 192 lines into four main species-specific groups (Shirasawa et al. 2013; Fig. 13.1). Additionally, investigation of population structure among the 192 lines using ADMIXTURE (Alexander et al. 2009) revealed $K = 11$, in accordance with ADMIXTURE's cross-validation error. This indicates that the 192 *Capsicum* lines cluster in four species-based groups, with 2, 2, 2, and 5 subclasses within the species *C. annuum*, *C. baccatum*, *C. frutescens*, and *C. chinense*, respectively (Fig. 13.1). Thus, the results of ADMIXTURE and dendrogram analyses were consistent.

Associations among genotypes and phenotypes have been analyzed using the mixed linear model (MLM) in TASSEL (Bradbury et al. 2007), taking into account the population structure and kinship matrix based on ADMIXTURE analysis and SNP data, respectively (Zhang et al. 2010). The threshold for association was set to $-\log10$ (P-value) > 5.46 at a significance level of 5% after Bonferroni multiple test correction. SNPs regulating *Capsicum* fruit orientation were identified in two loci: one at the 137,923,427-base position on chromosome 7 and another at the 508,932-base position on chromosome 10 (Fig. 13.2). Previously, the fruit orientation trait has been mapped on chromosome 12 based on QTL analysis (Lefebvre et al. 2002; Ogundiwin et al. 2005; Cheng et al. 2016) of biparental mapping populations derived from intra- and interspecific hybrids. Gopalakrishnan et al. (1989) showed that two recessive genes regulate upright fruit orientation, indicating that the genes conferring upright fruit habit depend on the genetic background. Further studies are needed to clarify the molecular mechanisms underlying these QTLs.

13.5 Future Perspectives

Owing to significant advances in sequencing technologies, high-throughput genomewide SNP analysis techniques, and molecular genetics approaches such as QTL analysis and GWAS, it is now possible to identify DNA markers tightly linked to agronomically important traits and to map the responsible genes. Accurate management of genetic resources is key for these analyses. Furthermore, genomic selection (GS), conceptualized by Meuwissen et al. (2001), facilitates rapid selection of superior genotypes, thus accelerating the breeding cycle (Crossa et al. 2017). In *Capsicum* spp., as in many other Solanaceae species (Yamamoto et al. 2016), GS has the potential to accelerate breeding programs and generate new cultivars with attractive agronomical traits using interspecific hybridization. To achieve this objective, breaking genetic barriers such as cross-incompatibility is essential. Alternatively, genome-editing techniques, such as clustered regularly interspaced palindromic repeats (CRISPR)-Cas9 technology and transcription activator-like effector nucleases (TALENs), have been effective in creating knockout/knockdown mutated alleles in the Solanaceae (Umemoto et al. 2016; Yamamoto et al. 2018; Van Eck 2018). Overall, genomics, genetics, and new plant breeding technologies promise the development of new attractive cultivars not only in peppers but also in other crop plants.

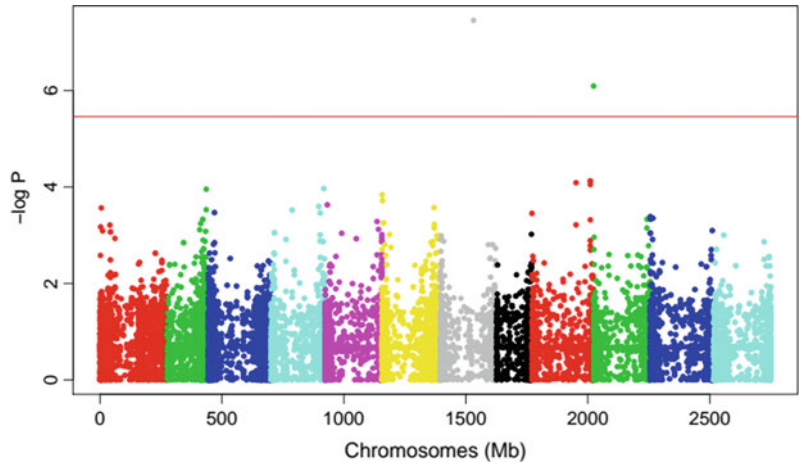


Fig. 13.2 Manhattan plot showing the genomewide association study (GWAS) of *Capsicum* fruit orientation. Different colors indicate 12 chromosomes of the *Capsicum* genome. Red line indicates a threshold of −log10 (*P*-value) = 5.46 at a significance level of 5% after Bonferroni multiple test correction

References

Ahn YK, Manivannan A, Karna S, Jun TH, Yang EY et al. (2018) Whole genome resequencing of Capsicum baccatum and Capsicum annuum to discover single nucleotide polymorphism related to powdery mildew resistance. Sci Rep 8:5188

Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. Genome Res 19:1655–1664

Bolger A, Scossa F, Bolger ME, Lanz C, Maumus F et al (2014) The genome of the stress-tolerant wild tomato species *Solanum pennellii*. Nat Genet 46:1034–1038

Bombarely A, Rosli HG, Vrebalov J, Moffett P, Mueller LA, Martin GB (2012) A draft genome sequence of *Nicotiana benthamiana* to enhance molecular plant-microbe biology research. Mol Plant Micr Interact 25:1523–1530

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633–2635

Cheng J, Zhao Z, Li B, Qin C, Wu Z, Trejo-Saavedra DL, Luo X, Cui J, Rivera-Bustamante RF, Li S, Hu K (2016) A comprehensive characterization of simple sequence repeats in pepper genomes provides valuable resources for marker development in *Capsicum*. Sci Rep 6:18919

Chhapekar S, Kehie M, Ramchiary N (2016) Advances in molecular breeding of *Capsicum* species. In: Deka PC (ed) Biotechnological tools for genetic resources. Daya Publishing House, New Delhi, pp 233–274

Crossa J, Pérez-Rodríguez P, Cuevas J, Montesinos-López O, Jarquín D et al (2017) Genomic selection in plant breeding: methods, models, and perspectives. Trends Plant Sci 22:961–975

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM et al (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nat Rev Genet 12:499–510

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K et al (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One 6:e19379

Gopalakrishnan TR, Gopalakrishnan PK, Peter KV (1989) Inheritance of clusterness and fruit orientation in chilli (*Capsicum annuum* L.). Indian J Genetics 49:219–222

Han K, Jeong HJ, Yang HB, Kang SM, Kwon JK et al (2016) An ultra-high-density bin map facilitates high-throughput QTL mapping of horticultural traits in pepper (*Capsicum annuum*). DNA Res 23:81–91

Han K, Lee HY, Ro NY, Hur OS, Lee JH et al (2018) QTL mapping and GWAS reveal candidate genes controlling capsaicinoid content in *Capsicum*. Plant Biotechnol J. https://doi.org/10.1111/pbi.12894

Hill TA, Ashrafi H, Reyes-Chin-Wo S, Yao J, Stoffel K et al (2013) Characterization of *Capsicum annuum* genetic diversity and population structure based on parallel polymorphism discovery with a 30 K unigene Pepper GeneChip. PLoS One 8:e56200

Hirakawa H, Shirasawa K, Miyatake K, Nunome T, Negoro S et al (2014) Draft genome sequence of eggplant (*Solanum melongena* L.): the representative *Solanum* species indigenous to the old world. DNA Res 21:649–660

Huang X, Feng Q, Qian Q, Zhao Q, Wang L et al (2009) High-throughput genotyping by whole-genome resequencing. Genome Res 19:1068–1076

Hulse-Kemp AM, Ashrafi H, Plieske J, Lemm J, Stoffel K et al (2016) A HapMap leads to a *Capsicum annuum* SNP infinium array: a new tool for pepper breeding. Hort Res 3:16036

Hulse-Kemp AM, Maheshwari S, Stoffel K, Hill TA, Jaffe D et al (2018) Reference quality assembly of the 3.5-Gb genome of *Capsicum annuum* from a single linked-read library. Hort Res 5:4

IPGRI, AVRDC, and CATIE (1995) Descriptors for *Capsicum* (*Capsicum* spp.). International Plant Genetic Resources Institute, Rome

Kim S, Park M, Yeom SI, Kim YM, Lee JM et al (2014) Genome sequence of the hot pepper provides insights into the evolution of pungency in *Capsicum* species. Nat Genet 46:270–278

Kim S, Park J, Yeom SI, Kim YM, Seo E et al (2017) New reference genome sequences of hot pepper reveal the massive evolution of plant disease-resistance genes by retroduplication. Genome Biol 18:210

Lefebvre V, Pflieger S, Thabuis A, Caranta C, Blattes A et al (2002) Towards the saturation of the pepper linkage map by alignment of three intraspecific maps including known-function genes. Genome 45:839–854

Liedl BE, Labate JA, Stommel JR, Slade A, Kole C (2013) Genetics, genomics and breeding of tomato. CRC Press, Boca Raton

Meuwissen TH, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829

Mimura Y, Inoue T, Minamiyama Y, Kubo N (2012) An SSR-based genetic map of pepper (*Capsicum annuum* L.) serves as an anchor for the alignment of major pepper maps. Breed Sci 62:93–98

Minamiyama Y, Tsuro M, Hirai M (2006) An SSR-based linkage map of *Capsicum annuum*. Mol Breed 18:157–169

Moscone EA, Baranyi M, Ebert I, Greilhuber J, Ehrendorfer F et al (2003) Analysis of nuclear DNA content in *Capsicum* (Solanaceae) by flow cytometry and feulgen densitometry. Ann Bot 92:21–29

Nimmakayala P, Abburi VL, Saminathan T, Alaparthi SB, Almeida A et al (2016a) Genome-wide diversity and association mapping for capsaicinoids and fruit weight in *Capsicum annuum* L. Sci Rep 6:38081

Nimmakayala P, Abburi VL, Saminathan T, Almeida A, Davenport B et al (2016b) Genome-wide divergence and linkage disequilibrium analyses for *Capsicum baccatum* revealed by genome-anchored single nucleotide polymorphisms. Front Plant Sci 7:1646

Ogundiwin EA, Berke TF, Massoudi M, Black LL, Huestis G (2005) Construction of 2 intraspecific linkage maps and identification of resistance QTLs for *Phytophthora capsici* root-rot and foliar-blight diseases of pepper (*Capsicum annuum* L.). Genome 48:698–711

Paran I, Ben-Chaim A, Kang BC, Jahn M (2007) Capsicums. In: Kole C (ed) Genome mapping and molecular breeding in plants, vol 5. Springer, New York, pp 209–226

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS One 7:e37135

Poland JA, Rife TW (2012) Genotyping-by-sequencing for plant breeding and genetics. Plant Genome 5:92–102

Qin C, Yu C, Shen Y, Fang X, Chen L et al (2014) Whole-genome sequencing of cultivated and wild peppers provides insights into *Capsicum* domestication and specialization. Proc Natl Acad Sci USA 111:5135–5140

Razali R, Bougouffa S, Morton MJL, Lightfoot DJ, Alam I et al (2017) The genome sequence of the wild tomato *Solanum pimpinellifolium* provides insights into salinity tolerance. Biorxiv. https://doi.org/10.1101/215517

Shirasawa K, Ishii K, Kim C, Ban T, Suzuki M et al (2013) Development of *Capsicum* EST-SSR markers for species identification and *in silico* mapping onto the tomato genome sequence. Mol Breed 31:101–110

Shirasawa K, Hirakawa H, Isobe S (2016) Analytical workflow of double-digest restriction site-associated DNA sequencing based on empirical and *in silico* optimization in tomato. DNA Res 23:145–153

Sierro N, Battey JN, Ouadi S, Bovet L, Goepfert S et al (2013) Reference genomes and transcriptomes of *Nicotiana sylvestris* and *Nicotiana tomentosiformis*. Genome Biol 14:R60

Sierro N, Battey JN, Ouadi S, Bakaher N, Bovet L et al (2014) The tobacco genome sequence and its comparison with those of tomato and potato. Nat Commun 5:3833

Sugita T, Semi Y, Sawada H, Utoyama Y, Hosomi Y et al (2013) Development of simple sequence repeat markers and construction of a high-density linkage map of *Capsicum annuum*. Mol Breed 31:909–920

Taranto F, D'Agostino N, Greco B, Cardi T, Tripodi P (2016) Genome-wide SNP discovery and population structure analysis in pepper (*Capsicum annuum*) using genotyping by sequencing. BMC Genom 17:943

The Potato Genome Sequencing Consortium (2011) Genome sequence and analysis of the tuber crop potato. Nature 475:189–195

The Tomato Genome Consortium (2012) The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485:635–641

Umemoto N, Nakayasu M, Ohyama K, Yotsu-Yamashita M, Mizutani M et al (2016) Two cytochrome P450 monooxygenases catalyze early hydroxylation steps in the potato steroid glycoalkaloid biosynthetic pathway. Plant Physiol 171:2458–2467

Van Eck J (2018) Genome editing and plant transformation of solanaceous food crops. Curr Opin Biotechnol 49:35–41

Varshney RK, Graner A, Sorrells ME (2005) Genomics-assisted breeding for crop improvement. Trends Plant Sci 10:621–630

Walsh BM, Hoot SB (2001) Phylogenetic relationships of Capsicum (Solanaceae) using DNA sequences from two noncoding regions: the chloroplast atpB-rbcL spacer region and nuclear waxy introns. Int J Plant Sci 162:1409–1418

Wu F, Eannetta NT, Xu Y, Durrett R, Mazourek M et al (2009) A COSII genetic map of the pepper genome provides detailed picture of synteny with tomato and new insights into recent chromosome evolution in the genus *Capsicum*. Theor Appl Genet 118:1279–1293

Yamamoto E, Matsunaga H, Onogi A, Kajiya-Kanegae H, Minamikawa M et al (2016) A simulation-based breeding design that uses whole-genome prediction in tomato. Sci Rep 6:19454

Yamamoto T, Kashojiya S, Kamimura S, Kameyama T, Ariizumi T et al (2018) Application and development of genome editing technologies to the Solanaceae plants. Plant Physiol Biochem. https://doi.org/10.1016/j.plaphy.2018.02.019

Yi G, Lee JM, Lee S, Choi D, Kim BD (2006) Exploitation of pepper EST-SSRs and an SSR-based linkage map. Theor Appl Genet 114:113–130

Zhang Z, Ersoz E, Lai CQ, Todhunter RJ, Tiwari HK et al (2010) Mixed linear model approach adapted for genome-wide association studies. Nat Genet 42:355–360

Sequence Databases and Online Tools for Capsicum Research

14

Feng Li, Minglei Liu and Yingjia Zhou

Abstract

Pepper is an important vegetable crop in terms of economic value and unique biological features, such as diversity in fruit morphology and capsaicin biosynthesis. Understanding pepper gene function is essential to explore the mechanisms how these unique features were developed in pepper and improve pepper agricultural traits with modern biotechnologies. Pepper has a very big genome of more than 3.5 Gb nucleotides, which is even more than the human genome. Maturation of high-throughput DNA sequencing technologies has made it possible for researchers to start decoding the pepper genome and investigate gene expression pattern genome-wide. In this chapter, we summarize various sequence databases generated for pepper studies and various online tools that are available for pepper research. Currently, six capsicum draft genomes and a pangenome with 383 pepper cultivars were published. These genomes were sequenced by Illumina-sequencing platform and provided very important foundation for further genome and transcriptome studies. Several hundred pepper transcriptome databases including both mRNA and small RNA were generated using sanger, 454 and Illumina platforms for different research objectives ranging from marker discovery, hormone signaling, pathogen resistance to database construction. Finally, we introduced several online resources for pepper genome, transcriptome data access and analyses. These sequence data and online tools will be useful for dissecting pepper gene function. Even though, significant progress has been made in generating pepper sequence databases and online tools in the past few years, large gaps exist comparing to the data and tools available for other plant species. More attention needs to be paid into this line of research in future for the pepper research community.

14.1 Capsicum Genome and Resequencing Databases

Pepper is an economically important vegetable crop and belongs to the *Capsicum* genus of Solanaceae family. Cultivated pepper was mainly domesticated from five species of the genus: *Capsicum annuum*, *C. baccatum*, *C. chinense*, *C. frutescens,* and *C. pubescens*. Their genome size was estimated ranging from 3.31 to 4.38 Gb (*Ca*3.31, *Cb*3.61, *Cc*3.35, *Cf*3.33, and *Cp*4.38) (Moscone et al. 2003). In the past four years,

F. Li (✉) · M. Liu · Y. Zhou
Key Laboratory of Horticultural Plant Biology (MOE), College of Horticultural and Forestry Sciences, Huazhong Agricultural University, Wuhan, China
e-mail: chdlifeng@mail.hzau.edu.cn

N. Ramchiary and C. Kole (eds.), *The Capsicum Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-319-97217-6_14

Table 14.1 Comparison of published pepper genomes

	Capsicum annuum				*C. baccatum*	*C. chinense*
Variety/accession	*UCD-10X-F1*	*CM334*	*Zunla-1*	*glabriusculum*	*PBC81*	*PI159236*
Genome version	UCD10X v1	CM334 v2.0	v2.0	v2.0	v1.2	v1.2
Est. genome size (Gb)		3.48	3.26	3.07	3.90	3.20
Total base (Gb)	3.21	3.06	3.35	3.48	3.20	3.00
Anchored (Gb)	2.67	2.90	2.64	2.42	2.82	2.81
N50 contig (kb)	123.00	30.00	55.44	52.23	39.00	50.00
N50 scaffold (Mb)	3.69	2.47	1.23	0.45	2.00	3.30
genes annotated		35,884	35,336	34,476	35,874	35,009
Availability	PRJNA376668	peppergenome.snu.ac.kr	peppersequence.genomics.cn		peppergenome.snu.ac.kr	
Chromosome 1 (Mb)	256	309	301	338	259	241
Chromosome 2 (Mb)	154	170	164	163	169	170
Chromosome 3 (Mb)	271	283	262	265	298	275
Chromosome 4 (Mb)	232	240	216	179	219	232
Chromosome 5 (Mb)	221	239	217	195	228	237
Chromosome 6 (Mb)	228	242	220	201	253	251
Chromosome 7 (Mb)	227	251	222	221	255	127
Chromosome 8 (Mb)	174	142	153	83	210	196
Chromosome 9 (Mb)	219	271	239	179	197	258
Chromosome 10 (Mb)	222	233	206	202	230	233
Chromosome 11 (Mb)	236	267	220	229	264	251
Chromosome 12 (Mb)	233	251	230	197	237	226
References	Hulse-Kemp et al. (2018)	Kim et al. (2014)	Qin et al. (2014)		Kim et al. (2017)	

draft genomes for three of these species were published (Table 14.1; Hulse-Kemp et al. 2018; Kim et al. 2014, 2017; Qin et al. 2014).

Four genomes were reported for *C. annuum* species including two cultivated varieties, Criollos de Morelos 334 (CM334) and Zunla1 (Kim et al. 2014; Qin et al. 2014), one F_1 hybrid from cross between CM334 and a nonpungent blocky pepper-breeding line (Hulse-Kemp et al. 2018), and a progenitor variety of Zunla1, Chiltepin (Qin et al. 2014). All these genomes are sequenced by a next-generation sequencing-based whole genome shotgun approach. The size of sequenced *C. annuum* genomes ranged from 3.06 to 3.48 Gb (Table 14.1), which is close to the previously estimated total nuclear genome size of 3.31 Gb (Moscone et al. 2003). The N50 contig size ranges from 30 to 123 kb, and N50 scaffold size ranges from 0.45 to 3.69 Mb (Table 14.1). Numbers of annotated protein-coding genes are all around 35,000 (Table 14.1). Majority of the contigs in these genomes were anchored to *12* chromosomes with the total anchored bases ranging from 2.42 to 2.9 Gb. The overall size distribution of chromosomes is similar among different assemblies (Table 14.1). The assembled chromosome *1* has the largest size in three of the four genomes with up to 338 Mb in *Chiltepin*. Chromosome *2* and *8* have much smaller sizes ranging from 154 to 170 Mb and 83 to 174 Mb, respectively (Table 14.1). Other chromosomes mostly consist of 220 to 280 Mb nucleotides (Table 14.1). In terms of chromosome sizes, the Zunla-1 and CM334 assemblies showed closest distribution in different chromosomes (CoV > 0.95).

One genome is published for *C. baccatum PBC81* (Kim et al. 2017). Its size is 3.20 and

about 2.82 Gb bases were anchored onto 12 chromosomes (Table 14.1). The sizes of N50 contig and N50 scaffold are 39 kb and 2.00 Mb, respectively. About 36,000 protein-coding genes are annotated (Table 14.1). One genome is published for *C. chinense PI159236*, with size of 3.00 Gb sequenced bases (Kim et al. 2017). The *C. chinense* genome database consists of 3.00 Gb nucleotides, among which 2.81 Gb were anchored to 12 chromosomes (Table 14.1). The N50 contig and N50 scaffold lengths are 50 kb and 3.30 Mb, respectively. About 35,000 protein-coding genes are annotated (Table 14.1). The chromosome *2* of both *C. baccatum* and *C. chinense* and the chromosome *7* of *C. chinense* are quite small while other chromosomes are quite even in length (Table 14.1). Both *C. baccatum* and *C. chinense* genomes are available at author's website (Table 14.1).

The assembled genomes are available through author's own websites or public databases, such as NCBI, Solgenomics website, or PepperHub (Table 14.1; Liu et al. 2017a). Homologous gene search for pepper by BLAST program is available at Solgenomics, NCBI, and PepperHub. NCBI and PepperHub also offer pepper gene structure browsing.

More recently, a pepper pangenome was published based on resequencing data using 355 cultivars from *C. annuum* and 28 cultivars from *C. chinense*, *C. baccatum* and *C. frutescens* species (Ou et al. 2018). The pepper pangenome contains 51,757 high-quality genes, among which 28840 genes were shared by all four *Capsicum* species and thus were considered core capsicum genes. A pepper pangenome website (http://www.pepperpan.org:8012/) was set up to allow users search the genetic variation at their locus of interest.

All the above *Capsicum* genomes were sequenced using Illumina next-generation sequencing platform with multiple-size insert genomic DNA libraries. Considering recent advancement in sequencing technologies in recent years, such as single molecular sequencing (Ardui et al. 2018), Bionanotechnology (Stankova et al. 2016) and Nanopore technology (Michael et al. 2018), which generate much longer sequencing reads, and new sequencing strategies, such as 10× genomics (Spies et al. 2017) and Hi-C (Korbel and Lee 2013), which significantly help improve sequence assembly; there is much room to improve the pepper genome sequencing in near future.

14.2 Capsicum Transcriptome Databases

To date, transcriptome study in pepper is very limited compared to other crops. There are about 30 published pepper transcriptome papers with different research objectives, which include nearly 30 pepper varieties (Table 14.2). In the early time, Sanger sequencing-based expressed sequence tag (EST) approach was adopted for a few studies, to generate pepper transcriptome databases for gene function study, marker discovery, and making microarray chip (Table 14.2). Most of these transcriptome studies were done in the past five-to-six years using next-generation sequencing techniques with aims for marker discovery or investigation of specific biological question (Table 14.2).

The first pepper transcriptome database was reported 14 years ago using EST approach. About 8200 ESTs were sequenced from pepper leaves infiltrated with nonhost pathogen *Xanthomonas axonopodis* pv. *glycines*, flower buds, and anthers. Based on these EST sequences, microarray chip was developed and applied to profile gene expression during nonhost resistance reaction. Hundred DEGs were identified including CDPK, bZIP transcription factor, and genes involved in hormone biosynthesis (Lee et al. 2004). Later, a large-scale EST sequencing project was reported with 122,582 sequenced ESTs and 116,412 refined ESTs from 21 pepper EST libraries to investigate the complexity of pepper transcriptome. An online pepper EST database (http://genepool.kribb.re.kr/pepper/) was set up for users to identify genes in pepper plants, analyze gene expression patterns, and comparing the ESTs with those of other members of the *Solanaceae* family (Kim et al. 2008). More recently, a Capsicum transcriptome database

Table 14.2 List of pepper varieties studied in transcriptome analysis

Pepper variety	Methodology	Availability	Research aim	References
Bukang	EST	NCBI-EST	Database	Kim et al. (2008)
Bukang and HangKeun	EST	NCBI-EST	Nonhost resistance	Lee et al. (2004)
Serrano Tampiqueno 74 and Sonora Anaheim	EST and 454 sequencing	NA*	Database	Gongora-Castillo et al. (2012a)
Bukang and ECW	Microarray	NA	Abiotic stress and hormone signaling	Lee and Choi (2013)
Super Bigarim	Microarray	NA	Abiotic stress and hormone signaling	Shin et al. (2017)
Bukwang	Microarray	NA	Pepper–white fly interaction, systemic signaling	Park and Ryu (2014)
Maor, Early Jalapeno and CM334	mRNA-seq	SRR495602-7	Marker discovery	Ashrafi et al. (2012)
Saengryeg 211 and 213	mRNA-seq	ERR179754-5	Marker discovery	Ahn et al. (2013)
Mandarin and Blackcluster	mRNA-seq	ERP001872/5	Marker discovery	Ahn et al. (2014)
Yolo Wonder and CM334	mRNA-seq	NA	Marker discovery	Nicolai et al. (2012)
TF68 (YCM334 × Taean)	mRNA-seq	GSE29215	Marker discovery	Lu et al. (2012)
C. frutescens	mRNA-seq	SRR387333	Capsaicinoid biosynthesis, marker discovery	Liu et al. (2013b)
Tampiqueno 74	mRNA-seq	SRR1119016-39	Fruit development	Martinez-Lopez et al. (2014)
CM334	mRNA-seq	SRR5506999-7025 SRR6109696-907 SRR6122654-703	Fruit development and pathogen infection	Kim et al. (2018)
MSL8214A	mRNA-seq	SRR6002840	Male sterility	Qiu et al. (2018)
CMS121A and 121C	mRNA-seq	SRR802896 SRR802919	Male sterility	Liu et al. (2013a)
Xiangyan-16	mRNA-seq	SRX1959970	Abiotic stress and hormone signaling	Li et al. (2016)
CM334 and ECW30R	mRNA-seq	PRJNA385363	Pepper-P. infestans nonhost resistance	Lee et al. (2017)
(C. f PI290972 × C.a Mazurka) × Walock (3×)	mRNA-seq	NA	CaCV resistant trait cloning	Widana-Gamage et al. (2016)
PI235047 and PI 585270	mRNA-seq	NA	Bs4C-R cloning	Strauss et al. (2012)
Sonora Anaheim	mRNA-seq	SRR653187-95	PGMV infection recovery	Gongora-Castillo et al. (2012b)
PI6214	mRNA-seq	PepperHub**	Database construction	Liu et al. (2017a)

NA*, not available; PepperHub**, available at pepperhub.hzau.edu.cn

(http://www.bioingenios.ira.cinvestav.mx:81/Joomla/) was established to allow users search, download pepper transcriptome sequences and annotations. This database consists of transcripts assembled from EST and pyrosequencing data derived from different tissues of *C. annuum* varieties, Serrano Tampiqueño 74, and Sonora Anaheim, including 32,314 high-quality contigs and 51,118 singletons (Gongora-Castillo et al. 2012a). These databases provided valuable resources for pepper research in early time, but unfortunately, they are no longer accessible.

Molecular markers are very useful in pepper breeding and mapping functional genes. To facilitate marker discovery, several research groups conducted transcriptome sequencing using several pepper cultivars (Table 14.2). *De novo* transcriptomes were sequenced by Illumina genome analyzer and assembled by de novo assembly approach for pepper varieties Maor, Early Jalapeno and Criollo de Morelos-334 (CM334). Around 76,000–83,000 contigs with N50 length about 1.5 kb were assembled for each variety. By alignment to a reference sequence assembly from over 125,000 EST sequences of an F_1-hybrid line Bukang, a total of 22,863 putative SNPs were obtained (Ashrafi et al. 2012). For *C. annuum* varieties Saengryeg 211, Saengryeg 213, Mandarin and Blackcluster, transcriptomes were sequenced by 454 pyrosequencing platforms, and 120 to 140 Mb nucleotide data were generated for each cultivar, assembled into 30,000–40,000 transcripts. Sequence variant analyses identified 3766 and 2431 simple sequence repeat markers for Saengryeg 211 and Saengryeg 213, respectively, 1025 and 1059 genotype-specific SNPs for Mandarin and Blackcluster, respectively (Ahn et al. 2013, 2014).

Transcriptome from Yolo Wonder, a big sweet pepper variety and CM334 were sequenced with 454 and IGA platform, respectively. Sequencing of Yolo Wonder yielded 175 Mb data which was assembled into 23,748 contigs and 60,370 singletons. CM334 transcriptome sequencing generated 2513 Mb raw data and assembled into 128,504 contigs and 90,185 singletons. Over 11,000 reliable SNPs were found in nearly 6000 genes (Nicolai et al. 2012). Similarly, transcriptome from a mature fruit of a red pepper TF68 (F_1 progeny of YCM334 and Taean) were sequenced using 454 sequencing techniques, which generated about 30 Mb data and assembled 33,530 unigenes, and identified about 2000 SNP and SSR markers (Lu et al. 2012).

Different plant tissues and organs are composed of cells differentiated from single stem cells, but those cells have different morphologies and functions. The difference was not due to different sets of genes they contain but was largely due to different sets of genes expressed in different cell types. For similar reason, same cell type may accumulate different metabolite at different developmental stages, for example, in pepper pericarp cells, different pigments accumulate at different stages during fruit maturation process. These phenomena suggest that developmental programs and biochemical pathways are controlled by unique pattern of gene expression in different tissues at different developmental stages, thus transcriptome analysis provides a powerful tool to dissect developmental and biochemical mechanisms by investigation of gene expression pattern at genome-wide scale. In pepper, a few transcriptome studies have been conducted aiming to understand male sterility, fruit development and capsaicinoid biosynthesis (Table 14.2).

To better characterize the capsaicinoid biosynthesis pathway, Liu and colleague conducted sequencing and *de novo* assembly of *C. frutescens* transcriptome and obtained a total 54,045 high-quality unigenes (transcripts) using Trinity software (Liu et al. 2013b). They predicted three new structural genes (*DHAD*, *TD*, *PAT*), which filled gaps of the capsaicinoid biosynthetic pathway predicted by Mazourek et al. (2009) and revealed new candidate genes involved in capsaicinoid biosynthesis based on KEGG (Kyoto Encyclopedia of Genes and Genomes) analysis (Liu et al. 2013b; Mazourek et al. 2009). For analysis of the dynamics of gene expression during pepper flower and fruit development, transcriptomes of flower buds and fruits at different developmental stages were

sequenced for *C. annuum* varieties Tampiqueno 74, Zunla1, CM334 and 6421 (Kim et al. 2014, 2018; Liu et al. 2017a; Martinez-Lopez et al. 2014; Qin et al. 2014). Martinez-Lopez et al. (2014) observed that genes related to capsaicinoid and ascorbic acid biosynthesis were significantly upregulated at 20 days after anthesis (DAA) while those related to carotenoid biosynthesis were highly expressed in the last period of fruit maturation (40–60 DAA). Possible Myb transcription factors involving regulation of capsanthin biosynthesis were also identified (Liu et al. 2017a).

Male sterile lines (MSLs) are commonly used in pepper seed production. Toward understanding the mechanism for male sterility, comparative transcriptome analyses were carried out for MSL 8214A and its maintainer line 8214B, as well as MSL 121A and its restorer line 121C. Differential gene expression analysis suggested that ubiquitin ligase genes, cell cycle regulators, ATP synthase subunits, cytochrome oxidase, and PPR gene family may involve in pepper male sterility (Liu et al. 2013a; Qiu et al. 2018).

Pepper production is challenged by abiotic stress, such as cold, drought and salt, and hormone signaling plays important role in plant response to abiotic stress. To investigate the mechanism by which pepper response to different abiotic stresses, pepper transcriptome analyses were performed using both microarray and mRNA sequencing methods on several pepper varieties that were subjected to stress and hormone treatments (Table 14.2). Diverse sets of differentially expressed genes (DEG) specific to each treatment or in common were identified (Lee and Choi 2013) and transcription factors, such as *NAC*, *WRKY40* and *ERFs*, were found potentially enhancing chilling tolerance (Shin et al. 2017). It was shown that 24-Epibrassinolide (EBR) treatment also increase pepper tolerance to chilling (Li et al. 2015). In a follow-up study, transcriptomes from EBR + Chill treated and chill treated control were sequenced and compared. The results showed that EBR treatment upregulated photosynthesis and redox-related genes which is consistent with its role in maintenance of photosystem II and enhancement of antioxidation system in chilling stress (Li et al. 2016). It would be interesting to investigate the aforementioned DEGs and transcription factor's involvement in the EBR-mediated chilling tolerance.

For transcriptome profiling of immune response to multiple pathogens, pepper leaves were inoculated with *Phytophthora infestans* (*Pi*), pepper mottle virus, and tobacco mosaic virus P0 strain. Inoculated leaves harvested at multiple time points from three biological replicates were ground in liquid nitrogen, which was used for total RNA purification and transcriptome sequencing. These data provided valuable resources for further investigation of interaction between pepper and these pathogens (Kim et al. 2014, 2018). Recently, time course study of transcriptomes was performed on detached leaves infiltrated with potato late blight (*Pi T30-4*). A subset of *EAS*/*EAH* gene was specifically induced which are involved in pepper phytoalexin capsidiol synthesis. *EAS*/*EAH* belongs to a multigene family, and these induced genes are members that specifically expanded in capsicum species (Lee et al. 2017). These results indicated that up-regulation of nutrient transporter and auxin response gene expression in root by whitefly infestation may contribute to the increase of root biomass (Park and Ryu 2014).

It was reported earlier that whitefly infestation on pepper (*C. annuum* L. cv. Bukwang) leaves led to an increase of biomass in root (Yang et al. 2011). Transcriptome analysis was conducted using microarray technology to investigate the mechanism involving communication between leaf and root in pepper–whitefly interaction. The results indicated that upregulation of nutrient transporters and auxin response genes in root by whitefly infestation may underlie increased root biomass (Park and Ryu 2014).

Transcriptome analysis has also been employed to identify disease resistance gene and study mechanism of plant–virus interaction in pepper. It was very successful in determining the candidate *Bs4C* gene that is activated by *AvrBs4* TAL effector from *Xanthomonas*, which set up an exemplary case using RNA-seq to clone TAL-specific R genes from large-genome crops

(Strauss et al. 2012). Similar approach was applied to identify candidate resistance gene against Capsicum chlorosis virus (CaCV) that was introgressed from *Capsicum chinense* into a bell pepper (*C. annuum*). Several candidate CNL genes were also identified which may underlie the CaCV resistance (Widana-Gamage et al. 2016). To investigate the mechanism of pepper recovery from pepper golden mosaic virus (PepGMV) infection, transcriptome analysis was conducted for healthy, symptomatic, and recovered pepper leaves following pepper golden mosaic virus infection. Differential expression gene analysis identified several pathogenesis-related (PR) genes and plant hormone signaling genes which may be important for the recovery phenotype (Gongora-Castillo et al. 2012b). Similar analyses were performed on CMV-infected pepper leaves and around 2000 DEGs were identified, including several key genes commonly seen induced upon pathogen inoculation such as chitinase, PR protein, TMV resistance protein, WRKY transcription factor, and jasmonate ZIM-domain protein. (Zhu et al. 2018)

14.3 Capsicum Small RNA Databases

Small (s)RNAs are 20–24 nt long single-stranded RNAs that associate and guide Argonaute protein to form RNA-induced silencing complex (RISC). RISC mediates target gene silencing via sequence complementarity between target RNA and sRNA. SRNAs include miRNAs that are processed from hairpin structure of single-stranded RNA precursors and siRNAs that are processed from double-stranded RNA precursors (Cui et al. 2017). miRNAs play important roles in plant development and response to biotic and abiotic stresses, and thus miRNA caught more and more attention in different research fields (Sunkar et al. 2012). miRNA discovery usually involves sequencing of small RNA population from a total RNA sample, mapping the sequencing reads to genome database, retrieving sRNA flanking sequences, and analyzing their secondary structures (Li et al. 2012; Wu et al. 2012). Targets of miRNA are usually predicted with bioinformatics tools based on sequence complementarity; however, this approach usually results in high false-positive rate. Usually, miRNA-programed RISC cleaves target RNA at a defined position and generates a cleavage product with 5′ monophosphate group, which can be captured by a modified 5′ RACE experiment (also called degradome, or dRNA sequencing) (Li et al. 2012; Wu et al. 2012). Predicted miRNA targets with dRNA reads mapped to the predicted cleavage site are usually of high confidence. Thus, miRNA discovery and functional characterization heavily depend on sequencing and bioinformatics analyses.

Pepper miRNA research was started relatively late compared to other plant species, and the first report on pepper miRNA was published seven years ago. Using a bioinformatics approach, Kim and colleague identified 11 miRNAs (miR156, 164, 172, 414, 472, 855, 1023, 1074, 1320, 1428, 2093) and their 54 targets from pepper EST databases (Kim et al. 2011). A few small RNA sequencing studies were conducted using several *C. annuum* varieties, such as Luosijiao and 06J19-1-1-1-2 (Liu et al. 2017c), CM334 (Hwang et al. 2013; Kim et al. 2014) and Zunla-1 (Qin et al. 2014). These studies characterized 43 families of conserved miRNAs (Table 14.3) and several hundred novel miRNAs (Liu et al. 2017c). The small RNA databases generated from these studies provided valuable resources to analyze tissue-specific expression and function of pepper miRNAs. More recently, MiRTrans, a multiomics approach was developed for miRNA and target analysis and applied for pepper miRNA-target identification. Lending support from regression analysis of miRNA-target expression and mapping dRNA reads to predicted cleavage sites, MiRTrans identified 58 miRNA-transcript pairs with high confidence from 18 miRNA families conserved in eudicots. Most of these targets were transcription factors (Zhang et al. 2017).

There are some integrated miRNA databases host pepper miRNAs, such as miRNEST and PNRD (Szczesniak and Makalowska 2014; Yi et al. 2015), where users can browse pepper

Table 14.3 Distribution of miRNA family on different chromosomes

miR family	Member no.	Number of family members on each chromosome											
		Ch01	Ch02	Ch03	Ch04	Ch05	Ch06	Ch07	Ch08	Ch09	Ch10	Ch11	Ch12
miR156	10		1	2			2	2		2			1
miR159	4	2		2									
miR160	2					1				1			
miR162	8			8									
miR164	5			1			1		1	1	1		
miR166	10	1		3	1		1		2		2		
miR167	4			3			1						
miR168	3	2											1
miR169	24	2	1	1	1		1	11	5		2		
miR171	11	1	3	1	1		2	1		2			
miR172	10				1	3	2				1	1	2
miR319	9	1		2			2					2	2
miR390	3						1			1	1		
miR393	2			1		1							
miR394	3				1				1				1
miR395	12											12	
miR396	4							2		2			
miR397	3							3					
miR398	5											4	1
miR399	8	1		3			1				3		
miR403	1	1											
miR408	1								1				
miR477	2								2				
miR482	10				3		4			1	2		
miR536	1	1											
miR827	1												1
miR1446	4			4									
miR1507	1				1								
miR2873	3							1					2
miR3627	1		1										
miR4376	1						1						
miR4414	2			1			1						
miR5300	2					1							1
miR5301	1				1								
miR5303	2		1				1						
miR6022	3	2							1				
miR6023	1												1
miR6025	1					1							
miR6026	1	1											
miR6149	1					1							
Total	180	15	7	32	10	8	21	20	13	10	12	19	13

miRNA, miRNA*, and pre-miRNA sequences, as well as their targets. However, the number of miRNA entries is very limited. For miRNA prediction, psRNATarget webserver (Dai et al. 2018) is a good option. In this sever, users can upload small RNA database and discover miRNA precursors from server integrated pepper EST and unigene databases. Users can also upload pepper transcriptome sequences and identify potential targets that are regulated by miRNAs registered in miRbase.

14.4 Online Resources for Capsicum Research

The National Center for Biotechnology Information website (www.ncbi.nlm.nih.gov) provides a general platform for all biological research disciplines, where we can find literature, nucleic acid and protein-sequence information, and gene expression datasets for pepper research. The online blast tools hosted on NCBI website allow users to search deposited pepper sequences by specifying "Capsicum" in the "Organism" option. NCBI genome database hosts all six published pepper genomes with download links for fasta format genome and protein sequences, and annotation files in GFF, Genbank or tabular format. NCBI also provides dedicated BLAST page for users to search the published genomes and a browser to browse gene structure in a genome region. Similar functions are also found in the dedicated genome website maintained by the authors (Table 14.1).

The Sol Genomics Network (www.solgenomics.net) is a more specialized online server that provides various genomic resources for major *Solanaceous* crops. It offers BLAST search for pepper genome, FTP download of genome sequences and annotation, as well as Jbrowse function for browsing gene structure in any genome region.

PepperHub is a dedicated informatics server for pepper research (Liu et al. 2017a). Its genome module offers Gbrowse function for browsing pepper gene structure and BLAST function for search pepper genome, mRNA, and protein sequences. In addition to the function provided by NCBI, authors' genome websites and Sol Genomics Network, PepperHub genome module provides a blastdbcmd function which allow users to retrieve pepper genome, mRNA and protein sequences with a given ID and specified region and strand. The PepperHub transcriptome module provides four user-friendly bioinformatics tools for users to analyze pepper gene expression and do coexpression analysis. It also hosts large volume of transcriptome data organized into different experiment sets. By choosing different sets of data, users can easily do tissue-specific gene expression analysis and analyze dynamic changes of gene expression in different treatments. The gene expression data can be downloaded as tabular format or visualized as plant structure cartoon, heatmap, or line chart. The small RNA module allows users to browse published pepper miRNAs and search for their targets (Liu et al. 2017c).

Genome editing with clustered regularly interspaced palindromic repeat (CRISPR) associated protein (CAS) becomes powerful tools in gene function studies (Schindele et al. 2018). Because CAS proteins require a PAM sequence adjacent to the target DNA sequence identical to the CRISPR spacer sequences, bioinformatics tools are needed to help users choose proper target sequences in a given genome and analyze their potential offtargets. Webserver CRISPR-P is created for such purpose in plant research, but Capsicum genome is currently not supported on this server (Liu et al. 2017b). Thus, we recently add a genome-editing module to PepperHub for users to design CAS9-mediated genome-editing site. This module contains three functional tools including "spacerselector", "offtargets," and "gRNAdesigner". On the spacerselector page, user can input a fasta format of pepper gene sequences and click "submit". And then, the server will return a list of spacer sequences with adjacent PAM sequences, as well as their position in query sequences and strandness. It also lists potential offtarget sites for each designed spacer sequences. The spacer sequences start

either with a 5′ A or G, which are designed for expression from a U3 or U6 promoter. The "offtarget" function provides option for users to check the potential offtargets for a given spacer sequences. Users can input the fasta format of the spacer sequences, and output is similar as the offtargets listed in the "spacerselector" function. The "gRNAdesigner" function can help user to design primers to synthesize a gRNA expression unit driven by tomato U6 promoter. The designed expression unit can be cloned into a binary vector with SpeI and XhoI sites, and multiple gRNA expression units can be cloned into the same vector.

The Kyoto Encyclopedia of Genomes and Genes (KEGG) recently added a pepper genome on its website (https://www.genome.jp/dbget-bin/www_bget?gn:T04646) (Kanehisa et al. 2018). It provides a very good source for annotation of pepper gene function. In total, 37483 pepper genes were annotated and 277 pathways were mapped. However, the downside is that KEGG uses its own gene ID to label each pepper gene, and it is not convenient to convert them into CM334 or Zunla gene IDs which are currently more widely used. Some tools should be developed to allow cross-reference among different genome annotation databases

14.5 Future Aspects

Pepper is an important vegetable crop, but the current research status in pepper does not match its significance in agriculture. Much work is needed to improve quality of the genome sequence database. More transcriptome data including mRNA and small RNA expression data are needed to dissect pepper gene function. For the better use of the published transcriptome data, it is necessary to put them all together in the server such as PepperHub to connect the data directly to analysis tools that everyone can use. Currently, the small RNA function in PepperHub cannot allow prediction of miRNA precursor or target. These functions were developed for other plant genomes in psRobot and SoMART webservers (Li et al. 2012; Wu et al. 2012), but pepper genome was not included in this server. These functions will be added in future update of PepperHub.

Acknowledgements This work is funded by the start fund provided by Huazhong Agricultural University to Li F (2013RC001).

References

Ahn YK, Tripathi S, Cho YI, Kim JH, Lee He, Kim DS, Woo JG, Cho MC (2013) De novo transcriptome assembly and novel microsatellite marker information in *Capsicum annuum* varieties Saengryeg 211 and Saengryeg 213. Bot Stud 54:58

Ahn YK, Tripathi S, Kim JH, Cho YI, Lee He, Kim DS, Woo JG, Cho MC (2014) Transcriptome analysis of Capsicum annuum varieties Mandarin and Blackcluster: assembly, annotation and molecular marker discovery. Gene 533:494–499

Ardui S, Ameur A, Vermeesch JR, Hestand MS (2018) Single molecule real-time (SMRT) sequencing comes of age: applications and utilities for medical diagnostics. Nucl Acids Res 46:2159–2168

Ashrafi H, Hill T, Stoffel K, Kozik A, Yao J, Chin-Wo SR, Van Deynze A (2012) *De novo* assembly of the pepper transcriptome (*Capsicum annuum*): a benchmark for *in silico* discovery of SNPs, SSRs and candidate genes. BMC Genomics 13:571

Cui J, You C, Chen X (2017) The evolution of microRNAs in plants. Curr Opin Plant Biol 35:61–67

Dai X, Zhuang Z, Zhao P (2018) psRNAtarget: a plant small RNA target analysis server (2017 release). Nucleic Acids Res 46(W1):W49–W51

Gongora-Castillo E, Fajardo-Jaime R, Fernandez-Cortes A, Jofre-Garfias AE, Lozoya-Gloria E, Martinez O, Ochoa-Alejo N, Rivera-Bustamante R (2012a) The capsicum transcriptome DB: a "hot" tool for genomic research. Bioinformation 8:43–47

Gongora-Castillo E, Ibarra-Laclette E, Trejo-Saavedra DL, Rivera-Bustamante RF (2012b) Transcriptome analysis of symptomatic and recovered leaves of geminivirus-infected pepper (*Capsicum annuum*). Virol J 9:295

Hulse-Kemp AM, Maheshwari S, Stoffel K, Hill TA, Jaffe D, Williams SR, Weisenfeld N, Ramakrishnan S, Kumar V, Shah P et al (2018) Reference quality assembly of the 3.5 Gb genome of *Capsicum annuum* from a single linked-read library. Hort Res 5:4

Hwang DG, Park JH, Lim JY, Kim D, Choi Y, Kim S, Reeves G, Yeom SI, Lee JS, Park M et al (2013) The hot pepper (*Capsicum annuum*) microRNA transcriptome reveals novel and conserved targets: a foundation for understanding microRNA functional roles in hot pepper. PLoS One 8:e64238

Kanehisa M, Sato Y, Furumichi M, Morishima K, Tanabe M (2018) New approach for understanding

genome variations in KEGG. Nucleic Acids Res. https://doi.org/10.1093/nar/gky962

Kim HJ, Baek KH, Lee BW, Choi D, Hur CG (2011) In silico identification and characterization of microRNAs and their putative target genes in Solanaceae plants. Genome 54:91–98

Kim HJ, Baek KH, Lee SW, Kim J, Lee BW, Cho HS, Kim WT, Choi D, Hur CG (2008) Pepper EST database: comprehensive in silico tool for analyzing the chili pepper (*Capsicum annuum*) transcriptome. BMC Plant Biol 8:101

Kim MS, Kim S, Jeon J, Kim KT, Lee HA, Lee HY, Park J, Seo E, Kim SB, Yeom SI et al (2018) Global gene expression profiling for fruit organs and pathogen infections in the pepper *Capsicum annuum* L. Sci Data 5:180103

Kim S, Park J, Yeom SI, Kim YM, Seo E, Kim KT, Kim MS, Lee JM, Cheong K, Shin HS et al (2017) New reference genome sequences of hot pepper reveal the massive evolution of plant disease-resistance genes by retroduplication. Genome Biol 18:210

Kim S, Park M, Yeom SI, Kim YM, Lee JM, Lee HA, Seo E, Choi J, Cheong K, Kim KT et al (2014) Genome sequence of the hot pepper provides insights into the evolution of pungency in Capsicum species. Nat Genet 46:270–278

Korbel JO, Lee C (2013) Genome assembly and haplotyping with Hi-C. Nat Biotechnol 31:1099–1101

Lee HA, Kim S, Kim S, Choi D (2017) Expansion of sesquiterpene biosynthetic gene clusters in pepper confers nonhost resistance to the Irish potato famine pathogen. New Phytol 215:1132–1143

Lee S, Choi D (2013) Comparative transcriptome analysis of pepper (*Capsicum annuum*) revealed common regulons in multiple stress conditions and hormone treatments. Plant Cell Rep 32:1351–1359

Lee S, Kim SY, Chung E, Joung YH, Pai HS, Hur CG, Choi D (2004) EST and microarray analyses of pathogen-responsive genes in hot pepper (*Capsicum annuum* L.) non-host resistance against soybean pustule pathogen (Xanthomonas axonopodis pv. glycines). Funct Int Genomics 4:196–205

Li F, Orban R, Baker B (2012) SoMART: a web server for plant miRNA, tasiRNA and target gene analysis. Plant J 70:891–901

Li J, Yang P, Gan Y, Yu J, Xie J (2015) Brassinosteroid alleviates chilling-induced oxidative stress in pepper by enhancing antioxidation systems and maintenance of photosystem II. Acta Physiol Plant 37:1–11

Li J, Yang P, Kang J, Gan Y, Yu J, Calderon-Urrea A, Lyu J, Zhang G, Feng Z, Xie J (2016) Transcriptome analysis of pepper (*Capsicum annuum*) revealed a role of 24-epibrassinolide in response to chilling. Front Plant Sci 7:1281

Liu C, Ma N, Wang PY, Fu N, Shen HL (2013a). Transcriptome sequencing and *de novo* analysis of a cytoplasmic male sterile line and its near-isogenic restorer line in chili pepper (*Capsicum annuum* L.). PLoS One 8:e65209

Liu F, Yu H, Deng Y, Zheng J, Liu M, Ou L, Yang B, Dai X, Ma Y, Feng S et al (2017a) PepperHub, an informatics hub for the chili pepper research community. Mol Plant 10:1129–1132

Liu H, Ding Y, Zhou Y, Jin W, Xie K, Chen LL (2017b) CRISPR-P 2.0: an improved CRISPR-Cas9 tool for genome editing in plants. Mol Plant 10:530–532

Liu S, Li W, Wu Y, Chen C, Lei J (2013b) De novo transcriptome assembly in chili pepper (*Capsicum frutescens*) to identify genes involved in the biosynthesis of capsaicinoids. PLoS One 8:e48156

Liu Z, Zhang Y, Ou L, Kang L, Liu Y, Lv J, Wei G, Yang B, Yang S, Chen W et al (2017c) Identification and characterization of novel microRNAs for fruit development and quality in hot pepper (*Capsicum annuum* L.). Gene 608:66–72

Lu FH, Cho MC, Park YJ (2012) Transcriptome profiling and molecular marker discovery in red pepper, *Capsicum annuum* L. TF68. Mol Biol Rep 39:3327–3335

Martinez-Lopez LA, Ochoa-Alejo N, Martinez O (2014) Dynamics of the chili pepper transcriptome during fruit development. BMC Genomics 15:143

Mazourek M, Pujar A, Borovsky Y, Paran I, Mueller L, Jahn MM (2009) A dynamic interface for capsaicinoid systems biology. Plant Physiol 150:1806–1821

Michael TP, Jupe F, Bemm F, Motley ST, Sandoval JP, Lanz C, Loudet O, Weigel D, Ecker JR (2018) High contiguity Arabidopsis thaliana genome assembly with a single nanopore flow cell. Nat Commun 9:541

Moscone EA, Baranyi M, Ebert I, Greilhuber J, Ehrendorfer F, Hunziker AT (2003) Analysis of nuclear DNA content in Capsicum (Solanaceae) by flow cytometry and Feulgen densitometry. Ann Bot 92:21–29

Nicolai M, Pisani C, Bouchet JP, Vuylsteke M, Palloix A (2012) Discovery of a large set of SNP and SSR genetic markers by high-throughput sequencing of pepper (*Capsicum annuum*). Genet Mol Res 11:2295–2300

Ou L, Li D, Lv J, Chen W, Zhang Z, Li X, Yang B, Zhou S, Yang S, Li W et al (2018) Pan-genome of cultivated pepper (Capsicum) and its use in gene presence-absence variation analyses. New Phytol 220 (2):360–363

Park YS, Ryu CM (2014) Understanding cross-communication between aboveground and belowground tissues via transcriptome analysis of a sucking insect whitefly-infested pepper plants. Biochem Biophys Res Commun 443:272–277

Qin C, Yu C, Shen Y, Fang X, Chen L, Min J, Cheng J, Zhao S, Xu M, Luo Y et al (2014) Whole-genome sequencing of cultivated and wild peppers provides insights into Capsicum domestication and specialization. Proc Natl Acad Sci USA 111:5135–5140

Qiu Y, Liao L, Jin X, Mao D, Liu R (2018) Analysis of the meiotic transcriptome reveals the genes related to the regulation of pollen abortion in cytoplasmic male-sterile pepper (*Capsicum annuum* L.). Gene 641:8–17

Schindele P, Wolter F, Puchta H (2018) Transforming plant biology and breeding with CRISPR/Cas9, Cas12 and Cas13. FEBS Lett 592:1954–1967

Shin SY, Park MH, Choi JW, Kim JG (2017) Gene network underlying the response of harvested pepper to chilling stress. J Plant Physiol 219:112–122

Spies N, Weng Z, Bishara A, McDaniel J, Catoe D, Zook JM, Salit M, West RB, Batzoglou S, Sidow A (2017) Genome-wide reconstruction of complex structural variants using read clouds. Nat Methods 14:915–920

Stankova H, Hastie AR, Chan S, Vrana J, Tulpova Z, Kubalakova M, Visendi P, Hayashi S, Luo M, Batley J et al (2016) BioNano genome mapping of individual chromosomes supports physical mapping and sequence assembly in complex plant genomes. Plant Biotechnol J 14:1523–1531

Strauss T, van Poecke RM, Strauss A, Romer P, Minsavage GV, Singh S, Wolf C, Strauss A, Kim S, Lee HA et al (2012) RNA-seq pinpoints a *Xanthomonas* TAL-effector activated resistance gene in a large-crop genome. Proc Natl Acad Sci USA 109:19480–19485

Sunkar R, Li YF, Jagadeeswaran G (2012) Functions of microRNAs in plant stress responses. Trends Plant Sci 17:196–203

Szczesniak M, Makalowska I (2014) miRNEST 2.0: a database of plant an animal mcroRNAs. Nucleic Acids Res 42:D74–D77

Widana Gamage SM, McGrath DJ, Persley DM, Dietzgen RG (2016) Transcriptome analysis of Capsicum chlorosis virus-induced hypersensitive resistance response in bell Capsicum. PLoS One 11:e0159085

Wu HJ, Ma YK, Chen T, Wang M, Wang XJ (2012) PsRobot: a web-based plant small RNA meta-analysis toolbox. Nucl Acids Res 40:W22–W28

Yang JW, Yi HS, Kim H, Lee B, Lee S, Ghim SY, Ryu CM (2011) Whitefly infestation of pepper plants elicits defence responses against bacterial pathogens in leaves and roots and changes the below-ground microflora. J Ecol 99:46–56

Yi X, Zhang Z, Ling Y, Xu W, Su Z (2015) PNRD: a plant non-coding RNA database. Nucleic Acids Res 43:D982-9

Zhang L, Qin C, Mei J, Chen X, Wu Z, Luo X, Cheng J, Tang X, Hu K, Li SC (2017). Identification of microRNA targets of capsicum spp. Using mirtrans —a trans-omics approach. Front Plant Sci 8:495

Zhu C, Li X, Zheng J (2018) Transcriptome profiling using Illumina- and SMRT-based RNA-seq of hot pepper for in-depth understanding of genes involved in CMV infection. Gene 666:123–133