Compendium of Plant Genomes
Series Editor: Chittaranjan Kole

Ezio Portis
Alberto Acquadro
Sergio Lanteri *Editors*

The Globe Artichoke Genome

Springer

Compendium of Plant Genomes

Series Editor

Chittaranjan Kole, Raja Ramanna Fellow, Government of India, ICAR-National Research Center on Plant Biotechnology, Pusa, 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 about 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

Ezio Portis · Alberto Acquadro · Sergio Lanteri
Editors

The Globe Artichoke Genome

Springer

Editors
Ezio Portis
DISAFA Plant Genetics and Breeding
University of Torino
Largo P. Braccini 2
Grugliasco, Torino 10095
Italy

Alberto Acquadro
DISAFA Plant Genetics and Breeding
University of Torino
Largo P. Braccini 2
Grugliasco, Torino 10095
Italy

Sergio Lanteri
DISAFA Plant Genetics and Breeding
University of Torino
Largo P. Braccini 2
Grugliasco, Torino 10095
Italy

ISSN 2199-4781 ISSN 2199-479X (electronic)
Compendium of Plant Genomes
ISBN 978-3-030-20011-4 ISBN 978-3-030-20012-1 (eBook)
https://doi.org/10.1007/978-3-030-20012-1

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

The globe artichoke (*Cynara cardunculus* var. *scolymus*), together with cultivated cardoon (*C. cardunculus* var. *altilis*) was domesticated in the Mediterranean region from the wild cardoon (*C. cardunculus* var. *sylvestris*). The species is a member of the *Compositae* (a.k.a. *Asteraceae*), which is the largest and one of the most ecologically successful botanical families, including species of great economic importance such as sunflower, lettuce and chicory. Italy, is the top globe artichoke producing country and harbours its richest primary gene pool, while Sicily Island is considered the place where globe artichoke domestication occurred. Over time, globe artichoke cultivation has spread to other Mediterranean countries and more recently to the Americas and China. The primary product of globe artichoke is the immature inflorescence, but, as well as other members of the *C. cardunculus* complex, it is a multi-use crop exploitable also a source of phenolics, sesquiterpene lactones and inulin as well as for the production of lignocellulosic biomass and seed oil for both edible and biofuel purposes.

In this book emphasis is given to the recent development of the globe artichoke ad cultivated cardoon genome sequences, made available to the scientific community by a Consortium including two Italian Universities, namely the University of Torino (DISAFA) and the University of Catania (Di3A) in collaboration with the Genome Center of Davis (USA). The availability of globe artichoke genome sequence has already made it possible to deepen the knowledge on the structure and function of its genome organization, to perform the resequencing of genotypes of globe artichoke varietal types, in order to dissect the path from sequence variation to phenotype, as well as to accomplish comparative genomic studies within the *Compositae* family. Furthermore, has laid the groundwork for future genomic assisted breeding programs and the application of genome manipulation technologies.

This book also provides exhaustive information on *C. cardunculus* botany, domestication, propagation, genetic resources, uses and pathways involved in the synthesis of bioactive compounds. Moreover, it includes a survey of the state of the art on the development of genetic maps and location of genes and QTLs affecting traits of economic interest.

Torino, Italy

Ezio Portis
Alberto Acquadro
Sergio Lanteri

Contents

1 *Cynara cardunculus* L.: Historical and Economic Importance, Botanical Descriptions, Genetic Resources and Traditional Uses

Gaetano Roberto Pesce and Giovanni Mauromicale

Abstract

The chapter begins with a brief etymological discussion, followed by an overview of ancient sources that speak of *Cynara cardunculus*. According to these testimonies, the domestication of the globe artichoke could have begun in the first century CE. Then the most recent statistics on globe artichoke in the world are explained; they show that, in recent years, its cultivation is spreading even in countries that are not traditionally producers. The chapter also provides the botanical classification and the description of the species, highlighting that three botanical varieties are distinguishable, namely globe artichoke, cultivated cardoon and wild cardoon. The latter is thought to be the progenitor of both the cultivated forms. There is also a remarkable diversity of forms and cultivars of globe artichoke and cardoon, as explained in the paragraph about genetic resources. The chapter ends with a description of the traditional uses of globe artichoke and cardoon.

1.1 Historical Outline, Spread in the World and Economic Importance

The English name «artichoke» derives from northern Italian «articiocco» (Skeat 1887), which in turn derives, like the Spanish «alcachofa», the Portuguese «alcachofra» and the Italian «carciofo», from the Arabic خرشوف (kharshuf) through the Hispanic Arabic «al-harsúf». The adjective 'globe' refers to immature inflorescence (head, capitulum) with the appearance of a pine cone and a globe shape, to distinguish it from two other vegetables, named artichoke: Jerusalem artichoke (*Helianthus tuberosus* L.) and Chinese artichoke (*Stachys sieboldii* Miq.) grown for their edible stem tubers. The genus name *Cynara* could derive from the Greek κῦων (kyon—dog) (Craig 1858). The reference to the dog has a negative connotation, because of the objectionable thorns, which are likened to a dog's teeth (Small 2009). The name *Cynara* could be also related to Κινάρος (Kinaros), an island of Aegean Sea, for a kind of globe artichoke native to that island (Quattrocchi 1999). However, as we shall see, the word κινάρα (kinara), less frequently κυνάρα (kynara), is found in literature to probably indicate the globe artichoke. The specific epithet *cardunculus* comes from Latin *carduus* with diminutive suffix-*unculus*, namely little cardoon. *Scolymus*, namely the former specific epithet of globe artichoke and its current botanical variety,

G. R. Pesce · G. Mauromicale (✉)
Di3A (Dipartimento di Agricoltura, Alimentazione e Ambiente), University of Catania, via Valdisavoia 5, 95123 Catania, Italy
e-mail: g.mauromicale@unict.it

E. Portis et al. (eds.), *The Globe Artichoke Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-030-20012-1_1

is the latinization of σκόλυμος (scolymos), which belongs to the semantic field of 'thorn' (σκόλοψ—scolops). On this, Hesiod (eighth–seventh century BCE), in his 'Works and Days', and Theophrastus (fourth–third century BCE), in his 'Enquiry into Plants', used the word σκόλυμος to refer to *Scolymus hispanicus* L. (Liddell and Scott 1996). Again in the cited work, Theophrastus referred to a plant calling it κάκτος (kaktos), maybe the cardoon, and stated that it only grows in Sicily and not in Greece. But if we take a leap forward by a few centuries, we read that Athenaeus (second–third century CE) in his 'Deipnosophists' challenged Theophrastus. He was indeed sure that κάκτος (kaktos) was what the κινάρα (kinara) was for the Greeks. He also stated that the plant called κάκτος (kaktos) is the same plant that the Romans called κάρδος (cardos, *carduus* in Latin). After all, Athenaeus wrote, κάκτος (kaktos) and κάρδος (cardos) differ only by two letters. Consequently, Athenaeus made no difference between κινάρα (kinara) and *carduus*. From this overview, a certain vagueness emerges in the ancient references, and for this reason De Candolle (1885) suggested that cultivated globe artichoke was unknown in classical times. However, Columella (first century CE) in his 'On Agriculture' talked about *cinara*, defining it *hispida* (spiny), and was the first to provide a description of what is probably the globe artichoke, for at least two reasons. The first is the fact that he focused his attention on the capitulum. The second is that he spoke about cultivation practices that seem more suited to the globe artichoke than to the wild cardoon. He also wrote about making cheese using the wild cardoon flower (*agrestis cardui flore*) as rennet. In Chap. 43 of book XIX of his 'The Natural History', Pliny the Elder (first century CE) wrote about *carduus*, cultivated in Carthage and especially in Cordoba (*apud Carthaginem Magnam Cordubamque praecipue*). He also wrote that this crop is very profitable, although he did not understand why it was so appreciated; in fact, he thought that it was a monstrosity that not even beasts would eat. According to Foury (2005), Pliny was speaking of the globe artichoke. Indeed, the use of the word 'monstrosity' could derive from the wonder in front of a novelty represented by the large capitulum. A little further on (book XXI, Chap. 59), Pliny wrote about the *cactos*, clearly copying what Theophrastus had written. Foury (1989) deduced, based on Columella and Pliny's writings, that the cultivation of globe artichoke started around the first century CE. Probably in the same period, the domestication of globe artichoke was ongoing, but not accomplished (Sonnante et al. 2007b). In this regard, Sicily is supposed to be one of the centres of domestication of the globe artichoke (Mauro et al. 2009). Pedanius Dioscorides (first century CE), who was a physician, pharmacologist and botanist, in his 'On Medical Material', gave just a quick nod to κινάρα (kinara), but Galen (second century CE), in his 'On the power of foods', spoke more widely about the globe artichoke, pointing out that some call it κινάρα (kinara) and others κυνάρα (kynara), blaming the latter as a solecism. He also noted that many people consumed the plant's capitulum, which they called 'vertebra' [σφόνδυλος/σπόνδυλος (sfondylos/spondylos)]. The word 'vertebra' to refer to the capitulum also appears in the Diocletian's Edict on Maximum Prices of 301 CE, both in Greek and in Latin [σφόνδυλοι κιναρῶν (sfondyloi kinaron), *sponduli*]. In the third or perhaps fourth century CE, a compilation of recipes under the name of Apicius, 'On the Subject of Cooking', was compiled. It contains three recipes for cardoons (*cardui*) and seven recipes for *funduli* or *sfondili*, which are, according to Foury (2005), receptacles. Some Roman mosaics conserved in the Bardo Museum in Tunis datable to the Imperial period (second and third century CE) show heads of globe artichokes (both spiny and spineless) (Fig. 1.1). We have little knowledge about globe artichoke in the Middle Ages, because of a surprising historical-documentary eclipse. After the fall of the Western Roman Empire (end of the fifth century), the Eastern Roman Empire continued to prosper and expand; there, vegetables were in great demand and their cultivation spread outside the city walls from the second half of sixth century (Foury 2005). Chapter 39 of XII book of 'Geoponica' (tenth century) speaks about the

cultivation of κινάρα (kinara). Among other things, the chapter contains some curious tips on how to get tasty 'fruits' (capitula); it is attributed to Marcus Terentius Varro, although you cannot find it in his 'Agricultural Topics in Three Books'. After the Umayyads conquered the Iberian Peninsula (eight century), agriculture improved from a technological point of view and was enriched with the introduction of new cultivated species. This would reasonably suggest that the cultivation of the globe artichoke was maintained in those territories, as a legacy of the Roman Empire, and that it was not introduced from the East (Foury 2005). In paragraph 658 of 'Compendium on Simple Medicaments and Foods', Ibn al-Baitar (first half of the thirteenth century), described a plant called «harshef», which actually includes two forms: one wild and the other cultivated. The latter, called «kenguer» or «kenarya» in Andalusian, is described in paragraph 1976 of the same work. Clearly, it is hard to miss the similarity between the Andalusian «kenarya» on the one hand and the Latin *cinara* or the Greek κινάρα (kinara) on the other. According to Foury (2005), it is quite probable that the *cinara* described by Columella, and therefore our globe artichoke, survived in the Hispanic–Berber refuge until the fifteenth century. The technical skill of the Andalusian agronomists and gardeners has probably led to the improvement and diversification of the globe artichoke, above all through the propagation by seed frequently used at the time. Mauro et al. (2009) hypothesize that globe artichoke was selected to become the plant we know today in the domestic vegetable gardens and in the monasteries, in the period between the ninth and sixteenth centuries. Around 1466, on the threshold of the Modern Age, globe artichokes had been transported from Naples to Florence, from where, according to Jean Ruel, they passed to France at the beginning of the sixteenth century. Hermolao Barbaro says that in 1473, in Venice, globe artichoke appeared as a novelty. In the early Modern Age (sixteenth century), globe artichoke spread in Europe, as witnessed by paintings not only from Italy, but also from Flanders and Bohemia (Fig. 1.2).

The most recent statistics on globe artichoke in the world show that the most important producer country is Italy (Fig. 1.3) with almost 44,000 ha and ~406,000 t year^{-1}. Italy alone accounts for about 27% of the world production (FAOSTAT 2018). Traditionally, globe artichoke is cultivated in Mediterranean countries, where it is almost 80% of the world's globe artichoke growing areas, with nearly three-quarters share of the output (Fig. 1.3), but other countries—

Fig. 1.1 Roman mosaics depicting globe artichokes (second and third centuries). Bardo National Museum, Tunis

Fig. 1.2 Giuseppe Arcimboldo, 'Vertumnus' (1590–1591). This painting portrays Emperor Rudolf II of Habsburg as Vertumnus, a Roman deity (circled in red, a globe artichoke). It is kept in Skokloster Castle in Håbo, Sweden

especially Peru, Argentina and China—produce significant amounts (FAOSTAT 2018). In this regard, Fig. 1.4 provides the charts of globe artichoke production in the current main producing countries. These charts just give us an idea of the production trends in the period from 1961 to 2016 (FAOSTAT 2018). It is noticeable that traditionally producing countries, such as France, Greece, Spain and even Italy, have a decreasing trend, while other countries, among which Egypt stands out, have increased their production in the last few years. In the same way, it should be noted that even countries outside the Mediterranean basin (such as Argentina, China and Peru) are increasing their production. In general, it can be said that even if the globe artichoke is still mainly produced in the countries of the Mediterranean basin, it is slowly spreading in the rest of the world. Bianco (2012) has summed up this situation aptly: 'Artichoke: an international soul, but a Mediterranean heart'. The FAO (2018) also provides data on the globe artichoke trade, which however are largely indicative. Figure 1.5 represents the main commercial flows of the globe artichoke; the values shown are the averages of the three-year period (2014–2016). In order to represent the most significant trade movements, only flows greater than US$500,000 were taken into consideration. Figure 1.5 clearly shows that France is the main importing country, Spain being the main exporting country. The most important trade flow is between these last two countries, more precisely from Spain to France. Spain is almost a net exporter (it imports very little product), while France and Italy are at the centre of a dense network of imports and exports. The figure also shows that globe artichoke exchanges in the American continent are separate from the rest of

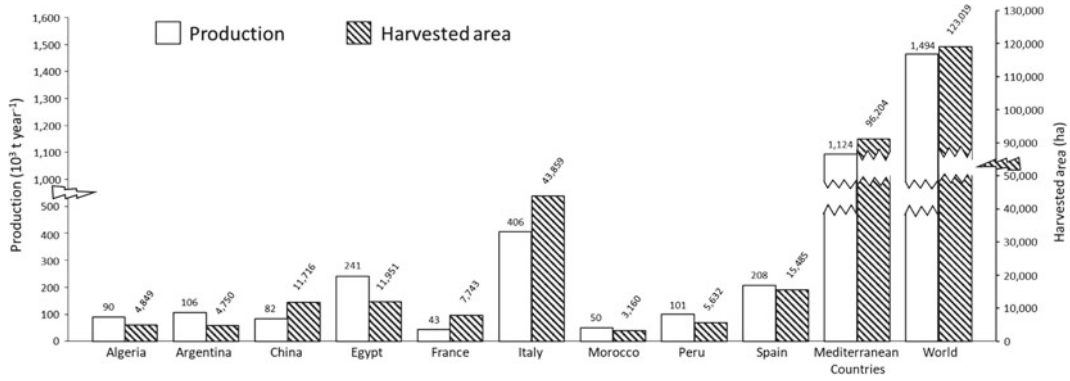


Fig. 1.3 Globe artichoke annual production and harvested areas in the main producing countries. Considered period 2014–2016. *Source* FAOSTAT (2018)

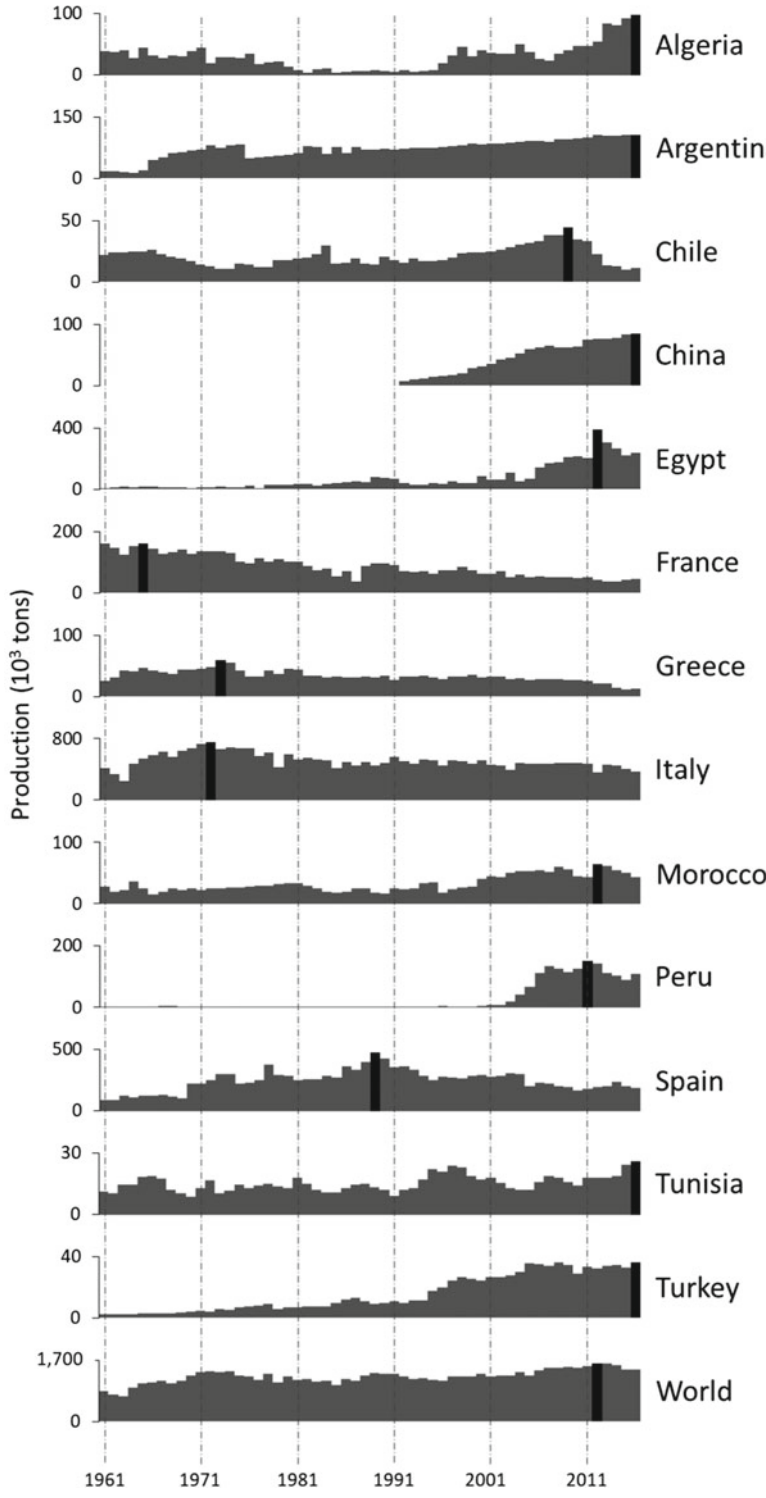


Fig. 1.4 Charts of globe artichoke production in the main producing countries. In each chart, the year of greatest production is coloured in black. Considered period 1961–2016. *Source* FAOSTAT (2018)

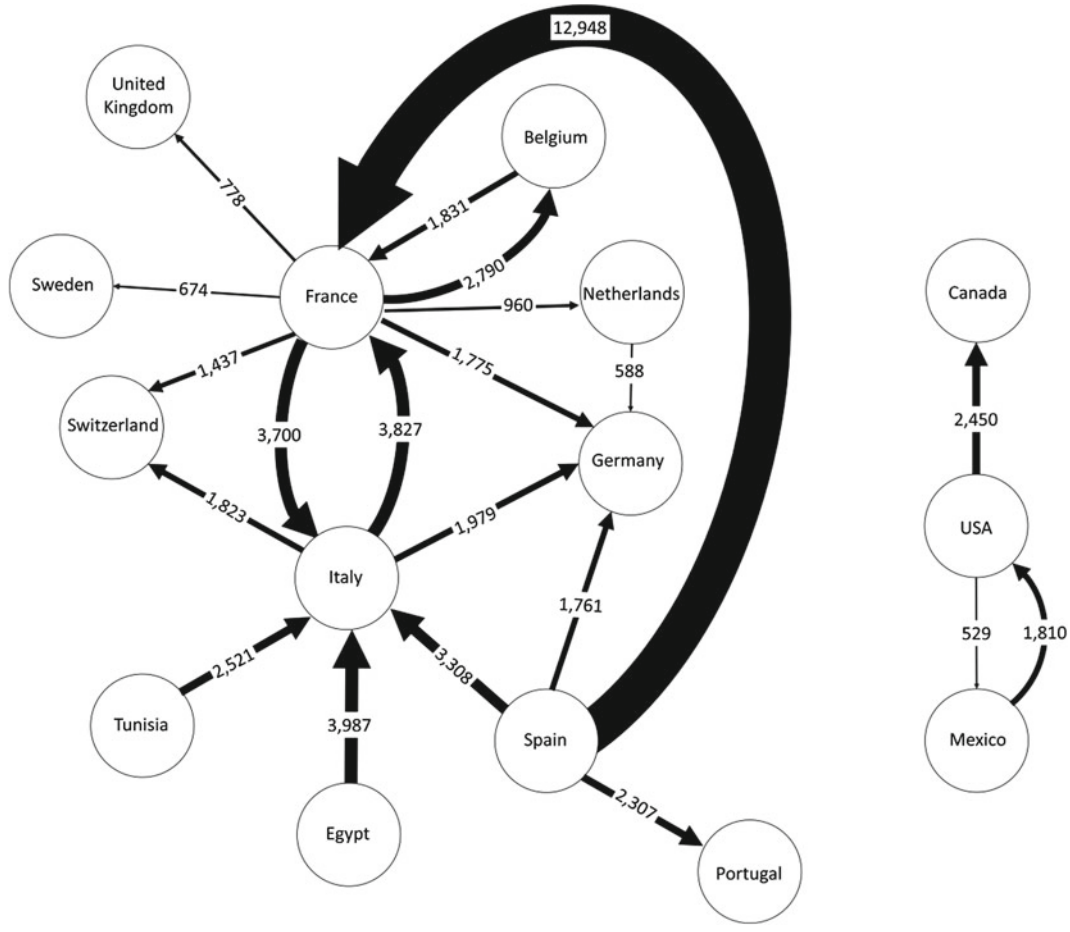


Fig. 1.5 Diagram of main trade flows of globe artichoke in the world (only flows exceeding $500,000 are considered). Values are expressed as thousands of US dollars, and the width of the arrows corresponds to the size of the flows. Considered period 2014–2016. *Source* FAOSTAT (2018)

the world, or at least the significant ones as defined above. This diagram does not include Algeria, Argentina, China, Morocco and Peru, which are among the major producer countries; they allocate their production mainly to domestic consumption or, in some cases, their exports are not taken into account by official sources. The comparison between Figs. 1.6 and 1.7 shows the variation of the gross production values of globe artichoke in the different continents over a period of twenty years (between 1994–96 and 2014–16). The total value does not change very much (from 1058 to 991 million US dollars), while it is clear that the hub of production moves a little from Europe to the other continents. In fact, the gross production value of Europe has decreased by 6%, while those of Africa, Americas and Asia have increased respectively by 252, 52 and 134%.

Regarding the cardoon, unfortunately there is not the same abundance of official statistics as for the globe artichoke. The areas devoted to cardoon cultivation (officially about 2–3000 ha, though this value is underestimated) are localized in Spain, Italy, France and Greece (Ierna and Mauromicale 2010).

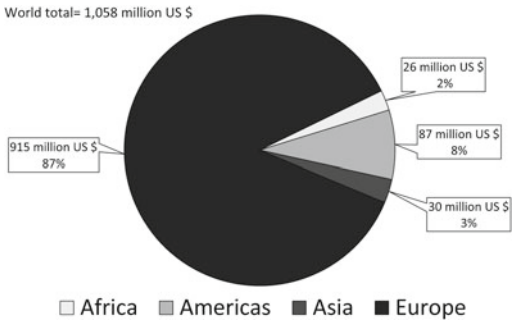


Fig. 1.6 Gross production values of globe artichoke expressed as constant 2004–2006 million US dollars in the different continents. Considered period 1994–1996. *Source* FAOSTAT (2018)

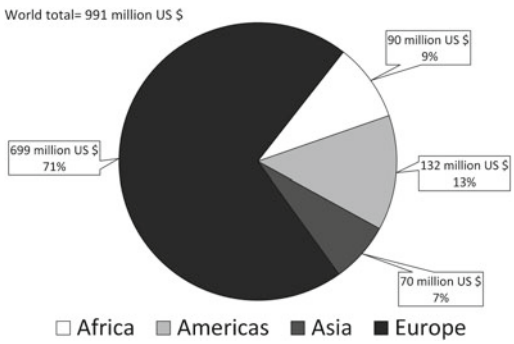


Fig. 1.7 Gross production values of globe artichoke expressed as constant 2004–2006 million US dollars in the different continents. Considered period 2014–2016. *Source* FAOSTAT (2018)

1.2 Botanical Classification and Description of the Species

The genus *Cynara*, belonging to the Asteraceae family, is native to the Mediterranean region, sharing its distribution with the olive (*Olea europea*). The members of the genus are the *Cynara cardunculus* L. species complex, consisting of the globe artichoke [var. *scolymus* (L.) Fiori], the cultivated cardoon (var. *altilis* DC.) and the wild cardoon [var. *sylvestris* (Lamk) Fiori], and other six wild species. The three *C. cardunculus* forms are fully cross-compatible with one another, and form fertile intervarietal hybrids (Basnizki and Zohary 1994). The predominant allogamy of *C. cardunculus* led to the formation of populations whose kinship is difficult to understand. Studies on the isoenzymatic relationships (Rottemberg et al. 1996), the intercross (Rottemberg and Zohary 1996), as well as on genetic and molecular relationships intra- and intervarietal (Lanteri et al. 2004a, b; Portis et al. 2005a, b, c; Mauro et al. 2009) lead to the conclusion that the cultivated botanical varieties of *C. cardunculus* L. are phylogenetically very close to the wild cardoon, which is probably the common ancestor of the first two. Presumably, from the latter species two divergent lines were implemented: one selected for the width of the leaf midrib, leading to the cultivated cardoon, the other for the head size, leading to the globe artichoke (Basnizki and Zohary 1994; Lanteri et al. 2004b).

1.2.1 Globe Artichoke

Globe artichoke (Figs. 1.8 and 1.9) is a perennial herbaceous plant that can be reproduced by achenes (improperly called ‘seeds’) or usually vegetatively propagated by side offshoots originating from underground buds, *ovoli* (underground semi-dormant dried offshoots with apical and lateral buds) or by division of rooted basal stem portions. Tissue culture procedures are occasionally used for nursery production of disease-free plantlets. However, these in vitro multiplication methods have not yet been optimized, and success is still poor, at least for early Mediterranean genotypes such as ‘Violet de Provence’, ‘Violetto di Sicilia’ and ‘Blanca de España’. Plants, when grown from achenes, have a primary taproot and numerous adventitious secondary roots, whereas plants from offshoots, *ovoli* or other vegetative parts have fibrous adventitious roots that enlarge over time (while the smallest disappear), losing their absorption function in order to become storage organs and provide support (Bianco 1990). Their total depth is about 40 cm. While the plant grows, a rhizomatous stem, commonly called *ceppaia* in Italian, becomes increasingly evident at the base of the main stem, and on its surface several buds differentiate, originating offshoots. Offshoots do

not all develop at the same time, but in sequence because of apical dominance, hence on the same *ceppaia* buds at different age and physiological stage are usually present. Apical dominance intensity varies depending on the cultivar (Mauromicale and Copani 1990) and decreases when the main shoot differentiates the main head. The latest offshoots, which did not produce heads, dry up their aboveground fraction at the onset of summer drought, becoming *ovoli* (Jannaccone 1967). During the vegetative phase, the caule is very short, so the plant has some whorls of leaves, clustered tightly at its base, giving it a typical 'rosette' form. Leaf size, shape and number are variable depending on the cultivar and phenological growth stages. They are ashy green coloured on the lower side, have thick midribs and are typically fewer in the earliest cultivars. The apical bud of each shoot switches from the vegetative phase to reproductive phase. The flower stem, variable in length depending on the cultivar, on the time of year and on eventually supplied hormone treatments, is cylindrical, ashy green coloured, longitudinally ribbed, erect and has ramifications that vary in number. It has trichomes and alternate lanceolate leaves. Stems of first, second, third (and so on) orders also have a capitulum at their distal end, whose size decreases as the order of ramification increases. The plant, depending on the cultivar, may produce from 3–4 to 20 capitula. The weight of a single capitulum can vary greatly depending on the cultivar and its position in the plant, but in general it can be said to range between 100 and 300 g. In most cultivars, a well-developed capitulum contains from 800 to 1400 florets (Bazniski and Zohary 1994). The latter are hermaphrodite, tubular, proterandrous and blue-violet coloured. The fruit is a tetragon-shaped or flattened cypsela or achene, dark-coloured or greyish, uniform or mottled, whose weight ranges between 30 and 70 mg. The calyx, metamorphosed into a feathery organ (pappus), favours dissemination. Commonly the two terms 'cypsela' and 'achene' are used interchangeably; however, in the course of this chapter, the term 'achene' is used. The three botanical varieties of *C. cardunculus* have the same floral biology, as explained later in the paragraph on the cultivated cardoon. Since the main product of the globe artichoke is the capitulum, it is worthwhile dwelling on the flowering transition of the vegetative apex. In this regard, Morone Fortunato et al. (1981) made microscopic observations during the formation of the capitulum on four cultivars of globe artichoke. The vegetative apex of the globe artichoke (Fig. 1.10) is formed by a *tunica*, with 8–9 cell layers, and a less homogeneous *corpus* than the *tunica*. The *tunica* is made up of two parts: one is the central apex, consisting of two initial cells recognizable because of their greater diameter. The other is located between the initial cells and the primordia of the leaves. This part is made up of smaller cells that divide more frequently.

The corpus instead is divided into three zones:

(a) the zone of the mother cells, which represent the initial cells of the corpus, located under the initial cells of the *tunica*;
(b) the rib-meristem, consisting of the cells arranged in columns. They are larger the further they are from the apex; in this zone, the divisions generally occur according to a plane perpendicular to the axis of the apex;
(c) the flank-meristem, consisting of cells that envelop the rib-meristem, this zone contributes to the development of the leaves primordia and procambium.

Fig. 1.8 Globe artichoke field (photograph taken by Mauromicale)

Fig. 1.9 Biodiversity in globe artichoke (photograph taken by Mauromicale)

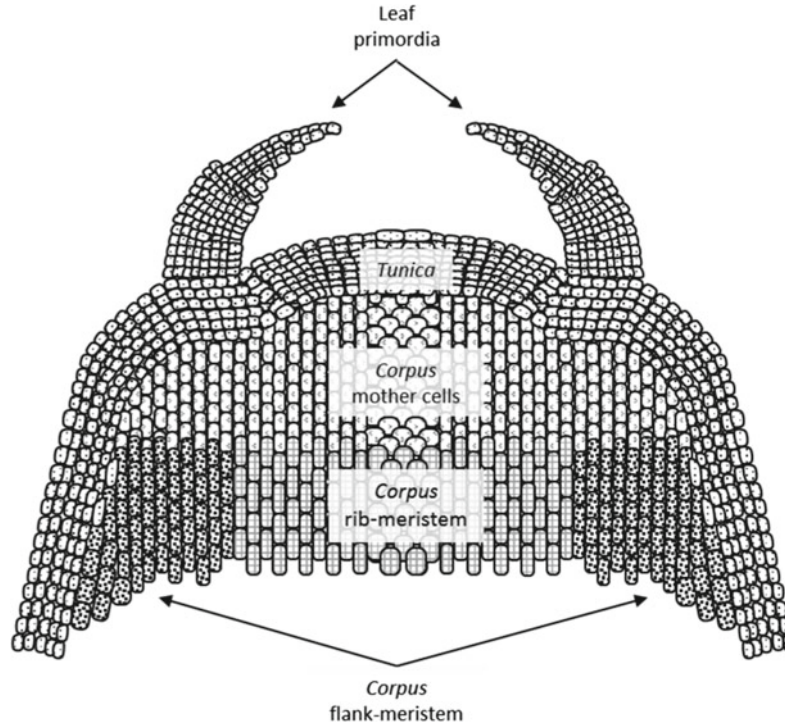


Fig. 1.10 Schematic representation of globe artichoke vegetative apex according to Morone Fortunato et al. (1981)

The meristematic apex in the vegetative phase appears as a small dome set within the hood of the young leaves. In the very early stages of the reproductive transition, there is an increase in mitotic activity at the border between the central area of the mother cells and the rib-meristem.

Gradually, this activity expands in the central zone of the mother cells, where the cells become small and rich in protoplasm. The mitotic activity does not appear in the cells of the rib-meristem and in the underlying pith. Therefore, the apex has a parenchymatous pith, surrounded by some layers of meristematic cells that gradually decrease to form a mantle–core structure (this is the reproductive transition). The apex, which was previously globose, flattens and expands progressively, highlighting the mantle formed by the two meristematic layers, which overlap the parenchymatous heart. In this stage, on the peripheral area of the apex surface, the flower primordia make their appearance and develop in a centripetal sense. When the florets have developed on the whole thalamus, then the differentiation is complete.

Macroscopically, the flower organogenesis of the head of the cultivar Violet de Provence is shown schematically through the following sequence of phenological stages (Foury 1967):

R: indicates the caulinar apex transition from the vegetative phase to the reproductive one;
A: the head is perceptible to the touch but it is completely enveloped in a leaf rosette that surrounds the caulinar structure;
B: the lengthening of the stem and the deployment of the leaves allow seeing the head at the centre of the leaf rosette;
C: the head is fully visible, and the inflorescence reaches a length of 2–4 mm;
D: the flower stem is fully stretched, and head has reached the optimum size for harvesting;
E: external bracts begin to diverge and the thalamus, which was concave, begins to flatten, while central flowers (florets) reach a length of about 2 cm;
F: the central bracts open, and in the centre the first florets appear;
G: appearance of the flowers and anthesis of peripheral flowers.

Baggio et al. (2011) describe 15 development scale stages of 'Nobre-UPF', a Brazilian cultivar. Seven of them are equivalent to the first observed by Foury (1967), while the other eight stages comprise the phases from the full-opening of the capitulum to the fruit dispersion.

1.2.2 Cultivated Cardoon

Commonly known as cynara, cardoon is a perennial ($\sim$10 years) field crop for numerous purposes and non-conventional use (Gominho et al. 2018). It is well adapted to the xerothermic conditions of southern Europe. The cardoon has an annual development cycle, in which the plants grow and develop during autumn, winter and spring, with summer dormancy. The underground plant part consists of the main taproot, of a variable number of secondary fibrous roots as well as of a rhizome, more or less expanded, containing buds, both single and gathered in groups. The one-year roots keep the absorption function until the spring, when they swell and gradually take on the reserve and support functions, while a new adventitious root system takes place adopting the absorption function. The cardoon's root system is highly developed, being able to reach a depth of more than 1 m, which allows the plant to explore a significant volume of soil. In plants older than a year, the underground stem becomes increasingly evident; it is a rhizome, also commonly called *ceppaia*, which contains conducting cells and reserve parenchyma. Several buds differentiate on the rhizome surface, giving rise to side shoots called offshoots, which can grow up to form a new plant, thus expanding the rhizomatous fraction. In the epigeic part, cultivated cardoon has a 'rosette', since its stalk is very shortened (3–4 cm) and it has a high number of leaves (over 40), alternate and pinnate. Leaves can reach a length of more than 1 m (especially those basal); they are characterized by a well-developed petiole and a large midrib. The leaf lamina, crossed by numerous veins, could assume different shapes in different cultivars as well as within the same plant (heterophyllia). It is dark green coloured or ashen on the top page, greyish on the bottom page, because of its thick hairiness. At flowering, the plant has one or more erect flower stems, of variable height (1.5–3 m

approximately), each of which has a diameter of about 2–4 cm (Fig. 1.11). Stems are longitudinally ribbed, tomentose and very branched; they also have small leaves, alternate and lanceolate. Each branch has an inflorescence (head or capitulum) at its distal end, which has a round or oval shape depending on the cultivar. The plant has a large number of heads (10–30), the largest and earliest of which (main head) is on the top of the main stem; the heads of the next order (first, second, third and so on), progressively emitted, are gradually smaller. Each capitulum has several hundreds of flowers (florets), which are hermaphrodite, tubular and fitted in a well-developed receptacle (thalamus). At the full anthesis, the florets have very long stigmata (∼5–8 cm). Florets are usually blue-violet coloured, although there are mutant genotypes having florets white or lilac coloured. Analogously to the globe artichoke, the fruit is an achene, endowed with a plumy structure—the pappus—which favours the dispersion by wind action (anemochorous dissemination). Thousand 'seeds' weight ranges between 20 and 40 g. From a biological standpoint, the cultivated cardoon is a perennial geophyte/hemicryptophyte herbaceous species, whose field duration is indefinite, thanks to the vitality of the rhizome. In the areas characterized by Mediterranean subarid climate, where the crop expresses its full production potential, the crop cycle is autumn–winter–spring, with a vegetative stasis phase in the summer, more or less prolonged, while in middle latitudes the vegetative stasis occurs between spring and autumn. In southern Italy, the cultivation cycle begins with the germination of achenes in the autumn, followed by a long vegetative phase, which lasts until the beginning of following spring. During this phase, the gradual transition to reproductive phase of the shoot apex occurs, and the first head starts to differentiate at late winter-early spring. The first head initially appears as a swelling in the centre of the leaf rosette and becomes more evident because of the main flower stem elongation. At the same time, other heads progressively differentiate on the top of the branches of main floral stem. The anthesis starts in the late spring (late May–early June) and, of course, affects first the main head, then the other heads according to their order. At advanced flowering stage, each head has several hundred florets at a different development stage. In every single head, the flower development starts from the most peripheral florets and proceeds centripetally until the centre. In *C. cardunculus*, the flowering is dicogamous and, specifically, proterandrous. Indeed, the stigma becomes receptive from 4–5 to 8 days after anthesis, that is when the pollen, whose germination lasts 3–4 days, has already lost its vitality. Pollination is entomophilous, and reproduction is mainly by cross-fertilization, because of the previously mentioned proterandry mechanism. However, a small amount of self-fertilization between different inflorescences of the same individual (geitonogamy) is inevitable, due to their progressive ripening. Ripening of the achenes takes place starting from 50 to 60 days after anthesis and is accompanied by the progressive desiccation of the aboveground biomass. When achenes reach ripeness, they begin to be disseminated by wind (anemochorous dispersal), thanks to the disintegration of the receptacle and the peculiar structure of the pappus. Given the lack of dormancy mechanisms, except in some special cases (temperatures higher than 29–30 °C accompanied by anoxia), the 'seed' is readily germinable when favourable environmental conditions are met (sufficient soil moisture and temperatures between 14 and 24 °C). The vegetative regrowth after the summer is assured by the underground rhizome buds, which have remained dormant during the hot and dry season.

Archontoulis et al. (2010) described the cultivated cardoon phenological growth stages based on the Biologische Bundesanstalt, Bundessortenamt, CHemische Industrie (BBCH) scale. They defined nine principal growth stages as listed below:

- Principal growth stage 0: germination/sprouting/bud development
- Principal growth stage 1: leaf development
- Principal growth stage 2: formation of side shoots/tillering
- Principal growth stage 3: rosette growth

Fig. 1.11 Development in height of cultivated cardoon (photograph taken by Mauromicale)

- Principal growth stage 4: development of vegetative plant parts
- Principal growth stage 5: inflorescence emergence and development
- Principal growth stage 6: flowering and capitulum formation
- Principal growth stage 7: development of capitulum and seeds
- Principal growth stage 8: capitulum and seed ripening
- Principal growth stage 9: senescence.

1.2.3 Wild Artichoke or Wild Cardoon

Wild cardoon is a typical geophyte species, perennial, capable of both sexual reproduction and agamic propagation. It has well-developed roots, with a taproot able to penetrate to a depth of about 1.5–2 m allowing the plant to tolerate water deficits, and a variable number of secondary roots. It has a rhizome, more or less expanded, containing buds, both single and gathered in groups. The stem is robust, simple and striated, it can be hairy or hairless, but without thorns. The basal leaves, arranged as a large rosette, are deeply engraved, pinnate, up to 35 cm long, white and hairy on the bottom surface, and with long spines at the edge (1–3 cm). Stem leaves are simple, alternate and spiny. The flowering axis is erect, branched, rugged, striated longitudinally and has alternate leaves. The main stem has branches with heads at their distal ends. The plant height is variable, from 40 cm up to 120–130 cm. The flowers (florets), hermaphrodite, tubular, characteristic of the Asteraceae, are grouped in an inflorescence (head or capitulum). Florets at full ripening can be 8 cm or longer, they are violet coloured, but there are mutant genotypes with white (Fig. 1.12) or pink florets. The fruits are tetragon-shaped or flattened achenes, dark greyish and mottled, combined with the calyx metamorphosed into a pappus, which promotes the anemochorous dissemination.

Fig. 1.12 Wild cardoon with white-coloured florets (photograph taken by Mauromicale)

1.3 Genetic Resources

Due to its long-term cultivation in Italy, France and Spain, it is here that the globe artichoke and cardoon germplasm is the richest. Many cultivar and landraces are cultivated in specific areas because very often they are well adapted to particular environmental local conditions. The CPVO/UPOV 'Protocol for tests on distinctness, uniformity and stability' (CPVO-TP/184/2), entered into force in 2013 and submitted to a proposal for a partial revision in 2017, applies to both globe artichoke and cultivated cardoon and provides the characteristics to be used in distinctness, uniformity and stability (DUS) tests for morphological characterization. The latter is fundamental for rationalizing germplasm collection, removing redundant germplasm and a large number of synonyms. In fact, the number of globe artichoke and cultivated cardoon cultivars grown in the world cannot be easily determined, as some of them are synonyms. A cultivar in one location is frequently known by other names in other localities. For example, the Italian globe artichoke cultivar Violetto di Sicilia has about 41 synonyms (Bianco 2005). In the bibliography of globe artichoke compiled by Bianco (1968), 286 cultivars are listed. Variability between cultivars was recorded for many traits, such as plant size, leaf shape, earliness of bolting, head number per stem, shape, size and colour of the head, ratio of edible parts, shelf-life and presence of thorns or spines (Dellacecca et al. 1976). However, according to Bazniski and Zohary (1994) and Mauromicale and Ierna (2000), only 11–12 vegetatively propagated cultivars can be considered to be of major commercial importance ('Violet de Provence', 'Violetto di Sicilia', 'Romaneschi' group, 'Blanca de Tudela', 'Gros Camus de Bretagne', 'Green Globe', etc.). The globe artichoke germplasm currently available is grouped and classified according to different criteria, which are based on the harvest time and on the morphometric features of heads. Regarding the harvest time, usually autumn-harvested cultivars and spring-harvested cultivars are distinguished in two different groups. The cultivars belonging to the first group, called 'early', provide an almost continuous production from autumn to spring (from October to April) and include 'Violetto di Sicilia' and various similar cultivars ('Masedu', 'Molese', etc.). Other early cultivars are 'Spinoso Sardo', 'Violetto Spinoso di Palermo', 'Violet de Provence' and 'Blanca de Tudela'. The cultivars belonging to second group, called 'late' or 'spring', yield only between March and May–June, and include 'Romanesco' types and 'Violetto di Toscana', widespread in Campania, Lazio and Tuscany, respectively. 'Camus de Bretagne' and 'Blanc Hyérois' are also included. On the other hand, from the standpoint of head characteristics, cultivars are usually grouped according to the colour of the outer bracts, the size, the shape and the presence of thorns on bracts. Regarding the latter characteristic, cultivars are divided into 'spineless' ('Violetto di Sicilia', 'Violet de Provence', 'Romanesco', 'Violetto di Toscana', etc.) and 'thorny' ('Violetto Spinoso di Palermo', 'Spinoso Sardo', 'Spinoso di Albenga', etc.) (Mauromicale 1987). The cultivars belonging to the early group are also called 'reflowering' (which bloom between autumn and spring), while the ones belonging to late group are called 'non-flowering' (which bloom only during spring). A first attempt to make a classification was carried out by Porceddu et al. (1976), who gathered the globe artichoke germplasm into four main groups (Fig. 1.13):

i. The 'Spinosi' group ('Violetto Spinoso di Palermo', 'Spinoso Sardo', 'Spinoso di Albenga', etc.);
ii. The 'Violetti' group ('Violetto di Toscana', 'Violetto di Chioggia', 'Nostrano', 'Violetto di Pesaro', etc.),
iii. The 'Romaneschi' group ('Campagnano', 'Castellammare', 'Tondo di Paestum', 'Camard', 'Blanc Hyérois', 'Camus de Bretagne', etc.),
iv. The 'Catanesi' group ('Violetto di Sicilia', 'Violet de Provence', etc.).

This classification, with appropriate corrections, has been updated by molecular investigations, carried out by amplified fragment length

Fig. 1.13 The four main types of globe artichoke, grouped on the head shape basis. **a** 'Thorny', **b** 'Violetti', **c** 'Romaneschi', **d** 'Catanesi'

polymorphism (AFLP) (Lanteri et al. 2004b). This study has shown that the genetic variability observed within accessions belonging to same varietal type was higher than that found between different accessions of the same varietal type, confirming the multiclonal composition of many seed-propagated cultivars. The globe artichoke genome was recently decoded (Scaglione et al. 2016). The assembly covers 725 of 1084 Mb that constitute the genome of the species. The sequence codifies for about 27,000 genes. The understanding of the genome structure of the globe artichoke is crucial to identify the genetic basis of agronomic characteristics and the future application of selection programs. In the Mediterranean areas, globe artichoke cultivation is characterized by the diffuse use of local genotypes, which are often marked by a multiclonal composition, with a very broad genetic base and also with a high degree of heterozygosity (this is the reason of the widespread recourse to the agamic propagation). Despite the availability of such a wide biodiversity, globe artichoke cultivation in these areas is still mostly based on the large-scale cultivation of a few cultivars. In recent years, however, a strong germplasm innovation is taking place, thanks to the introduction, in some areas, of new agamic propagated or seed-propagated cultivars. The increase in biodiversity, indeed, is the best way to: (i) meet the market requirements; (ii) better satisfy the productive chains (fresh consumption and processing industry); (iii) give temporal continuity to product quality; and (iv) improve farmers' incomes.

The globe artichoke breeding program established in many countries, especially in France (since 1965), Israel, USA, Italy and Spain by some seed companies, has focused on the creation of seed-propagated cultivars. Some of the seed-propagated cultivars are open-pollinated, while most are F_1 hybrids. Currently, many seed-propagated cultivars ('Opal', 'Concerto', 'Madrigal', 'Harmony', 'Opera', 'Capriccio', 'Romolo', 'Imperial Star', 'Green Globe', etc.) are grown in the world both for fresh products and for industrial processing, as they allow the crop to be treated as annual, reduce the cost of planting, the use of pesticides, the use of fertilizers and needs for watering. Significantly, seed propagation allows new varietal types to be both propagated and diffused more rapidly than vegetatively propagated ones. The development of in vitro culture techniques, to obtain healthy, genetically uniform and rapidly propagated material, has recently allowed, in Italy, to spread some cultivars ('Apollo', 'Romanesco Clone C3', 'Exploter', etc.). Regarding cultivated cardoon, there are not many cultivars on the market today. They appear in the catalogues of a few seed companies, and they are the result of selection for 'leaves' production for human consumption. The objectives of such selection are the size of the plant, the colour and the texture of the petioles and leaf lamina. These genotypes, unlike globe artichoke, are marked by a lower degree of heterozygosity (from which derives the possibility of gamic propagation) and a minor genetic variability (Portis et al. 2005b). Commercial cultivated cardoon genotypes are obtained mostly through mass selection based on the maternal parent only; this is the reason why

within each cultivar there is some genetic and phenotypic variation, unlike the globe artichoke F_1 hybrids. Genetic improvement has until now marginally concerned this crop, especially regarding energy uses. Cultivated cardoon for energy purposes may be chosen among traditional genotypes characterized by large biomass production. This is the reason why specific programs of cultivar constitution are needed, in order to implement cardoon cultivars specifically suitable for biomass and achenes production. In this regard, some researchers of the University of Catania have selected 'Altilis 41', a genotype of cultivated cardoon suitable for the production of biomass for energy purposes. In Italy, the most common cultivars of cultivated cardoon are 'Bianco avorio' (vigorous, spineless) and 'Gobbo di Nizza', both grown in Piedmont. 'Bianco pieno migliorato', 'Pieno inerme' and 'Gigante di Romagna' are cultivated in Emilia Romagna. In Spain, unlike Italy, cultivars are distinguished based on the use, so there are cultivars for fresh consumption and others for the processing industry. In France, the most used cultivars are 'Blanc améliorée' and 'Rouge d'Alger'.

The 36th complete edition of EU common catalogue of varieties of vegetable species (European Commission 2017) and its 4 supplements (2018) include 56 globe artichokes and 33 cultivated cardoon cultivars.

1.4 Congener Species

The taxonomy of the species belonging to genus *Cynara* is not so simple, because of the morphological similarity within *Cynareae* tribe, which requires the use of distinctive, not always evident, characteristics (Foury 2004). According to Rottenberg and Zohary (2005), *Cynara* genus comprises the botanical varieties of *C. cardunculus* L., as well as other six (maybe seven) species, all native to the Mediterranean Basin. The other species of the genus belong to the Mediterranean native flora and, like the wild cardoon, are perennial herbaceous plants with bushy habitus, thorny and with diploid chromosome number $2n = 2x = 34$ (at least in the accessions studied to date); some of these species have highly endemic distribution (Bazniski and Zohary 1994).

According to Wiklund (1992), other species of the genus *Cynara* are:

- *Cynara humilis* L.;
- *C. cyrenaica* Maire & Weiller in Maire;
- *C. algarbiensis* Coss. ex Mariz;
- *C. cornigera* Lindley in Sibthorp & Lindley;
- *C. baetica* (Spreng.) Pau;
- *C. syriaca* Boissier;
- *C. auranitica* Post in Post & Autran.

In the above list, *C. syriaca* and *C. auranitica*, which in other classifications are considered as synonymous, are regarded as different species. Furthermore, *C. tournefortii* Boiss. & Reuter is not in this list, as Wiklund (1992) has treated it as a separate taxon.

Reproductive barriers separate the *C. cardunculus* complex from the other *Cynara* species. The crosses between *C. cardunculus* and any of the other species *C. syriaca*, *C. algarbiensis*, *C. baetica* or *C. humilis* all produce few seeds, and the hybrids are generally sterile (Rottenberg and Zohary 1996). These four wild *Cynara* species are therefore regarded as members of the secondary wild gene pool of globe artichoke and cardoon (Rottenberg and Zohary 2005). On both morphological (Wiklund 1992) and cytogenetic (Rottenberg et al. 1996) grounds, the closest of these species to the cultivated complex is *C. syriaca*. The monophyly and evolution of the *Cynara* spp. have been investigated between various internal transcribed spacer (ITS) regions (Robba et al. 2005; Sonnante et al. 2007a), leading to the suggestion that the *C. cardunculus* complex is more differentiated and evolved than other wild species (Mauro et al. 2009). *Cynara* species are generally robust, herbaceous plants. Plant size ranges between 0.5 and three metres, in some samples. Leaf colour is green with varying intensity. *C. cornigera* has variegated leaves, similar to those of *Silybum marianum*. The typical head is common to all

species and is constituted of glabrous bracts forming an involucre that contains small flowers. Flower is generally blue-purple, but can also be white.

1.5 Traditional Uses

1.5.1 Globe Artichoke

The main product from globe artichoke plant is the head, mostly used for fresh consumption and to a small extent destined for the processing industry. The Commission Implementing Regulation (EU) No 543/2011, following the Commission Regulation (EC) No 1221/2008 (repealed), establishes a general marketing standard for globe artichoke and for other 25 fruit and vegetable products. According to this standard, globe artichokes for fresh consumption must have minimum quality requirements and must therefore be intact, sound (free affected by rotting or deterioration), clean, practically free from pests, practically free from damage caused by pests, free of abnormal external moisture and free of any foreign smell and/or taste. Globe artichoke is also marketed in various types of processed products. The processing industry mainly uses small capitula harvested at the end of the season (April and May in Mediterranean countries), when the market price is low. Physical, chemical and biotechnological processes may be involved in the transformation. Globe artichoke can physically be treated with high or low temperature (heat treatments, freezing), or chemically processed with organic acids—to lower the pH of the governing liquid—or salt—to lower the activity of water. The biotechnological processes consist in the production of lactic acid through the spontaneous or controlled fermentation carried out by bacteria. The stabilizing heat treatment often heavily affects the nutritional and organoleptic characteristics of the processed product, which, in the end, differ significantly from the fresh ones. With low temperature treatments, products with organic characteristics similar to fresh ones are obtained, but they are characterized by a shorter shelf-life compared to the heat-treated product and, furthermore, require an adequate cold chain for their distribution. Among the traditional processed products, globe artichokes in oil represent the most important. They can be obtained from semi-finished products. In this case, the previously calibrated capitula are blanched in a citric solution and then subjected to turning and trimming to obtain ‘hearts’ or ‘bottoms’. The latter may or may not undergo lactic fermentation, depending on the brine concentration. The semi-finished products thus obtained, stored in drums and sent to the industries for further processing, are then subjected to a washing to eliminate the excess salt before being placed in glass cans that are filled by a governing liquid consisting mainly of oil. Finally, the hermetically sealed jars are subjected to heat treatment. Artichokes in oil with better organoleptic characteristics can be obtained without the previous semi-processing step. In this case, the turned and trimmed capitula undergo a blanching in a citric solution and a short treatment with an acetic acid solution. Subsequently, the glass cans are filled and the oil is added to the product at a temperature of about 90 °C. Pasteurization follows the vacuum sealing. During storage, the globe artichokes in oil may show white inulin crystals (Le Roux et al. 1978; Almela et al. 2004). The deep-frozen hearts of globe artichoke are also widespread industrial products. Their production starts with the turning and trimming of capitula, followed by their immediate immersion in citric or ascorbic acid solution, in order to prevent the browning on the cutting surfaces. The so obtained hearts are then immersed in boiling water. After the blanching, the product is pre-chilled and then deep-frozen at −40 °C. Other globe artichoke processing products are dehydrated, freeze-dried, pre-cooked, creams, pickles and vegetable soups. There are also new products that are still under study, such as minimally processed and ready-to-use globe artichokes. Nowadays, there is an increasing demand for fresh and ready-to-use products. However, minimal processing operations can cause undesirable changes in the sensorial, nutritional and health-promoting properties of the

product (Shahidi 1997). Furthermore, globe artichoke has a high polyphenol content, and this feature makes it very susceptible to browning. Most studies on minimally processed globe artichokes have focused on efficient ways to reduce browning and the growth of microorganisms, as well as on the use of innovative packaging and the choice of the most appropriate genotype (Del Nobile et al. 2009; Amodio et al. 2011; Restuccia et al. 2014; Pandino et al. 2017).

1.5.2 Cultivated Cardoon

Cultivated cardoon features in many Italian dishes, but its use is not so common. Nevertheless, there are several cultivated cardoon cultivars, often linked to the areas where they are mostly produced. The offshoots represent the edible product, which have thick and fleshy leaf veins. To this end, offshoots undergo the 'whitening' technique before their harvest. This technique is carried out during the autumn months (from September to November) and aims to promote the etiolation of leaf veins in order to make them tender and juicy. Traditionally, this technique consists in tying the leaf rosette in the upper third, then covering it with opaque sheets (plastic, paper etc.). When offshoots are 'ripened', they are gradually collected from November to April. The manual harvest is carried out by cutting the plant at the collar. The marketable yield, i.e. without the outer leaves and the upper blade, reaches 15–20 t ha^{-1}. As previously mentioned, the cultivated cardoon is consumed mainly fresh or after cooking, but it is also consumed preserved in oil, while the frozen product loses its characteristic taste and is therefore not recommended. Like its closest relatives, the cultivated cardoon is also used to obtain vegetable rennet.

1.5.3 Wild Cardoon

The wild cardoon is traditionally used as vegetable rennet for production of typical cheeses in some countries such as Portugal, Spain, Morocco and Italy. The mechanisms and the procedures of cheese production using *C. cardunculus* rennet are discussed in the Sect. 4.7.

References

Almela P, Fernández-López JA, Almela L, Muñoz JA (2004) Glucidic content in canned heart artichoke. Acta Hortic 660:563–567

Amodio ML, Cabezas-Serrano AB, Peri G, Colelli G (2011) Post-cutting quality changes of fresh-cut artichokes treated with different anti-browning agents as evaluated by image analysis. Postharvest Biol Technol 62:213–220

Archontoulis SV, Struik PC, Vos J, Danalatos NG (2010) Phenological growth stages of *Cynara cardunculus*: codification and description according to the BBCH scale. Ann. Appl. Biol. 156(2):253–270

Baggio MI, Palla F, Boscardin DS, Mantovani N, Grando MF, Augustin L, Suzin M, Donida B, Lombardo S, Mauromicale G (2011) Floral biology of globe artichoke (*Cynara cardunculus* var. *scolymus*) 'Nobre-UPF', a Brazilian cultivar. Acta Hortic 942:297–302

Basnizki Y, Zohary D (1994) Breeding of seed planted artichoke. Plant Breed Rev 12:253–269

Bianco VV (1968) Bibliography on artichokes. University of California, Davis, USA

Bianco VV (1990) Carciofo (*Cynara scolymus* L.). In: Bianco VV, Pimpini F (eds) Orticoltura. Pàtron Editore, Bologna, Italy, pp 209–247

Bianco VV (2005) Present situation and future potential of artichoke in the Mediterranean basin. Acta Hortic 681:39–58

Bianco VV (2012) The artichoke: a travelling companion in the social life, traditions and culture. Acta Hortic 942:25–40

Craig J (1858) A new universal etymological, technological and pronouncing dictionary of the English language: embracing all terms used in art, science, and literature, vol 1. Routledge, London, UK

CPVO (2013) Protocol for tests on distinctness, uniformity and stability - *Cynara cardunculus* L.—artichoke, cardoon (CPVO-TP/184/2). http://cpvo.europa.eu/sites/default/files/documents/cynara_2.pdf

De Candolle A (1885) Origin of cultivated plants. D. Appleton, New York, USA

Del Nobile MA, Conte A, Scrocco C, Laverse J, Brescia I, Conversa G, Elia A (2009) New packaging strategies to preserve fresh-cut artichoke quality during refrigerated storage. Innov Food Sci Emerg Technol 10:128–133

Dellacecca V, Magnifico V, Marzi V, Porceddu E, Scarascia Mugnozza GT (1976) Contributo alla conoscenza delle varietà di carciofo coltivate nel mondo. Atti II Congresso Internazionale Carciofo. Edizioni Minerva Medica, Torino, pp 199–316

European Commission (2011) Commission implementing regulation (EU) No 543/2011 of 7 June 2011 laying down detailed rules for the application of Council Regulation (EC) No 1234/2007 in respect of the fruit and vegetables and processed fruit and vegetables sectors. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32011R0543&from=EN

European Commission (2017) Common catalogue of varieties of vegetable species—36th complete edition. https://publications.europa.eu/en/publication-detail/-/publication/3f3f1355-d987-11e7-a506-01aa75ed71a1/language-en/format-PDF/source-74108706

FAOSTAT (2018) FAO Statistical Database. Faostat. http://www.fao.org/faostat/en/#home

Foury C (1967) Étude de la biologie florale de l'artichaut (*Cynara scolymus* L.); application a la sélection (1er partie). Ann Amelior Plant 17:357–373

Foury C (1989) Ressources génétiques et diversification de l'artichaut (*Cynara scolymus* L.). Acta Hortic 242:155–166

Foury C (2004) Origine, evoluzione e miglioramento delle piante da orto. Tecn Agric 56:75–120

Foury C (2005) Some historical sketches on artichoke and cardoon. Acta Hortic 681:29–38

Gominho J, Curt MD, Lourenço A, Fernández J, Pereira H (2018) *Cynara cardunculus* L. as a biomass and multi-purpose crop: a review of 30 years of research. Biomass Bioenergy 109:257–275

Ierna A, Mauromicale G (2010) *Cynara cardunculus* L. genotypes as a crop for energy purposes in a Mediterranean environment. Biomass Bioenergy 34 (5):754–760

Jannaccone A (1967) Ricerche sperimentali sulla biologia del carciofo (*Cynara scolymus* L.). Riv Agron 1:24–30

Lanteri S, Acquadro A, Saba E, Portis E (2004a) Molecular fingerprinting and evaluation of genetic distances among selected clones of globe artichoke (*Cynara cardunculus* L. var. *scolymus* L.) 'Spinoso sardo'. J Hortic Sci Biotechnol 79:863–870

Lanteri S, Saba E, Cadinu M, Mallica GM, Baghino L, Portis E (2004b) Amplified fragment length polymorphism for genetic diversity assessment in globe artichoke. Theor Appl Genet 108:1534–1544

Le Roux P, Le Meur JF, Cleret JJ, Bourgeois C (1978) Prévention de la cristallisation de l'inuline dans le conserve d'artichaut. Indus Agric Aliment 95:597–605

Liddell H, Scott R (1996) A Greek-English Lexicon, 9th edn. Clarendon Press, Oxford, UK

Mauro R, Portis E, Acquadro A, Lombardo S, Mauromicale G, Lanteri S (2009) Genetic diversity of globe artichoke landraces from Sicilian small-holdings: implications for evolution and domestication of the species. Conserv Genet 10:431–440

Mauromicale G (1987) Panorama varietale del carciofo e sua possibile evoluzione. Inf Agrar 43:69–75

Mauromicale G, Copani V (1990) Tecniche per regolare la produzione degli organi di moltiplicazione in *Cynara scolymus* L. Tecn Agric 42:5–14

Mauromicale G, Ierna A (2000) Characteristics of heads of seed-grown globe artichoke [*Cynara cardunculus* L. var. *scolymus* (L.) Fiori] as affected by harvest period, sowing date and gibberellic acid. Agronomie 20(2):197–204

Morone Fortunato I, Vanadia S, Lattanzio V (1981) Osservazioni sulla transizione a fiore dell'apice vegetativo di quattro varietà di *Cynara scolymus* L. Studi sul carciofo. Atti III Cong. Int. di Studi sul Carciofo. Laterza, Bari, pp 255–264

Pandino G, Barbagallo RN, Lombardo S, Restuccia C, Muratore G, Licciardello F, Mazzaglia A, Ricceri J, Pesce GR, Mauromicale G (2017) Quality traits of ready-to-use globe artichoke slices as affected by genotype, harvest time and storage time. Part I: Biochemical and physical aspects. LWT Food Sci Technol 76A:181–189

Porceddu E, Dellacecca V, Bianco VV (1976) Classificazione numerica di cultivar di carciofo. Atti II Congresso Internazionale Carciofo. Edizioni Minerva Medica, Torino, pp 1105–1119

Portis E, Acquadro A, Comino C, Mauromicale G, Saba E, Lanteri S (2005a) Genetic structure of island populations of wild cardoon [*Cynara cardunculus* L. var. *sylvestris* (Lamk) Fiori] detected by AFLPs and SSRs. Plant Sci 169:199–210

Portis E, Barchi L, Acquadro A, Macua JI, Lanteri S (2005b) Genetic diversity assessment in cultivated cardoon by AFLP (amplified fragment length polymorphism) and microsatellite markers. Plant Breed 124:299–304

Portis E, Mauromicale G, Barchi L, Mauro R, Lanteri S (2005c) Population structure and genetic variation in autochthonous globe artichoke germplasm from Sicily Island. Plant Sci 168:1591–1598

Quattrocchi U (1999) CRC world dictionary of plant names: common names, scientific names, eponyms, synonyms, and etymology. CRC Press, Boca Raton, USA

Restuccia C, Lombardo S, Pandino G, Licciardello F, Muratore G, Mauromicale G (2014) An innovative combined water ozonisation/O3-atmosphere storage for preserving the overall quality of two globe artichoke genotypes. Innov Food Sci Emerg Technol 21:82–89

Robba L, Carine MA, Russell SJ, Raimondo FM (2005) The monophyly and evolution of *Cynara* L. (Asteraceae) *sensu lato*: evidence from the internal transcribed spacer region of nrDNA. Plant Syst Evol 253 (1–4):53–64

Rottemberg A, Zohary D (1996) The wild ancestry of cultivated artichoke. Genet Resour Crop Evol 43:53–58

Rottemberg A, Zohary D (2005) Wild genetic resources of cultivated artichoke. Acta Hortic 681:307–314

Rottemberg A, Zohary D, Nevo E (1996) Isozyme relationship between cultivated artichoke and the wild relatives. Genet Resour Crop Evol 43:59–62

Scaglione D, Reyes-Chin-Wo S, Acquadro A, Froenicke L, Portis E, Beitel C, Tirone M, Mauro R,

Lo Monaco A, Mauromicale G, Faccioli P, Cattivelli L, Rieseberg L, Michelmore R, Lanteri S (2016) The genome sequence of the outbreeding globe artichoke constructed de novo incorporating a phase-aware low-pass sequencing strategy of F_1 progeny. Sci Rep 6, Article number:19427

Shahidi F (1997) Natural antioxidants: an overview. In: Shahidi F (ed) Natural antioxidants: chemistry, health effects and applications. AOCS Press, Champaign, USA, pp 1–7

Skeat W (1887) An etymological dictionary of English language. Clarendon press, Oxford, UK

Small E (2009) Top 100 food plants—the World's most important Culinary crops. NRC Research Press, Ottawa, Canada

Sonnante G, Carluccio AV, Pignone D, Vilatersana R (2007a) Variable rDNA regions provide suggestions on artichoke domestication and evolutionary history. Acta Hortic 730:123–125

Sonnante G, Pignone D, Hammer K (2007b) The domestication of artichoke and cardoon: from Roman times to the genomic age. Ann Bot 100:1095–1100

Wiklund A (1992) The genus *Cynara* L. Bot J Linn Soc 109:75–123

Cynara cardunculus Propagation

2

Nicola Calabrese, Vanina Cravero and Mario A. Pagnotta

Abstract

The propagation in *Cynara cardunculus* is carried out following different procedures of both vegetative and seed propagation. In the countries of the Mediterranean basin, where the globe artichoke is grown on about 80% of the world's surface, the propagation is usually made by vegetative methods using offshoots, 'ovoli' or rhizome parts. Conversely, wild and cultivated cardoons are commonly propagated by seed. The vegetative propagation of artichoke is often cause of agronomic and pathological disadvantages; to overcome these problems, new genotypes coming from 'in vitro' multiplication or from seed propagation methods have been successfully introduced in the last twenty years in all producing countries, and now their spread is constantly increasing. Recently, in the seed propagation group, F1 hybrids are commercially used. In the present chapter, the vegetative and seed propagation methods, the techniques to obtain healthy seedlings for multiplication, the influence of environmental factors on the seed production and seed germination parameters are widely described. The genetic at the base of new varieties development and the methods applied to obtain globe artichoke F1 hybrids are also reported together with their procedures and advantages versus disadvantages.

2.1 Introduction

Globe artichoke (*Cynara cardunculus* var. *scolymus* L.) is a species that is consumed from 850 BC. The first cultural methods date back to twelfth century (Coulter 1941). It is an allogamous species, with two alternative propagation systems: by seeds (sexual) and vegetative (also called asexual, agamic or clonal).

Globe artichoke is native to the Mediterranean region; however, during migration process, Italian, Spanish and French immigrants introduced this culture in other European areas and even in America (especially in Argentina, Peru, Chile and California).

In the main producing countries, Italy, Spain, Egypt, France, Tunisia, Algeria and Turkey, where cultivation dates back to several centuries ago, the globe artichoke propagation is usually

N. Calabrese
Institute of Science of Food Production—National Council of Research, Via G. Amendola 122/o, 70126 Bari, Italy

V. Cravero
Instituto de Investigaciones en Ciencias Agrarias de Rosario (IICAR, CONICET-UNR), Campo Experimental Villarino (S2125ZAA), Zavalla, Santa Fe, Argentina

M. A. Pagnotta (✉)
Department of Agricultural and Forestry Sciences, Tuscia University, Via S. C. de Lellis, snc, 01100 Viterbo, Italy
e-mail: pagnotta@unitus.it

E. Portis et al. (eds.), *The Globe Artichoke Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-030-20012-1_2

carried out by vegetative methods, by means offshoots, 'ovoli' (underground dormant shoots) or rhizome parts, often self-produced by the farmers who use poorly selected plant material taken directly from the commercial globe artichokes open fields (Ledda et al. 2004; Mauromicale et al. 2004; Macua et al. 2005; Ouselati and Ghezal 2009). The multiplication by seed is common in China, USA, Peru and other countries where the commercial cultivation of the artichoke has occurred more recently.

Cultivated cardoon is usually propagated by seed; the current cultivars have been obtained by public institutions or private companies starting from mass selection on the basis of the phenotypic conformity of the maternal plant phenotype. The commercial cultivars show high genetic heterogeneity as shown by Portis et al. (2005).

Wild cardoon is not a commercial product. It is naturally propagated by seed, especially in environments that are little modified by men.

The plant material used for the artichoke vegetative multiplication is characterized by large variability in terms of age, physiological stage, shape, size, position on the mother plant and number of buds present on rhizome (La Malfa and Foury 1976; Bianco 1990). This method, very simple to apply, has caused over time serious agronomic and pathological problems, with also economic negative effects for the producers. Several studies have reported a progressive decreasing of yield and quality characteristics of heads and the spread of diseases caused mainly by viruses and fungi (Rana et al. 1992; Gallitelli et al. 2004; Amenduni et al. 2005; Cirulli et al. 2010).

The 'in vitro' multiplication enables the production of healthy seedlings with higher multiplication rate (Ancora et al. 1981), solving the sanitary problems associated with the traditional propagation methods (Pécaut et al. 1983). Nevertheless, the high cost of micro-propagated plants, and in addition to the early cultivars, the loss of precocity of production are still a limit to their greater diffusion (Saccardo 2009). As to these reasons, seed-propagated cultivars, open pollinated or hybrids, are currently available for the growers and represent a good alternative to the traditional ones, and their diffusion is rapidly increasing all over the world.

Studies on new seed-grown cultivars underlined the high yield and good heads quality (Mauromicale et al. 1989; Calabrese et al. 2004; Miguel et al. 2004; Baixauli et al. 2007; Macua and Lahoz 2016a). Seed-propagated plants are usually grown as annual crops with positive agronomic and environmental effects.

The new F1 hybrid is characterized by contemporary harvests, high productivity, good quality of heads and suitability for processing (Calabrese et al. 2007, 2011; Lombardo et al. 2012; Di Venere et al. 2016; Macua and Lahoz 2016b).

2.2 Vegetative Propagation

2.2.1 An Overview

Vegetative propagation is any form of asexual reproduction (without fertilization process that involves gametes or sex cells fusion) occurring in plants in which a new plant grows from a fragment of the mother plant or from a specialized reproductive structure. It can occur naturally or be induced. This kind of reproduction is based on the totipotential cells which allow that genetically identical copies are generated from a single mother plant. The set of descendant plants obtained by vegetative propagation from a single plant is known as **clone**.

Cloning allows to perpetuate high-performance genotypes, selected for their superiority in some traits of men interest, without worrying of the segregation of possible heterozygote genes. Members of a clone show phenotypic uniformity, and they will have similar appearance, size, flowering, ripening and harvest dates (Gutierrez Caro and Ipinza Carmona 1998). All variability observed between plants has environmental origin.

2.2.2 Globe Artichoke Vegetative Propagation

In globe artichoke, different organs are used for vegetative propagation depending on the geographical region customs: offshoots, ovoli or stumps (Fig. 2.1).

Offshoots are shoots developed from buds formed on the underground stem which are able to develop roots, and are the most used procedure of globe artichoke vegetative propagation. A single mother plant develops many offshoots over time, so they may have differences on size. Using these offshoots to establish a new plantation or to replace missing plants could have negative impact as low implantation rate and variations in plant development (Harwood and Markarian 1967). Offshoots selected for propagation should have 2–3.5 cm basal diameter (García 2005), 15–20 cm length (Tesi et al. 2003) and a well-developed root system (Bianco 1990).

Ovoli are underground parts of shoots with one apical and several lateral buds. They are collected during summer when the aerial part of the mother plant is dried (Bianco 1990). A single plant could produce more than 20 ovolis; nevertheless, by cutting the apical bud or by tamping operation, this number can be increased. Ovolis that reach 13–14 cm length and 1–3.5 cm diameter gives the best results for propagation use (Ciancolini 2012).

Stumps are also obtained at the end of the productive cycle of the plant. They are basal stem pieces with attached root sections which have several buds able to generate new plants. This propagation system generates high heterogeneity between plants and late production (Bianco 1990).

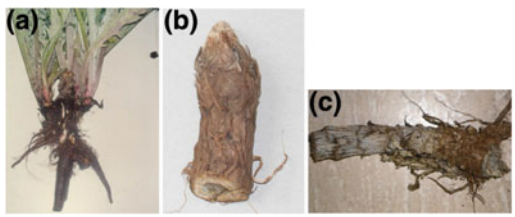


Fig. 2.1 Vegetative propagation organs: **a** offshoots, **b** ovoli and **c** stump

According to records, both kinds of propagation (sexual and asexual) were used equally and simultaneously at the beginning of the culture; however, sexual multiplication was turned down during the seventeenth century due to the low quality and heterogeneity observed between plants obtained by seed (see Sect. 2.4.2). The increase of vegetative propagation gave rise to clones with good adaptive characters which originated most of the cultivars currently used (Pécaut 1993).

2.2.3 Clones Around the World

Many artichoke clonal populations are grown in different geographical areas, where they are usually named according to the locality of cultivation (Bianco 1990). Like this, a unique cultivar could be called differently depending on the cultivation region, while, in the same way, different cultivars could take the same name. An example is the cultivar "Violetto di Sicilia" which was cultivated in Sicily, Apulia and Sardinia and for which 41 alternative names are known. Currently, also 23 different names are known around the world for "Romanesco", 14 for "Spinoso Sardo", 10 for "Spinoso di Palermo" and seven for "Violetto di Provenza" and "Violetto di Toscana". Likewise, "Blanca de Tudela" (originated in Tudela, Spain) was called "Blanco de San Juan" when it was introduced in Argentina (due to the culture location in San Juan Province). Later, it was introduced in Chile where it was named "Argentina". Similar situation was observed for "Romanesco" and "Precoce di Jesi"; both Italian cultivars were called "Frances" and "Precoz Italiano", respectively, when they were introduced in Argentina (García et al. 2005).

2.2.4 Clones Improvement and Cultural Practices

While only few cultivars were spreading all around the world, new intraclonal variability has been generated throughout all the culture history due to the allogamous nature of the species and the emergence of mutations with positive or neutral effects on the culture, which remained over the time. This variability is showed as heterogeneity in the culture and in the production. Several farmers and breeders have made efforts to practise intraclonal selection in order to restore the uniformity of the crops. In this way, positive results have been obtained in France over the cultivar "Violeta de Provenza" (Pochard et al. 1969), in Spain over "Blanca de España" (Trigo-Colina 1980), in Italy over "Violetto di Sicilia" (Abatte and Noto 1979; Mauromicale 1987) and in Argentina over "Francés" (Cointry et al. 1999). In all cases, traits related to yield and earliness were improved.

In most production areas in which the asexual multiplication system is applied, the cultural practices carried out include soil mechanical tillage between lines, manual weeding between plants (especially during the culture establishment), phytosanitary treatments, plant growth regulators application, irrigation (by flooding, dripping or spray), fertilization, and autumn and spring thinning. However, to determine the best dates of implantation and harvest, it is necessary to establish two main groups of countries: those belonging to the northern hemisphere and those belonging to the southern one. In Mediterranean countries (Spain, Italy, France, Morocco, Algeria, Greece and Egypt), plantations are carried out in summer (July–September) to obtain spring production; nevertheless, depending on the climate and the cultivars earliness, farmers can apply specific cultural practices in order to change the harvest period obtaining early production (autumn–winter). In USA, the culture is implanted during spring obtaining harvest in autumn in the first culture year. In the southern hemisphere, there is a greater variability with respect to the planting and harvest dates. In Chile, planting is carried out in summer (December–February) when it is made from stumps while, when offshoots were used, it is carried out at the end of this season (March). In Peru, it is necessary to distinguish the coast to the mountain areas; in the first one, implantation is carried out at the beginning of autumn (end of March–April) and harvest is performed in spring (August–November), whereas in the mountain areas the culture is implanted at the beginning of spring (September) to harvest between January and March (summer–autumn). In Argentina, also different areas should be taken into account; in La Plata, offshoots are planted in spring obtaining harvest from late autumn, whereas in Rosario plantations are carried out in autumn (March–April) also from offshoots and in Cuyo, stumps are planted in summer (January–February) (García et al. 2005).

2.2.5 Clones Propagation and Agronomical Issues

Gibberellic acid application is the most common practice adopted worldwide to obtain early production, especially in late cultivars (Pochard 1964).

Traditionally, globe artichoke is growing as a perennial crop that is maintained between three and six (or more) years in the field. A rosette of leaves is formed during the vegetative period, followed by a floral stem elongation. Roots and the underground part of the stem become fleshy and form the storage organ. At the end of the productive cycle, the leaves rosette dry and die. The aboveground biomass is cut to stimulate new buds development. Some of them grow as soon as the stem apex becomes reproductive, others when the harvested stem is cut; however, most buds stay dormant (Pécaut 1993). The culture is usually handled as single plant in order to avoid a high-density plantation. Year to year, offshoots are removed and used to replace any losses or increase the plantation size. Sometimes, two

offshoots per plant are left to allow staggered harvest, since the most vigorous offshoot will begin the productive phase before than the less developed one (García et al. 1998).

This kind of clonal multiplication is rather easy and inexpensive and has the advantage of maintaining over the time, genotypes with desirable qualitative traits (such as spine absence) of more complex ones (such head size) (Porceddu et al. 1976). Also, it is possible to perpetuate certain gene combinations generated by crosses between clones which show high heterosis levels (Hammer 1988; Balloux et al. 2003).

However, show a series of disadvantages such as:

- Low multiplication rate—most cultivars produce three–five offshoots/plant/year, and the most prolific eight–ten (Harbaoui and Debergh 1980; Pécaut et al. 1983).
- Great heterogeneity in plant vigour and production due to differences in the propagation organs development and their rooting ability (La Malfa and Foury 1971; Mauromicale et al. 1989).
- High transplanting cost (Mauromicale et al. 1989), in the fixed costs composition, transplanting represents 17% of the total direct costs and 7–10% of the incomes, regardless of offshoot quality (Calabrese et al. 2004).
- High thinning costs: require 12 daily wages/ha (15% of variable costs) (Hang 1993).
- Fungi, bacteria and viruses spreading which produce several plant damages such as chlorosis or necrosis on leaves, stunting, wilting and collapsing of plants (Cardarelli et al. 2005a), causing drastic performance reductions and great economic losses.

Costs generated by thinning could be eliminated handling the culture as bush. Nevertheless, this practice would result in an increase of the culture density and could be detrimental for the culture performance, causing decrease in head size and number (Pécaut and Foury 1992; Schrader 1992; García et al. 1998).

2.2.6 Micropropagation

Micropropagation techniques allowed overcome most of these issues (Basniski and Zohary 1987; Calabrese and Bianco 2000). Moreover, in general, plants obtained by micropropagation tend to be more vigorous, earlier and more productive (Cointry et al. 1999). Nevertheless, this kind of propagation is expensive; efficient propagation protocols are only available for some few cultivars and cause somaclonal variants, especially in early cultivars like "Violeto di Provenza" (Pécaut and Martin 1992; Cardarelli et al. 2005b) which often show heteroblasty. Three variants have been described for this cultivar; "Pastel" with pinnate leaves and late production, "Bull" with entire leaves, similar precocity to wild type but with spherical heads and "Pastel–Bull" with pinnate leaves, spherical heads and late production. "Bull" and "Pastel–Bull" forms are tetraploid, whereas "Pastel", as well as the wild type, is diploid (Cointry et al. 1999). These variations, observed in "Violeto di Provenza" micropropagation, were later confirmed in other cultivars like "Catanese" (Rossi and De Paoli 1990), "Niscemese", "Liscio Sardo Prococissimo" and "Tudela" (Pécaut and Martin 1992). The somaclonal variation is different across genotypes and across subculture (Rey et al. 2013b). A micropropagation strategy has been proposed in order to proceed with variety registration (Rey et al. 2013b) or to make efficient and economical the in vitro germplasm conservation (Tavazza et al. 2015).

2.3 Seed Propagation of Globe Artichoke and Cultivated Cardoon

2.3.1 Overview

The seed propagation of artichoke and cardoon occurs through achenes, a monocarpellate, indehiscent dry fruit that contains a single seed (Bianco 1990). Commonly, artichoke and cardoon achenes are called, marketed and used as seeds because of their similarity to the seeds.

This kind of propagation represents a valid alternative to the vegetative reproduction, because it allows a greater rationalization of the cultivation techniques, increases the yield and improves the plants health status (Foury and Martin 1976; Mauromicale et al. 1989; Welbaum 1994; Calabrese and Bianco 2000; Calabrese et al. 2007).

In particular, the seed propagation promotes: (i) the use of healthy and certified transplants (in accordance with European and international rules) useful also for the traceability of fresh and processed products throughout all supply chain; (ii) the development of the nursery activities; (iii) the possibility to grow artichoke and cardoon as annual crops that, compared to the multi-year cropping cycle, enables a best crop rotation farm programmes reducing the environmental impact; (iv) the introduction of new genotypes more suitable for specific purposes (fresh market, processing, pharmaceutical and industrial products); (v) the cultivation in unfavourable pedo-climatic conditions as salinity of water irrigation and soil, use of unconventional irrigation waters, high/low temperatures during harvesting; and (vi) the development of long vertical taproots in direct-seeded plants, which penetrate into the soil deeper and utilize moisture and fertilizers more efficiently. Seed propagation prevents the transfer and the spread of soil-borne pathogens and a large number of virus infections that are common using offshoots, 'ovoli' or rhizome parts. In short words, the seed propagation makes the artichoke cultivation friendlier both to the growers and to the environment.

The studies to obtain seed-propagated cultivars began in France in the 1960s (Foury 1967, 1969; Pochard et al. 1969; La Malfa and Foury 1976) and later in Israel (Basnizki 1985; Basnizki and Zohary 1987). The first commercial seed-propagated cultivar was Talpiot, obtained from repeated self-pollination of the Italian cultivar Precoce di Jesi (Basnizki and Zohary 1987) followed by H044, H223, H137, H271, Orlando, Giorgio and Agribas (Mauromicale et al. 1989; Calabrese et al. 1994; 2004). In the same years in the USA, new seed cultivars as Emerald, Desert Globe, Imperial Star, Imperial Condor, etc., were marketed.

In recent years, a considerable number of new seed-propagated cultivars, open pollinated or hybrids, have been developed and successfully introduced in many countries, and they are gaining in popularity. Currently in the European plant variety database, a total of 55 varieties of globe artichoke are listed, 30 of which are hybrids, 18 registered from Italy, 7 from the Netherlands, 5 from Spain (http://ec.europa.eu). In the same database, 18 varieties of cardoon (9 from Italy, 6 from France and 3 from Spain) are also registered.

Estimates made in Italy in the 2017 indicate a surface of about 2000 ha (equal to 5% of the total artichoke harvested area) destined to seed-propagated cultivars, instead they reach about 2% of the total cultivated area in Spain and Egypt. In China (11750 ha), Peru (6100 ha) and USA (2750 ha), all surface is covered by seed-propagated cultivars (FAO 2018).

The achenes germination, the contemporary of the emergence and the morphological characteristics of the seedlings are influenced by numerous factors like: environmental conditions that occur during the achenes formation (Foury et al. 1978; Ortega 2002); parents genotype (Foury 1994); mother plant and seed age (Damato et al. 2007b); germination temperature (Mauromicale and Licandro 2002; Damato and Calabrese 2005b; Vannella et al. 2005); presence/absence of light during germination (Basnizki and Mayer 1985); chemical and physical characteristics of the germination substrate (Raccuia et al. 2004); seeds osmotic conditioning (Ierna et al. 2004; Damato and Calabrese 2007); treatment with growth regulators (Schrader 1990); control of parasites present on the seed teguments (Vilchez et al. 2005; Ziani et al. 2010).

Studies carried out in nursery to optimize the growing techniques of seedlings to be used for transplanting have shown that fertilization (Elia and Santamaria 1994; Santamaria et al. 2007), addition of mycorrhizae to the substrate (Campanelli et al. 2011, 2014) and the size of alveolate

trays greatly influence the biometric characteristics and the growth of seedlings.

The results of research concerning the artichoke and cardoon seed germination, as well as the transplants growing techniques, are reported below.

2.3.2 Factors Affecting Seed Germination

2.3.2.1 Temperature

The achenes germination is strongly influenced by the temperature. In general, the results of numerous studies have shown that, despite the diversity of the individual experimental conditions, the optimal germination temperature is 20 °C; however, it remains at satisfactory levels when the temperature is between 15 and 25 °C (Basnizki 1985; Foury 1994; Ierna et al. 2004; Damato and Calabrese 2005b; Lekić et al. 2011).

Basnizki and Mayer (1985) report that the rate and the speed of germination are higher at 15–20 °C, while with alternate temperatures 25–15 °C, the germination is slightly slower and it is considerably reduced at 10 and 30 °C. Light suppresses germination, but the light sensitivity differs among genotype.

Roots formation at 35 °C shows an abnormally thickening, and their ability to function is questionable. Germination is greater if it occurs in the dark; the presence of light stops or slows down considerably the germination. At temperatures above 30 °C, the achenes exuded a greyish mucilage which acts as barriers to gaseous diffusion, especially to O_2, blocking the germination. The excretion of mucilage might be a way to prevent germination in non-optimal conditions. In fact, by eliminating the mucilage and bringing the temperature to the optimum germination values, the achenes germinate normally (Basnizki and Mayer 1985).

The percentage of non-germinated seeds increases progressively with the increase in temperature and reaches 42% at 30 °C. The percentage of abnormal seedlings is lower at temperature of 20 °C (Damato and Calabrese 2005b). The achenes of hybrid cultivars, compared to open pollinated cultivars, show higher germination rate (84 vs. 65%) and two days earlier germination (Calabrese et al. 2007).

These indications are useful because the artichoke sowing and the seedlings growing cycle in the nursery often take place in late spring–summer. In these seasons, it is essential to control the temperature with appropriate cooling systems both in the seed germination rooms and in the nurseries.

2.3.2.2 Salinity

Salinity of irrigation water or soil is a widespread stress factor limiting seed germination, seedlings growth and productivity of cultivated plants. According to the study on salt tolerance of Mass and Hoffman (1977), the artichoke in terms of seed germination can be classified as moderately tolerant, even if the tolerance depends on genotype. Seed germination and seedling growth may be adversely influenced by both the high temperature (>30 °C) and the use of saline irrigation water (Basnizki 1985).

Mauromicale and Licandro (2002) observed that artichoke germination declined with decreasing osmotic potential of the germination medium, but with lower magnitude at 20 °C than at 30 °C. This indicates that artichoke significantly increases its sensitivity to salt stress at no-optimal germination temperature. The threshold of osmotic potential that reduced germination by 50% was about −1.70 MPa at 20 °C and increased to −0.90 MPa at 30 °C. Seedling emergence was 96% using tap water and decreased to 48% when the osmotic potential of water was −0.5 MPa; no emergence was detected at −1.0 MPa. More than 50% of emerged seedlings irrigated with water at osmotic potential of −0.5 MPa died after 4–5 days from emergence. The most sensitive phase to salinity during seed germination is immediately after the appearance of the cotyledon leaves.

Seed germination responses of cardoon genotypes subjected to different salt stress treatments with NaCl or polyethylene glycol (0.3–0.6–0.9 MPa) showed a severe reduction of germination percentage in relation to the genotype and the increasing of conductivity (Argento

et al. 2016). Similar results on seed germination of wild cardoon genotypes were found by Raccuia et al. (2004); the values at which germination percentage was reduced by 50% ranged between −0.18 and −0.51 MPa in NaCl solution, and between −0.18 and −0.68 MPa in polyethylene glycol (PEG). The germination percentage was higher in PEG than in NaCl and varied among genotypes.

2.3.2.3 Seed Priming

The 'priming' techniques allow seeds to germinate more rapidly under unfavourable environmental conditions, and it is an effective method to overcome thermo-dormancy in many vegetable species including artichoke. The priming techniques are carried out on artichoke seeds before the sowing with the aim to increase contemporaneity and speed of germination, synchronize seedling emergence, and improve seedling vigour, and it is based on seeds hydration treatments in liquid or humidified solid medium. Osmotic priming is a controlled hydration technique that brings seed to have water content which permits pre-germination metabolic activity to spin off and prevent radical emergence.

Positive results on artichoke achenes germination were obtained with matriconditioning by Damato and Calabrese (2005a, 2007). The germination was higher (64%) with the achenes conditioned for four days at −1.2 MPa. With the sowing performed immediately after conditioning, the highest germination value (67%) and the lowest abnormal seedlings (5%) were observed, while when the sowing carried out 30 days after conditioning, the germination value was only 36% and the abnormal seedling 9%.

2.3.2.4 Chemicals and Biological Treatments

Soaking artichoke seeds in 500–1000–2000 mg kg^{-1} ethephon solution for 5 min significantly increased the rate and uniformity of germination and did not affect the subsequent seedlings growth. Gibberellins and cytokinin seed treatments do not influence the germination rate (Schrader 1990). Fluctuating temperatures (25/15 °C) enhance seed germination in the presence of abscisic acid (ABA), but ABA content is not different in seeds submitted at fluctuating or constant (20 °C) temperatures (Huarte and Benech-Arnold 2010).

As effects of coating treatments with chitosan, a biopolymer with antifungal properties, on germination of artichoke achenes, microbial growth control and seedlings development were reported by Ziani et al. (2010). Highest germination rate (78%) was obtained using a 4% chitosan solution with pH values between 5.3 and 6 (also optimal range of antifungal activity). Chitosan solutions formed a protective film around the seeds that reduced the number of fungal parasites except *Rhizopus* spp. All chitosan treatments improved the seedlings roots growth, whereas combination of chitosan and thiram (TMDT) at low concentration gave a strong antifungal protection and improved seed germination and seedlings growth.

Artichoke seed treatment with 8-hydroxy quinoline sulphate, thiram, captan and sodium hypochlorite gave excellent control against *Rhizopus*, *Alternaria* and *Aspergillus* spp.; hot water (57 °C) had intermediate effect compared to the control. Better germination rate resulted in seeds treated with 8-hydroxy quinoline sulphate, thiram and captan (Vilchez et al. 2005).

Artichoke seed inoculation by different combinations of plant growth promoting rhizobacteria (PGPR) on germination and seedling growth characteristics was investigated by Jahanian et al. (2012). The highest germination percentage and the mean time of germination were observed with seeds inoculated by combination of *Pseudomonas putida*, *Azotobacter* and *Azospirillum*.

Seed germination and seedlings emergence of different cardoon genotypes under cadmium (Cd) and arsenic (As) stresses decreased with the increasing of As concentration in the medium from 80% at 10 and 50 μM to 30% at 200 μM. High concentrations of As, even in combination with Cd, stopped the radicle emergence from the germinated seed. Low concentration of As and all concentrations of Cd caused a reduction of the seedling growth. The response to heavy metals stresses also varied among genotypes (Pappalardo et al. 2016).

2.3.3 Seed Yield

Artichoke and cardoon are largely cross-pollinated crops; honey bees are the principal pollination vectors (Foury 1967). The flowering in artichoke and cardoon plants follows a regular pattern and can last even 40 days. The main head flowers about three weeks before third-order heads (Basnizki 2007). Head flowering is centripetal, and allogamy is favoured by protandry.

The fertilization occurs 18–24 h from pollination. The initials of cotyledons become visible 6–8 days later; they continue to develop for the next 25–30 days while the formation of the pericarp is delayed; at the end of cotyledons growth, the pericarp cell walls begin to thicken, to lignify and to produce pigmentation (Foury 1967). The achene physiological maturity usually occurs 40–50 days after pollination, and then they can germinate. The colour of pericarp varies from uniform light grey to dark brown with more or less pronounced streaks; the average size is 7 mm in length and 4 mm in width. The achene weight depends on genotype and environmental factors during flowering (Bianco 1990); usually, the weight of 1000 achenes ranges from 30 to 70 g.

Seed production in artichoke usually is not high. The number of seeds per heads is extremely variable and depends on the genotype, the climatic conditions during the flowering period, the presence of pollinating insects, the plant growing techniques and, especially, on the cultivation methods of the mother plants (Bianco 1990; Foury 1994).

In optimal pollination situation (best results are under dry weather conditions), only about the half of the florets of the head produce seeds (Basnizki and Zohary 1994). Under open pollination, the number of seeds per head ranges from 105 to 700 for cv. Balady and Violetto, respectively (Bianco 1990). The seed yield per head also differs between primary, secondary and tertiary heads even if the experimental results are disagreeing. Basnizki and Zohary (1994) reported that seed set was greater in the primary heads than in heads of secondary order and in well-developed plants, and the seed yield from primary heads commonly accounts about 50% of total plant seed yield. Foury et al. (1978) observed no differences in seed yield between primary and secondary heads while Ortega (2002), by cutting of primary heads before bolting, found that seed yield was higher on quaternary and tertiary than secondary heads.

Damato et al. (2007a) referred that artichoke seed yield can be also affected by plant density: the seed production from main head was highest at 1.3 plant m^{-2} (71 kg ha^{-1}), whereas at 1.0 plant m^{-2} the highest seed yield (35 kg ha^{-1}) was harvested from lateral heads. Two-year artichoke plants produced 2.4 and 1.7 t ha^{-1} with plant density of 2.8 and 4.8 plant m^{-2}, respectively (Jevdjovic et al. 2001). At the same plant density, the seed yield of one-year crop plants was 0.8 and 0.5 t ha^{-1}, respectively. Similar results were described by Damato et al. (2007b) that noticed a progressive increase of seed yield 49, 158 and 173 kg ha^{-1} from one-, two- and three-year crops, respectively. Harvest dates, removal central or lateral heads at different timing and GA_3 rate did not influence seed yield. Ierna and Mauromicale (2010) also reported that the crossing of artichoke with cultivated and wild cardoon improved its performance, allowing a twofold grain production.

Several studies carried out in Italy and other Mediterranean countries demonstrated the high variability of wild and cultivated cardoon response in terms of seed production performance. Genotype, plant density, weed control, fertilization, age of the crop, moderate drought during the growing and flowering season can affect cardoon seed yield (Foti et al. 1999; Piscioneri et al. 2000; Gherbin et al. 2001; Curt et al. 2002; Raccuia and Melilli 2007; Ierna et al. 2012; Raccuia et al. 2012; Vasilakoglou and Dhimab 2014; Boari et al. 2016). In cardoon, seed yield ranged from 0.2 to 2.8 t ha^{-1}, seed yield per plant varied from 7 to 126 g, and the 1000 seed weight from 26 to 46 g.

2.3.4 Strategies to Improve Nursery Growing Techniques

The use of seeds allows the new plantings to be set by direct seeding, but transplanting, due to the high cost of the hybrid cultivars seeds, is usually used to improve stand establishment, plant uniformity, growth performance, sanitary status of fields, marketable yield and to reduce the irrigation water consumption. The use of healthy and vigorous seedlings is a prerequisite for successful transplanting and optimal production.

In nursery, the sowing is mechanically carried out in alveolar plastic, polystyrene or cardboard containers. Transplantation usually occurs after 35–50 days from sowing when the seedlings have reached the stage of 3–4 true leaves and have a good root system.

The optimization of fertilization contributes greatly to the production of vigorous seedlings. Elia and Santamaria (1994) stated that the production of quality artichoke seedlings requires nutrient solutions that contain at least 130 mg L^{-1} of N and rates of 100 and 250 mg L^{-1} of P and K, respectively. Studies carried out by Santamaria et al. (2007) on seedlings of Concerto hybrid growing in nursery up to 44 days from sowing showed that the transplant height, leaf area and fresh weight were higher with urea (between 200 and 275 mg L^{-1}) than ammonium nitrate. With *N* dose increasing, plant height, number of leaves, leaf area, fresh weight, specific leaf area (SLA) and SPAD index increased. Leaf number was higher at 200 mg L^{-1} of N-rate, while higher dry matter content was at 50 mg L^{-1}. The highest N-rate had a positive effect on height, leaf area, fresh weight and SLA of the transplants. Leskovar and Othman (2016a) observed that transplants fertilized with low *N* level (75 mg L^{-1}), as compared to high N level (150 mg L^{-1}), positively improved root surface area, root length, root branching and thinner root diameter.

Growth and physiology under abiotic stresses in artichoke transplants were studied by Shinohara and Leskovar (2014). Heat stress inhibited shoot growth, while drought stress inhibited root growth; seedling physiological parameters were reduced by heat and drought stress combined. Abscisic acid at 1000 mg L^{-1} enhanced drought tolerance of transplants that was associated with the maintenance of shoot water status via stomatal closure.

The effects of exogenous applications of ethylene regulators at different concentrations on germination and early root growth of globe artichoke were evaluated by Shinohara et al. (2017). They suggest that exogenous ethylene could be useful to alleviate heat stress on artichoke seedlings, with increasing root hair density, root area and lateral roots, which in turn may improve early growth during stand establishment.

The use of mycorrhizal inoculation in the early stages of nursery cultivation enhances the vigour and health of plants. The practice of mycorrhizal inoculum is effective in improving the intrinsic quality of the artichoke plantlets produced in nurseries (Ruta et al. 2005). Mycorrhizal (*Glomus viscosum*) plants showed higher leaves number and leaf areas, greater shoot length, shoot, root density, root fresh and dry mass. This also corresponded with increased photosynthetic rates and stomatal conductance of mycorrhizal plants. Mycorrhizal colonization improves relative water content and increases proline concentration in plant tissue (Campanelli et al. 2011).

Grafting techniques have been studied both to improve artichoke resistance to diseases, especially *Verticillium* wilt, and tolerance to abiotic stress conditions. Seed-propagated artichoke cultivars could be successfully grafted onto cultivated cardoon rootstocks; the splice grafting technique proved to be more suitable than cleft grafting for globe artichoke. Under *Verticillium*-infested soils, the disease incidence in grafted plants was lower in comparison to ungrafted plants, and the agronomical performance of grafted globe artichoke was higher. The head yield can also be increased under non-infested soils, but in a different way depending on scion–rootstock combination (Temperini et al. 2013).

2.4 F1 Hybrids in Globe Artichoke

2.4.1 Overview of Crossing System and the Genetic at the Base of F1 Globe Artichoke Hybrids

After the great success in terms of yield of F1 hybrids in maize, starting in the early 1930s and continuing over years (Durvick 1984), F1 hybrid procedures were tested in other crop plants from rice to sunflower. The principle is to utilize hybrid vigour, also called heterosis, that results in a hybrid organism with higher value of characteristics like size, growth rate, fertility and yield, but also resistance to biotic and abiotic environmental rigours, comparing to those of its parents. More and higher are the differences between the uniting gametes (i.e. the parents) greater, on the whole, would be the hybrid vigour. **Heterosis** is often represented as the opposite of inbreeding depression, in which in an allogamous organism a series of self-crosses (or crosses among related individuals) results in the homozygosity of recessive alleles carrying negative characters. Hence, outcrossing should result in heterosis. As much as the parents are different as stronger will be the heterosis since different alleles will be mean and the number of heterozygote loci would be higher.

Recently, thanks to new tools such as genomics, transcriptomics, proteomics and metabolomics, questions about the causes of heterosis have been readdressed. It has been set that hybrid vigour is also due to 'epistatic interactions' and 'epigenetics factors' (Baranwal et al. 2012). Hybrids gain advantages from the control of circadian-mediated physiological and metabolic pathways, leading to growth vigour and increased biomass (Ni et al. 2009). DNA methylation machinery, which maintains and regulates epigenomic status, of the parental lines interacts to obtain hybrids. The siRNA population changes and non-additive expression of more than half of the miRNA were observed in *Arabidopsis* (Ha et al. 2009).

The Mendelian F1 generation is the result of a cross between two pure lines (i.e. lines homozygote at all their loci). In outcrossing plants, the prolonged selfing face with plant sterility, therefore sometimes the so-called F1 hybrids are just the first filial generation of offspring of distinctly different parental types (Runge and Patterson 2006) and are not true Mendelian F1. In fact, the hybrids are not obtained by crossing two pure lines, but crossing inbred lines, i.e. lines obtained after a series of self-pollination. Another advantage of the Mendelian F1 hybrids, in agreement with the first law of Mendel, is their uniformity. Uniformity is a requirement characteristic in all the varieties, but it is not always "automatically" reached in the globe artichoke F1 hybrids.

In choosing the parental lines for a F1 hybrid, together with their characteristics, it is important to evaluate their **combining ability** (Fasahat et al. 2016). The combining ability, which could be specific or general, was firstly introduced by Sprague and Tatum (1942). Combining ability is the estimation of a line (parents) ability to combine with another, or others, line/s (parents) on the basis of its offspring performance (Allard 1960). The **general combining ability** is due to the activities/properties of the genes as well as by the additive × additive interactions (Sprague and Tatum 1942), while the **specific combining ability** is mainly due to dominance effects and eventually to epistatic interaction, including additive × dominance and dominance × dominance interactions (Fasahat et al. 2016). General combining ability is calculated, for each trait, as the deviation of the mean offspring performance of a genotype from the grand mean of all offsprings included in the particular mating design. Specific combining ability is defined as the deviation of the performance of hybrid combinations from the performance expected on the basis of the general combining ability of the parental inbred lines. In hybrid breeding, the particular combination of inbred lines out of many possible combinations is selected which exhibits the highest F1 performance. Consequently, inbred lines to be used as parental are generally selected in order to have the highest specific combining ability. Different matting designs are set up to determine the general

combining ability (polycross or topcross) and the specific combining ability (factorials or diallels) (Zhang et al. 2015).

2.4.2 Historical Events in the Globe Artichoke F1 Hybrids Set up

The globe artichoke, cultivated in open fields, produces seeds that are the result of random crosses, and each single flower present in the inflorescence (head) could be crossed with pollen originated from different individuals. Therefore, the open crosses, without any external control, result in the production of extremely variable and different seeds individuals (or plants), also within the same head. The outcrossing habit of globe artichoke is not due to self-incompatibility but it is due to protandry, with a difference of about 2–3 days between anthers and stigma maturation. Pollen remains viable for 2–3 days depending on environmental conditions. Moreover, in the inflorescence, which could be containing 800–1400 flowers, the flower maturity is centripetal (i.e. peripheral flowers start first than central flowers). Hence, self-pollination is possible both within the same flower and within the same inflorescence.

The development and utilization of globe artichoke F1 hybrids have several milestones that are summarized in Fig. 2.2 (Saccardo et al. 2013). The first passage which made possible to start speaking about hybrids in globe artichoke is due to the research of Foury and Pécaut that beginning to set methods for control the globe artichoke pollination and evaluating germplasm more suitable for crosses (Foury 1969; Pécaut 1985, 1993). Pécaut and Foury (1992) were about the first to report heterosis phenomena in globe artichoke, even if it was already observed by several workers. They examined 21 cross

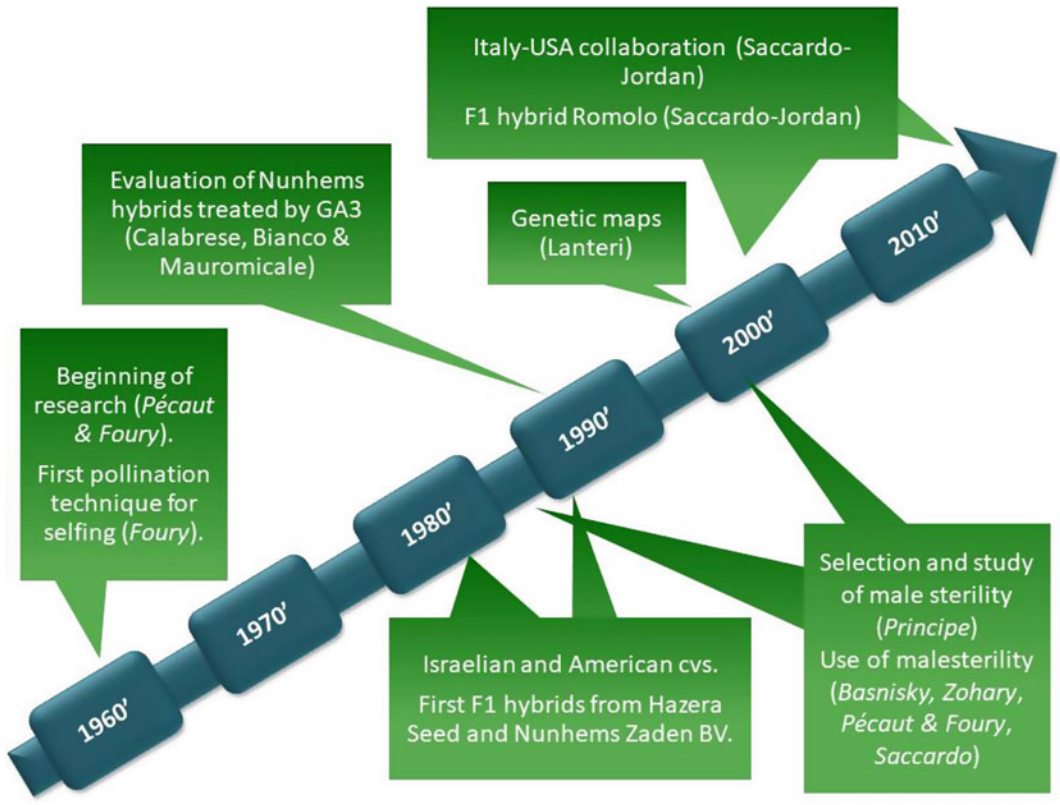


Fig. 2.2 Milestones of the breeding for globe artichoke hybrids

combinations among inbred (I_3–I_4) parental lines and recorded an increase, of more than 80%, in total yield of the globe artichoke F1 hybrids compared with their parents (Pécaut and Foury 1992).

The first globe artichoke hybrids were utilized in Israel and in USA thanks to the work of Basnizki and Zohary (1987, 1994). Important finding that made easier the crossing processes avoid pollen contamination was the availability of a male sterile gene(s). The procedures used by Basnizki and Zohary could be summarized into the following five steps: (a) selection of uniform homozygous lines by selfing commercial lines; (b) evaluation of the inbred lines' combining ability; (c) incorporation of the genetic male sterility into some of the inbred lines; (d) testing the F1 hybrids for commercial traits; and finally, (e) registering the F1 hybrid varieties.

The first globe artichoke F1 hybrids, called HU#044, HU#137 and HU#223, have been registered in Israel by Hazera Seed Co and the HU#271 by the Nunhems Zaden BV in the Netherlands. The hybrids were then evaluated in several globe artichoke producers' countries, especially France and Italy. Generally, the F1 hybrids were found to be superior, both in terms of yield and quality, with respect to traditional clones (Mauromicale et al. 1989). Globe artichoke suffers strong inbreeding depression so the inbreed lines are not highly homozygote, and this results in a high risk of uniformity lack in F1 hybrids; hence, the hybrid uniformity is one of the most important characteristics to be tested prior to the quality and morphological traits evaluation (Pagnotta et al. 2016). The hybrids have been also tested for their ability to respond to gibberellic acid treatment (Mauromicale and Ierna 2000) and to processing procedures (Calabrese et al. 1990).

Thanks to the molecular markers and the globe artichoke linkage map developed by Lanteri et al. (2006), hybrids lines have been evaluated also molecularly (Lo Bianco et al. 2011; Saccardo et al. 2013; Pagnotta and Noorani 2014; Sharaf-Eldin et al. 2015; Pagnotta 2016).

Globe artichoke hybrids and crosses were also used to evaluate, for some specific traits, the **inheritance** ration and determine the number of loci involved to coding the traits or to determine the **hereditability**. Cravero et al. (2005) studied the inheritance for *head colour* suggesting that two loci with simple recessive epistasis effects are involved and *head tightness* suggesting that two loci with simple dominant epistasis are involved in the expression of the different types. Pagnotta et al. (2016) determined the hereditability values of *earliness*, *plant height*, *spiny*, *mucron*, *head shape*, *head colour*, and *head compactness* with values ranging from 19 to 54%, and with differences between maternal and paternal hereditability.

As mentioned above, the genetic ***male sterility*** in globe artichoke made the crossing process much easier avoiding selfing. The first male sterile plants were found in a segregant population developed in France (Pécaut 1985). The genetic base was firstly attributed to a single recessive gene *ms1* (Principe 1984), and subsequently, a polygenic control of three recessive genes (*ms1*, *ms2 and ms3*) was suggested (Basnizki and Zohary 1994; Stamigna et al. 2004). Apparently, two male sterile genes were recently successfully transferred from globe artichoke to cultivated cardoon with a two loci segregation (15:1) ratio (Lopez Anido et al. 2016). Male sterile genes did not affect the floral morphology, the anther dimension or the timing of flower but only the pollen fertility, which seems to be due to a low nutritional activity of the anther tapetal cells resulting in a post-meiotic block (Lo Bianco et al. 2012). On the contrary, female reproductive organs are different; the male sterile lines have the ovaries and ovules more elongated with respect to the male fertile ones (Lo Bianco et al. 2012).

Nowadays, several globe artichoke hybrids are produced and cultivated around the world. The most widely cultivated are from Nunhems (Opal, Concerto, Madrigal, Harmony, Symphony), Agriseeds (Rinaldo, Amos, Napoleon), Semiorto/Topseeds (Romolo), among others. The **join collaboration** between the Tuscia University (Viterbo, Italy) and the Big Heart Seed Company (Brawely, CA-USA) started at the beginning of the century in order to obtain

new F1 hybrids stable and with characteristics of interest for Italian market (Saccardo 2009; Saccardo et al. 2016; Rey et al. 2016). Thanks to the environmental conditions present in the Imperial Valley (CA-USA), seed production was higher. The dry atmospheric condition without drought stress thanks to abundant water availability ensure good fertility, while the lack of wild and cultivated *Cynara* species ensure isolation by distance avoiding pollen contamination. Some F1 hybrids have now been registered and will be soon ready for commercial production (Pagnotta and Noorani 2014).

The joint programme resulted in several hundred hybrid produced every year which are generally evaluated in California and Italy environments, using proper descriptors as reported by Ciancolini et al. (2012), in order to select the most suitable for the market needs. Out of 150 cross-combinations produced in 2010, the 10 best hybrids (uniform and good shape), derived from crosses between seven male sterile clones and five fertile homogeneous parents, were evaluated molecularly and for agro-morphological traits (Rey et al. 2013a). Some of the produced hybrids are also evaluated in other countries; among other in: Chile, Argentina, Peru, Texas (Leskovar and Othman 2016b), Turkey, Egypt, Morocco, Saudi Arabia (Sharaf-Eldin et al. 2015). Currently, the joint programme produces more than 500 crosses; around 30 of them, with interesting agro-morphological traits, uniform and early producers.

Selected hybrids are under evaluation to test their reproducibility and possibility of success in the global seed market before proceeding with their registration. The different market requests in terms of globe artichoke typologies and quality characteristics, which vary between the fresh market (green, purple, spiny, ovate, round heads) and the processing industry market characterized for green ovate typologies, allow the production/registration of a wide range of globe artichoke hybrids covering the different market requests.

2.4.3 Globe Artichoke F1 Hybrids: Advantages and Disadvantages of Their Obtaining and Growing

The globe artichoke F1 hybrids have several advantages over traditional cultivars, among the others:

(a) Conversion of the globe artichoke cultivation from perennial into annual (or biennial), with an easier management of farming crop rotations and higher flexibility.
(b) Reduced cost (up to one-tenth) of the plantlet obtained by seeds compared with the micro-propagated ones.
(c) Facilities for plant production nurseries.
(d) Reduced cost of labour during annual operations and treatments.
(e) Uniformity in plants development since it is not dependent of the size of the propagation organ.
(f) High plant sanitary. Plantlets are virus-free and being annually, this would persist during all the cultivation.
(g) High resistance to pests. Plants are more vigorous and resistant.
(h) Efficient use of both water and fertilizers, thanks to a deeper roots system.
(i) High potential for organic cultivation due to vigorous, healthy and efficient use of soil fertilizer despite low chemical inputs.

However, there are some important points to take into consideration during hybrid selection and production, including:

(a) Difficult to obtain good parental lines. Correct management of the selfing avoiding high inbreeding depression.
(b) Avoid pollen contamination and right size of cages to let pollinator insects working with low mortality.

(c) Difficult to produce high seed yields and hence economical production of F1 hybrid seed.
(d) Correct management of the production cycle for seed-propagated plants.

2.4.4 Procedures to Obtain Globe Artichoke F1 Hybrids

Parents choice is done looking to specific combining ability and traits characteristics, but there are great limitation to this principle since, as mentioned elsewhere, the line used as mother should carry the male sterile gene (MS) while the line used as father (MF) should support selfing. In order to obtain a more homozygote line after a series of selfing, some outcrosses are also performed to avoid a too high inbreeding depression and sterility. When parental lines are obtained and chosen, they could be multiplied by micropropagation. Another important point to be considered for parental lines selection is the timing of flower. In case of great difference, the pollen should be collected and stored in refrigeration, but this implies limitation for an industrial seeds production.

The amount of seeds produced by each plant is an essential key factor which affects seeds cost. It is genotypically determined with a great variation among genotypes (also under open condition) (see Sect. 2.3.3). It could range from 105 to 700 seeds per head (Bianco 1990). Seeds production is also affected by climatic conditions, rain and humidity affect negatively the pollen vitality, and by the different heads order, primary heads produce more than secondary and tertiary heads.

The choice of parental lines could be helped by their morphological and molecular characterization which contribute to an accurate classification of plant genetic resources to be used in globe artichoke F1 hybrid (Crinò et al. 2008; Mondini et al. 2009; Lo Bianco et al. 2011; Crinò and Pagnotta 2017; Pagnotta et al. 2016, 2017).

To perform the crosses, there are several procedures which could be following depending also the amount of seeds production would be performed. Male sterile lines are extremely important to avoid selfing, while to avoid cross-pollination with undesirable pollens usually a net is put around the head(s) or plants are placed in cages. Procedures go from crossing of single heads which could be done manually by bring the pollen collected on the paternal flowers to the maternal one (close under net) and disperse it using a brush, or cutting the paternal head and put it directly on the maternal one. Of course, during all the steps, from pollen collection to dispersion, it should be avoided pollen contamination. In natural conditions, the main pollination system is entomophily so MS and MF plants could be closed into cages with net and in the cages could be introduced pollinator insect, generally *Bombus*, cut-bees or honey bees. The species of insect to be used is mainly determined by the cost and their availability in the area of seeds production.

Pollination efficiency could be further optimized by establishing the most suitable pollination time over the day. Significant differences in seed production were found between the morning (8–13 h) and the afternoon (14–19 h) pollination times.

References

Abatte V, Noto G (1979) Variabilità ambientale e genotipica in popolazione siciliane de *Cynara scolymus* ed isolamento di nuovi cloni di Violetto de Sicilia. In: Studi sul Carciofo. Atti III Cong. Int. Studi sul Carciofo. Laterza Bari, pp 843–852

Allard RW (1960) Principles of plant breeding. Wiley Inc, New York, USA

Amenduni M, Cirulli M, D'Amico M, Colella C (2005) Verticillium wilt of artichoke caused by *Verticillium dahliae* Kleb. Acta Hortic 681:603–606. https://doi.org/10.17660/ActaHortic.2005.681.83

Ancora G, Belli-Donini ML, Cuozzo L (1981) Globe artichoke plants obtained from shoot apices through rapid in vitro micropropagation. Sci Hort 14:207–213

Argento S, Puglia G, Pappalardo H, Pulvirenti M, Melilli MG, Raccuia SA (2016) Seed germination

responses to salt stress in wild and cultivated Sicilian cardoon genotypes. Acta Hortic 1147:9–14

Baixauli C, Giner A, Miguel A, López S, Pascual B, Maroto JV (2007) Agronomic behavior of seed propagated artichoke cultivars in the Spanish Mediterranean area. Acta Hort 730:143–147

Balloux F, Lehmann L, De Meeûs T (2003) The population genetics of clonal and partially clonal diploids. Genetics 164:1635–1644

Baranwal VK, Mikkilineni V, Zehr UB, Tyagi AK, Kapoor S (2012) Heterosis: emerging ideas about hybrid vigour. J Exp Bot 63(18):6309–6314

Basnizki Y (1985) *Cynara scolymus*. In: Halevy AA (ed) CRC handbook of flowering, vol II. CRC Press, Boca Raton, FL, USA, pp 391–399

Basnizki Y (2007) Growth and ripening of globe artichoke achenes. Ital J Agron 4:373–376

Basnizki J, Mayer M (1985) Germination in *Cynara* seeds: effects of light and temperature on the function of the endosperm. Agronomie 5:529–532

Basnizki J, Zohary D (1987) A seed-planted cultivar of globe artichoke. Hort Sci 22:678–679

Basnizki J, Zohary D (1994) Breeding of seed planted artichoke. Plant Breed Rev 12:253–269

Bianco VV (1990) Carciofo (*Cynara scolymus* L.). In: Bianco VV, Pimpini F (eds) Orticoltura. Patron Editore, Bologna (in Italian), pp 209–251

Boari F, Schiattone MI, Calabrese N, Montesano F, Cantore V (2016) Effect of nitrogen management on wild and domestic genotypes of cardoon for agro-energy purpose. Acta Hortic 1147 ISHS

Calabrese N, Bianco VV (2000) Effect of gibberellic acid on yield and quality of seed grown artichoke (*Cynara cardunculus* L. var. *scolymus* Fiori). Acta Hortic 514:25–32

Calabrese N, Carito A, Boari F, Cantore V, De Palma E, Damato G (2011) Agronomical evaluation of artichoke cultivars propagated by seed. Acta Hortic 942:153–158

Calabrese N, De Palma E, Bianco VV (2004) Yield and quality of new commercial seed grown artichoke hybrids. Acta Hortic 660:77–80

Calabrese N, De Palma E, Damato G (2007) Harvest time and yield of artichoke cultivars propagated vegetatively or by seed. Acta Hortic 730:345–350

Calabrese N, Elia A, Bianco VV (1990) Epoca di raccolta e caratteristiche qualitative di capolini di carciofo (*Cynara scolymus* L.) fresco e surgelato. Riv Agron 24:192–196

Calabrese N, Elia A, Sarli G (1994) Yield and quality of new artichoke cultivars propagated by 'seed'. Acta Hortic 371:189–193

Campanelli A, Ruta C, Tagarelli A, Morone-Fortunato I (2011) Nursery inoculation with the arbuscular mycorrhizal fungus *Glomus viscosum* and its effect on the growth and physiology of hybrid artichoke seedlings. Ital J Agric 6:159–164

Campanelli A, Ruta C, Tagarelli A, Morone-Fortunato I, De Mastro G (2014) Effectiveness of mycorrhizal fungi on globe artichoke (*Cynara cardunculus* L. var. *scolymus*) micropropagation. J Plant Interact 9 (1):100–106

Cardarelli M, Rouphael Y, Saccardo F, Colla G (2005a) An innovative vegetative propagation system for large-scale production of globe artichoke transplants. Part I. Propagation system setup. Hort Technol 15 (4):812–816

Cardarelli M, Rouphael Y, Saccardo F, Colla G (2005b) An innovative vegetative propagation system for large-scale production of globe artichoke transplants. Part II. Propagation system validation. Hort Technol 15(4):817–819

Ciancolini A (2012) Characterization and selection of globe artichoke and cardoon germplasm for biomass, food and biocompound production. These Doctorat de l'Universite de Toulouse, France, p 251

Ciancolini A, Rey NA, Pagnotta MA, Crinò P (2012) Characterization of Italian spring globe artichoke germplasm: morphological and molecular profiles. Euphytica 186(2):433–443

Cirulli M, Bubici G, Amenduni M, Armengol J, Berbegal M, Jiménez-Gasco M, Jiménez-Díaz RM (2010) *Verticillium* wilt: a threat to artichoke production. Plant Dis 94(10):1176–1187

Cointry EL, López Anido FS, García SM, Firpo IT (1999) Mejoramiento genético del alcaucil (*Cynara scolymus* L.). Avances en Hortic 4(1):51–60

Coulter FC (1941) The story of garden vegetables: artichoke. Seed world 49:10–11

Cravero VP, Picardi LA, Cointry EL (2005) An approach for understanding the heredity of two quality traits (head color and tightness) in globe artichoke (*Cynara Scolymus* L.). Genet Mol Biol 28(3):431–34

Crinò P, Tavazza R, Rey Munoz NA, Trionfetti Nisini P, Saccardo F, Ancora G, Pagnotta MA (2008) Recovery, morphological and molecular characterization of globe artichoke 'Romanesco' landraces. Genet Resour Crop Evol 55:823–833

Crinò P, Pagnotta MA (2017) Phenotyping, genotyping, and selections within Italian local landraces of romanesco globe artichoke. Diversity 9(14):1–15

Curt MD, Sanchez G, Fernandez J (2002) The potential of *Cynara cardunculus* L. for seed oil production in a perennial cultivation system. Biomass and Bioenergy 23(1), 33–46

Damato G, Calabrese N (2005a) Solid matrix priming influences germination on artichoke achenes. Acta Hortic 681:323–328

Damato G, Calabrese N (2005b) Temperature sensitive phase during germination of artichoke 'seed'. Acta Hortic 681:369–374

Damato G, Calabrese N (2007) Osmoconditioning and germination temperature in "seed" of two artichoke cultivars. Acta Hortic 730:331–336

Damato G, Cascarano MA, Casciaro L (2007a) First results of some cultural practices on yield and quality of artichoke 'seed'. Acta Hortic 730:201–209

Damato G, Sarli G, Calabrese N (2007b) Age of mother plant, head removal, GA3, yield and quality of artichoke achenes. Acta Hortic 730:337–344

Di Venere D, Pieralice M, Linsalata V, Gatto MA, Sergio L, Calabrese N (2016) Biochemical evaluation of artichoke cultivars propagated by seed. Acta Hortic 1147:89–94

Durvick DN (1984) Genetic contribution to yield grains of U.S. hybrid mayze, 1930 to 1980. In: Fehr WR (ed) Genetic contribution to yield grains of five major crop plants CSSA Special publication no 7. American Society of Agronomy, Medison, WI, USA, pp 15–47. (Crop Sci Soc Am)

Elia A, Santamaria P (1994) Influence of nitrogen, phosphorus, potassium on artichoke transplant growth. Agr Med 124:106–111

FAO (2018) http://www.fao.org/faostat/en/#data/QC; visited 30 Jan 2018

Fasahat P, Rajabi A, Rad JM, Derera J (2016) Principles and utilization of combining ability in plant breeding. Biom Biostat Int J 4(1):1–24

Foti S, Mauromicale G, Raccuia SA, Fallico B, Fanella F, Maccarone E (1999) Possible alternative utilization of *Cynara* spp. I. Biomass, grain yield and chemical composition of grain. Ind Crops Prod 10:219–228

Foury C (1967) Étude de la biologie florale de l'artichaut (*Cynara scolymus* L.). Application a la sélection. Ann Amélior Plantes 17(4):357–373

Foury C (1969) Étude de la biologie florale de l'artichaut (*Cynara scolymus* L.). Application a la sélection 2° partie. Étude des descendances obtenues en fécondation contrôlée. Ann Amélior Plantes 19 (1):23–52

Foury C (1994) Artichaut. In: Chaux et Foury, Productions légumières, 2°. Lavoisier Ed., Paris, pp 405–438

Foury C, Martin F (1976) Etude des possibilitès de creation etd'utilisation de varètès d'artichaut issues de semences. Atti 2° Congr. Int Carciofo, 1973 Bari, Ed. Minerva Medica, Torino, pp 667–679

Foury C, Martin F, Imperiali M (1978) Remarques sur la production des semences d'artichaut (*Cynara scolymus* L.). Ann Amélior Plantes 28:45–60

Gallitelli D, Rana G, Vovlas C, Martelli G (2004) Viruses of globe artichoke: an overview. J Plant Pathol 86 (4):267–281

García SM (2005) Estrategias alternativas para incrementar la producción y precocidad en el cultivo de alcaucil (*Cynara scolymus* L.). Tesis Doctorado en Ciencias Agrarias Agrarias, Universidad Nacional de Rosario, Argentina, 180 p

García SM, Firpo IT, Cointry EL, López Anido FS, Cravero VP (2005) Artichoke situation in Argentina. Acta Hortic 681:195–200

García SM, Firpo IT, López Anido FS, Cointry EL (1998) Incidencia de técnicas culturales sobre la rentabilidad del alcaucil (*Cynara scolymus* L.) en Argentina. Avances en Horticultura 3(1) (online)

Gherbin P, Monteleone M, Tarantino E (2001) Five years evaluation on cardoon (*Cynara cardunculus* L. var *altilis*) biomass production in a Mediterranean environment. Ital J Agron 5(1–2):11–19

Gutierrez Caro B, Ipinza Carmona R (1998) La Multiplicación Clonal en el Mejoramiento Genético Forestal. In: Ipinza R, Gutierrez B, Emhart V (eds) Mejora Genética Forestal Operativa, First edn, pp 201–218

Ha M, Lu J, Tian L, Ramachandran V, Kasschau KD, Chapman EJ, Carrington JC, Chen X, Wang X-J, Chen ZJ (2009) Small RNAs serve as a genetic buffer against genomic shock in Arabidopsis interspecific hybrids and allopolyploids. Proc Natl Acad Sci USA 106:17835–17840

Hammer K (1988) Preadaptations and the domestication of crops and weeds. Biologische Zentralblatt 107:631–636

Hang G (1993) Análisis económico del alcaucil. Boletín Hortícola. Facultad de Ciencias Agrarias y Forestales. UNLI. 1(1):414

Harbaoui Y, Debergh P (1980) Application du culture in vitro pour l´amèlioration des plantes potagères. Reun. Eucarpia. Section Legumes: Multiplication in vitro des clones sélectionnées d´artichaut (*Cynara scolymus* L.). Versailles 1–7

Harwood RR, Markarian D (1967) Annual culture of globe artichoke (*Cynara scolymus* L.). Preliminary report. Curtice-Burns Inc. Am Soc Hortic Sci 92:400–409

Huarte H, Benech-Arnold R (2010) Hormonal nature of seed responses to fluctuating temperatures in *Cynara cardunculus* (L.). Seed Sci Res 20(1):39–45

Ierna A, Mauro RP, Mauromicale G (2012) Biomass, grain and energy yield in *Cynara cardunculus* L. as affected by fertilization, genotype and harvest time. Biomass Bioenerg 36:404–410

Ierna A, Mauromicale G (2010) *Cynara cardunculus* L. genotypes as a crop for energy purposes in a Mediterranean environment. Biomass and Bioenergy 34(5):754–760

Ierna A, Restuccia A, Mauromicale G (2004) Effects of seed osmopriming on germination of *Cynara cardunculus* under low, optimal and high temperatures. Acta Hortic 660:333–338

Jahanian A, Chaichi MR, Rezaei K, Rezayazdi K, Khavazi K (2012) The effect of plant growth promoting rhizobacteria (PGPR) on germination and primary growth of artichoke (*Cynara scolymus*). Int J Agric Crop Sci 4(14):923–929

Jevdjovic R, Maletic R, levdjovic J (2001) Influence of vegetation space and usage of artichoke plant on yield and quality of artichoke seed (*Cynara scolymus* L.). J Agric Sci 46(2):97–104

La Malfa G, Foury C (1971) Aspects de la multiplication végétative de l´artichaut dans le bassin occidental de la méditerranée. Pép Hort Mar 114:19–29

La Malfa G, Foury C (1976) Caratteristiche morfologiche e biologiche delle piante in *Cynara scolymus* L. in rapporto agli organi di moltiplicazione. Atti 2° Congr. Int Carciofo, 1974 Bari. Ed. Minerva Medica, Torino, pp 407–418

Lanteri S, Acquadro A, Comino C, Mauro R, Mauromicale G, Portis E (2006) A first linkage map of globe artichoke (*Cynara cardunculus* var. *scolymus* L.) based on AFLP, S-SAP, M-AFLP and microsatellite markers. Theor Appl Genet 112(8):1532–1542

Ledda L, Mameli MG, Milia M, Marras GF (2004) Influence of plant shading and ovoly typology on globe artichoke development, early production and head atrophy: preliminary results. Acta Hortic 660:365–371

Lekić S, Stojadinović J, Todorović G, Jevdjović R, Draganić I, Djukanović L (2011) Effects of substrates and temperatures on *Cynara cardunculus* L. seed germination. Romanian Agric Res 28:223–227

Leskovar DI, Othman Y (2016a) Low nitrogen fertigation promotes root development and transplant quality in globe artichoke. Hort Sci 51(5):567–572

Leskovar DI, Othman Y (2016b) Morpho-physiological characteristics and yield of early and mid-season globe artichoke. Acta Hortic 1147:155–158

Lo Bianco C, Fernández JA, Migliaro D, Crinò P, Egea-Gilabert C (2011) Identification of F1 hybrids of artichoke by ISSR markers and morphological analysis. Mol Breed 27(2):157–170

Lo Bianco C, Saccardo F, Olimpieri I, Mazzucato A, Crinò P (2012) Floral biology in male sterile clones of globe artichoke [*Cynara cardunculus* subsp. *scolymus* (L.) Hegi]. Acta Hortic 942:159–164

Lombardo S, Pandino G, Mauromicale G, Carle R, Knödler M, Schieber A (2012) New seed propagated cultivars of globe artichoke suitable for processing uses. Acta Hortic 942:139–146

López Anido FS, Martin EA, García SM, Cointry EL (2016) Successful Transferring of male sterility from globe artichoke into cultivated cardoon. Acta Hortic 1147:163–166

Macua JI, Lahoz I (2016a) Seed artichoke in Navarre. Acta Hortic 1147:177–182

Macua JI, Lahoz I (2016b) Annual and multiannual production of different artichoke cultivars in northern Spain. Acta Hortic 1147:171–176

Macua JI, Lahoz I, Santos A (2005) Agronomic and qualitative influence of different artichoke varieties in stumps fields with one-year and multi-year crops. Acta Hortic 681:151–158

Mass EV, Hoffman GJ (1977) Crop salt tolerance-current assessment. J Irrig Drain Div ASCE 103:115–134

Mauromicale G (1987) Panorama varietale del carciofo e sua prevedibile evoluzione. L'inf Agr 43(4):69–76

Mauromicale G, Basnizki J, Cavallaro V (1989) Primi risultati sperimentali sulla propagazione del carciofo (*Cynara scolymus* L.) per seme. Riv Agron 23 (6):417–423

Mauromicale G, Ierna A (2000) Characteristics of heads of seed-grown globe artichoke [*Cynara cardunculus* L. var. *scolymus* (L.) Fiori] as affected by harvest period, sowing date and gibberellic acid. Agronomie 20(2):197–204

Mauromicale G, Licandro P (2002) Salinity and temperature effects on germination, emergence and seedling growth of globe artichoke. Agronomie 22:443–450

Mauromicale G, Licandro P, Ierna A, Morello N, Santoiemma G (2004) Planning of globe artichoke plantlets production in nursery. Acta Hortic 660:279–284

Miguel A, Baixauli C, Aguilar JM, Giner A, Maroto JV, Lòpez S, Pascual B (2004) Cultivar trials of seed propagated artichoke. Acta Hortic 660:111–116

Mondini L, Noorani A, Pagnotta MA (2009) Assessing plant genetic diversity by molecular tools. Diversity 1:19–35

Ni Z, Kim ED, Ha M, Lackey E, Liu J, Zhang Y, Sun Q, Chen ZJ (2009) Altered circadian rhythms regulate growth vigour in hybrids and allopolyploids. Nature 457(7227):327–331

Ortega RG (2002) Effect of head position and climatic conditions on seed yield of globe artichoke (*Cynara cardunculus* L. var. *scolymus* L. Fiori), cv. Imperial Star. Sci Hortic 93(2):187–192

Ouselati T, Ghezal I (2009) Il carciofo in Tunisia. In: Calabrese N (ed) Il carciofo e il cardo. Script Bologna, pp 386–391. ISBN 978-96301-05-0

Pagnotta MA (2016) Artichoke, an uncivilised crop rich of diversity to preserve and valorize. Acta Hortic 1147:243–256

Pagnotta MA, Fernández JA, Sonnante G, Egea-Gilabert C (2017) Genetic diversity and accession structure in European *Cynara cardunculus* collections. PloS ONE 12(6)

Pagnotta MA, Noorani A (2014) Genetic diversity assessment in European Cynara collections. In: Genomics of Plant Genetic Resources. Springer, The Netherland, pp 559–584

Pagnotta MA, Rey NA, Mondini L, Aringoli R, Jordan R, Saccardo F (2016) Assessment of artichoke hybrids under USA and Italian conditions and the hereditability of some important traits. Acta Hortic 1147:257–264

Pappalardo H, Leonardi C, Genovese C, Toscano V, Melilli MG, Raccuia SA (2016) Effects of heavy metals on seedlings germination and growth in different cardoon genotypes. Acta Hortic 1147:281–288

Pécaut P (1985) Recherches sur l'artichaut. Rapport d'activité 1983–1984 de la Station d'amélioration des plantes maraîchères d'Avignon, INRA, France, pp 21–24

Pécaut P (1993) Globe artichoke *Cynara scolymus* L. In: Kalloo G, Bergh BO (eds) Genetic improvement of vegetable crops. Pergamon Press, pp 737–746

Pécaut P, Dumas de Vaulx R, Lot H (1983) Virus-free clones of globe artichoke (*Cynara scolymus*) obtained after in vitro propagation. Acta Hortic 131:303–310

Pécaut P, Foury C (1992) L'artichaut. In: Gallais A, Bannerot H (eds) Amélioration des espèces végétales cultivées. INRA, Paris, France, pp 460–469

Pécaut P, Martin F (1992) Non-conformity of in vitro propagated plants of early Mediterranean varieties of globe artichoke (*Cynara scolymus* L.). Acta Hortic 300:363–366

Piscioneri I, Sharma N, Baviello G, Orlandini S (2000) Promising industrial energy crop, *Cynara cardunculus*: a potential source for biomass production and alternative energy. Energy Convers Manage 41 (10):1091–1105

Pochard F (1964) Modifications de la croissance et du développement de l'artichaut provoquées par la gibberelline. Ann Amelior Plantes 14(2):219–225

Pochard E, Foury C, Chambonner D (1969) Il miglioramento genetico del carciofo. Atti 1° Cong. Int di Studi sul Carciofo, Bari, pp 117–143

Porceddu E, Dellacecca V, Bianco VV (1976). Classificazione numerica di cultivar di carciofo. Atti II Congresso Internazionale Carciofo, Bari. Ed. Minerva Media, Torino, pp 1105–1119

Portis E, Barchi L, Acquadro A, Macua JI, Lanteri S (2005) Genetic diversity assessment in cultivated cardoon by AFLP (amplified fragment length polymorphism) and microsatellite markers. Plant Breed 124:299–304

Principe JA (1984) Male-sterility in artichoke. Hort Sci 19 (6):864–865

Raccuia SA, Cavallaro V, Melilli MG (2004) Intraspecific variability in *Cynara cardunculus* L. var. *sylvestris* Lam. Sicilian populations: seed germination under salt and moisture stresses. J Arid Environ 56(1):107–116

Raccuia SA, Gallo G, Melilli MG (2012) Effect of plant density on biomass and grain yields in *Cynara cardunculus* var. *altilis* cultivated in Sicily. Acta Hortic 942:303–308

Raccuia SA, Melilli MG (2007). Biomass and grain oil yields in *Cynara cardunculus* L. genotypes grown in a Mediterranean environment. Field Crops Res 101 (2):187–197

Rana GL, Elia A, Nuzzaci M, Lafortezza R (1992) Effect of latent artichoke virus on the production of artichoke heads. J Phytopathol 135:153–159

Rey NA, Jordan R, Jordan A, Saccardo F, Pagnotta MA (2013a) Hybrids agromorphological assessment. Acta Hortic 983:237–242

Rey NA, Jordan R, Saccardo F, Pagnotta MA (2016) A successful strategy to obtain artichoke hybrids. Acta Hortic 1147:357–368

Rey NA, Papacchioli V, Tavazza R, Pagnotta MA (2013b) Gauging the genetic changes occurring across globe artichoke micropropagation towards an appropriate variety registration and nursery production. Scientia Hortic 156:121–126

Rossi V, De Paoli G (1990) Micropropagazione per un carciofo di qualità. L'informatore agrario 9:207–209

Runge MS, Patterson C (2006) Principles of molecular medicine. Humana Press, 58 p. ISBN 978-1-58829-202-5

Ruta C, Tagarelli A, Morone Fortunato I (2005) Mycorrizathion on micropropagated artichoke. Acta Hortic 681:407–412

Saccardo F (2009) Miglioramento genetico in Il carciofo e il cardo. In: Calabrese N (ed) Script Bologna, pp 286–297. ISBN 978-96301-05-0

Saccardo F, Jordan JR, Jordan A, Crinò P, Micozzi F, Lo Bianco C, Temperini A, Rey NA, Pagnotta MA (2013) Innovative strategy to obtain F1 hybrids of globe artichoke. Acta Hortic 983:159–172

Saccardo F, Jordan JR, Mancuso G, Crinò P, Pagnotta MA (2016) 'Romolo' and 'Istar', First artichoke seed propagated hybrids enrolled into Italian national variety register. Acta Hortic 1147:381–384

Santamaria P, Ayala O, Buttaro D, Parente A, Kirkayak B (2007) N-rate and N-source for producing artichoke cv. 'Concerto transplants' in float bed under greenhouse conditions. Acta Hortic 730:315–319

Schrader WL (1990) Artichoke seed germination responses to chilling and growth regulator treatments. Hort Sci 25(9):1084

Schrader WL (1992) Growth regulator effects or earliness and yield in artichokes grown as annuals from seed. Hort Sci 27:643

Sharaf-Eldin MA, Al-Tamimi A, Alam P, Elkholy SF, Jordan JR (2015) Genetic relatedness of artichoke (*Cynara scolymus* L.) hybrids using random amplified polymorphic DNA (RAPD) fingerprinting. Genet Mol Res 14:18431–18439

Shinohara T, Leskovar DI (2014) Effects of ABA, antitranspirants, heat and drought stress on plant growth, physiology and water status of artichoke transplants. Sci Hort 165:225–234

Shinohara T, Martin EA, Leskovar DI (2017) Ethylene regulators influence germination and root growth of globe artichoke seedlings exposed to heat stress conditions. Seed Science and Technology 45(1):167–178

Sprague GF, Tatum LA (1942) General versus specific combining ability in single crosses of corn. J Am Soc Agr 34:923–932

Stamigna C, Micozzi F, Pandozy G, Crinò P, Saccardo F (2004) Artichoke hybrids production by use of male sterile clones [*Cynara scolymus* L.]. *Italus Hortus* (Italy)

Tavazza R, Rey NA, Papacchioli V, Pagnotta MA (2015) A validated slow-growth in vitro conservation protocol for globe artichoke germplasm: a cost-effective tool to preserve from wild to elite genotypes. Sci Hortic 197:135–143

Temperini O, Calabrese N, Temperini A, Rouphael Y, Tesi R, Lenzi A, Carito A, Colla G (2013) Grafting

artichoke onto cardoon rootstocks: graft compatibility, yield and *Verticillium* wilt incidence. Sci Hort 149: 22–27

Tesi R, Lombardi P, Lenzi A (2003) Nursery production of rooted offshoots of globe artichoke (*Cynara scolymus* L.). Acta Hortic 660:399–404

Trigo-Colina L (1980) Estudio del comportamiento clonal de la población de alcachofa Blanca de España cultivada en el valle del Ebro. Inia 13(3):49–57

Vannella S, Damato G, Calabrese N (2005) Influence of temperature and substrate on the germination of artichoke achenes. Acta Hort 681:361–367

Vasilakoglou I, Dhimab K (2014) Potential of two cardoon varieties to produce biomass and oil under reduced irrigation and weed control inputs. Biomass Bioenerg 63:177–186

Vilchez M, Paulus AO, Mayberry KS (2005) Globe artichoke seed treatment to control seedborne fungi and seed quality evaluation. Acta Hortic 681:581–586

Welbaum GE (1994) Annual culture of globe artichoke from seed in Virginia. Hort Technol 4(2):147–150

Zhang X, Lv L, Lv C, Guo B, Xu R (2015) Combining ability of different agronomic traits and yield components in hybrid barley. PLoS ONE 10(6):e0126828

Ziani K, Ursúa B, Maté JI (2010) Application of bioactive coatings based on chitosan for artichoke seed protection. Crop Prot 29:853–859

Globe Artichoke Tissue Culture and Its Biotechnological Application

3

Cinzia Comino, Andrea Moglia, Annamaria Repetto and Raffaela Tavazza

Abstract

Plant tissue culture techniques have made significant contributions to the advance of agricultural sciences in recent times, and today, they constitute an indispensable tool in modern agriculture. In particular, micropropagation is one of the most widely used techniques for rapid vegetative in vitro propagation. This chapter aims to review advances and current knowledge in key areas of plant tissue culture and biotechnology as applied to globe artichoke.

3.1 Introduction

Plant tissue culture is based on the principle of plant cellular totipotency that is the inherent potentiality of a plant cell to regenerate a whole plant. This capacity is retained even by a differentiated cell. Single cells, protoplasts, pieces of leaves, stems or roots can thus be used to generate a new plant on a nutrient culture *media* of known composition. Several factors considerably influence the phenomenon of morphogenesis during in vitro culture. These include genotype, explant type, growth regulators, nutrients, other additives and physical factors. The establishment of efficient in vitro tissue culture protocols is an essential prerequisite to fully exploit the advantage of the plant cell totipotency for genetic improvement. Globe artichoke conventional breeding and propagation methods, although well developed, could not meet up with the ever-increasing industry demands. Biotechnology techniques involving plant tissue culture and recombinant DNA technologies are powerful tools that can complement conventional breeding and expedite globe artichoke improvement. Unfortunately, the rate of progress in these biotechnology techniques in globe artichoke is relatively slower than in other members of *Asteraceae*, due to its high genotypic dependence and recalcitrant nature. Nevertheless, considerable success has been achieved in the area of micropropagation, virus eradication, clonal fidelity assessment, germplasm storage, protoplast culture, production of secondary metabolite, genetic transformation and cryopreservation.

In this chapter, we present the current state of globe artichoke cell culture regarding its application, problems, and potential.

C. Comino · A. Moglia
DISAFA Plant Genetics and Breeding, University of Torino, 10095 Grugliasco, Torino, Italy

A. Repetto
Agency for Agricultural Research (Agris), Viale Trieste 111, 09123 Cagliari, Italy

R. Tavazza (✉)
Italian National Agency for New Technologies, Energy and Sustainable Development (ENEA), Casaccia Research Center, Via Anguillarese 301, 00123 Rome, Italy
e-mail: raffaela.tavazza@enea.it

E. Portis et al. (eds.), *The Globe Artichoke Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-030-20012-1_3

3.2 In Vitro Propagation

3.2.1 Overview

Micropropagation is the tissue culture technique used for the rapid vegetative in vitro propagation of ornamental, crop plants, and fruit trees. This method produces multiple numbers of vegetative shoots from a single clone within a short time and space. Commonly, there are four necessary phases for obtaining a successful micropropagation (i) establishment of aseptic in vitro culture, (ii) shoots' multiplication, (iii) rooting, and (iv) acclimatization. Since the traditional propagation of globe artichoke through "ovoli," offshoots or stumps produce a low rate of multiplication, heterogeneity of plants, and transmission of diseases (Pecaut et al. 1983), in vitro culture remains the only way out for speed mass propagation of globe artichoke. Besides, micropropagation has many advantages over conventional vegetative propagation: production of disease-free material, independence from seasonal cycles, high rate of plant propagation, homogeneity and vigor of the obtained plant material, conservation of germplasm. Conversely, the disadvantages include cost factors, possible occurrence of genetic instability and possibility of contamination of cultures.

De Leo and Greco (1973) were the first to apply this technique to globe artichoke, establishing suitable culture conditions for rapid propagation of élite materials. Since then, several reports have been published describing the use of different plant tissue and in particular of shoot apices to produce homogeneous and disease-free globe artichoke plants (Murashige 1974; Ancora 1986; Rossi and De Paoli 1990; Frau et al. 2004). The propagation methodology was afterward adjusted to improve efficiency using new culture *media* (Tavazza et al. 2004; Elia et al. 2006), growth regulators (Tavazza et al. 2004) and mycorrhization (Ruta et al. 2005), thus facilitating the large-scale production of numerous cultivars and their clones.

Micropropagation proved to be a valid technique particularly for the propagation of spring cultivars (Saccardo et al. 2007), and today, it allows the multiplication and distribution to the farmers of clones of different cultivars (Ancora 1986; Ancora and Saccardo 1987). Clones obtained in this way have shown improved field performance for both qualitative and quantitative traits. Besides, the use of the micropropagated materials has led to change the globe artichoke cultivation from poliannual to annual (Saccardo and Ancora 1984). The application of the micropropagation technique in autumn cultivars, such as the 'Spinoso sardo,' is hindered by the loss of earliness of micropropagated plants occurring at an increasing rate with the subculture number (Cadinu et al. 2004, 2006; Tavazza et al. 2004; Elia et al. 2006). These plants also show a specific "*wild*" phenotype because of the presence of highly pinnate-parted leaves and the late budding of the inflorescence, in contrast to the normal phenotype (Fig. 3.1). The causes that determine the appearance of the *off-type* plants in autumn cultivars are still under study. Phenotypic, cytological, biochemical deviations expressed in the progeny of plants regenerated from in vitro culture are known as 'somaclonal variation' (Larkin and Scowcroft 1981). Genetic (e.g., single-base-pair changes, chromosome modifications, and changes in ploidy) as well as epigenetic alterations (e.g., DNA methylation) are often

Fig. 3.1 'Spinoso sardo' wild and normal phenotype: **a** in vitro; **b** in the field plants

associated with in vitro propagation. DNA methylation analysis of plants showing these phenotypes performed using a restriction-enzyme-based method has been recently reported (Comino et al. 2017).

3.2.2 Sterilization of the Explants

Surface sterilization of explants is a tricky step since it has to remove bacterial and fungal contaminants with minimal damage to plant cells (Husain and Anis 2009). The most widely used sterilization procedures for globe artichoke are given in Table 3.1. Before sterilization, vegetative apices of shoots from field plants are washed in running water and, after elimination of outer leaves, immersed in the antioxidant solution. This latter step is essential, since sterilization procedures may elicit a wound response in addition to that caused by excision, and a consequent production of phenolic compounds that can be deleterious for the in vitro growth leading to tissue browning.

3.2.3 Culture *Media*

Composition of plant tissue culture *medium* should supply all cell growth needs and can vary depending on the type of plant tissues or cell used for culture. A typical nutrient *medium* contains inorganic salts (micro- and macro-elements), vitamins, organic supplements, growth regulators, carbon source and gelling agents in the case of a solid *medium*. The Murashige and Skoog (MS) *medium* (1962) is the most widely used for in vitro propagation of many plant species. Some *media* have been devised for specific tissues and organs.

Plant growth regulators play an essential role in determining the development route of cells and plant tissues in vitro. The type and concentration of hormones depend mainly on the plant species, the tissue or the organ chosen and on the goal to be achieved. Auxins, cytokinins, and gibberellins are the most commonly used growth regulators. The use of *media* with a low hormone concentration is usually required during the globe artichoke culture establishment (Tavazza et al. 2009). Moreover, a high

Table 3.1 Surface sterilization procedures used for globe artichoke explants

Procedure	References
Immersion in ethanol 70% 5 s, sodium hypochlorite 20% (w/v) 20 min, three rinses in sterile water containing ascorbic acid 0.01% (w/v)	Ancora (1986)
Immersion in mercuric chloride 0.5% (w/v) 3 min, sodium hypochlorite 20% (w/v) 20 min, three rinses in distilled water containing citric acid 0.015% (w/v) and ascorbic acid 0.01% (w/v)	Ordas et al. (1990)
Rinse under running water 15 min, immersion in ethanol 70% 30 s, sodium hypochlorite 20% (w/v) 30 min, and three rinses in distilled water containing citric and ascorbic acid	Iapichino (1996)
Immersion in sodium hypochlorite 10% (w/v) 20 min, ethanol 70% 2 min, three rinses in distilled water	Repetto et al. (1996)
Immersion in mercuric chloride 0.5% (w/v) 5 min, sodium hypochlorite 17% 10 min, three rinses with distilled water	Brutti et al. (2000)
Rinse under running water, immersion in sodium hypochlorite 2.5% (w/v) containing 2 ml of "Tween 20" 20 min, three rinses in distilled water containing citric acid and ascorbic acid 0.15% (w/v) 30 min	Elia et al. (2006)
Rinse under running water 30 min, immersion in ethanol 70% 5 s, Sodium hypochlorite 2% (w/v) with two drops of "Tween 20" for 20 min	El-Zeiny et al. (2013)
Immersion in mercuric chloride 0.5% (w/v) 3 min, sodium hypochlorite 0.5% (v/v) with 3–4 drops of "Tween 80" 20 min, three rinses with distilled water	Tavazza et al. (2013a, b)
Washing with a solution of citric acid and ascorbic acid 2% (w/v) 15 min, immersion in sodium hypochlorite 1.5% (w/v) 15 min, three rinses with a solution of citric acid and ascorbic acid 2%(w/v)	Barba and Taglienti (2015)

concentration of cytokinins promotes the shoots regeneration while a high concentration of auxins favors rhizogenesis (Tavazza et al. 2004; Morone Fortunato et al. 2005). Cyclodextrins are also used for rhizogenesis (Brutti et al. 2000; Cavallaro et al. 2004). A balance between auxin and cytokinin leads to the development of a mass of undifferentiated cells called *callus* (Rout 2004).

The pH of the *media* is critical as it can influence both the growth of plants and the activity of growth regulators. It must be set between 5.4 and 5.8.

3.2.4 Meristem Culture

Culture is usually initiated by dissecting away the shoot tip until obtaining only the apical dome, together with one or two leaves primordial (Fig. 3.2a); the exposed meristem is then excised and placed onto a suitable culture *medium*. The size of the excised explants is critical for virus elimination: The percentage of virus-free plants obtained is inversely proportional to the size of the meristem explant used.

A meristem or a shoot apex can form a culture according to many factors: (i) age and physiological status of the mother plant; (ii) position of the explant within the plant; (iii) size of the explant used to initiate the culture. Furthermore, in vitro behavior is a characteristic linked to the genotype, so that often cultivars of the same species have a different attitude in in vitro environments (Hartmann and Kester 1983). If suitable culture conditions are adopted, elongated shoots develop from the meristem (Fig. 3.2b).

The plants are grown in climatic cells with controlled environmental parameters: temperature of 22 °C, photoperiod of 16 h of light and 8 of darkness and light intensity of 4000 lx. A summary of *media* used thus far for the establishment of globe artichoke in vitro cultures is reported in Table 3.2.

3.2.5 Shoot Multiplication

Once an actively growing culture is established, plantlets are transferred to a proliferation *medium*. During the proliferation phase, through appropriate hormonal concentrations, the development of axillary buds from the original shoot is promoted (Fig. 3.3). Axillary-derived shoots can then be subcultured onto the same *medium* to promote their elongation and the proliferation of new buds. This process may be repeated several times until a large number of plants are obtained or can go on indefinitely. The prolongation of the number of subcultures may, however, lead to the appearance of somaclonal variations (Cadinu et al. 2006).

In plant tissue culture, the rate of shoot multiplication can be determined by enhanced axillary branching. Multiplication rates depend on cultivar, clone, and explant source. In globe artichoke, several *media* have been used to stimulate shoot multiplication, usually supplemented with a low auxin/cytokinin ratio (Table 3.3).

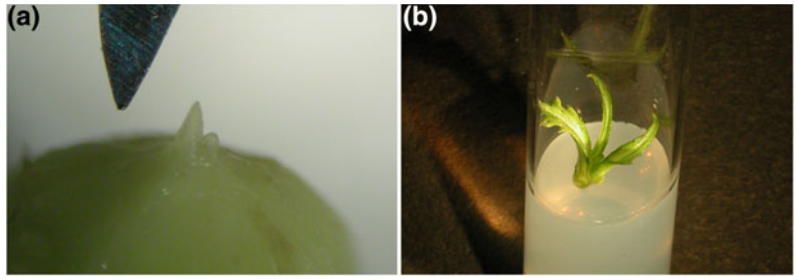


Fig. 3.2 Establishment of globe artichoke aseptic culture: **a** excision of a meristematic apex from offshoots; **b** meristematic apex, after 1 month of culture

Table 3.2 *Media* used for the establishment of globe artichoke in vitro cultures

Medium	References
MS + NaH_2PO_4 50 mg L^{-1} + Myo-inositol 100 mg L^{-1} + Tyrosine 100 mg L^{-1} + Adenine 40 mg L^{-1} + IAA 0.5 mg L^{-1} + Kin 10 mg L^{-1}	Ancora (1986)
Macro- and micro-DKW + MS vitamins + BAP 0.1 mg L^{-1} + IBA 0.01 mg L^{-1}	De Paoli and Rossi (1987)
MS + IBA 0.1 mg L^{-1} + BAP 0.5 mg L^{-1} + 2iP 0.01 mg L^{-1}	Repetto et al. (1996)
MS macronutrient with 50% reduction of NH_4NO_3 and KNO_3 + Heller micronutrients + B5 vitamins + Myo-inositol 100 mg L^{-1} + NAA 0.1 mg L^{-1}	Brutti et al. (2000)
TL macro- and microelements + B5 vitamins + 25 mg L^{-1} + Myo-inositol 100 mg L^{-1} + Thiamine HCl 0.4 mg L^{-1} + 2iP 1 mg L^{-1} + IAA 1 mg L^{-1} + GA_3 0.025 mg L^{-1}	Castiglione et al. (2006)
MS + B5 vitamins + IBA 1 mg L^{-1} + GA_3 0.1 mg L^{-1}	EI-Boullani et al. (2012)
MS with a quarterly nitrate concentration + BA 0.03 mg L^{-1} + GA_3 0.05 mg L^{-1}	Bedini et al. (2012)
MS + Kin 5.0 mg L^{-1} + IAA 0.5 mg L^{-1}	El-Zeiny et al. (2013)
Gik (basal medium described by Tavazza et al. 2004) + Kin 2 mg L^{-1} + IBA 0.1 mg L^{-1}	Rey et al. (2013a, b)
Base substrate (Morone Fortunato et al. 2005) + 2iP 1 mg L^{-1} + IAA 1 mg L^{-1} + GA_3 0.025 mg L^{-1}	Ruta et al. (2013)

Fig. 3.3 Axillary shoot formation

Table 3.3 *Media* used for the proliferation of globe artichoke

Media	References
MS + Kin 5 mg L^{-1} + IAA 0.5 mg L^{-1}	Okasha and Ragab (1994)
MS + IBA 0.1 mg L^{-1} + Kin 0.5 mg L^{-1} + 2iP 0.1 mg L^{-1}	Repetto et al. (1996)
BM + Kin 2 mg L^{-1} + 2iP 1 mg L^{-1} + NAA 10 mg L^{-1}	Brutti et al. (2000)
MS + Kin 2 mg L^{-1} + IBA 0.1 mg L^{-1}	Tavazza et al. (2004)
MS + Kin 0.1 mg L^{-1} + GA_3 0.05 mg L^{-1}	Babes et al. (2004)
MS with a quarter nitrate concentration + BA 0.03 mg L^{-1} + GA_3 0.05 mg L^{-1}	Bedini et al. (2012)
Substrate (Morone Fortunato et al. 2005) + BAP 0.05 mg L^{-1}	Ruta et al. (2013)

3.2.6 Root Induction

The rooting phase is the final step of micropropagation with the obtaining of a seedling-like shoots morphologically complete (Fig. 3.4). Recalcitrance to form adventitious roots can be a major obstacle in vegetative propagation. In fact, the formation of a functional root system is fundamental for the acquisition of a plantlet's ability to manage water and mineral nutrients under the post in vitro conditions.

Rooting of shoots has been a major problem in globe artichoke propagation. According to some studies, in globe artichoke rhizogenesis is stimulated in vitro by the auxin (Harbaoui and Debergh 1980; Draoui et al. 1993; Tavazza et al. 2004; Morone Fortunato et al. 2005) or by gibberellic acid (Morzadec and Hourmant 1997) and facilitated by complementary conditions such as (i) percentage of sucrose (Morone Fortunato and Ruta 2003; Tavazza et al. 2004), (ii) addition of activated carbon (Bigot and Foury 1984; El-Zeiny et al. 2013) or of cyclodextrine (Brutti et al. 2000; Cavallaro et al. 2004), (iii) type of the inert component, and (iv) size of the seedling (Morone Fortunato and Ruta 2003; Morone Fortunato et al. 2005). Besides, it has been suggested that the phenotypic features and the quality of in vitro shoots could play a key role in the globe artichoke-rooting phase (Ancora et al. 1981; Morzadec and Hourmant 1997).

Rooting of about 60% of the proliferated shoots (cv. 'Romanesco') has been obtained on MS (half strength) supplemented with NAA (Ancora et al. 1981). The use of NAA in place of IBA or IAA gives rise to short, thick roots, which are not easily damaged upon transfer to soil. An increase in the percentage of rooted plants (cv.'Brindisino') has also been observed by Rossi and De Paoli (1990) exposing the shoots to a liquid medium containing NAA and IAA at high concentrations before transfer into rooting medium. IBA is sometimes preferred for adventitious root induction in globe artichoke since it is more stable than IAA during the sterilization process of the medium (Hartmann and Kester 1983) and encouraged the root formation (number and length) more than in IAA (El-Zeiny et al. 2013); in time IAA is also broken down by light exposure.

The auxin liability often limits expected effect on in vitro culture of plant tissue. To overcome this problem, additives like cyclodextrins can be used. These additives are cage molecules that have the property of including organic molecules such as auxins and protecting them from photodegradation or oxidation, thus improving the rhizogenesis of globe artichoke. A double rooting phase, i.e., a preculture medium (MS medium with half nitrate concentrations, IAA, and paclobutrazol (PBZ) followed by a rooting medium (MS medium, cyclodextrin, and NAA), was used by Okasha and Ragab (1994) to facilitate rooting of globe artichoke plantlets. By using this last rooting medium, Repetto obtained in 'Spinoso sardo' 80% of rooting plants (unpublished data). The enhancement of in vitro rooting in cv. Early French was observed in a medium supplemented with NAA and cyclodextrins (Brutti et al. 2000). An increase in rooting percentage (100 and 62%) following treatment with cyclodextrins was also achieved by Dridi (2003) and Cavallaro et al. (2004), respectively.

Fig. 3.4 In vitro rooted plants ready for acclimatization

Light generally has a negative effect on root formation. High irradiances inhibit root formation more than low irradiation. By using a 5-day culture in darkness, in the presence of IBA, López-Pérez and Martínez (2015) obtained a high percentage of rooting from globe artichoke plants of cvs. 'Blanca de Tudela' (73.3%), 'Calico' (80%), and 'Salambo' (93.3%). Decreasing light intensity from 40 to 20 μE m^{-2} $s^{-1}$and reducing explant density from 6–7 to 3–4 shoots per 132 cm^2 area of culture medium EI-Boullani et al. (2012) succeed to increases the multiplication rate of a Moroccan artichoke accession by twofold.

3.2.7 Plant Establishment

Once a satisfactory root system has formed (Fig. 3.4), plantlets are transferred to sterile compost and acclimatized under greenhouse condition. The acclimatization phase of micropropagated plantlets is one of the most delicate steps of micropropagation and can create high levels of stress since the leaves of the in vitro plantlets, photo-synthetically not very active, are not well adapted to an in vivo climate. Acclimatization can take place by allowing the in vitro plants to get used to lower relative humidity gradually. After 7–10 days, hardened plants are transferred individually in pots before moving to field condition. The in vitro plants, once transplanted in open field, showed a vegetative vigor and productivity far superior respect to the plants raised traditionally. "Ovoli" of 'Spinoso sardo' obtained from micropropagated mother plants, developed plants able to provide commercial productions 31% greater respect to those obtained from non-micropropagated ones and characterized by a greater precocity and a contemporaneity in the emission of the head. In the first four weeks of harvesting, 9% of the micropropagated plants were in production, versus only 2.3% of non-micropropagated ones. The quality of the product also showed differences such as a longer stem length in the first-order flower heads, which is a valuable factor for raw consumption, and a greater caliber and weight of the third-order heads. Even though micropropagation ensures the production of the high-quality commercial plant, acclimatization of in vitro globe artichoke plantlets still represents a critical phase of micropropagation process due to the remarkable loss of seedlings during transfer from in vitro to ex vitro environment. To increase the survival capacity of the plants and overcome transplant stress, the technique of mycorrhizal inoculation seems to be the most promising. There are progressively more reports that demonstrated the effectiveness of mycorrhization on micropropagated plantlets during the delicate acclimatization phase and its positive impact on their post-transplanting performance (Morone Fortunato et al. 2005; Ruta et al. 2005).

3.3 Production of Virus-Free Plants

3.3.1 Overview

The shoot tip culture is the most used method to produce virus-free plants. It is based on the assumption that pathogen concentration is not uniform throughout the infected plant. Cells non-invaded by pathogens are often found in the apical meristem of a rapidly elongating stem. In addition, in most cases, this explant maintains the genetic stability of the genotype. Shoot meristem tip culture alone or in combination with thermotherapy, or cryotherapy is being widely used to arise virus-free plants from infected stock plants.

3.3.2 Virus Eradication in Globe Artichoke

To date, twenty-four plant viruses, belonging to 13 genera, are known to infect globe artichoke (Gallitelli et al. 2004), most of which were recorded from Europe and the Mediterranean basin. In the last fifty years, fifteen viruses belonging to Nepovirus, Potyvirus, Tospovirus, Tombusvirus, Fabavirus, and Cucumovirus were reported, identified and characterized in Italy (Barba et al. 2013). The current health status of

globe artichoke germplasm is severely compromised by numerous viral infections (Foddai et al. 1983; Pasquini et al. 2003; Rana and Martelli 1983) in particular by the artichoke latent virus (ArLV), the artichoke Italian latent virus (AILV), the artichoke mottled crinkle virus (AMCV) and, in recent years, the tomato spotted wilt virus (TSWV).

Viral diseases significantly reduce the yield and quality of this crop affecting globe artichoke industry. The nursery activity demands commercialization of pathogen-free propagation materials as required by EU legislation (Gallitelli and Barba 2003). It is, therefore, imperative the production and maintenance of virus-free planting material.

Two approaches, often used in combination, namely in vitro thermotherapy and meristem tip culture, have been widely used to produce vegetatively propagated virus-free plants. The application of these techniques in early cultivars of globe artichoke is underutilized due to the loss of earliness. Nevertheless, in vitro/vivo thermotherapy followed by meristem tip culture proved to be a suitable technique to obtain ArLV and AILV-free plants of the early flowering cultivars cv. 'Brindisino' without the risk of running into this problem (Gallitelli et al. 2006; Acquadro et al. 2010). In the 'Spinoso sardo', out of the 1512 plants arising from meristem tip culture, 52% have been shown to be free of the most pathogenic viruses, i.e., ALV, AILV, and AMCV. For the late globe artichoke ecotypes, meristem tip culture allowed to obtain ArLV, AILV, TSWV-free mother plants belonging to the most important cultivars grown in Latium ('C3', 'Grato 1', 'Grato 2', 'Castellammare') and in Campania regions ('Tondo di Paestum', 'Bianco di Pertosa') (Barba et al. 2004; Morone Fortunato et al. 2007). In addition to meristem tip cultures and thermotherapy techniques for virus elimination, a new method 'the cryotherapy' has come to attention (see Sect. 3.5.4). In cryotherapy, plant pathogens such as viruses, phytoplasmas, and bacteria are eradicated from shoot tips by exposing them briefly to liquid nitrogen; the lethal injury of the cells during cryo-treatment is utilized to kill infected tissue. Cryotherapy has proven to be a novel and efficient biotechnological tool for the eradication of some globe artichoke viruses (Tavazza et al. 2013a; Taglienti et al. 2013). An advantage of this technique is that pathogen eradication is independent of the size of shoot tips used and applies to a virus infecting meristem, where meristem tip culture fails. Moreover, cryotherapy allows retrieving virus-free plants in a shorter time compared to meristem tip cultures and thermotherapy in both globe artichoke ecotypes (Taglienti et al. 2013). However, so far, this technique is available only for a limited number of globe artichoke cultivars, and further studies are needed to expand cryotherapy to additional genotype and to increase the proportion of pathogen-free regenerants.

3.4 Plant Tissue Culture for Biotechnology

3.4.1 Overview

The application of plant tissue culture technologies for agricultural biotechnology is vast. Generally, biotechnology utilizes living systems for creating or developing valuable products. Plant tissue culture is the most prominent technique of plant biotechnology, with diverse applications in several fields. Plant tissue cultures are used as an important source of primary and secondary metabolites, and in recent past, plant genetic transformation has gained importance for increasing quality and yield of plants, and in manipulating them for improved agronomic performance.

3.4.2 *Callus* Induction and Culture

In nature, unorganized cell masses, known as *callus* or tumors, are generated in plants in response to stresses, such as wounding or pathogen infections. The discovery that the exogenous application of auxin and cytokinin allows the induction of *callus* (callogenesis) in various plant species, determined the widely use of *callus* in both basic research and industrial applications (Ikeuchi et al. 2013).

Callus induction and its in vitro culture have been widely described (Ordas et al. 1990; Kchouk et al. 1997a; El-Bahr et al. 2001; Ruta et al. 2013; Menin et al. 2013).

In 1990, Ordas et al. described *callus* induction using in vitro culture of bracts. A range of *callus* types, i.e., friable, green or light brown in color, was induced with varying frequencies depending on both the hormone combination and the size of the flower buds used for the explant. In particular, it was found out that supplementation with a combination of 2 mg L^{-1} BAP and 5 mg L^{-1} NAA promoted the quickest response and the maximum amount of *callus* formation on bracts from flower buds 2 cm long.

Ruta and colleagues (2013) reported on callogenesis from explants of bracts of Catanese type. They confirmed the callogenic formation in response to both hormonal concentration and combination. The highest percentage of callogenesis was observed from explants cultured on basal *medium* with the combination of the three growth regulators: NAA, BAP, and GA_3 (at the concentration of 5, 0.2, 1 mg L^{-1}, respectively), and with a photoperiod of 16 h/8 h (light/dark). No *callus* was observed with explants cultured in basal-free hormone *media* or basal *medium* added only with BAP and GA_3.

Setting up of callogenesis from leaves has been described in Menin et al. (2013) and Bekheet et al. (2014a). Menin and colleagues chose leaves as starting material because less prone to necrosis (Kchouk et al. 1997a; Moglia et al. 2012), and so they represent the most suitable explants for tissue culture in vitro. The leaf explants used for callogenesis originated from virus-free, meristem culture-derived plantlets. In their study, they compared more than 100 combinations of *media* supplements (e.g., different growth regulators, presence of absorbers of polyphenols, and inhibitors of polyphenol oxidase), along with various light regimes, on three globe artichoke genotypes to define the optimal conditions for *callus* induction from leaf explants. The optimum combination of plant growth regulators to induce *callus* formation in all varietal types was, after 1 week of culture, 1.0 mg L^{-1} BAP plus 3 mg L^{-1} NAA and constant darkness (no other supplements like as absorbers of polyphenols or inhibitors of polyphenol oxidase). The *calli* were pale green in color with a watery consistency.

Bekheet et al. (2014a) obtained *calli* from leaf explants of globe artichoke Green Globe variety cultured on MS *medium* supplemented with 5 mg L^{-1} NAA, 2 mg L^{-1} Kin, 0.1 mg L^{-1} GA_3. Effects on *callus* growth after addition to the culture *medium* of gelling agents (gelrite or agar), picloram, salicylic acid (SA), and jasmonic acid (JA) were also described. *Callus* on *medium* gelled with gelrite usually showed higher growth and better quality in comparison to that grown on the same *medium* gelled with agar; supplementation of culture *medium* with picloram generally enhanced *callus* growth values compared to the control. The addition of SA (most effective concentration 75 μM) or JA (50 μM) induced significant increase of *callus* growth (presented as fresh weight and dry matter), with SA being more effective on the enhancement of *callus* growth compared with JA.

Finally, Bekheet et al. (2014b) reported on the induction of callogenesis from leaf, petiole, and stem explants of in vitro grown shootlets. Results indicated that *callus* was induced on all explants within five weeks of culturing; however, among the three explants examined, leaf segments were the most suitable for *callus* induction giving the best results of *callus* frequency with all combinations of growth regulators used. The highest value of *callus* frequency (80%) was registered when leaf explant was cultured on *medium* containing 5 mg L^{-1} Kin and 0.5 mg L^{-1} IAA.

Overall, all these results confirmed that the efficiency of callogenesis in globe artichoke varies widely according to the cultivar, the starting tissues, and the combinations of growth regulators.

3.4.3 Protoplast Isolation and Culture

Protoplasts are naked cell without the cell wall; the plasma membrane is the only barrier between the living cytoplasm and the exterior environment. The removal of the cell wall opens up a

wide range of possibilities for the in vitro manipulation of plant cells. Thus, they are considered the most promising tools for genetic manipulation of plants. Normally, once cultured on appropriate media, protoplasts can synthesize new cell walls (Willison and Cocking 1972) and regenerate complete plants.

Several techniques have been devised for isolation and culture of isolated protoplasts (Cocking 1972). Theoretically, they can be isolated from a wide variety of plant tissues and organs that include leaves, roots, fruits, and embryos. Among these, the mesophyll tissues of young plants are the most frequently used. Cell suspension cultures also provide excellent source materials for isolating protoplasts in species in which attempts to induce sustained division of mesophyll protoplasts have failed.

Cynara scolymus L. is one of those species in which protoplasts from leaves of in vitro plants failed to divide under a wide range of culture conditions. Only one report on the successful isolation of protoplasts from *C. scolymus* (Ordas et al. 1991) is available. Viable protoplasts were isolated from a cells suspension culture of cv. Romanesco (Fig. 3.5a). The isolation method yields up to 3 × 106 protoplasts per ml packed cell volume (PCV). After three days in culture, up to 85% of the protoplasts regenerated a new cell wall; and in 1–3 additional days, they start to divide. It is noteworthy that there is a direct relationship between wall formation and cell division (Razdan 2003).

As already reported for many plant species (Nagata and Takebe 1971; Mbanaso and Roscoe 1982), also in globe artichoke, the high plating density (2×10^5 prot. mL^{-1}) and the use of agarose plating culture technique enabled cell division and stimulated the development of protoplast-derived colonies, respectively (Fig. 3.5b).

Indeed, the use of soft agar allowed to avoid clumping of protoplasts and to facilitate plating efficiency determination (Vasil and Vasil 1987). The latter, determined after four weeks, generally was in the range of $1 \pm 0.5\%$.

Protoplast yield, viability, and cell division were influenced by several factors, e.g., age of the cell culture, enzyme composition, culture density (Loudon et al. 1989; Roest and Gilissen 1989).

The development of finely divided fast-growing suspension was found to be essential for efficient protoplast isolation and division (Abdullah et al. 1986). Notably, the effectiveness of enzyme solutions and the protoplast yield and quality were varied considerably by the age of the cell suspension culture. Concerning enzymes, pectinase used to disperse cells

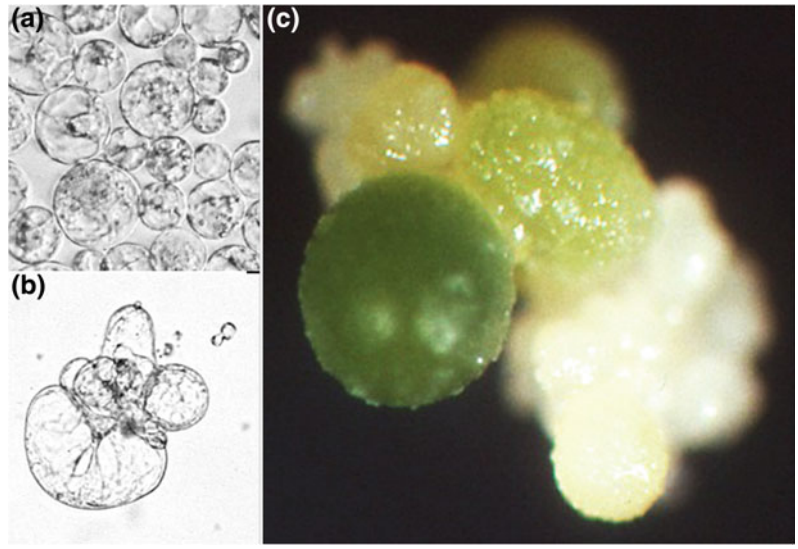


Fig. 3.5 Development of protoplasts of globe artichoke cv. 'Romanesco': **a** freshly isolated protoplasts from 10-day-old suspension culture; **b** protoplast-derived colony 21 days after isolation; **c** *callus* formation after 3 months of culture

aggregates seems to be needful to release good quality protoplasts.

Growing calli (5–7 mm in size), once transferred on medium containing 0.5 mg L^{-1} BAP and 0.5 mg L^{-1} IBA, proliferated into compact green calli which occasionally develop globular structures, which, however, failed to regenerate shoots (Fig. 3.5c).

Although the establishment of a cell suspension culture of globe artichoke by Ordas et al. (1991) paved the way for isolation and culture of protoplasts, the bottlenecks in protoplast-to-plant regeneration still persist, highlighting once again the recalcitrant nature of this species. Overcoming the regeneration bottleneck in globe artichoke would make feasible the use of biotechnological techniques involving the actual manipulation of genetic material.

3.4.4 Plant Regeneration

Plant regeneration is an essential step in the success of any crop improvement program utilizing in vitro tissue culture. A well-defined, reproducible, and highly efficient plant regeneration procedure is a prerequisite for most of the biotechnological approaches.

Plant tissue culture relies on the capacity of a plant cell or tissue to develop an entire plant if suitably stimulated (totipotency). Regeneration begins with a process of morphogenesis. The two primary morphogenic pathways leading to whole plant regeneration involve either somatic embryogenesis, or shoot organogenesis followed by root organogenesis. Both developmental pathways can occur either directly without an intermediate *callus* stage, or indirectly following an unorganized *callus* stage (Gamborg and Phillips 1995). Several factors influence the phenomenon of morphogenesis during in vitro culture. These include genotype, physiological stage of the donor plant, explant source, explant age, explant size and growth regulators.

Globe artichoke has been shown to be recalcitrant to plant regeneration under in vitro conditions. In fact, in the face of several reports published on in vitro propagation, advanced biotechnological approaches are still lacking with only a few works available.

Regeneration in globe artichoke culture has been observed nearly exclusively via indirect organogenesis. In comparison to organogenesis, embryogenesis has been unfortunately neglected. Scaramella Petri and Ricci (1979) occasionally observed embryoid formation on floral stem explants cultured on Bourgin and Nitsch (1967) *medium* supplemented with 1 mg L^{-1} kinetin and 0.5 mg L^{-1} NAA. The embryos were however unable to grow further.

Different globe artichoke explants were explored for regeneration capability. Since young tissues and organs have higher regeneration capacity than the older ones, in globe artichoke the capitulum was expected to be a suitable material for shoot regeneration, based on the positive results reported on related species, such as Gerbera (Murashige et al. 1974; Pierik et al. 1975) and Chrysanthemum (Roest and Bokelmann 1975). However, the *callus* induced on receptacle section of young capitula (Ancora et al. 1979; Morone Fortunato et al. 1979) failed to differentiate into shoots.

Among the explants, young, immature flower buds (1.5–2 cm in size), collected from 2-year-old field-grown plants, have been found to be more responsive to indirect shoot organogenesis (Ordas et al. 1990). Bract-derived *callus* gives rise to shoot differentiation (31%) within 5–6 weeks (Fig. 3.6a) when cultured on MS *medium* containing 30 g L^{-1} sucrose and a combination of auxin and cytokinin (2 mg L^{-1} NAA and 5 mg L^{-1} BAP). Morphogenesis was initiated by cell proliferation in both epidermal and sub-epidermal layers of the cut ends of the bracts followed by shoot formation. De novo differentiation of shoot primordia from *callus* tissue was substantiated by histological analysis. Shoot elongation and proliferation of axillary buds were obtained following transfer on MS *medium* containing 0.5 mg L^{-1} IBA and 0.5 mg L^{-1} BAP (Fig. 3.6b).

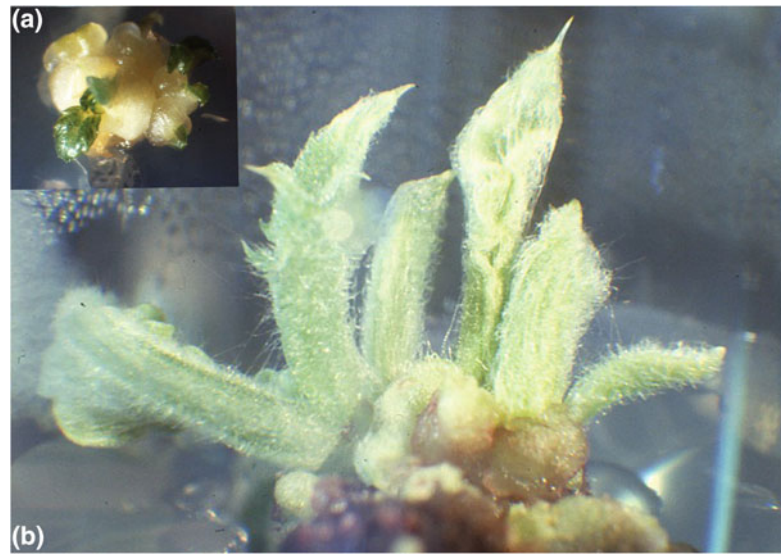


Fig. 3.6 Plant regeneration from bracts of globe artichoke cv. 'Romanesco': **a** development of shoot buds into a leafy-shoot after 5–6 weeks of culture; **b** regenerated plantlets in vitro

The size of the flower buds was found to significantly affect *callus* induction and regeneration while no substantial differences among bracts of the same flower bud were observed.

Regeneration competence of other explants such as hypocotyls, shoot apices, leaves, roots was also explored without success. A sudden but casual induction of shoots in *calli* derived from cotyledon and receptacle explants was reported by Devos et al. (1975) after several subcultures on a *medium* containing 10^{-6} M BAP and 10^{-7} M 2,4-D.

Differential regeneration capacity of adventitious shoots on different explants from the same plant has been documented several times. This differential response observed between tissue sources may be due to tissue competency, which in turn could be due to different levels of pre-existing inductive factors in the tissue.

So far, the immature flower buds, collected from field-grown plants, represent the only explant competent for regeneration and this is a limit for their routine use in in vitro studies. In fact, the seasonality of the species (one biological cycle/year), the specific developmental stage of the immature flower buds, and the low regeneration frequency still represent a bottleneck.

3.4.5 Genetic Transformation

The difficulty in obtaining regenerated globe artichoke plants has hampered the possibility to fully exploit transgenesis as a tool for genetic improvement (Ordas et al. 1991; Kchouk et al. 1997b). In fact, although globe artichoke cotyledons and leaf tissues represent the most suitable explants for in vitro culture (Menin et al. 2012, 2013), with cotyledons giving better rates of neoformation than leaves in a shorter time (Kchouk et al. 1997a), the obtainment of whole regenerated plants is a very inefficient process.

Transgenic *calli* of globe artichoke were obtained from leaf explants of in vitro plantlets (cv. 'Romanesco' and cv. 'Violet d'Hyeres') by using *Agrobacterium tumefaciens* strain C58pGV2260 containing GUS-INT and NPT-II genes (Gonzalès et al. 1994; Kchouk et al. 1997b) (Fig. 3.7a). However, no shoot regeneration was obtained from GUS-positive *calli* using regeneration media containing different BAP, NAA, and IBA concentrations. In another attempt, leaf explants of in vitro clones of the varietal type 'Romanesco' were agroinfected with the *A. tumefaciens* strain AGL0 01-124 containing the binary vector pCAMBIA 2301

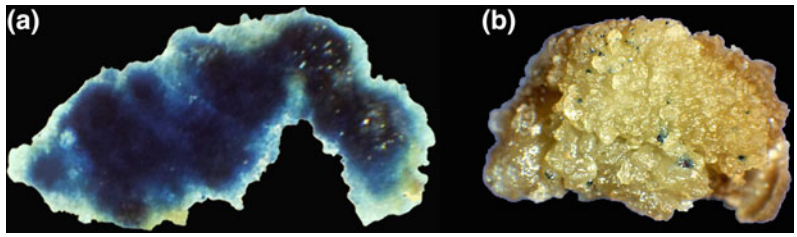


Fig. 3.7 Transformation of globe artichoke. **a** GUS expression in stably transformed, leaf-derived callus of cv. 'Violet d'Hyéres' four weeks after *Agrobacterium* transformation. **b** Histochemical localization of GUS activity in morphogenetic *callus*, 48 h after particle bombardment of cv. 'Romanesco' with the plasmid pBI121

(Menin et al. 2012). The *calli* originated from leaf explants were subjected to histochemical analysis, and about 30% of them resulted positive to the β-glucuronidase gene (GUS) assay. The detection of blue spots in the transformed *calli* showed that stable transformation took place unevenly in the transformed tissues. No shoots were obtained from these transformed *calli* due to the low regeneration rate.

Finally, to bypass tissue culture-related regeneration difficulties, particle bombardment approach of gene delivery was investigated in globe artichoke by Gonzalès et al. (1994). Transfer and expression of the GUS gene were obtained into three-target tissue: leaf explants, axillary buds, and morphogenic bract-derived *callus*. 66.6% of axillary buds, 32.3% of morphogenic *calli* (Fig. 3.7b) and 20% of leaf explants resulted positive to GUS assay 24 h after bombardment although no transgenic plants were recovered from the organized tissue.

3.4.6 Secondary Metabolite Production from In Vitro Culture

Pharmaceutical interest in globe artichoke is related to its high content in polyphenolic compounds, mono- and dicaffeoylquinic acids (1,5-O-dicaffeoylquinic acid) which possess antioxidative, hepatoprotective, diuretic, and choleretic activity (Lattanzio et al. 2009; Wang et al. 2003). Globe artichoke leaves can accumulate up to 50 g CQA kg^{-1} dry weight (DW), mainly as chlorogenic acid (CGA; 5-O-caffeoylquinic acid) while, in mature inflorescences, the CGA content reaches only 1–2 g kg^{-1} DW (Wang et al. 2003; Lattanzio et al. 2009).

The extraction of secondary metabolites requires a large amount of plant material, and their production by plants may greatly vary in relation to the physiological stage of the tissues as well as the genotype and its interaction with environmental and seasonal factors. The biotechnological production of high-value metabolites by the plant in vitro system may be considered an attractive alternative to classical extraction technologies.

A previous study compared the content of CGA and 1,5-diCQA in leaf tissues and *calli* obtained from leaf explants of two globe artichoke genotypes (SAROM and C3-RR) belonging to the varietal type 'Romanesco' (Moglia et al. 2012; Menin et al. 2013). In both varietal clonal types, the content of 1,5-diCQA of the *calli* was three to fourfold greater than in the leaf, while the content of CGA accumulates preferentially in globe artichoke leaf tissue, in accordance with what was reported in cultivated cardoon (Trajtemberg et al. 2006).

The synthesis of plant secondary metabolites is induced by pathogen attack, physical wounding, or UV irradiation. An increase in diCQAs content in foliar disks of globe artichoke was detected following UV-C exposure (Moglia et al. 2008). The effect of UV-C irradiation was thus tested on *callus* tissues of Romanesco SAROM and slight enhancement of both mono- and diCQAs acids accumulation was observed, suggesting

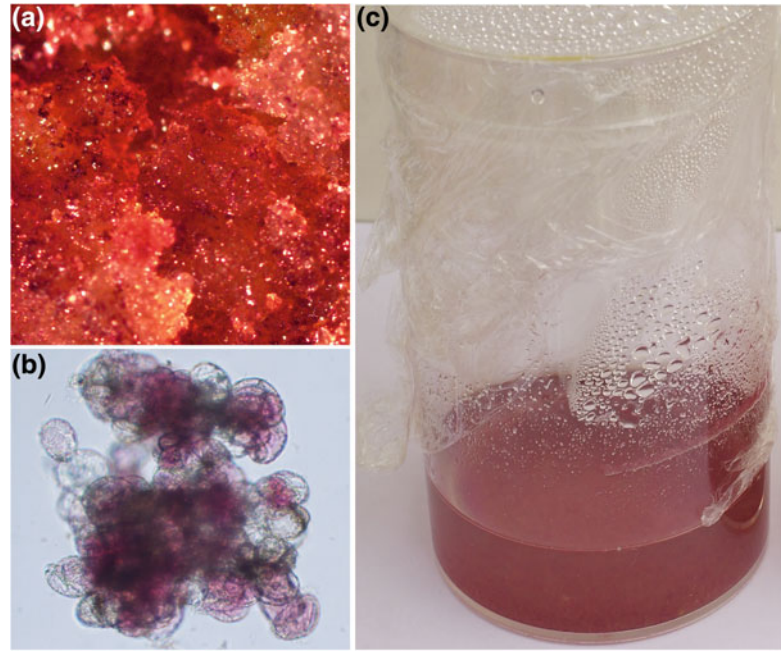


Fig. 3.8 Establishment of anthocyanin-producing suspension culture of globe artichoke cv. 'Grato 1': **a** *callus* formation from in vitro leaves after 2 months of culture; **b** cell-line selection; **c** *callus*-derived suspension culture after 4 months of culture

the possibility to use UV-C-irradiated globe artichoke *callus* culture for the production of phenolic esters (Moglia et al. 2012; Menin et al. 2013).

Two suspension cultures, an anthocyanin-producing and a not-producing cultures, and the sourced *callus* of globe artichoke 'Grato 1' genotype (Fig. 3.8) were evaluated in terms of their polyphenol content (TP), total anthocyanin content (TA), and antioxidant activity (Meneghini et al. 2016; Pandino et al. 2017a). The cell suspension cultures showed enhanced antioxidant activities compared to *calli*. Anthocyanin-producing suspension cultures accumulated higher TP and exhibited very strong DPPH free radical scavenging activity and reducing power compared to the non-producing suspension culture, due to the presence of anthocyanins in the cultures. 1,5-O-diCQA together with 5-O-CQA were the predominant caffeic acid derivatives underlined in plant cell extracts. Pandino et al. (2017a) highlighted the similarity in polyphenol profiles between globe artichoke 'Grato 1' suspension cultures and the leaves from field-grown plants, and that cell suspension cultures of 'Grato 1' accumulated significantly higher levels of total phenolic content than leaves.

In another study (Bekheet et al. 2014a, b), the in vitro hypolipidemic and antioxidant effects of extracts of globe artichokes *callus* cultures have been analyzed in comparison with that of milk thistles: Globe artichoke *callus* extract showed more potent hypolipidemic and antioxidant effect.

It was reported that mycorrhizal inoculation might represent an efficient strategy to increase plant phenolic compounds in globe artichoke plants (Palermo et al. 2013). Micropropagation/mycorrhization lead to a higher content of CQAs in globe artichoke plants as compared to control plants, which have been raised using conventional vegetative propagation (Pandino et al. 2017b).

3.5 Germplasm Maintenance and Conservation

3.5.1 Overview

The basic principle of the germplasm conservation technology is to preserve the maximum gene pool of a determined plant species, considering

their actual and potential use for the future (Razdan 2003).

A broad range of materials can be conserved, including traditional varieties, landraces, modern cultivars, breeding lines, wild and weedy species, alongside the new category of biotechnology products, as cell lines with special attributes, and genetically transformed materials. This new germplasm is often of high added value and very difficult to produce (Engelmann et al. 1994).

The development of efficient techniques to ensure their safe conservation is therefore of paramount importance.

There are essentially two complementary basic strategies for the conservation: namely in situ and ex situ. The former is the preservation of species of living organisms in the habitat where they naturally occur; the latter involves a set of conservation techniques for the transfer of a target species outside their natural habitats. It entails the preservation and maintenance of samples of living organisms in the forms of whole plants, seed, pollen, and tissue or cell culture.

The most widely used method for conserving plant genetic resource depends on seed. However, several categories of crops present problems related to seed storage, thus these species are mainly propagated vegetatively to maintain clonal genotype.

The globe artichoke is an example of such a crop where conservation in seed form is unsuitable due to its high heterozygosity level; plants are usually conserved in field genebanks. Although the latter strategy provides easy access to conserved material for use, its maintenance requires greater space, labor, and cost. Besides, field genebanks run the risk of destruction by environmental and biological hazards (Bianco 1990; Babes et al. 2004).

In vitro conservation has been proposed as a safer alternative to the field genebank (De Langhe 1984; Withers 1989), and, more generally as an efficient technique complementary to in situ practices to preserve recalcitrant and vegetatively propagated species (Dodds 1991). This technique has the advantage of requiring less space for storage. The main purpose of in vitro conservation technology is the maintenance and exchange of germplasm in the disease-free and genetically stable state through tissue culture. The maintenance of the true-to-type nature of in vitro plants is an important requisite to uphold certain agronomic and horticultural traits when using elite genotype, not just for conservation, but also for commercial purposes.

The use of shoot tips or axillary buds as a base for multiplication and conservation greatly reduces the risk of genetic variation (Murashige 1974). However, the possibility of occurrence of somaclonal variants even in such cultures cannot be ruled out (Devarumath et al. 2002). Somaclonal variation can result from either preexisting variation in explant tissues or induced variation during tissue culture (Skirvin et al. 1994). Factors such as genotype, *media* composition, phytohormones, explant source and number of subcultures may affect the extent of somaclonal variation during tissue culture. Therefore, a rigorous analysis of the genetic stability of the plantlets produced through tissue culture is essential.

In vegetatively propagated crops, morphological criteria have been applied to assign a genotype; however, such differences are difficult to detect in in vitro propagated plantlets.

For conservation purpose, molecular markers have come up as the most desirable tool for establishing genetic similarity or dissimilarity of in vitro propagated plants (Rey et al. 2013a, b; Tavazza et al. 2015).

Different in vitro conservation methods can be employed, depending on the use and storage duration required (Engelmann 1997, 2011). These fall into three categories: (i) short-term, (ii) medium-term, and (iii) long-term conservation. For short-term (normal growth) and medium-term (slow growth) storage, used for germplasm evaluation and distribution, the aim is to reduce the growth and to increase the intervals between subcultures. Conversely, long-term conservation (cryopreservation) foresees the conservation of the plant material for an unlimited period but not its distribution.

Each conservation method has its advantages and disadvantages, and complementary strategies

are required to maximize the conservation of genetic diversity, which varies between species.

3.5.2 Short-Term Conservation

Globe artichoke genetic resources comprise a diversity of genetic material contained in traditional varieties, modern cultivars, and crop wild relatives.

Since 2007, globe artichoke landraces have been collected, characterized and preserved to conserve these gene pools both in the field and in in vitro collections (Trionfetti et al. 2007; Pagnotta 2012; Tavazza et al. 2013b).

The techniques to set up an in vitro collection such as meristem tip culture and micropropagation and the procedures/conditions for the establishment and multiplication of axillary shoot cultures have been previously described.

In vitro conservation of globe artichoke germplasm can partly be achieved by regular subculture on fresh *media*, making possible to maintain cultures virtually indefinitely under normal growth conditions. However, on normal growth medium, in vitro plantlets have to be frequently transferred (3–4 weeks) which involves not only manpower and high costs, but also the possible hazard of contamination, and, sometimes, the loss of the entire sample. For these reasons, micropropagation has been considered only for short-term storage of globe artichoke germplasm. Nonetheless, such method of storage has the advantages that the stored material is readily available for evaluation and distribution can be easily evaluated to be alive, and cultures may be quickly replenished when necessary.

Globe artichoke is a perennial crop, and the morphological variations generated in tissue culture could be visible only at the maturity. Hence, an early detection of genetic variations among the tissue culture raised plantlets is considered paramount for conservation purposes.

Rey et al. (2013a, b) evaluated the usefulness of molecular markers, such as inter-simple sequence repeats (ISSR) and simple sequence repeats (SSR), in assessing the differences between mother and in vitro plants of six clones of globe artichoke and the possible variation occurring during the cycles of micropropagation. ISSRs proved to be useful in evaluating the genetic stability/variability of both in vitro and in vivo propagated plants reflecting the findings of other studies (Leroy et al. 2001; Martins et al. 2004; Martin et al. 2006; Joshi and Dhawan 2007).

Notably, ISSRs showed that the greatest number of changes occurred in the early stages of subculturing. Statistically significant differences were revealed among globe artichoke clones, suggesting either that some clones are more susceptible to somaclonal variation, or that the in vitro instability is a consequence of a clone versus culture *medium* interaction. Nevertheless, the general trend of changes across the subculture was similar, reaching a maximum rate between the second and fourth subcultures (Rey et al. 2013a, b).

It should be noted that the changes detected by molecular markers are not necessarily impacting on morphological aspects since could be located all over the genome including not coding regions.

These results provide an important basis for giving evidence that the clonal fidelity of micropropagated plants of globe artichoke cannot always be assured even when organized meristem tissues are used as explant sources. Assessment of genetic fidelity of micropropagated plants at an early stage helps fine-tuning of conservation protocol parameters and judges the suitability of multiplication protocol for large-scale applications.

3.5.3 Medium-Term Conservation: Slow Growth

For medium-term conservation, the aim is to reduce the growth of in vitro plantlets and to increase the interval between subcultures. This is achieved by modifying the physical environment of culture, the *medium* composition or both (Engelmann 1997).

Several parameters influence the efficiency of in vitro slow growth storage protocol such as the

type of the explants, their physiological state, the time elapsed after the last subculture, and the type of culture vessel used.

Choice of treatment is largely species-dependent and substantiated by the ability of specific cultures to endure the stresses incurred; thus, it is essential to optimize regimes for each species for the timing of subculture and regeneration.

The most successful strategies have involved temperature reduction, often combined with a decrease in light intensity or dark growth. The storage temperature is strictly dependent on the cold sensitivity of the species. At low temperatures, the aging of the plant material is slowed down but not completely stopped as in cryopreservation. Lowering temperature and light exposure have physiological consequences, such as the reduction of respiration and metabolism, water loss, and wilting, which allow for safe conservation (Hughes 1981). In globe artichoke, in vitro slow growth was achieved by combining temperature reduction with the absence of light. Shoot buds of cv. Balady (Egyptian local cultivar) and cv. Violetto di Sant'Erasmo, could be stored up to 3-months at 6 °C (Bekheet 2007; Bekheet and Usama 2007) or 4-months at 4 °C (Benelli et al. 2010) respectively without any detrimental effect on viability and regrowth of the shoots. Although the rates of regrowth slightly decrease as storage duration increased (6, 9 and 12 months), the recovery percentage was enough to obtain high frequencies of healthy shoots (Bekheet 2007; Bekheet and Usama 2007) genetically identical to the control as highlighted by RAPD analysis.

The inclusion of non-metabolisable sugar alcohol mannitol or sorbitol along with an increased or decreased sucrose concentration in the culture *medium* was reported to be effective in plant growth reduction and in extending the subcultural interval (Wang and Hu 1985; Roca et al. 1989). Mannitol or sorbitol increases the osmotic pressure of the *medium*, and as a result, the water availability to the growing cultures is reduced (Dodds and Roberts 1985; Thompson et al. 1986). The application of controlled osmotic stress has been used with success for the conservation (12 months at 18 ± 2 °C) of seven late type's globe artichoke genotype (Tavazza et al. 2015). In this work, two concentrations of sucrose (2 and 3% w/v) in combination with either mannitol (4% w/v) or sorbitol (2 and 4% w/v) were tested. After 12 months of storage, culture survival across genotypes ranged from 65 to 85% and all the *media* supported 100% regrowth (Fig. 3.9). No culture abnormalities were observed throughout and after minimal growth maintenance. On the contrary, Bekheet (2007) reported that under osmotic stress, a high percentage of globe artichoke cultures suffered from leaf etiolation.

The best storage parameters for slow-growing globe artichoke germplasm were defined in terms of their genetic and field performance once retrieved from storage. The *medium* supplemented with 20 g L^{-1} sucrose and 20 g L^{-1} sorbitol proved to be that where the seven genotypes displayed phenotypes similar to the control, coupled with the lowest percentage of changes (2.43%) at a molecular level (Tavazza et al. 2015).

Slow growth strategies are applicable for medium-term conservation (Reed 1995; Reed and Change 1997) and are now routinely used for the storage of shoot culture in many industrial laboratories and international/regional centers (e.g., CIAT, CIP, CATIE). Such method enables increasing the efficient use of resources, which can be used for multiplication, evaluation, and distribution (Williams 1984) and for reduction of labor cost and staff time.

3.5.4 Long-Term Conservation: Cryopreservation

Cryopreservation, i.e., storage of biological material at an ultra-low temperature, typically that of liquid nitrogen (−196 °C), currently offers the only safe and cost-effective option for long-term conservation of plant germplasm (Engelmann 1997; Wang and Perl 2006). Under cryostorage condition, all cellular divisions and metabolic processes are completely arrested (Reed 2008); plant material can be stored without any change for a theoretically unlimited period.

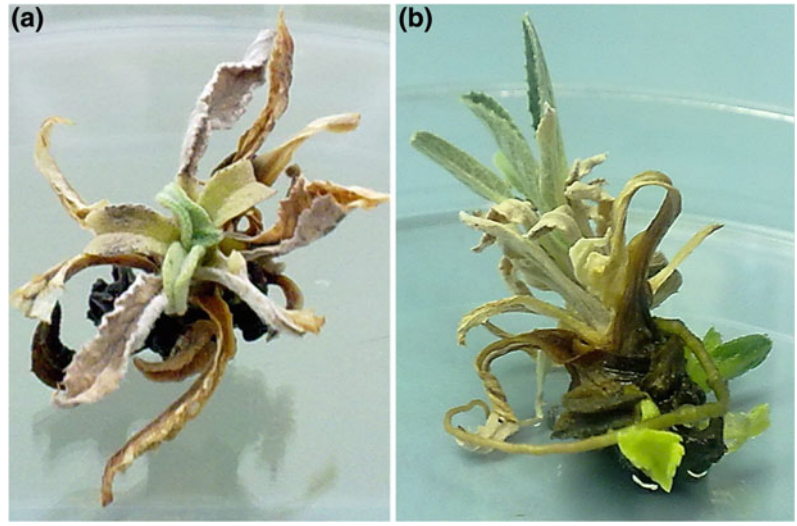


Fig. 3.9 In vitro development of retrieved globe artichoke plants after 12 months of conservation by slow growth culture: **a** 12-month-old conserved plantlet grown in the presence of sorbitol; **b** plants recovery and axillary bud proliferation after 4 weeks

The overall process for cryopreservation consists of three phases: conditioning plants, cryogenic phase, and recovery processes.

To date, cryopreservation has been successfully applied to many horticultural plants, e.g., garlic (Keller 2002; Volk et al. 2004), asparagus (Mix-Wagner et al. 2000), chicory (Benelli et al. 2011), and potato (Kaczmarczyk et al. 2011). For cryopreservation of vegetatively propagated species, shoot-meristematic tissues are usually preferred over non-organized tissues, such as *calli* and cell suspension, because they are genetically more stable (Helliot et al. 2002; Benelli et al. 2012). Among different cryogenic techniques (Feng et al. 2011), vitrification is by far the most widely used cryopreservation protocol; it is easy, high reproducible and can be successfully applied to a wide range of tissue and plant species. Recently, the "vitrification/one-step freezing" method (Sakai et al. 1990) was applied to cryopreserve globe artichoke shoot tips, dissected from in vitro plantlet (Taglienti and Barba 2012; Tavazza et al. 2013a; Taglienti et al. 2013). This technique involves the following steps: cold-hardening, preculture of samples on *medium* enriched with cryoprotective substances, treatment with a loading solution (LS) (e.g., a mixture of 2 M glycerol and 0.4 M sucrose; Matsumoto et al. 1995), dehydration with a highly concentrated Plant Vitrification Solution 2 (PVS2) (Sakai et al. 1990), rapid freezing and thawing, removal of cryoprotectants and recovery. To make this protocol suitable for globe artichoke, it is essential to perform preliminary experiments with every clone to assess their viability and regrowth capacity after each preparation step (Reed 2000).

In late genotypes of globe artichoke, cold hardening of in vitro stock culture (4–5 weeks), as well as the shoot tip size, is regarded as the two factors affecting globe artichoke shoot tip survival after cryopreservation and the key for attaining post-LN growth (Tavazza et al. 2013a). In fact, the authors reported that without cold acclimation step, shoot tips did not survive dehydration and subsequent LN exposure. The size of the shoot tips was also critical, and medium size tips (3–4 mm in length) proved to be the best responding starting material. Conversely, the preculture procedure did not affect regrowth of cryopreserved shoot tips.

The latter result is in contrast to previous findings (Taglienti and Barba 2012) reporting that it was necessary to preculture shoot tips (1.0 mm in size) for 48 h on a *medium* supplemented with 0.3 M sucrose before freezing. This difference could be ascribed to the shorter cold acclimation of plantlets used by the authors.

The temperature of treatment and duration of exposure to the vitrification solution are

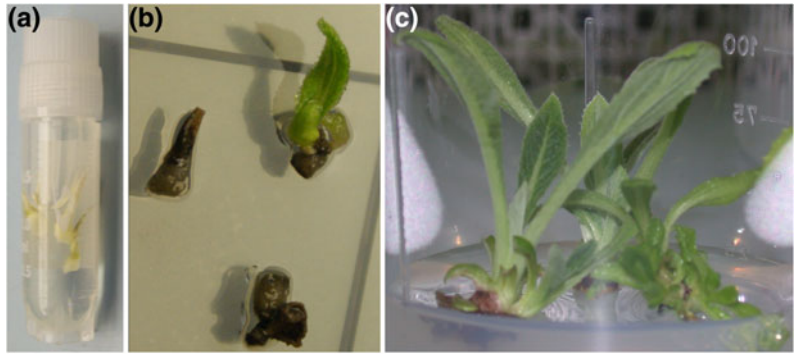


Fig. 3.10 Plant regeneration from cryopreserved shoot tip of globe artichoke cv. 'Grato 1': **a** dehydration of shoot tips with the PVS2 solution; **b** shoot tip regrowth after 8 weeks of post-thaw culture following cryopreservation; **c** shoot proliferation of cryopreserved plantlet after 11 weeks of post-thaw culture

important factors affecting survival and regrowth of cryopreserved explant (Niino et al. 1997). Exposure to a PVS2 at a low temperature (0 °C) has been shown to minimize toxicity (Fig. 3.10a). For the two late genotypes used, the best survival percentages were obtained with an exposure time of 90 min at 0 °C. Under these conditions, shoot tips' survival after cryopreservation achieves 55–61% depending on the genotype (Fig. 3.10b). Moreover, the exposure time to PVS2 solution seems to be associated with the size of the excised shoot tips and appears to be genotype-specific at the same temperature.

A rapid growth recovery after warming denotes that the structural and physiological integrity of shoot apices has been preserved during the cryopreservation process. Globe artichoke surviving shoot tips had a slight delay in the development after the incubation on the recovery *medium*. Then, they showed the same regrowth rate (31–36%) as the control and developed shoots directly into plantlets, which proved to be stable regarding chromosome number (Fig. 3.10c).

Taglienti et al. (2013) by modifying the vitrification technique (Sakai et al. 1990) established a protocol for cryopreservation of early globe artichoke genotype obtaining a high regeneration rate (70–95%).

Finally, cryopreservation of shoot tips can also be used to eliminate pathogens (Wang et al. 2009; Wang and Valkonen 2009). The "vitrification/one-step freezing" method was notably effective (100%) in eliminating the endemic artichoke latent virus (ArLV) from both early and late globe artichoke genotypes (Tavazza et al. 2013a; Taglienti et al. 2013). Variations in survival, regeneration rates and virus eradication efficiency were observed either between late and early genotypes or within each genotype when tested under the same culture conditions (Taglienti et al. 2013). It is therefore essential that procedures be optimized for genotypic differences in dehydration and chemical sensitivity (Reed 2000).

References

Abdullah R, Cocking EC, Thompson JA (1986) Efficient plant regeneration from rise protoplasts through somatic embryogenesis. Bio/Technology 4:1087–1090

Acquadro A, Papanice MA, Lanteri S, Bottalico G, Portis E, Campanale A, Finetti-Sialer MM, Mascia T, Sumerano P, Gallitelli D (2010) Production and fingerprinting of virus-free clones in a reflowering globe artichoke. Plant Cell, Tiss Organ Cult 100 (3):329–337

Ancora G (1986) Globe artichoke (*Cynara scolymus* L.) In: Bajai YPS (ed) Biotechnology in agriculture and forestry, crops I. Springer, Berlin, pp 471–484

Ancora G, Saccardo F (1987) Carciofo: nuove tecniche di propagazione. L'Informatore Agrario 43(4):77–78

Ancora G, Belli-Donini ML, Cuozzo L (1979) Moltiplicazione in vitro di apici vegetativi di carciofo. In: Atti 3rd Congr Int Stud Carciofo, Bari Ind Grafica Laterza, Bari, pp 239–246

Ancora G, Belli-Donnini ML, Cuozzo L (1981) Globe artichoke plants obtained from shoots apices through

rapid in vitro micropropagation. Sci Hortic 14:207–213
Babes G, Lumia V, Pasquini G, Di Lernia G, Barba M (2004) Production of virus free artichoke germplasm. Acta Hortic 660:467–472
Barba M, Taglienti A (2015) Risanamento di carciofo tramite coltura dell'apice meristematico e/o termoterapia in vitro. Colture in vitro: note di laboratorio, pp 119–120
Barba M, Di Lernia G, Babes G, Citrulli F (2004) Produzione e conservazione di germoplasma di carciofo di tipo romanesco esente da virus. Italus Hortus 11(5):5–10
Barba M, Tiberini A, Taglienti A (2013) Collezione di germoplasma di carciofo sanitariamente qualificata. Acta Italus Hortus 88:31–38
Bedini L, Lucchesini M, Bertozzi F, Graifenberg A (2012) Plant tissue cultures from four Tuscan globe artichoke cultivars. Cent Eur J Biol 7(4):680–689
Bekheet SA (2007) In vitro preservation of globe artichoke germplasm. Plant Tissue Cult Biotechnol 17(1):1–9
Bekheet S, Usama IA (2007) In vitro conservation of globe artichoke (*Cynara scolymus* L.) germplasm. Int J Agric Biol 9(3):404–407
Bekheet SA, El-Bahr MK, Ali SA, Hamed MA (2014a) Callus production of globe artichoke and milk thistle: in vitro hypolipidemic and antioxidant activities. World J Pharm Res 3:1–17
Bekheet SA, Gabr AMM, El-Bahr MK (2014b) Development a protocol for in vitro morphogenesis and multiplication of globe artichoke (*Cynara scolymus* L.). Int J Acad Res Part A 6(6):297–303
Benelli C, De Carlo A, Previati A, Roncasaglia R (2010) Recenti acquisizioni sulla conservazione in vitro in crescita rallentata. In: Proceedings of 'IX Giornate Scientifiche SOI'. Italus Hortus 17:91
Benelli C, Previati A, Lambardi M, De Carlo A (2011) Shoot-tip vitrification protocol for red chicory (*Cichorium intybus* L.) lines. Adv Hortic Sci 25:44–50
Benelli C, De Carlo A, Engelmman F (2012) Recent advances in the cryopreservation of shoot-derived germplasm of economically important fruit trees of *Actinidia*, *Diospyros*, *Malus*, *Olea*, *Prunus*, *Pyrus* and *Vitis*. Biotechnol Adv 31:175–185
Bianco VV (1990) Carciofo (*Cynara scolymus* L.) In: Bianco VV, Pimpini F (eds) Orticoltura. Patron Editore, Bologna, pp 209–251
Bigot C, Foury L (1984) In vitro propagation of globe artichoke (*Cynara scolymus* L.) from seeds: field comparison of some clones with their parent lines. Agronomie 4(8):699–710
Bourgin JP, Nitsch JP (1967) Obtention de Nicotiana haploides a partier d'etamines cultivees in vitro. Ann Physiol Veg 9(4):377–382
Brutti C, Apostolo N, Ferrarotti SA, Lorente BE, Krymliewic Z (2000) Micropropagation of *Cynara scolymus* L. employing cyclodextrins to promote rhizogenesis. Sci Hortic 83:1–10
Cadinu M, Repetto A, Frau A, Beneventi S, Meloni S (2004) Influence of the explant type on the phenotypic changes in micropropagated plants of artichoke. Acta Hortic 660:373–380
Cadinu M, Mallica GM, Baghino L, Repetto A, Frau A, Meloni S (2006) Progressi nella coltura in vitro e nel miglioramento genetico del carciofo "Spinoso sardo". In: Atti del Convegno Progetto MiPAF Carciofo, Roma, pp 70–72
Castiglione V, Cavallaro V, Di Silvestro I, Melilli MG (2006) Influence of different substrates on in vitro initiation of some early and late cultivars of globe artichoke (*Cynara cardunculus* L. subsp. *scolymus* L. Hayek). Acta Hortic 730:107–112
Cavallaro V, Castiglione V, Avola G, Finocchiaro E, (2004) Influence of different substrates on the in vitro rhizogenesis process of early artichoke (*Cynara cardunculus* L. subsp. *scolimus* (L.) *Hegi*). Acta Hortic 660:267–272
Cocking EC (1972) Plant cell protoplasts-isolation and development. Ann Rev Plant Physiol 23:29–50
Comino C, Cerruti E, Acquadro A, Marconi G, Repetto A, Pisanu AB, Pilia R, Albertini E, Portis E (2017) DNA methylation profile of micropropagated globe artichoke. In: Abstract book of the 40th new phytologist symposium. The New Phytologist Trust—Lancaster University, Lancaster, Vienna (Austria), p 58
De Langhe EAL (1984) The role of in vitro techniques in germplasm conservation. In: Holden JHW, Williams JT (eds) Crop genetic resources: conservation and evaluation. George Allen and Unwin, Boston, pp 131–137
De Leo P, Greco B (1973) Nuova tecnica di propagazione del carciofo: coltura in vitro di meristemi apicali. In: 2° Congr Int sul Carciofo, pp 657–665
De Paoli G, Rossi V (1987) Carciofo in vitro. Terra e Vita 12:55–61
Devarumath RM, Nandy S, Rani V, Marimuthu S et al (2002) RAPD, ISSR and RFLP fingerprints as useful markers to evaluate genetic integrity of micropropagated plants of three diploid and triploid elite tea clones representing *Camellia sinensis* (China type) and *C. assamica* ssp. *assamica* (Assam-India type). Plant Cell Rep 21:166–173
Devos P, De Bruijne E, De Langhe E (1975) Influence of 2,4-D on the propagation of *Cynara scolymus* L. in vitro. Meded Fac Landbouwhogesch Rijksuniv Gent 40:829–836
Dodds JH (ed) (1991) In vitro methods for conservation of plant genetic resources
Dodds JH, Roberts LW (1985) Experiments in plant tissue culture, 2nd edn. Cambridge University Press, Cambridge, UK
Draoui N, Ghorbel A, Kchouk ME (1993) In vitro culture of globe artichoke (*Cynara scolymus* L.) in Tunisia: utilization of in vitro methods in artichoke improvement. Agric Mediterr 123(2):139–145
Dridi B (2003) Un système intégré de micropropagation de l'artichaut (*Cynara scolymus* L.). Thèse de doctorat

en sciences agronomiques et biologiques appliquées: biotechnologie cellulaire et génétique. Belgique. 217p

EI-Boullani R, Elmaslih A, Fini AEL (2012) Improved in vitro micropropagation of artichoke (*Cynara cardunculus* var. *scolymus* L.). Eur J Sci Res 80(4):430–436

El-Bahr MK, Okasha KA, Bekheet SA (2001) In vitro morphogenesis of globe artichoke (*Cynara scolymus* L.). Arab J Biotechnol 4:119–128

Elia A, Conversa G, Montervino C, Di Brita S, Lotti C (2006) Micropropagation of the Artichoke cultivar Violet du Provence. Acta Hortic 730:127–134

El-Zeiny OAH, El-Behairy UA, Zocchi G, Rashwan MM (2013) Commercial production of globe artichoke (*Cynara scolymus* L.) in vitro. Egypt J Agric Res 91 (3):933–1007

Engelmann F (1997) In vitro conservation methods. In: Callow JA, Ford-Lloyd BV, Newbury HJ (eds) Biotechnology and plant genetic resources. CAB International, Wellingford UK, pp 119–161

Engelmann F (2011) Use of biotechnologies for the conservation of plant biodiversity. Vitro Cell Dev Biol Plant 47:5–16

Engelmann F, Dambier D, Ollitrault P (1994) Cryopreservation of cell suspensions and embryogenic calluses of citrus using a simplified freezing process. Cryo-Letters 15:53–58

Feng C, Yin Z, Ma Y, Zhang Z, Chen L, Wang B, Li B, Huang Y, Wang Q (2011) Cryopreservation of sweetpotato (*Ipomoea batatas*) and its pathogen eradication by cryotherapy. Biotechnol Adv 29:84–93

Foddai A, Corda P, Idini G (1983) Influenza del potyvirus latente del carciofo Spinoso sardo sulla produttivita delle piante in pieno campo. I risultati del primo anno d'impianto. Rivista di Patologia Vegetale 19(1):29–35

Frau A, Mallica G, Baghino L, Cadinu M, Repetto A (2004) La micropropagazione del carciofo Spinoso sardo: un valido strumento per aumentare la produttività degli impianti. Italus Hortus 11(5):38–41

Gallitelli D, Barba M (2003) Normativa vivaistica in Italia. In: Atti Giornate nazionali di studi sul carciofo, vivaismo e strategie di sviluppo del carciofo, Samassi (CA), pp 47–51

Gallitelli D, Rana GL, Vovlas C, Martelli GP (2004) Viruses of globe artichoke: an overview. J Plant Pathol 86(4):267–281

Gallitelli D, Papanice M, Campanile A, Bottalico G, Sumerano F (2006) Risanamento del carciofo Brindisino. In: Convegno conclusivo Progetto MiPAF Carciofo, pp 22–24

Gamborg OL, Phillips GC (eds) (1995) Plant cell, tissue and organ culture-fundamental methods. Springer, Heidelberg

Gonzalès MV, Kchouk ML, Tavazza R, Ancora G (1994) Trasformazione genetica del carciofo (*Cynara scolymus* L.) workshop: Biotecnologie avanzate applicate alle piante ortoflorofrutticole. II giornate scientifiche SOI. San Benedetto del Tronto, pp 82–84

Harbaoui Y, Debergh P (1980) Multiplication in vitro de clones selectionnes l'artichaut (*Cynara scolimus* L.) In application de la culture in vitro a l'ameloration des plantes patogenes. Reun. Eucarpia, Versailles, pp 1–7

Hartmann HT and Kester DE (1983) Plant propagation principles and practices, section 16, 4th edn. Prentice-Hall lnc., Englewood Cliffs New Jersy

Helliot B, Panis B, Poumay Y, Swennen R, Lepoivre P, Frison E (2002) Cryopreservation for the elimination of cucumber mosaic and banana streak viruses from banana (*Musa* spp.). Plant Cell Rep 20:1117–1122

Hughes KW (1981) In vitro ecology: exogenous factors affecting growth and morphogenesis in plant culture system. Environ Exp Bot 21:281–288

Husain MK, Anis M (2009) Rapid in vitro multiplication of *Melia azedarach* L. (a multipurpose woody tree). Acta Physiol Plant 31(4):765–772

Iapichino G (1996) Micropropagation of globe artichoke (*Cynara scolymus* L.) from underground dormant buds ("ovoli"). In vitro Cell Dev Biol Plant 32:249–252

Ikeuchi M, Sugimoto K, Iwase A (2013) Plant callus: mechanisms of induction and repression. Plant Cell 25:3159–3173

Joshi P, Dhawan V (2007) Assessment of genetic fidelity of micropropagated *Swertia chirayita* plantlets by ISSR marker assay. Biol Plant 51(1):22–26

Kaczmarczyk A, Rokka VM, Keller ERJ (2011) Potato shoot tip cryopreservation. A review. Potato Res 54:45–79

Kchouk M, Mliki A, Chatibi A, Ghorbel A (1997a) In vitro organogenesis and transgenosis aspects in globe artichoke (*Cynara scolymus* L.). Cell Mol Biol 43:399–408

Kchouk ML, Gonzalès MV, Mliki A, Chatibi A, Tavazza R, Ancora G, Ghorbel A (1997b) L'artichaut (*Cynara scolymus* L.): plante rècalcitrante. In: Quel Avenir pour l'Amélioration des Plantes? AUPELF-UREF, John Libbey Eurotext, Paris, pp 87–95

Keller ERJ (2002) Cryopreservation of *Allium sativum* L. (Garlic). In: Towill LE, Bajaj YPS (eds) Biotechnology in agriculture and forestry. Cryopreservation of plant germplasm II, vol 50. Springer, Berlin, pp 37–47

Larkin PJ, Scowcroft WR (1981) Somaclonal variation—a novel source of variability from cell cultures for plant improvement. Theory Appl Genet 60(4):197–214

Lattanzio V, Kroon P, Linsalata V, Cardinali A (2009) Globe artichoke: a functional food and source of nutraceutical ingredients. J Funct Foods 1:131–144

Leroy XJ, Leon K, Hily JM, Chaumeil P, Branchard M (2001) Detection of in vitro culture-induced instability through inter-simple sequence repeats analysis. Theory Appl Genet 102:885–891

López-Pérez AJ, Martínez JA (2015) In vitro root induction improvement by culture in darkness for different globe artichoke cultivars. In Vitro Cell Dev Biol Plant 51:160–165

Loudon PT, Nelson RS, Ingram DS (1989) Studies of protoplast culture and plant regeneration from

commercial and rapid-cycling *Brassica* species. Plant Cell Tissue Organ Cult 19(3):213–224

Martins M, Sarmento D, Oliveira MM (2004) Genetic stability of micropropagated almond plantlets, as assessed by RAPD and ISSR markers. Plant Cell Rep 23:492–496

Martin KP, Pachathundikandi SK, Zhang CL, Slater A, Madassery J (2006) RAPD analysis of a variant of banana (*Musa* sp.) cv. grande naine and its propagation via shoot tip culture. In Vitro Cell Dev Biol Plant 42(2):188–192

Matsumoto T, Sakai A, Yamada K (1995) Cryopreservation of in vitro-grown apical meristems of lily by vitrification. Plant Cell Tissue Organ Cult 41:237–241

Mbanaso ENA, Roscoe DH (1982) Alginate: an alternative to agar in plant protoplast culture. Plant Sci Lett 25(1):61–66

Meneghini M, Pandino G, Lombardo S, Mauromicale G, Tavazza R (2016) Exploration to use of globe artichoke cell cultures as bioactive compounds production. Acta Hortic 1147:413–420

Menin B, Moglia A, Comino C, Lanteri S, Van Herpen T, Beekwilder J (2012) In vitro callogenesis and agrobacterium-mediated transformation of globe artichoke. Acta Hortic 961:267–271

Menin B, Moglia A, Comino C, Hakkert JC, Lanteri S, Beekwilder J (2013) In vitro callus-induction in globe artichoke (*Cynara cardunculus* L. var. *scolymus*) as a system for the production of caffeoylquinic acids. J Hortic Sci Biotechnol 88(5):537–542

Mix-Wagner G, Conner AJ, Cross RJ (2000) Survival and recovery of asparagus shoot tips after cryopreservation using the droplet method. NZ J Crop Hortic Sci 28:283–287

Moglia A, Lanteri S, Comino C, Acquadro A, de Vos R, Beekwilder J (2008) Stress-induced biosynthesis of dicaffeoylquinic acids in globe artichoke. J Agric Food Chem 56:8641–8649

Moglia A, Menin B, Comino C, Lanteri S, Beekwilder J (2012) Globe artichoke callus as an alternative system for the production of dicaffeoylquinic acids. Acta Hortic 961:261–265

Morone Fortunato I, Ruta C (2003) Piantine micropropagate di carciofo precoce con inoculo di funghi micorrizici arbuscolari. Italus Hortus 10:213–218

Morone Fortunato I, Lattanzio V, Vanadia S (1979) Primi risultati sulla micropropagazione del carciofo. In: Atti 3rd Congr Int Stud Carciofo, Bari Ind Grafica Laterza, Bari, pp 247–254

Morone Fortunato I, Ruta C, Castrignanò A, Saccardo F (2005) The effect of mychorrizal symbiosis on the development of micropropagated artichoke. Sci Hortic 106:472–483

Morone Fortunato I, Barba M, Cardarelli M, Gallitelli D, Mauromicale G, Micozzi F, Ruta C, Tavazza R (2007) Micropropagazione, garanzia di sanità. L'Informatore Agrario 22:50–54

Morzadec JM, Hourmant A (1997) In vitro rooting improvement of globe artichoke (cv. Camius de Bretagne) by GA3. Sci Hortic 70:59–62

Murashige T (1974) Plant propagation through tissue culture. Ann Rev Plant Physiol 25:135–166

Murashige T, Skoog F (1962) A revised medium for rapid growth and bioassays with tabacco tissue culture. Physiol Plant 15:473–497

Murashige T, Jones JB, Serpa M (1974) Clonal propagation of gerbera through tissue culture. HortScience 9 (3):175–180

Nagata T, Takebe I (1971) Plating of isolated tobacco mesophyll protoplasts on agar medium. Planta 99:12–20

Niino T, Tahiro K, Suzuki M, Ohuchi S, Magoshi J, Akihaa T (1997) Cryopreservation of in vitro grown shoot tips of cherry and sweet cherry by one-step vitrification. Sci Hortic 70(2–3):155–163

Okasha KA, Ragab ME (1994) Studies on propagation of globe artichoke through tissue culture technique. HortScience 29(5):563

Ordas RJ, Tavazza R, Ancora G (1990) In vitro morphogenesis in the globe artichoke (*Cynara scolymus* L.). Plant Sci 77:253–259

Ordas RJ, Tavazza R, Ancora G (1991) Callus formation from isolated globe artichoke (*Cynara scolymus* L.) suspension protoplasts. Plant Sci 71(2):233–237

Pagnotta MA (2012) The European project: European genetic resources of *Cynara* spp. Acta Hortic (ISHS) 942:41–46

Palermo M, Colla G, Barbieri G, Fogliano V (2013) Polyphenol metabolite profile of artichoke is modulated by agronomical practices and cooking method. J Agric Food Chem 61:7960–7968

Pandino G, Meneghini M, Tavazza R, Lombardo S, Mauromicale G (2017a) Phytochemicals accumulation and antioxidant activity in callus and suspension cultures of *Cynara scolymus* L. Plant Cell Tissue Organs 128:223–230

Pandino G, Lombardo S, Lo Monaco A, Ruta C, Mauromicale G (2017b) vitro micropropagation and mycorrhizal treatment influences the polyphenols content profile of globe artichoke under field conditions. Food Res Int 99:385–392

Pasquini G, Barba M, Papanice M, Gallitelli D (2003) Produzione di germoplasma di carciofo virus esente. In: Atti "Giornate nazionali di studio sul carciofo: vivaismo e strategie di sviluppo" Samassi (CA), pp 60–63

Pecaut P, Dumas de Vaulx R, Lof H (1983) Virus-free clones of globe artichoke (*Cynara scolymus* L.) obtained after in vitro propagation. Acta Hortic 131:303–309

Pierik RLM, Binnendijk CM, Jansen JLM, Maasdam A (1975) Optimalization of gerbera plantlet production from excised capitulum explants. Sci Hortic 3:351–357

Rana GL, Martelli GP (1983) Virosi del carciofo. L'Italia agricola 1:27–38

Razdan MK (2003) Germplasm conservation. In: Introduction to plant tissue culture, 2nd edn. Science Publishers Inc

Reed BM (1995) Factors affecting the in vitro storage of strawberry. HortScience 30:871

Reed BM (2000) Cryopreservation of tropical plant germplasm. In: Hengelman F, Takagi H (eds). JIRCAS, Japan/IPGRI, Rome, Italy, pp 200–204

Reed BM (2008) Cryopreservation—practical consideration. In: Reed BM (ed) Plant cryopreservation: a practical guide. Springer, New York, pp 3–13

Reed BM, Change Y (1997) medium-and long-term storage of in vitro cultures of temperate and fruit crops. In: Razdan MK, Cocking EC (eds) Conservation of plant genetic resources in vitro, vol 1. General Aspects, Science Publishers Inc, Enfield, USA, pp 67–105

Repetto A, Cadinu M, Leoni S, Gallitelli D, Saldarelli P, Barbarossa L, Grieco F (1996) Effetto della coltura in vitro di apici meristematici sull'ottenimento di piante di carciofo "Spinoso sardo" e "Masedu" virus esenti. Notiziario sulla Protezione delle Piante 7:189–195

Rey NA, Tavazza R, Papacchioli V, Pagnotta MA (2013a) Genetic fidelity of micropropagated globe artichoke plantlets. Acta Hortic (ISHS) 983:225–229

Rey NA, Papacchioli V, Tavazza R, Pagnotta MA (2013b) Gauging the genetic changes occurring across globe artichoke micropropagation towards an appropriate variety registration and nursery production. Sci Hortic 156:121–126

Roca WM, Chavez R, Martin ML, Arias DI, Mafla G, Reyes R (1989) In vitro methods of grem-plasm conservation. Genome 31:813–817

Roest S, Bokelmann GS (1975) Vegetative propagation of *Chrysanthemum cineraraefolium* Ram. In Vitro Sci Hortic 3:317–330

Roest S, Gilissen LJW (1989) Plant regeneration from protoplasts: a literature review. Act Bot Neerl 38(1): 1–23

Rossi V, De Paoli G (1990) Micropropagazione per un carciofo di qualità. Informatore Agrario 9:201–209

Rout GR (2004) Effect of cytokinins and auxin on micropropagation of *Clitoria ternatea* L. Biol Lett 41(1):21–26

Ruta C, Morone Fortunato I, Tagarelli A (2005) Mycorrhization on micropropagated artichoke. Acta Hortic 681:407–410

Ruta C, Tagarelli A, Morone Fortunato I (2013) Callogenesis capability of artichoke (*Cynara cardunculus* var. *scolymus* L. Fiori). Acta Hortic 983:377–380

Saccardo F, Ancora G (1984) Contributo della micropropagazione al miglioramento della coltura del carciofo. L'Informatore Agrario, pp 24–26

Saccardo F, Micozzi F, Di Lernia G, Piccioni C, Barba M, Pagnotta MA (2007) Virus free artichoke germoplasm: quali-quantitative response of globe artichoke. Acta Hortic 730:375–379

Sakai A, Kobayashi S, Oiyama I (1990) Cryopreservation of nucellar cells of navel orange (*Citrus sinensis* var. *brasiliensis* Tanaka) by vitrification. Plant Cell Rep 9:30–33

Scaramella Petri P, Ricci A (1979) Micropropagation di *Cynara scolymus* L. var. sativa, da tessuti in vitro. In: Atti 3rd Congr Int Stud Carciofo, Bari Ind Grafica Laterza, Bari, pp 231–238

Skirvin RM, McPheeters KD, Norton M (1994) Sources and frequency of somaclonal variation. HortScience 29(11):1232–1237

Taglienti A, Barba M (2012) Preliminary experiments for a cryopreservation protocol in artichoke. Acta Italus Hortus 6:253–256

Taglienti A, Tiberini A, Barba M (2013) Cryotherapy: a new tool for the elimination of artichoke virus. J Plant Pathol 95(3):597–602

Tavazza R, Papacchioli V, Ancora G (2004) An improved medium for in vitro propagation of globe artichoke (*Cynaras colymus* L.) cv. *Spinoso sardo*. Acta Hortic 660:91–97

Tavazza R, Crinò P, Ancora G, Pagnotta MA (2009) Biotecnologie. Il carciofo e il cardo. Coltura & cultura. Bayer Crop Science, pp 304–313

Tavazza R, Lucioli A, Benelli C, Giorgi D, D'aloisio E, Papacchioli V (2013a) Cryopreservation in artichoke: towards a phytosanitary qualified germplasm collection. Ann Appl Biol 163:231–241

Tavazza R, Crinò P, Ciancolini A, Papacchioli V, Pagnotta MA, Rey NA, Mariotti R, Saccardo F (2013b) Salvaguardia e conservazione di risorse genetiche di carciofo primaverile. Acta Italus Hortus 8:15–23

Tavazza R, Rey NA, Papacchioli V, Pagnotta MA (2015) A validated slow-growth invitro conservation protocol for globe artichoke germplasm: a cost-effective tool to preserve from wild to elite genotypes. Sci Hortic 197:135–143

Thompson MR, Douglas TJ, Obata-Sasamoto H, Thorpe TA (1986) Mannitol metabolism in cultured plant cells. Physiol Plant 67:365–369

Trajtemberg SP, Apòstolo NM, Fernandez G (2006) Calluses of *Cynara cardunculus* var. *Cardunculus* cardoon (*Asteraceae*): determination of cynarine and chlorogenic acid by automated high-performance capillary electrophoresis. In Vitro Cell Dev Biol Plant 42:534–537

Trionfetti NP, Pagnotta MA, Crinò P, Cuozzolo L, Tavazza R, Ancora G (2007) Recovery and characterization of Italian artichoke traditional landraces of "Romanesco" type. Acta Hortic ISHS 730:101–106

Vasil V, Vasil IK (1987) Formation of callus and somatic embryos from protoplasts of a commercial hybrid of maize (*Zea mays* L.). Theoret Appl Genet 73:793–798

Volk GM, Maness N, Rotindo K (2004) Cryopreservation of garlic (*Allium sativum* L.) using plant vitrification solution 2. Cryo Lett 25:219–226

Wang PJ, Hu CY (1985) Potato tissue culture and its application in agriculture. In: Li PH (ed) Potato physiology. Academic Press, London, New York, pp 503–577

Wang QC, Perl A (2006) Cryopreservation in floricultural crops. In: da Silva JT (ed) Floricultural, ornamental and plant biotechnology: advances and topics. Global Science Books, London, UK, pp 523–539

Wang QC, Valkonen JPT (2009) Cryotherapy of shoot tips: novel pathogen eradication method. Trends Plant Sci 14:119–122

Wang M, Simon J, Aviles I, He K, Zheng Q, Tadmor Y (2003) Analysis of antioxidative phenolic compounds in artichoke (*Cynara scolymus* L.). J Agric Food Chem 51:601–608

Wang QC, Panis B, Engelmann F, Lambardi M, Valkonen JPT (2009) Cryotherapy of shoot tips: a technique for pathogen eradication to produce healthy planting materials and prepare healthy plant genetics resources for cryopreservation. Ann Appl Biol 154:351–363

Williams JT (1984). A decade of crop genetic resources research. In: Holden JHW, Williams JT (eds) Crop genetic resources: conservation and evaluation. George Allen & Unwin, London, pp 1–17

Willison JHM, Cocking EC (1972) The production of microfibrils at the surface of isolated tomato fruit protoplasts. Protoplasma 75:397–403

Withers LA (1989) In-vitro conservation and germplasm utilization. In: Brown ADH, Marshall DR, Frankel OH, Williams JT (eds) The use of plant genetic resources. Cambridge University Press, Cambridge, pp 309–334

Cynara cardunculus as a Multiuse Crop

4

Giovanni Mauromicale, Gaetano Roberto Pesce, Maria Dolores Curt, Jesús Fernández, Javier González, Jorge Gominho, Rafael Tabla, Isidro Roa and Ezio Portis

Abstract

The main use of *Cynara cardunculus* L. is food, but there are also alternative uses, of which this chapter gives an overview. All the possible alternative uses of *C. cardunculus* have emerged only in recent decades, with the exception of the florets as vegetable rennet, which has a very ancient tradition behind it. The need for renewable energy sources has led to studying *C. cardunculus* as an energy plant, from which biomass for direct combustion, biomethane, bio-ethanol and oil can be obtained. The latter is also suitable for human consumption. The aboveground biomass could also provide fiber for paper pulp production, as well as various livestock feed resources. The chapter, lastly, includes the above-mentioned use as vegetable rennet and the possible use as an ornamental plant.

G. Mauromicale · G. R. Pesce
Di3A (Dipartimento di Agricoltura, Alimentazione e Ambiente), University of Catania, via Valdisavoia 5, 95123 Catania, Italy

M. D. Curt · J. Fernández · J. González
Departamento de Producción Agraria, Universidad Politécnica de Madrid, Campus Ciudad Universitaria, Avda. Puerta de Hierro 2-4, 28040 Madrid, Spain

J. Gominho
Centro de Estudos Florestais, Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017 Lisbon, Portugal

R. Tabla · I. Roa
Dairy Department, Technological Institute of Food and Agriculture—Scientific and Technological Research Centre of Extremadura (INTAEX—CICYTEX), Junta de Extremadura, Avda. Adolfo Suarez s/n, 06071 Badajoz, Spain

E. Portis (✉)
DISAFA Plant Genetics and Breeding, University of Torino, Largo P. Braccini 2, 10095 Grugliasco, Torino, Italy
e-mail: ezio.portis@unito.it

4.1 Biomass for Energy

Increased use of renewable energy may offer significant opportunities for Europe to reduce greenhouse gas emissions and secure its energy supplies. In the long term, bio-energy crops could play a key role in achieving these objectives and, among them, perennial crops could provide the greatest potential (Bentsen and Felby 2012). This development will be driven by further productivity increases of agri-energy systems and the introduction of energy crops with higher yields (Mauro et al. 2015). The crops intended for the production of bio-energy differ from conventional food and fodder crops, as they are optimized to have high-energy efficiency together with a low environmental impact (Wiesenthal and Mourelatou 2006). The adaptability of wild and cultivated cardoons to the climatic conditions of Southern Europe is the main reason for their good performance in marginal lands and in dry conditions, namely in

E. Portis et al. (eds.), *The Globe Artichoke Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-030-20012-1_4

environments that are less suitable for traditional cultivations. Thanks to the autumn–winter–spring cycle, plants can exploit the autumn–winter rainfall available in the Mediterranean environment (Ierna et al. 2012). Moreover, the oldest roots, which have a reserve function, support the sprouting of the dormant buds after the summer dormancy. Furthermore, there is a side effect in the cultivation of *C. cardunculus* as an energy crop, i.e., its positive effect on soil fertility, especially in marginal lands. *C. cardunculus* has been shown to significantly increase soil organic matter, total nitrogen, assimilable P_2O_5 and exchangeable K (Mauromicale et al. 2014). The increase in soil organic matter may be due to the dry leaves detachment during late harvest in both cultivated and wild cardoons (Ierna et al. 2012). Additionally, the residues of the deep and well-developed roots increase the organic matter in the soil. The main roots are able to develop at depth and to produce profuse branches, while secondary roots, originated by annually formed stems, run horizontally at different soil depths, giving rise to a tangled root mass (Fernández and Curt 2005). In addition, plants at later stages of development would have a greater predisposition for the translocation of N to belowground parts (Fike et al. 2006). Deep roots are able to extract (down to 40 cm of depth and beyond) deep nutrient reserves (N, P, K) and return them to the top 30 cm of the soil as root residue biomass or use them for aboveground biomass production. The long time (9–10 months per year) of *C. cardunculus* seasonal growth cycle, with extensive living of aboveground vegetation and root systems, is able to reduce losses of dissolved phosphorus, nitrogen, potassium and other nutrients by soil erosion and by leaching with drainage water.

4.1.1 Aboveground Biomass Yield

Being a perennial species, *C. cardunculus* is suitable for growing in a short cycle (up to 3 years) or in a multiyear long cycle (over 10 years). In the first case, the crop can be included in a rotation plan with grasses or legumes; in the second, the crop may be appropriately grown in abandoned marginal lands, with benefits to both the soil and environment. Ierna and Mauromicale (2010) evaluated six genotypes of *C. cardunculus*, including one cultivated cardoon cultivar, one globe artichoke line, one wild cardoon ecotype and three F1 progenies: 'globe artichoke × wild cardoon', 'globe artichoke × cultivated cardoon' and 'cultivated cardoon × wild cardoon'. Their experiment was carried out under low-input conditions. Results ranged from ~12.0 t ha^{-1} $year^{-1}$ of globe artichoke to ~25.0 t ha^{-1} $year^{-1}$ of F1 globe artichoke × wild cardoon, with cultivated cardoon slightly less at ~21.5 t ha^{-1} $year^{-1}$ (Table 4.1). The productivity of genotypes has been further studied also in relation to different levels of mineral fertilization (Ierna et al. 2012). Interesting here is that different genotypes responded differently to mineral fertilization, but botanical varieties and, within these, genotypes proved to be factors determining notable variations in yield per hectare (Table 4.1). Weather conditions, in particular rainfall, can affect biomass production. This was clearly shown by Fernández et al. (2005) in their 9-year experiment in which biomass productivity of cultivated cardoon, in rain-fed conditions, was extremely variable from one year to the next according to rainfall (Table 4.1). Again, regarding the relationship between precipitation and biomass production, Pesce et al. (2017), in their 3-year experiment on three genotypes of cardoon under minimal input regime, found that, contrary to expectations, the production of biomass from cardoon in the third year was significantly lower than that in the second year, precisely because of the scarce rainfall (Table 4.1). Like all perennial biomass crops, cardoon has low yields in the first year, which can be considered a crop stabilization stage (about 60% of maximum yield), and reaches the maximum values at the 5th or 6th year of cultivation. Angelini et al. (2009) carried out a long-term (11 years) trial on two genotypes of cultivated cardoon to which, every year, mineral fertilizers (N, P_2O_5 and K_2O) were applied, as well as hoeing to keep the soil

Table 4.1 Aboveground biomass yield of *C. cardunculus* as affected by genotype and crop age

Genotype	Crop age											Sparkline charts	References
	1st year	2nd year	3rd year	4th year	5th year	6th year	7th year	8th year	9th year	10th year	11th year		
	Biomass yield (t DM ha^{-1} $year^{-1}$)												
3/10 V.S. (GA)	15.4[a]	13.3[a]	12.2[a]										Foti et al . (1999)
374 F1 (GA)	12.8[a]	9.1[a]	7.7[a]										
Gigante di Lucca (CC)	25.4[a]	35.8[a]	30.1[a]										
Bianco Avorio (CC)	28.4[a]	36.2[a]	27.2[a]										
Landrace of wild cardoon	19.3[a]	22.0[a]	15.1[a]										
Cultivated cardoon	17.0	13.0	11.3	3.6	19.4	17.8	25.3	4.2	14.0				Fernandez et al . (2005)
Cultivated cardoon	19.5	21.3	23.6										Ierna and Mauromicale (2010)
Globe artichoke	13.2	11.1	12.8										
Wild cardoon	7.4	17.9	16.6										
Cultivated cardoon × wild cardoon	17.3	21.1	17.2										
Globe artichoke × cultivated cardoon	11.8	19.4	27.7										
Globe artichoke × wild cardoon	27.0	27.6	20.5										
Bianco Avorio (CC)	22.7	31	34.2										Ierna et al . (2012)
Gigante di Lucca (CC)	20.7	24.8	29.3										
Cardo di Nizza (CC)	19.9	25.5	27.3										
Wild cardoon (CC)	7.8	19.4	15.8										
Line 307 (GA)	5.8	8.9	7.6										
Cultivated cardoon	5.2[b]	18.0[b]	18.1[b]	22.0[b]	18.5[b]	12.9[b]	8.0[b]	10.2[b]	8.4[b]	8.2[b]	9.8[b]		Angelini et al . (2009)
Bianco Avorio (CC)	9.7	16.0	15.5	16.4	22.0	12.1	10.6						Mauromicale et al. (2014)
Wild cardoon	5.5	6.1	7.0	8.0	8.8	8.3	8.1						
Altilis 41 (CC)	11.9	26.1	19.4										Pesce et al . (2017)
Bianco Avorio (CC)	11.8	21.3	17.2										
Sylvestris Marsala (WC)	7.0	15.4	13.1										

GA = Globe artichoke; CC = Cultivated cardoon; WC = Wild cardoon

[a] Value not including achenes' weight

[b] Average value between two cultivated cardoon genotypes

weed-free. The crop yield for each cultivar was very poor in the first year (mean 5.2 t ha^{-1}), but from the first to the second year the production of dry biomass increased considerably, remaining high until the fifth year (on average ~19 t ha^{-1} $year^{-1}$). Thereafter—from the sixth year onward—it decreased (Table 4.1). Unlike in cardoon, the production of biomass in globe artichoke seems to be greater in the first year (Foti et al. 1999; Ierna and Mauromicale 2010). Another long-term experiment (7 years) was conducted in Sicily (Mauromicale et al. 2014), but in this case under minimal input and also taking into account the wild cardoon. In this trial, results are in line to those found by Angelini et al. (2009). Under minimal input, wild cardoon, although less productive than the cultivated one, showed greater stability in annual biomass production (Table 4.1). The most important feature of a biomass plant should be the maximization of the biomass yield, in order to minimize the area for the production of energy crops (de Wit et al. 2011). In this direction, Portis et al. (2018) made efforts to identify the genetic basis of biomass accumulation in *C. cardunculus*. However, if on the one hand the lower productivity of the wild cardoon is a drawback, on the other, its constancy is an advantage not to be underestimated. Indeed, a more stable biomass production over time allows a more reliable quantification of the land for the cultivation of energy plants (Bentsen and Felby 2012), as well as the possibility of ensuring a programmable biomass production, in order to reliably supply energy production plants.

4.1.2 Biomass for Direct Combustion: Heating Power

Some scientific papers have explored the possibility of using the biomass (leaves, stems, flower heads and achenes) of *C. cardunculus* as solid biofuel, through direct combustion or pyrolysis (Damartzis et al. 2011; Ierna et al. 2012; Karampinis et al. 2012). This species is suitable for producing solid biofuel, because of its high biomass production with a relatively small demand for energy inputs, the low moisture content of its biomass at harvest, the lignocellulosic nature of its biomass and a high heating value. As a solid fuel, *C. cardunculus* can be used for heating applications or power generation (e.g., cofiring blended with coal) (Fernández et al. 2006; Ierna and Mauromicale 2010). Wild cardoon and cultivated cardoon are more suitable for this purpose than globe artichoke, for their greater productivity and hardiness. Combustion is the complete oxidation of a substance, burning in the presence of oxygen, which acts as a comburent. Combustion is ultimately an exothermic reaction, which causes the release of the fuel's chemical energy as heat. The most common technologies for the generation of heat through solid fuels are stoves, thermofireplaces and boilers. Foti et al. (1999) found that there is no significant difference in the higher heating value (HHV) belonging to the different botanical varieties of *C. cardunculus;* indeed, they ranged from 16 to 17 MJ kg^{-1} of dry matter (these values do not take the grain into account) (Table 4.2). Dahl and Obernberger (2004) found that the HHV of domestic cardoon pellet is 20.3 MJ kg^{-1} (Table 4.2). Fernández et al. (2006) found in their experiments that HHV of domestic cardoon amounts to ~17.0 MJ kg^{-1}, while its lower calorific value (LHV) was equal to ~15.9 MJ kg^{-1} (Table 4.2). Grammelis et al. (2008) studied *C. cardunculus* as a plant for the production of solid biofuel in the Mediterranean environment. They observed that the domestic cardoon biomass has a HHV of 13.7 MJ kg^{-1}, without the achenes. This value increases up to 16.3 MJ kg^{-1} if the achenes are also taken into account (Table 4.2). Angelini et al. (2009) measured the HHV of the biomass of two varieties of domestic cardoon. The value per unit of mass reported by these authors amounts to 14.9 MJ kg^{-1}, a value very close to that (14.8 MJ kg^{-1}) found by González et al. (2004), who studied cardoon pellets (Table 4.2). As shown by Mauromicale

Table 4.2 Higher heating value (HHV) and lower heating value (LHV) of *C. cardunculus* biomass in relation to botanical variety

Var. *scolymus*		Var. *altilis*		Var. *sylvestris*		References
HHV (MJ kg^{-1})	LHV (MJ kg^{-1})	HHV (MJ kg^{-1})	LHV (MJ kg^{-1})	HHV (MJ kg^{-1})	LHV (MJ kg^{-1})	
17.0[a]		16.3[a]		16.0[a]		Foti et al. (1999)
		20.3[b]				Dahl and Obernberger (2004)
		15.0[b]				González et al. (2004)
		17.0	15.9			Fernández et al. (2006)
		13.7[a]				Grammelis et al. (2008)
		16.3				
		14.9				Angelini et al. (2009)
		18.2		18.0		Mauromicale et al. (2014)
			From 17.4 to 19.8		From 17.7 to 19.0	Mauro et al. (2015)

[a]Not including achenes
[b]Pelletized biomass

et al. (2014), there are no significant differences in HHV among cultivated and wild cardoon (18.2 and 18.0 MJ kg^{-1} DM, respectively) (Table 4.2). Mauro et al. (2015) found that LHV of biomass including achenes was very similar in cultivated and wild cardoon, but they showed a different pattern of variability during their seven-year trial. In cultivated cardoon, LHV increased from the first year (17.4 MJ kg^{-1} DM) to the fourth (19.8 MJ kg^{-1} DM) and then showed a decrease from the fifth onward (up to 17.4 MJ kg^{-1} DM). On the contrary, wild cardoon reported a more stable LHV throughout the entire experimental period, which ranged from 17.7 MJ kg^{-1} in the third year to 19.0 MJ kg^{-1} DM in the seventh (Table 4.2). The biomass of *C. cardunculus* is also characterized by high ash content (13.9%). Angelini et al. (2009) found that the cultivated cardoon ashes contain SiO_2 (14.5%), CaO (17.7%), K_2O (21.5%) and Na_2O (10.0%). The alkali content is directly related to the presence of alkali salts in the biomass chemical composition, and they are unfavorable. Among the alkali, potassium stands out. It is an essential element for plant growth and is found in high percentages in fast-growing crops. During gasification, which is one of the phases of combustion, the alkali salts vaporize, forming products such as KOH and KCl. Subsequently, when the gas cools, these compounds condense in the form of very small solid particles that must be removed as their deposition can cause corrosion of the metal surfaces and a decrease in plant performance. This suggests its use in mixtures with lignocellulosic biomasses of different origin. The way the culture is managed can affect the quality characteristics of the biomass of *C. cardunculus*. For instance, if during the harvesting operations the soil excessively contaminates the biomass, this has the same effect as large ash content, with the known consequence of slagging problems. In addition, an excessive fertilization with KCl can lead to deposit formation in the boiler (Dahl and Obernberger 2004). Cardoon biomass has a high N content (1.5–2%) (González et al. 2004; Angelini et al. 2009), which however is not a good feature, because fuel-bound nitrogen is responsible for most nitrogen oxide (NO_x) emissions produced from biomass combustion. Cardoon biomass has a volatile matter content of about 73% DM that can be considered a desirable value for a good regulation of combustion or gasification processes in large-scale plants (Angelini et al. 2009).

4.1.3 Energy Yield Per Hectare

If, on one hand, the difference in heating values between the botanical varieties is small, on the other, it appears more evident when considering the quantity of biomass produced per hectare. Foti et al. (1999) found that these values range from 195.5 GJ ha^{-1} recorded for globe artichoke, 311 GJ ha^{-1} for the wild cardoon, to 505 GJ ha^{-1} of the cultivated cardoon (Table 4.3). Angelini et al. (2009) showed that the energy yield of cultivated cardoon per hectare amounted to ~193 GJ ha^{-1} $year^{-1}$ (Table 4.3). This value derives from the average between two cultivated cardoon genotypes, one more and the other less productive, as mentioned before. It is also the average over an 11-year period, which means it takes account of the low productivity of the last years. In Ierna and Mauromicale (2010), energy yield ranged between 204 GJ ha^{-1} reached by globe artichoke and 413 GJ ha^{-1} of F1 hybrid globe artichoke × wild cardoon (Table 4.3). Ierna et al. (2012) continued the work of investigation on *C. cardunculus* as an energy crop, comparing not only different genotypes, but also the response to different levels of fertilization. Three commercial varieties of cardoon ('Gigante di Lucca', 'Bianco Avorio' and 'Cardo di Nizza'), an ecotype of wild cardoon taken from Sicilian coastal areas and a synthetic globe artichoke cultivar ('Line 307') developed from 'Spinoso di Palermo' were compared. Three fertilization levels were also compared: low (50–50–50 kg ha^{-1} of N–P_2O_5–K_2O, respectively), medium

Table 4.3 Energy yield per hectare of *C. cardunculus* biomass in relation to genotype

Genotype	Total biomass energy yield (GJ ha^{-1} $year^{-1}$)	References
3/10 V.S. (GA)	228	Foti et al. (1999)
374 F1 (GA)	163	
Gigante di Lucca (CC)	505	
Bianco Avorio (CC)	505	
Landrace of wild cardoon	311	
Cultivated cardoon	354	Ierna and Mauromicale (2010)
Globe artichoke	204	
Wild cardoon	230	
Cultivated cardoon × wild cardoon	306	
Globe artichoke × cultivated cardoon	323	
Globe artichoke × wild cardoon	413	
Bianco Avorio (CC)	501[a]–554[b]–480[c]	Ierna et al. (2012)
Gigante di Lucca (CC)	400[a]–466[b]–420[c]	
Cardo di Nizza (CC)	420[a]–400[b]–458[c]	
Wild cardoon (CC)	260[a]–220[b]–283[c]	
Line 307 (GA)	140[a]–138[b]–141[c]	
Cultivated cardoon	193	Angelini et al. (2009)
Bianco Avorio (CC)	275	Mauromicale et al. (2014)
Wild cardoon	138	

GA Globe artichoke; *CC* Cultivated cardoon; *WC* Wild cardoon
[a]Low-level fertilization (50–50–50 kg ha^{-1} of N–P_2O_5–K_2O)
[b]Medium-level fertilization (100–150–150 kg ha^{-1} of N–P_2O_5–K_2O)
[c]High-level fertilization (200–300–200 kg ha^{-1} of N–P_2O_5–K_2O)

(100–150–150 kg ha^{-1} of N–P_2O_5–K_2O) and high (200–300–200 kg ha^{-1} of N–P_2O_5–K_2O). The highest energy yield was obtained by 'Bianco Avorio', the lowest by 'Line 307'. The performance of the tested genotypes proved to be fertilization level-dependent. Indeed, in 'Bianco Avorio' and 'Gigante di Lucca', energy yield peaked at medium fertilization level (554 and 466 GJ ha^{-1}, respectively), while in 'Cardo di Nizza' and wild cardoon energy yield peaked at high levels (458 and 283 GJ ha^{-1}, respectively), 'Line 307' being unresponsive to the fertilization level (Table 4.3). In the Mauromicale et al. (2014) experiment, the cultivated cardoon accession doubled the energy yield of wild cardoon over the whole seven-year cropping period (275 vs. 138 GJ ha^{-1} $year^{-1}$).

4.2 Biomethane Production

Anaerobic digestion (AD) is a technology to obtain energy from agricultural biomass; its objective is to produce biomethane, but it can also produce positive secondary effects. The agricultural sector, indeed, is the largest source of non-CO_2 emissions, like methane (CH_4) and nitrous oxide (N_2O) (Environmental Protection Agency 2012), and AD can reduce these emissions, for example from slurry storage. It can also lead to reduced impacts of pollution by waste disposal. Furthermore, the process upgrades the waste into a product and makes it a valuable organic fertilizer, thus closing the cycle from soil to crop, to product, to waste and back to the soil

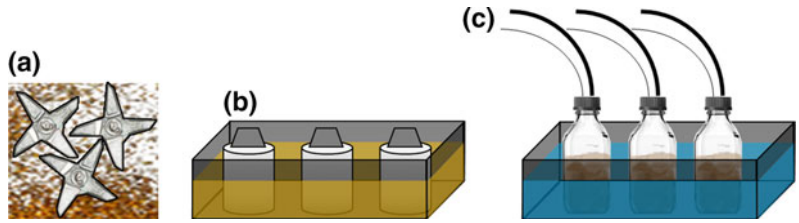


Fig. 4.1 Scheme of biomass treatment and anaerobic digestion adapted from Oliveira et al. (2012). Trial I: **a** grinding biomass. Size particles ranging from 60 to 40 mesh; **b** thermal pretreatment carried out with distilled water using a fixed temperature of 160 °C, a total solids content (TS) of 12.5% and a reaction time of 30 min; **c** putting pretreated biomass samples together with the inoculum for 33 days. Trial II: **a** grinding biomass. Size particles 10 mm; **b** thermal pretreatment carried out with NaOH (1.4%, w/v) solution using a fixed temperature of 160 °C, a total solids content (TS) of 12.5% and a reaction time of 30 min; **c** putting pretreated biomass samples together with the inoculum for 33 days

(Wellinger et al. 2013). In AD, several groups of bacteria work together in converting organic material into biogas. They feed on different substrates and give rise to different types of products, going through different phases.

The high polysaccharide and low lignin content of cardoon biomass provides opportunities for biomethane production via AD; this is currently achieved largely from maize silage, although bread wheat and triticale biomass is also exploited on a smaller scale (Berglund and Börjesson 2006; Dressler et al. 2012). The use of crop residues or biomass obtained from a crop such as cardoon, which has a minimal environmental impact, improves environmental performances of energy production through anaerobic digestion (Duca et al. 2015).

In many studies, the specific yield of the mass unit was measured and some authors pointed out that biomass pretreatment can affect it. In fact, Oliveira et al. (2012) have measured the yield in biomethane of the stems of the cultivated cardoon in the context of two trials in which two different pretreatments were compared (Fig. 4.1). In one, fine grinding (40–60 mesh) was combined with a thermal treatment in distilled water to give rise to self-hydrolysis. In the other, a coarser grinding (10 mm) of the biomass was combined with a thermochemical treatment, with the use of a NaOH solution (1.4%, w/v) instead of distilled water. In both trials, pretreatments were carried out inside stainless steel reactors tightly closed and placed together in a rotating structure within an oil heated bath (160 °C) (Fig. 4.1). The resulting pretreated material was mixed with the inoculum and fed directly into the reactors under a mesophilic temperature of 37 °C. The results of this experiment have shown that the reduction in the biomass particle size and thermal pretreatment with NaOH have a positive effect on the methane yield (Table 4.4). A full-scale experiment was also carried out by Ros et al. (2013), using an anaerobic continuously stirred tank reactor (8–10 rpm) with 300 L working volume capacity. It was at first filled with urban sewage sludge (215 L), while a daily amount (20 L) of fruit and vegetable processing sludge was added in order to reach steady-state conditions in the reactor. When the latter were achieved (after 55 days), a codigestion of the fruit and vegetable processing sludge with chopped fresh globe artichoke waste (60/40) was started (Fig. 4.2). Thanks to the addition of globe artichoke biomass, an increase in the average daily biogas production was achieved (from 244 to 354 L kg^{-1} dry DM day^{-1}) together with an improvement in the quality of the latter, because the percentage of methane increased from 40 to more than 70%. These results give us the clear idea that globe artichoke biomass can be used in biogas production. Other authors have followed the path of codigestion (Kalamaras and Kotsopoulos 2014), blending the silage biomass of the domestic cardoon along with the cattle manure (Fig. 4.3). Once again, the result was encouraging: cardoon silage in codigestion induced a specific

Table 4.4 Specific methane yield and methane yield per hectare obtained from *C. cardunculus* in relation to treatment and genotype

Treatment/genotype	Specific methane yield	Methane yield per hectare	References
40–50 Mesh grinding + thermal pretreatment	620 L kg^{-1} VS	3560 m^3	Oliveira et al. (2012)[a]
40–50 Mesh grinding untreated	500 L kg^{-1} VS	2871 m^3	
10-mm Grinding + thermochemical pretreatment	590 L kg^{-1} VS	3388 m^3	
10-mm Grinding untreated	310 L kg^{-1} VS	1769 m^3	
Cattle manure + ensiled biomass (ratio 60:40)	308 L kg^{-1} VS	–	Kalamaras and Kotsopoulos (2014)
Inoculum/substrate ratio 0.5	273 L kg^{-1} VS		Fabbri et al. (2014)
Inoculum/substrate ratio 1.54	294 L kg^{-1} VS	–	
Inoculum/substrate ratio 2	409 L kg^{-1} VS	–	
'Madrigal' (GA)	293 L kg^{-1} DOM	7656 N m^3	De Menna et al. (2016)
'Spinoso Sardo' (GA)	305 L kg^{-1} DOM	1648 N m^3	
'C3' (GA)	302 L kg^{-1} DOM	2960 N m^3	
'Tema' (GA)	305 L kg^{-1} DOM	4367 N m^3	
'Violetto' (GA)	263 L kg^{-1} DOM	1289 N m^3	
'Altilis 41' (CC)	205 L kg^{-1} DM	4074 N m^3	Pesce et al. (2017)
'Bianco Avorio' (CC)	246 L kg^{-1} DM	4162 N m^3	
'Sylvestris Marsala' (WC)	249 L kg^{-1} DM	2867 N m^3	

GA Globe artichoke; *CC* Cultivated cardoon; *WC* Wild cardoon
[a]These authors evaluated only the stems

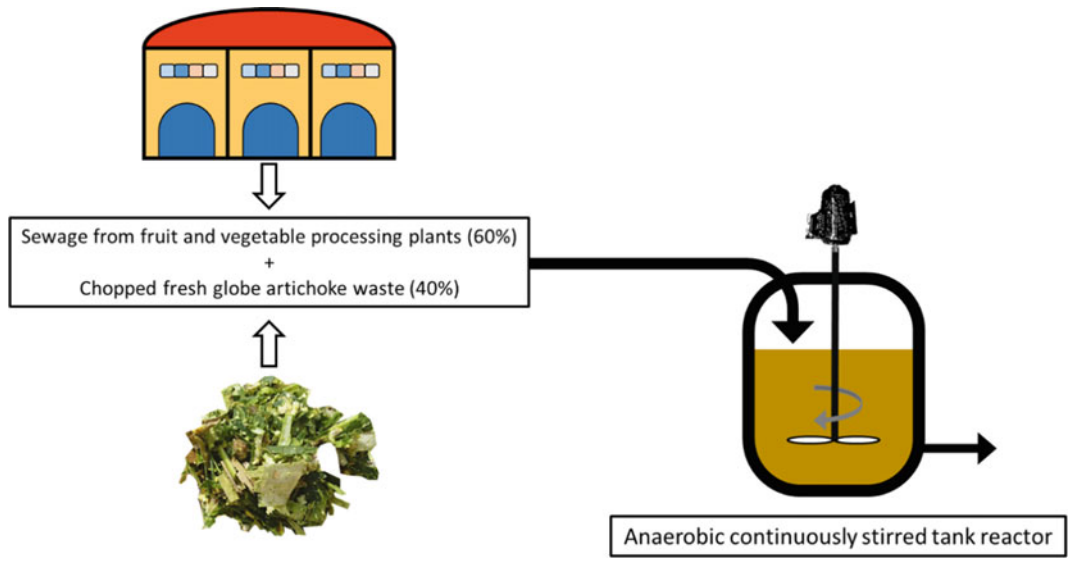


Fig. 4.2 Scheme of biomass treatment and anaerobic digestion adapted from Ros et al. (2013)

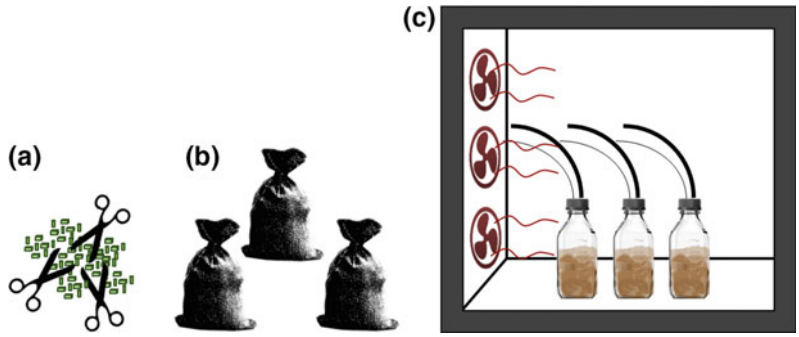


Fig. 4.3 Scheme of biomass treatment and anaerobic digestion adapted from Kalamaras and Kotsopoulos (2014). **a** Cutting biomass by scissors. Size particles of less than 2 cm; **b** packing biomass into anaerobic bags for 20 days; **c** putting substrate (biomass + manure) samples together with the inoculum in an oven at 37 °C for 45 days

production of biomethane even higher than that of maize silage (308 vs. 267 L kg^{-1} of volatile solids, Table 4.4). Still on the subject of industrial by-products of vegetable processing, interesting considerations on the effect of the inoculum/substrate ratio come from the work of Fabbri et al. (2014). They studied globe artichoke waste resulting from industrial transformation, comparing three different levels of inoculum/substrate ratio (0.5:1, 1.54:1 and 2:1), on the basis of volatile solids. Results showed that biogas and methane production increase as the inoculum/substrate ratio increases (Table 4.4).

Globe artichoke cultivation is characterized by a large amount of field residues, comprised between 80 and 85% of the aboveground biomass. Most of this aboveground biomass is either burned or buried, as it is not considered valuable. In order to enhance this large mass of field residues for energy purposes, De Menna et al. (2016) did some tests to evaluate the potential specific biomethane productivity of field residues coming from some varieties grown in Sardinia (Italy) and their productivity per hectare. They also evaluated the effect of an enzymatic pretreatment with cellulase, xylanase, beta-glucanase and other carbohydrase activities originating from a selected strain of *Trichoderma reesei*. If on the one hand the enzymatic pretreatment did not bring about significant differences in the specific biomethane production, on the other, the cultivar factor had a marked effect. 'Madrigal' and 'Tema' proved more suitable than 'Spinoso sardo' or 'Violetto', especially because of their greater yield per ha (Table 4.4). Fresh residual biomass could be harvested and utilized in biogas plants, allowing the integration of food and fuel production and the optimization of land use.

The potential of the cardoon as a biomethane crop has also been investigated by assessing the outcome of anaerobically digesting two cultivated forms and one wild one (Pesce et al. 2017). Fermentation laboratory tests were performed on ensiled biomass, as in the real-scale conditions. Fresh material was first chopped at a size ranging between 5 and 15 mm and then packed into a sealed black polythene bag. The samples of silage were fermented without any pretreatment or additives. Unstirred laboratory-scale fermenters were kept at 40 °C within a water bath. In each fermenter, the inoculum/substrate ratio was 2:1 on a DOM basis (Fig. 4.4). The yield of biomethane ranged from ~200 ('Altilis 41') to ~245 ('Bianco Avorio' and wild cardoon) N m^3 per t DM (Table 4.4). But when calculated on a per hectare basis, thanks to its higher per unit area biomass production, the biomethane yield of 'Altilis 41' equaled that of 'Bianco Avorio' (respectively, 4074 and 4162 N m^3, Table 4.4); the wild cardoon accession was less productive (2867 N m^3, Table 4.4). This data, obtained under very low crop input conditions, is sufficiently high to justify considering *C. cardunculus* a promising candidate bio-energy crop in the Mediterranean environment.

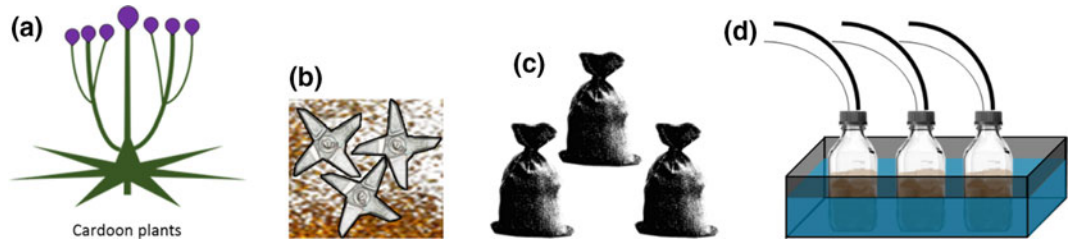


Fig. 4.4 Scheme of biomass treatment and anaerobic digestion adapted from Pesce et al. (2017). **a** Collection of cardoon biomass from a 3-year experimental field with three genotypes; **b** grinding biomass. Size particles ranging from 5 to 15 mm; **c** packing biomass into sealed black polythene bags for 90 days; **d** putting biomass samples together with the inoculum at 40 °C for 27 days

4.3 Bio-ethanol Production

Bio-ethanol has great potential as an alternative fuel to diesel, but is currently more expensive than non-renewable energy resources. Despite this, in recent decades the production of bio-ethanol has significantly increased. Brazil and the USA are the main bio-ethanol producing countries using sugarcane and corn, respectively, as raw material (Pirzadah et al. 2014). However, the bio-ethanol production could create conflicts between human food use and industrial use of these crops, especially in developing countries. Lignocellulosic biomass is a promising alternative to avoid the existing competition of food versus fuel caused by grain-based bio-ethanol production, and the biomass of *C. cardunculus* perfectly meets this need.

Ethanol production is normally performed in three steps: (1) obtaining a fermentable sugar solution, (2) fermentation of these sugars into ethanol and (3) separation and purification of the ethanol, usually by distillation–rectification–dehydration (Mussatto et al. 2010). The lignocellulosic biomass consists essentially of cell walls, which are made up of cellulose and hemicelluloses that are combined with lignin with lesser amounts of extractives, protein, starch and inorganics. However, these polysaccharides are recalcitrant to enzymatic hydrolysis; for this reason, the lignocellulosic biomass must be subjected to pretreatments. Currently, numerous pretreatment technologies have been studied to process biomass feedstock for bio-ethanol production and some of them have already been applied to the biomass of *C. cardunculus*. One of the pretreatment technologies is dilute acid hydrolysis. The severity factors of dilute acid hydrolysis, i.e., concentration of the acid and pretreatment temperatures, must be carefully measured, in order to maximize the quantity of fermentable sugars, without generating inhibitors (Pirzadah et al. 2014). All this is clearly shown by the work of del Campo et al. (2006) in which the dilute acid hydrolysis has been studied on wastes from processing globe artichoke and cultivated cardoon (Fig. 4.5). They observed that the globe artichoke biomass responds smoothly to the acid hydrolysis, giving rise to significant differences in the concentration of sugars in the liquid fraction, according to whether or not the pretreatment required an acidic solution (0% vs. 0.5% H_2SO_4). On the contrary, the biomass of cultivated cardoon does not seem to have given equally clear answers to the presence of sulfuric acid. Ballesteros et al. (2007) measured the monosaccharide yields in prehydrolyzate (liquid fraction) from dilute acid pretreatment experiments of cultivated cardoon stems (Fig. 4.5). In their experiment, pretreatments with different combinations of temperature (160, 180 and 200 °C), acid concentrations (0, 0.1 and 0.2% H_2SO_4) and solid concentrations [5, 7.5 and 10% of (w/v)] were compared. They have found that the most abundant simple sugar in the liquid fraction is xylose (accounting for up to 65–70% of total sugars) and that in any case its content is heavily dependent on the pretreatment conditions, as evidenced by the fact that the highest temperature (200 °C)

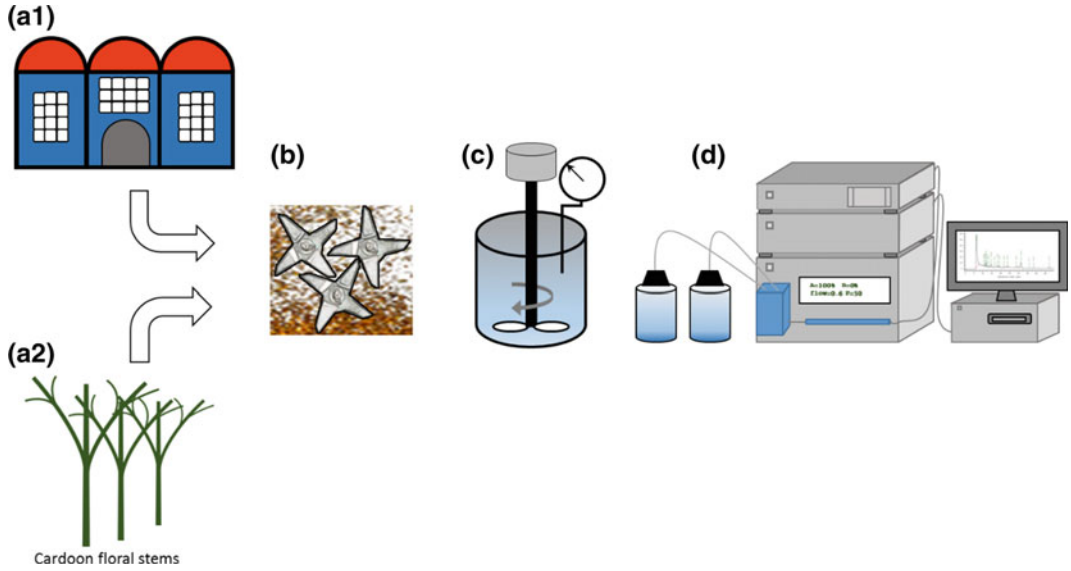


Fig. 4.5 Scheme of biomass treatment and analysis adapted from del Campo et al. (2006) and Ballesteros et al. (2007). **a1** Collecting waste from processing artichoke and cardoon (del Campo et al. 2006); **a2** collecting cardoon floral stems (Ballesteros et al. 2007); **b** milling of biomass; **c** pretreatments in autoclave at high temperature; **d** analysis of liquid fraction

caused its lowering in the liquid fraction. Glucose was found in significant amounts at the highest temperature of 200 °C in experiments performed with acid. This fact suggests a slight increase in cellulose hydrolysis under increasingly severe conditions. As the concentration of the acid increased, the content of xylose and glucose in the liquid fraction also increased, even though the concentration of the acid was in any case low. Dry matter concentration does not seem to have had a particular effect on the concentration of xylose or glucose in the liquid fraction.

Shatalov and Pereira (2011) employed response surface methodology for statistical modeling and optimization of low-temperature dilute sulfuric acid hydrolysis of the hemicellulose fraction of cultivated cardoon stems (Fig. 4.6). They used a central composite rotatable design to assess the effect of the principal independent process variables (reaction time, temperature and acid concentration) on efficiency and selectivity of heteroxylan conversion to xylose. In other words, they sought the pretreatment conditions that maximized the recovery of xylose in the liquid fraction, without affecting the cellulose and without giving rise to inhibitors. Of course, even moderate reaction conditions applied for xylan hydrolysis caused the partial hydrolysis of cellulose and decomposition of monosaccharides to furans. Under optimum reaction conditions (138.5 °C, 51.7 min and 1.28% acid concentration), 86% of xylose recovery was achieved, with limited cellulose degradation and furans formation (glucose 2.03 g and furfural 1.04 g per 100 g of cardoon). The hydrolysis of hemicellulose destroys the complex that it forms together with the lignin and protects the cellulose fibers in the cell walls. For this reason, the cellulose contained in the pretreated biomass was found to be more easily attackable in the enzymatic hydrolysis, with the conversion of about 76% of the cellulose into glucose (useful for ethanol production), against 19% for the untreated cardoon. A step forward was made by Cotana et al. (2015), who performed fermentation after steam explosion pretreatments at different severity conditions on cultivated cardoon biomass. They tested two different processes: separate hydrolysis and fermentation (SHF) and semi-simultaneous saccharification and

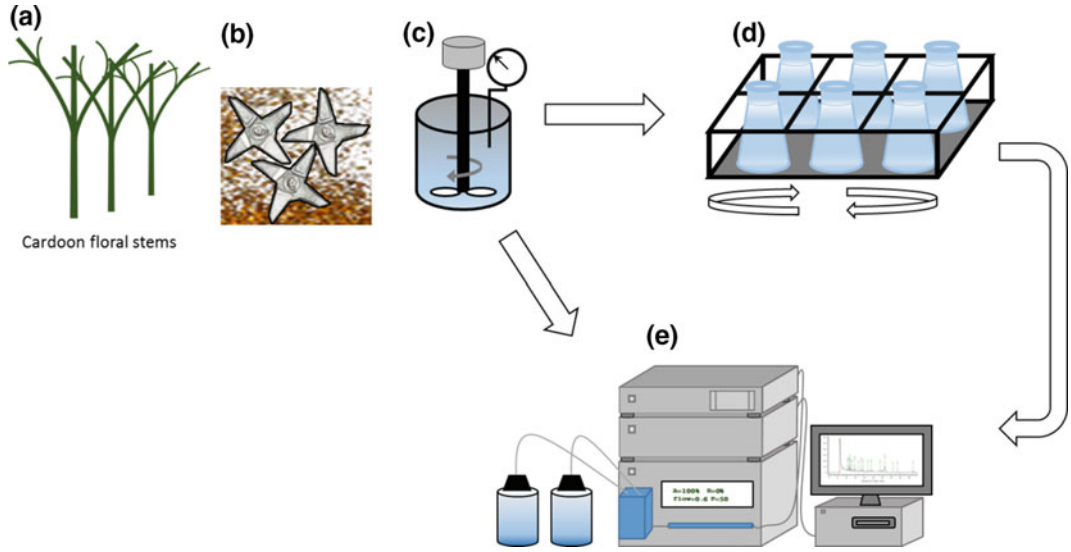


Fig. 4.6 Scheme of biomass treatment and analysis adapted from Shatalov and Pereira (2011). **a** Collecting cardoon floral stems; **b** milling of biomass; **c** pretreatments in autoclave at high temperature; **d** enzymatic hydrolysis performed at 50 °C for 72 h under 150 rpm rotation; **e** analysis of hydrolyzates coming from both pretreatments and enzymatic hydrolysis

fermentation (SSSF) (Fig. 4.7). The fermentations were carried out with *Saccharomyces cerevisiae* yeast. In SHF, there were two separate steps: pure enzymatic hydrolysis and pure fermentation, while in SSSF a short enzymatic prehydrolysis preceded the fermentation, in order to ensure a glucose concentration high enough to activate the yeast fermentation and good biomass liquefaction (necessary to guarantee an efficient mass transfer for the yeast). The results showed slightly better performances achieved by SSSF process compared to SHF, especially at low severity pretreatment conditions, with a maximum overall ethanol yield of 13.64 g of ethanol 100 g^{-1} of raw material, equal to 68.94% of the theoretical maximum yield (after 48 h of fermentation). For the SHF process, the best overall ethanol yield achieved was 13.17 g of ethanol 100 g^{-1} of raw materials, equal to 66.56% of the theoretical maximum yield (after 72 h of fermentation). Also, Fernandes et al. (2015) carried out the saccharification and fermentation of the cultivated cardoon biomass, after having subjected the latter to steam explosion at 235 °C (vapor pressure of 3.2 MPa) for 1 min using a severity factor of 3.97 (Fig. 4.7). As showed by scanning electronic microscope (SEM), steam explosion pretreatment caused the disruption of interfibrillar surfaces, improving the accessibility of carbohydrate complex toward enzymatic saccharification. Solid recovery from steam explosion was 75.8%, and the loss of solid matter corresponded essentially to the removal of hemicelluloses solubilized during the treatment. Cellulose content also slightly decreased, indicating that it was affected by the pretreatment. The biomass so obtained was subjected to alkaline extraction with 2% NaOH solution, in order to remove lignin, tannins, saponifiable extractives, part of hemicelluloses and other degradation products arisen during the steam explosion pretreatment. Biomass was finally subjected to bio-ethanol fermentation by *S. cerevisiae* yeast, carrying out both separate hydrolysis followed by fermentation (SHF) and simultaneous saccharification and fermentation (SSF). SSF process reached the highest maximum ethanol concentration (18.7 g L^{-1} using 8% of solid loading), productivity (26.6 g ethanol per 100 g of dry biomass, after steam explosion and alkali extraction) and fermentation efficiency (66.6%) during 24 h against the 96 h (72 h for

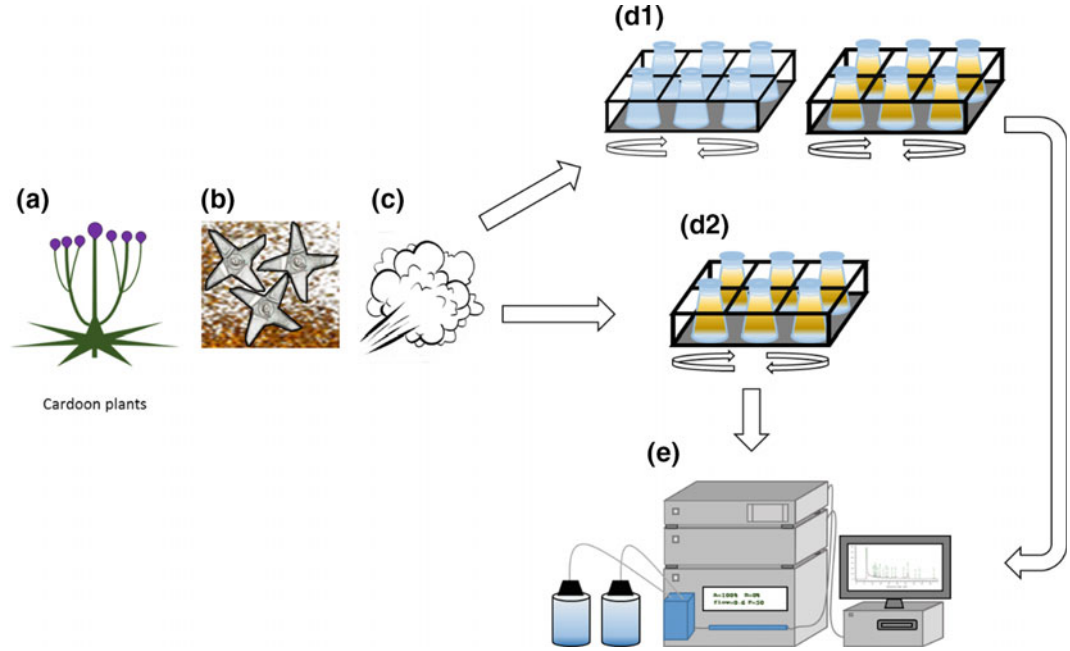


Fig. 4.7 Scheme of biomass treatment and analysis adapted from Cotana et al. (2015) and Fernandes et al. (2015). **a** Collecting cardoon biomass; **b** milling of biomass; **c** pretreatments with steam explosion; **d1** separate hydrolysis and fermentation (Fernandes et al. 2015) or semi-simultaneous saccharification and fermentation (Cotana et al. 2015); **d2** simultaneous saccharification and fermentation (Cotana et al. 2015; Fernandes et al. 2015); **e** analysis of products coming from enzymatic hydrolyses and fermentations

saccharification and 24 for fermentation) of the two-step process needed for SHF. The same saccharification and fermentation conditions described above have been used by Fernandes et al. (2018) to study the production of ethanol from cultivated cardoon after pretreatment with dilute acid hydrolysis (H_2SO_4 6.7%, 130 ° C, 55 min) (Fig. 4.8). After 24 h of fermentation using 8% of solid loading, the overall fermentation efficiency and ethanol concentration for SHF with these pretreatment conditions were respectively 52.5% and 11.5 g L^{-1}, while overall fermentation efficiency and ethanol concentration obtained in SSF were higher than those obtained in SHF mode (55.8% and 12.2 g L^{-1}, respectively). Compared to steam explosion, dilute acid hydrolysis rendered a lower ethanol concentration and a overall fermentation efficiency than those reported by Fernandes et al. (2015). Pesce (2017) evaluated the potential bio-ethanol production of globe artichoke crop residues, coming from two different genotypes, not subjected to any pretreatment. He provided an estimation of the potential production of bio-ethanol, both on a dry matter and a per hectare basis of non-pretreated biomass after enzymatic saccharification (Fig. 4.9). The genotypes studied were 'Opera F1', an early purple F1 hybrid specifically bred for annual production, and 'Spinoso sardo', a traditional landrace belonging to the thorny typology. 'Spinoso sardo' was more productive than 'Opera F1', both on a dry matter ($\sim$112 kg of bio-ethanol per t DM vs. $\sim$87) and per hectare basis ($\sim$1553 kg ha^{-1} vs. $\sim$1481 kg ha^{-1}), in spite of the higher biomass productivity of 'Opera F1'. The reason for this difference lies in the greater content of free sugars in the 'Spinoso sardo' than in the 'Opera F1' biomass.

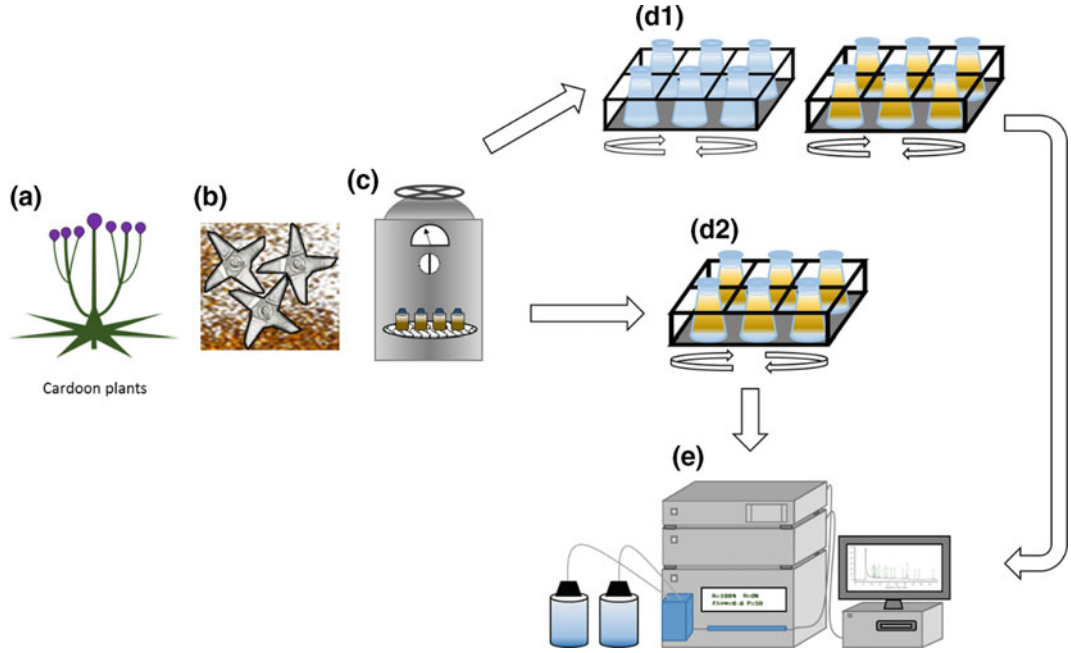


Fig. 4.8 Scheme of biomass treatment and analysis adapted from Fernandes et al. (2018). **a** Collecting cardoon biomass; **b** milling of biomass; **c** dilute acid hydrolysis pretreatments in autoclave at high temperature; **d1** separate hydrolysis and fermentation; **d2** simultaneous saccharification and fermentation; **e** analysis of products coming from enzymatic hydrolyses and fermentations

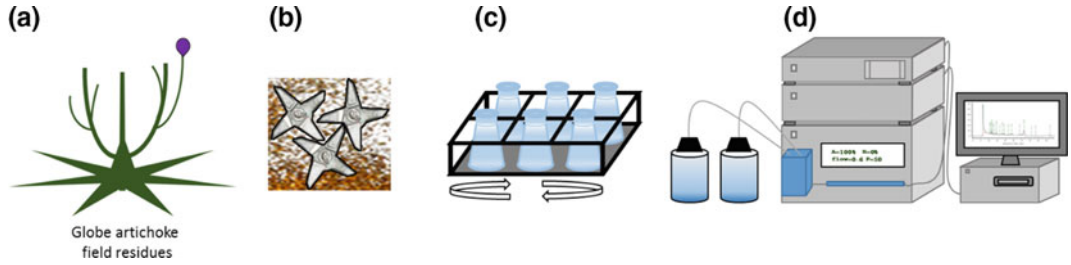


Fig. 4.9 Scheme of biomass treatment and analysis adapted from Pesce (2017). **a** Collecting globe artichoke biomass; **b** milling of biomass; **c** enzymatic hydrolysis performed at 50 °C for 72 h under 150 rpm rotation; **d** analysis of products coming from enzymatic hydrolysis

4.4 Oil Production (From Seeds)

The proposal of using *C. cardunculus* seeds for oil production was first made in the decade of 1990, in the framework of a European project (CORDIS-JOUB0030 1990), and it was based on the fact the species is botanically related to other oil crops such as sunflower or safflower (Curt et al. 2002). As a member of the Compositae family, the fruit of *C. cardunculus* is a cypsela, i.e., an achene originating from an inferior ovary (Beck 1891 cited in Marzinek et al. 2008). Due to the fact that the cypsela is a dry fruit, indehiscent, unilocular, and with a single seed not adnate to the pericarp, this type of fruit is commonly known as 'seed', although the seed is only the kernel that is inside the cypsela. Like other species in the Compositae family, *C. cardunculus* accumulates oil in the seed endosperm (Fig. 4.10). Different

Fig. 4.10 Fruits (cypselae or achenes) of *C. cardunculus,* commonly known as cynara 'seeds' or cynara 'grains'

procedures of oil extraction from cynara seeds have been tested with success such as screw vegetable oil expeller (Curt et al. 2002), solvent extraction (*n*-hexane) and supercritical fluid extraction (SFE) (Alexandre et al. 2012).

With a view to the exploitation of *C. cardunculus* as a poly-annual and multipurpose crop, the potential of this species for seed oil production has been studied by different authors for a wide range of crop conditions. This way the crop costs could be allocated to two crop products, the lignocellulosic biomass and the seed oil, and the economic balance be improved (Fernández and Curt 2004). Like other oil crops, the yield components of the oil yield are seed oil content (%) and seed yield (t seeds ha^{-1}) which, in the case of cynara, is a function of the number of heads (capitula) developed, the seed proportion in plant heads and the seed size. Oil quality is assessed by determining the fatty acid profile and oil properties, as related to different applications. Studies of *C. cardunculus* have been conducted to determine these parameters as well as the effect of a number of variables such as genotype (sensu lato) and environmental conditions. In large field experiments, procedures for the mechanical harvest of cynara seeds have been experimented as well (Pari et al. 2016). Compilation of bibliographic data on oil yield components of *C. cardunculus* is provided in Table 4.5.

Values reported for seed oil content range from 14.5% (Francaviglia et al. 2016) to 32.5% (Curt et al. 2002), depending on various factors such as crop conditions, plant growth, seed size and ecotype. Generally, low seed oil contents have been recorded for crops or individuals with poor growth, for whatever the reason that might be (tough crop conditions, marginal soils, wild genotypes, etc.). High oil contents have been related to seed size since the kernel proportion is usually higher for larger seeds than for small seeds. Kernel weight represents about 55.5% of the seed weight for seeds sizing 7.5–7.3 mm length (Curt et al. 2002). Under unconstrained conditions, 23–25% oil content can be expected for seeds of *C. cardunculus*, with a specific weight of 35–40 g/1000 seeds. Oil content and specific weight of cynara seeds have been reported rather uniform, suggesting that these variables have a small effect on cynara's oil production (Lag-Brotons et al. 2014). Nevertheless, genetically improved varieties have not been released for seed oil yet; possibly, large seeds with high kernel proportion and specific weight content could be developed for high oil yield by plant breeding.

C. cardunculus, as a multipurpose crop (biomass and oil), is intended for drylands under Mediterranean climates, where it was originated. Similar to other crops grown in the Mediterranean region, the most limiting factor for plant growth and, hence, for higher seed yield is the availability of water. In drylands, yields depend on the precipitation recorded during the growth cycle and soil water-holding capacity (Fernández et al. 2006). The importance of these factors on seed yield was clearly shown in the studies by Vasilakoglou and Dhima (2014) and Archontoulis et al. (2010a, b), which were carried out in different environments in Greece. Thus, irrigation increased 46% the seed yield (Archontoulis et al. 2010a) and seed yield in aquic soils was 2.5 times the seed yield recorded in dry soils (productive crop cycles) (Archontoulis et al. 2010b). It has been shown that high growth generally

Table 4.5 Bibliographic data on components of oil yield in *C. cardunculus*

Variable	Seed yield (t ha^{-1} y^{-1})	Moisture content (% w/w)	Seed oil content (% w/w)	Specific weight (g per 1000 seeds)	Seed/head (% w/w)	References
Location, head size and seed size	1.0–4.8	5.24–6.39	17.0–23.2	10–57	18–39	Archontoulis et al. (2010b)
Genotype and crop season	1.09–2.44	–	–	30.9–45.3	–	Bolohan et al. (2014)
Population, seed size and crop season	–	7.81	20.6–32.5	–	–	Curt et al. (2002)
Genotype	0.72–1.2	–	15.4–25.2	–	14.3–27.3	Curt et al. (2014)
Genotype and crop season	0.5–2.8	5.45–6.35	21.7–26.6	19–37	8–44.5	Foti et al. (1999)
Plant growth	–	–	–	–	18.6–28.7	Fernández et al. (2007)
Location	–	7.8%	21.5–26.1	–	–	Ferreira-Dias et al. (2018)
Genotype and crop season	0.57–3.69	–	14.5–20.3	–	–	Francaviglia et al. (2016)
Genotype and crop season	0.39–0.95	–	–	–	–	Gherbin et al. (2001)
Soil fertility	0.10–1.63	–	–	25–35	–	Gominho et al. (2011)
Genotype and crop season	0.29–1.69	–	–	–	–	González et al. (2011)
Genotype, fertilization, harvest time and crop season	0.2–1.4	–	–	–	–	Ierna et al. (2012)
Genotype and crop season	0.26–1.7	–	–	22–39	–	Ierna and Mauromicale (2010)
Organic fertilization (sewage sludge compost)	0.85–4.33	–	20.1–23.7	40	–	Lag-Brotons et al. (2014)
Genotype and crop season	–	–	21.1–26.6	–	–	Maccarone et al. (1999)
–	–	–	20.5 ± 1.0	–	–	Miceli and De Leo (1996)
Genotype and input conditions	0.5–3.3	–	–	20.2–34.7	11.5–20.4	Neri et al. (2017)
Genotype and crop season	0.13–0.84	–	18.5–23.6	–	18.3–38.2	Piscioneri et al. (2000)
Genotype and crop season	0.20–1.82	–	19.7–27.3	18.6–45.3	–	Raccuia and Melilli (2007)
Genotype	–	8–9	18.5–25.3	–	–	Raccuia et al. (2011)
Genotype, irrigation and weed competition	1.10–2.58	–	–	24.9–35.9	–	Vasilakoglou and Dhima (2014)

results in greater amount of heads, and in heads of larger size which contain higher proportion of seeds (% w/w) (Archontoulis et al. 2010b; Fernández et al. 2007). This plasticity of cynara explains the wide range of values reported in the literature for seed yields, from 0.2 to 4.3 t ha^{-1} $year^{-1}$ (values from small field experiments).

Among the physicochemical characteristics of cynara oil, it has been reported: 910–933 g/L density (Alexandre et al. 2012; Benjelloun-Mlayah et al. 1997), 0.3–2.0 acidity (% of oleic) (Carvalho et al. 2006; Maccarone et al. 1999), 0.9–3.2% unsaponifiable matter (Benjelloun-Mlayah et al. 1997; Maccarone et al. 1999; Miceli and De Leo 1996), 108.4–121.9 iodine value (Carvalho et al. 2006; Miceli and De Leo 1996) and 1.47 refractive index (Benjelloun-Mlayah et al. 1997; Miceli and De Leo 1996). Concerning the fatty acid profile, *C. cardunculus* oil has been found similar to common sunflower oil in most studies; the main components are linoleic (C18:2, ranging from 50 to 61%), oleic (C18:1, 18–34%), palmitic (C16:0, 3–14%) and stearic (C18:0, 3–12%) (Benjelloun-Mlayah et al. 1997; Carvalho et al. 2006; Curt et al. 2014, 2002; Lapuerta et al. 2005; Miceli and De Leo 1996). It is worth mentioning that Raccuia et al (2011) reported higher oleic content for some domestic cardoon varieties (four commercial varieties and four lines selected by self-pollination), in line with the composition reported by Piscioneri et al. (2000) for three commercial varieties of cardoon; both studies were conducted in Italy (Table 4.6).

In the context of bio-energy applications, the suitability of seed oil of *C. cardunculus* for biodiesel production has been studied (Alexandre et al. 2012; Bouriazos et al. 2014; Fernández and Curt 2004; Martínez et al. 2014; Sengo et al. 2010). Martínez et al. (2014) tested basic transesterification of cynara oil to produce biodiesel (fatty acid methyl esters, FAME) and reported that the product met the quality standards of density, kinematic viscosity, water content, acid value, iodine value, flash point and cold filter plugging point (CFPP) as set by the European Norm (EN14214), but failed in the FAME content (92.3% < 96.5% in EN14214) and cetane number (43.2 < 51.0). However, some biodiesel properties were found to be more influenced by transesterification and purification processes than by the raw material. In this respect, Alexandre et al. (2012) found that the two-step transesterification improved FAME contents. The performance of cynara biodiesel has been studied with success in engines (Lapuerta et al. 2005); results have shown a significant reduction on particulate emissions.

According to the ratio of unsaturated acids (about 5.7), linoleic-to-oleic ratio (around 1.8) and the virtual absence of erucic acid, cynara oil has been reported suitable for human consumption (Carvalho et al. 2006; Castejón et al. 2017). Thus, the nutritional compositional data of cynara oil is similar to common sunflower oil (Castejón et al. 2017). These authors showed that the composition of unoxidized cynara oil (crude unrefined oil) was 55.5 polyunsaturated fatty acids (PUFA), 27.4% monounsaturated fatty acids (MUFA) and 17.1% saturated fatty acids (SFA); during the frying process, cynara oil suffered a decrease in PUFA (8%) and an increase in SFA (5%) and MUFA (3%), with respect to the initial composition; however, its oxidative stability was better than sunflower oil. The oil composition in triacylglycerols, sterols and tocopherols was studied by Ferreira-Dias et al. (2018) in relation to crop location; it was found that the oil composition was more dependent on edapho-climatic conditions of the site than on cynara genetic variability. Sterols and tocopherols that confer antioxidant and bioactive properties to vegetable oils varied from 4254 to 5205 mg kg^{-1} and from 32.5 to 542.1 mg kg^{-1}, respectively. Cynara oil has not been marketed for culinary purposes yet, but it is currently advertised on the Internet as a nutraceutical product (pearl format) for improving the status and functioning of liver, based on the presence of silymarin in the cynara oil, a hepatoprotective plant compound (http://www.hipernatural.com/en/venhepacyn.html). In this regard, the content of silymarin in *C. cardunculus* seeds (unbred seeds) has been reported about 1% (w/w, dry basis) (Curt et al. 2005). On other possible applications, it is worth mentioning that cynara oil has been suggested for the production

Table 4.6 Literature data on the fatty acid profile of *C. cardunculus* seed oil

Fatty acid (common name and shorthand)															References
Myristic (C14:0)	Myristoleic (C14:1)	Palmitic (C16:0)	Palmitoleic (C16:1)	Margaric (C17:0)	Margaroleic (C17:1)	Stearic (C18:0)	Oleic (C18:1)	Linoleic (C18:2)	Linolenic (C18:3)	Arachidic (C20:0)	Gadoleic (C20:1)	Behenic (C22:0)	Erucic (C22:1)	Lignoceric (C24:0)	
r		3.4				12.1	23.6	60.9							Benjelloun-Mlayah et al. (1997)
0.14		11.10–15.47	0.23	0.05		3.20–4.84	24.89–28.83	47.66–59.12	0.06	0.4		0.13		0.23	Archontoulis et al (2010b)
Trace		13.9 ± 0.1	<0.1	Trace		3.4 ± 0.1	30.0 ± 0.9	51.2 ± 1.8		0.3 ± 0.1		<0.1		0.3 ± 0.1	Miceli and De Leo (1996)
0.12		8.96	0.23			2.63	27.65	45.04	4.09				10.33		Martínez et al. (2014)
0.11		10.62	0.14			3.7	24.95	59.87	0	0.36	0.15	0		0	Lapuerta et al. (2005)
0.10–0.13	0–0.03	10.21–11.20	0.13–0.20	0.02–0.06	0–0.04	3.23–4.49	21.96–27.98	56.42–62.84	0–0.04	0.31–0.42	0.09–0.14	0–0.07			Curt et al. (2002)
0.1		10.9	0.2			3.3	23.1	61.2	1.0	0.4	0.1	0.1		0.3	Carvalho et al. (2006)
0.1		10.4–11.7	0.1–0.2	0.1	0.0	3.2–3.6	19.9–25.5	58.4–64.9	0.06–0.09	0.3–0.4	0.1	0.1		0.1–0.2	Ferreira-Dias et al. (2018)
		10.1–11.7				3.1–4.1	18.0–34.5	49.9–66.8							Maccarone et al. (1999)
		9.6				3.1	28.9–47.2	38.7–54.1							Raccuia et al. (2011)
		6.1–10.8				1.8–3.6	26.1–83.6	4.0–54.1		0–0.3					Piscioneri et al. (2000)

of soaps and cosmetics (Sharaf-Eldin 2016, https://www.chateaudesplantes.com/en/le-cynara/, http://www.bionap.com/20-cosmetica-inglese/80-cardo-oil-en) and cynara seed cakes for pharmaceutical and cosmetics industry (Petropoulos et al. 2018), fertilizer and animal feed (Curt et al. 2010).

In conclusion, therefore, it can be said that yields of nearly 1 t seeds ha^{-1} $year^{-1}$ or 250 kg oil ha^{-1} $year^{-1}$ seem realistic in most circumstances, for the poly-annual crop of *C. cardunculus* in drylands of the Mediterranean region; cynara seed oil is similar to common sunflower oil and is suitable for the production of biodiesel and for food and nutraceutical applications.

4.5 Cellulose, Pulp and Paper

One of the possible utilization for the *C. cardunculus* aboveground biomass is as a fiber source for paper pulp production. A recent review published by Gominho et al. (2018) showed that different delignification processes were tested in cynara biomass to evaluate the pulp aptitude (Table 4.7); Soda, in situ soda by Benjelloun-Mlayah et al. (1997), soda-AQ by Antunes et al. (2000), kraft by Benjelloun-Mlayah et al. (1997), Poveda et al. (1999), Gominho et al. (2001), Gominho and Pereira (2006), Abrantes et al. (2007), Lourenço et al. (2017), Kraft-AQ by Abrantes et al. (2007),

Table 4.7 Total yield and pulp Kappa number of the different delignification processes used to study the pulping aptitude of different biomass fractions of *C. cardunculus*

Biomass fraction	Process	Total yield (%)	Kappa number	References
Stalks + leaves	Kraft	48–51	26–22	Benjelloun-Mlayah et al. (1997)
Stalks		45	–	
Stalks + leaves	Soda and in situ soda	42–78	17–50	
Stalks		41	22	
Stalks	Kraft	45–49	22–48	Poveda et al. (1999)
Stalks	Soda-AQ	35–40	17–22	Antunes et al. (2000)
Stalks	ASAM	49	15–20	Gominho and Pereira (2000)
Depithed stalks		52	15	
Stalks	Kraft	43.5	15.0	Gominho et al. (2001)
Depithed stalks		47.0	11.0	
Depithed stalks	Ethanol–water	40	28	Oliet et al. (2005)
Stalks	Kraft	44–49		Gominho and Pereira (2006)
Stalks	Kraft	37–39	14–21	Abrantes et al. (2007)
Stalks	Kraft-AQ	33–39	17–22	
Depithed stalks	Acetosolv	37–82	7–43	Ligero et al. (2007)
Depithed stalks	MILOX	47–76	29–10	Ligero et al. (2008)
Hairs and pappi	Kraft	48–63	6–38	Gominho et al. (2009)
Stalks	Ethanol–alkali	44–60	5–68	Shatalov and Pereira (2014)
Stalks	Kraft	45	–	Lourenço et al. (2017)
Stalks	Ethanol–water	61	–	

Adapted from Gominho et al. (2018)

ASAM by Gominho and Pereira (2000), ethanol–water by Oliet et al. (2005), Lourenço et al. (2017), Acetosolv and MILOX by Ligero et al. (2008) and ethanol–alkali by Shatalov and Pereira (2014). The major of these delignification studies used stalks and the depithed stalks as raw material, and only one dealt with hairs and pappi, both fractions from the cynara capitula (Gominho et al. 2009).

These studies proved that generally the cynara stalks could be delignified with high total yields (over 33%), with low rejects (between 1.9 and 3.5%) and low residual lignin content (Kappa numbers ranging from 5 to 22). However, using the whole cynara biomass is not advised due to the higher content in parenchyma cells in the pith of the stalks and branches. Gominho et al. (2001) using kraft as delignification process observed that when the pith was removed (mechanically or manually) the delignification yields increased in 2–4% and the Kappa numbers decreased in 30%. In addition, the presence of parenchyma cells created several difficulties during the delignification process, i.e., increased the chemicals consumption and the content on pulp fines and reduced the pulp yields (Ek et al. 2009). For this reason, these non-fibrous thin-walled cells have no interest for paper properties, and they are called as the zero fibers for the 'papermakers'. So, it is advised to remove the pith fraction as pretreatment to improve the pulp quality. Another possibility is to pretreatment the biomass, and different procedures have been tested to promote the disruption of parenchyma cells from the cynara biomass; Shatalov and Pereira (2014) used a prehydrolysis with sulfuric acid (1.28%) and Lourenço et al. (2017) steam explosion. In both cases, the hemicelluloses released can be valorized as by-products in a biorefinery context. A clear evidence of the positive achievement of the pith destruction is the increase in the average fiber length (0.64 mm vs. 0.83 mm) and the decrease in the coarseness (0.12 mg m^{-1} vs. 0.10 mg m^{-1}), respectively, for kraft pulps made with whole stalks and depithed stalks (Gominho et al. 2001). Gominho et al. 2001 observed that cynara pulps obtained with whole stalks refined easily (25° SR, 45° SR and 80° SR at 0, 1 000 and 4 000 PFI rev), but unfortunately, the drainage time increased very strongly. On the other hand, depithed stalk pulp presented higher strength properties (high bulk and significant strength properties, especially in relation to tensile; 104 N m g^{-1} at 750 PFI rev) than the whole stalk pulp.

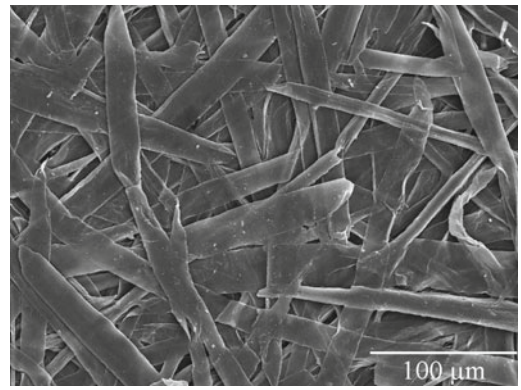


Fig. 4.11 General aspect of a sheet paper made with hairs, obtained by SEM analysis

Hair and pappi detached from cynara capitula biomass were also tested as raw material for pulp. The kraft pulps showed high pulp yields (63% for hairs and 48% for pappi) with low residual lignin (Kappa number 7 for hairs and 11 for pappi) and adequate pulp properties for paper (Fig. 4.11) (40 and 42 N m g^{-1} tensile index; 3.6 and 3.4 kPa m^2 g^{-1} burst index in unrefined pulps of hairs and pappi, respectively) (Gominho et al. 2009).

4.6 Use in Animal Feeding

Since the petioles of lower leaves of this plant are used as human food, it is logical to consider that their use could be of interest in different types of animals. Given the contents of fructans (inulin) and flavonoids of the plant, the most interesting option is its use in ruminants, because the ruminal microbial digestion should eliminate the possible disadvantages associated with an excessive supply of these compounds. Most of the studies carried out on the nutritive use of this plant have been done in these species. The cultivation of

C. cardunculus can give rise to various feed resources such as green forage, harvest residues (after the cut for biomass production), seeds and its presscake or meal after oil extraction.

4.6.1 Green Forage and Harvest By-Products

The cultivation of this species for industrial purposes can be compatible with obtaining green forage providing that the cut for this last purpose is made in autumn–winter, since, after this cut, the plant springs up from its rosette; however, this double use limits the biomass yield obtained (about 46%, after Fernandez et al. 2006). According to the edaphic and climatic conditions, the yield of green forage in the indicated period varied from 4 to 9.6 t of DM ha^{-1} (Fernández and Manzanares 1989; Romero et al. 1997; Delgado et al. 2005) increasing with the delay of the cut date. At the indicated time, the plant is in vegetative state and the cut provides only leaves, so that very high-quality forage is obtained in a period of scarce availability of grazing resources. This high quality is a consequence of a high concentration of nutrients (see Table 4.8), combining an adequate concentration of crude protein

Table 4.8 Mean values of chemical composition (g/kg of DM), ruminal degradability (%), digestibility (%) and net energy content (milk (UFL) and meat (UFV) forage units) of different feedstuffs from *C. cardunculus*

	Green forage	Dry leaves	Straw	Ensiled forage[a]	Seed
References[b]	1, 2, 3, 4	2	2	2	5, 6, 7, 8
Dry matter (DM; %)	12.5	91.5	92.2	18.0	93.6
Organic matter (OM)	839	746	866	834	956
Ether extract (EE)	18.7	31.6	14.3		258
Crude protein (CP)	151	78.4	72.0	135	190
NDCP (g/kg CP)[c]	285	271	181		137
Crude fiber (CF)	138	260	341	138	226
Neutral detergent fiber (NDF)	256	408	607	259	374
Acid detergent fiber (ADF)	170	301	438		295
Acid detergent lignin (ADL)	28.0	67.5	69.1		94.6
Soluble carbohydrates	270			147	
Crude energy (CE; kcal/kg DM)	3799	3488	3881		5558
DM degradability	81.4	56.5	40.9		56.8
OM degradability	79.9	53.4	36.2	87.5	
CP degradability	79.7				82.9
DM digestibility	79.7	53.7	55.6		55.3
OM digestibility	85.6	65.9	56.8		56.3
CE digestibility	82.7	65.8	51.6		59.2
NDF digestibility	78.9	66.9	50.7		20.9
ADF digestibility	80.3	69.2	51.4		21.1
EE digestibility					83.0
UFL/kg DM	0.970	0.661	0.548		0.901
UFV/kg DM	0.979	0.61	0.46		0.789

[a]Soluble N: 10.1; NH_3–N: 2.37

[b]1: Romero et al. (1997); 2: Cajarville et al. (1999); 3: Delgado et al. (2005); 4: Fernández et al. (2006); 5: Fernández and Manzanares (1989); 6: Cajarville et al. (2000); 7: Riahi (2015); 8: Genovese et al. (2016)

[c]CP linked to neutral detergent fiber

(CP) with a high content of soluble carbohydrates and a relatively low content in fiber and lignin, of which approximately 60% corresponds to cutin (Cajarville et al. 1999). There is controversy in the literature data about its fibrous contents. If the neutral detergent fiber (NDF) is not made by carefully disintegrating the insoluble residue, an artifact is formed, which is not disintegrated in the acid detergent fiber analysis (ADF); in this manner, fiber values are overestimated (Cajarville et al. 1999). In agreement with above indicated characteristics, this forage presents a high digestibility for both the total organic matter (OM) and the fiber. Thus, Cajarville et al. (1999) observed in sheep values of 86.1, 78.9 and 80.3% for OM, NDF and ADF, respectively. This OM digestibility value is two percentage points higher than that indicated in the INRA tables (2007) for the best forages. Given its high content in digestible nutrients, its energy content is close to that of concentrates. Thus, the indicated sample presented a net energy content of 0.97 UFL and 0.98 UFV, despite the fact that the OM content of this sample (812 g/kg DM) was lower than normal. Another clear index of the extraordinary quality of this forage is its low contents in undegradable materials in the rumen (6 and 4% of the OM and CP, respectively, for this sample). The digestion of this forage takes place basically in the rumen as shown by the apparent rumen degradability values of 80% obtained for the MO and CP of this same sample. This degradation is performed simultaneously for these both fractions at a medium rate. All the cited positive characteristics would allow a high efficiency of microbial protein synthesis in the rumen and, therefore, a high supply of amino acids to the animal. In addition, this forage has a high buffer capacity (about 640 meq OHNa/kg DM; Cajarville et al. 1999) that may contribute to contain the possible fall of ruminal pH due to its high content of soluble carbohydrates (270 g/kg DM, Cajarville et al. 1999) of fast degradation. In summary, the contribution of fermentable nutrients of this forage is close to that of the concentrates avoiding the risks of acidosis that involves the ingestion of high amounts of these feeds in high productive animals.

Given its vegetative state at the cut, the main limitation of the wintergreen forage is its low DM content (between 11 and 15%); in addition, it shows a strong difficulty to moisture loss during prehaymaking as a result of the cutinized surface of its leaves and the winter weather conditions. Consequently, the most valid conservation option is silage. Despite its high buffer capacity and high moisture content, this forage ensiled very well due in part to its high capacity of compaction (leading to 0.81 tm/m^3), which may be increased by 10% with a prehaymaking of one or two days. However, its high ensilability is mainly due to its high content of soluble carbohydrates that allow intense homolactic fermentation with moderate formation of acetic acid and practically no other volatile fatty acids. This fact, together with the low generation of ammonia during silage maturation, is a clear indicator of few possibilities for the action of undesirable flora during silage. The turn silage over is practically impossible since the residual soluble carbohydrates after the maturation of the silo represent something more than 50% of the initial ones. Direct silage involves significant losses (close to 20%, according to Cajarville et al. 1999), mainly by elution; however, these losses can be greatly reduced with prehaymaking (up to 75% with a period of 94 h). Likewise, liquid losses can be reduced placing in the silo-floor a layer of absorbent materials, such as straw, dehydrated beet pulp and harvest residues. The chemical composition of silages made at capitulum and seed ripening stage (Pesce et al. 2017) shows that the good facility for silage of this plant is maintained in the latest growth stages. These silages have enough CP content and lead to a reduction in elution and fermentation losses due to its high DM content; however, their nutritional value is reduced by its large levels of NDF (between 450 and 480 g/kg DM) and by a non-homolactic fermentation.

The basal leaves that remain under the cut in the biomass harvest can be collected and used as dry forage to present a considerable nutritional value (see Table 4.8). Thus, their CP content is a non-limiting factor for the ruminal microbial fermentation and their fiber contents are not

excessive. Also, OM and fiber digestibility and in special its ingestion level are high for a dry crop by-product. Their main limitation is its high ash content, due to an important contamination with soil. In spite of this, this by-product maintains contents in net energy of the order of a hay of medium–good quality. The quality of this by-product would benefit from some mechanical treatment that would reduce soil contamination. The seed harvesting allows obtaining biomass that, if not used for industrial purposes, can be used in the ruminant feeding. By including the stems and the remains of the chapters, their nutritional value is considerably reduced. However, this by-product presents values of ingestion, digestibility and energy content 20% higher than those of wheat or barley straws and similar to those of leguminous straws (Cajarville et al. 1999).

The important changes that occur throughout the phenological cycle of this plant allow a good prediction of the digestibility of OM, energy or fibers of these forages based on their fiber or lignin contents (Cajarville et al. 1999).

4.6.2 Seeds

This seed has large contents as in both lipids and CP as in fiber (see Table 4.8), since its small size implies a low almond/hull ratio. The almond is rich in the first two fractions, while the hull, which represents almost a half of the seed (45%, according to Cajarville et al. 2000), is very rich in lignified fiber and poor in CP, being the latter mainly linked to fiber. Thus, in monogastric species, its use would be restringed to limited quantities to avoid the high supply of indigestible fiber. In addition, its nitrogen value is only medium, because its profile of essential amino acids, although rich in branched-chain amino acids and tryptophan, has a low concentration of sulfur-containing amino acids and lysine and a non-negligible part of lysine is provided by the hull, which reduces its digestibility. Its main nutritional asset would be its lipid content, which, as noted in other chapters, is highly unsaturated, highlighting in them the high proportions of oleic (24–28%) and especially linoleic acids (57–61%), as well as the main oleaginous seeds most commonly used in animal nutrition. According to Cajarville et al. (2000), its OM digestibility in ruminants is 56–57%, as a result of a high digestibility of lipids and proteins (82–84%) and a low fiber digestibility (around 21% for NDF and ADF). The nitrogen value of the seeds is also intermediate, because in addition to the already mentioned limitations, the ruminal fermentation of this seed will support a poor ruminal microbial protein synthesis due to its large content in both lipid and poorly degradable fiber; also, its protein has a high and fast degradation, as in most oilseeds. Thus, considering the data of Cajarville et al. (2000) and that the ruminal undegradable fraction must be in turn indigestible in the small intestine, the intestinal digested protein of this seed would be less than 3% of its total protein. These last authors show also that this seed may be included at 30% of the diet of ruminants without negative effects on the fiber ruminal digestion of diet's forages; however, they recommended an inclusion level of 10% to avoid an excessive fiber supply and to take advantage of synergistic effects.

The use of extraction whole meals presents the limitation of excessive indigestible fiber content. Thus, after solvent oil extraction, its levels of NDF, ADF and lignin will be in the order of 50, 40 and 13.5%, respectively, which limits the possibilities of inclusion in the diet of feeds of greater nutritional concentration. The decortication of the seed would allow improving its nutritional value; however, the previously mentioned limitations of the seed nutritive value would not justify the cost of this process. Therefore, the production of presscake or solvent meals should be conditioned by the interest of oil production and the inclusion levels in diets should be limited.

4.7 Use of Florets in Dairy Industry

Coagulation is a key property for milk conservation because it makes possible the transformation of milk into cheese. According to

St-Gelais and Collet (2010), milk coagulation may be defined as the destabilization of the casein micelles, which flocculate and aggregate to form a gel enclosing fat and soluble milk components. The loss of the colloidal stability in micelles can be caused by acidification, enzymatic action or a combination of the two. In acid coagulation, the H^+ ions released in milk gradually neutralize the electronegative charges of the casein micelles, decreasing their electrostatic repulsion until they disappear. At room temperature, micelles begin to aggregate at pH 5.2, although total flocculation occurs at pH 4.6, when the isoelectric pH of casein is reached (Troch et al. 2017). Enzymatic coagulation is based on the action of proteolytic enzymes. The cleavage of the micelle-stabilizing protein κ-casein at or near of the peptide bond Phe_{105}–Met_{106} provokes the aggregation and gelation of the rennet-altered micelles. This particular zone of the κ-casein is very susceptible to aspartic proteinases due to the especial conformation of the molecule in the region of residues 98–111 (McSweeney 2007a). Most proteinases are able to coagulate milk under the proper conditions; however, few of them are specific enough to not overhydrolyze the coagulum, which provokes the reduction in cheese yields. The best rennets are characterized by high milk-clotting activity (MCA) on κ-casein and low non-specific proteolytic activity (PA) (Pires et al. 1994). Calf rennet is considered to be the most suitable milk-clotting protease, due to its high MCA/PA ratio (Mazorra-Manzano et al. 2013). Its principal proteinase is chymosin, although it also contains pepsin, a protease of broader specificity which accounts for ∼10% of the milk-clotting activity by calf rennet (McSweeney 2007b).

Several factors such as high price of rennet, vegetarian diet or religious dietary laws (e.g., Halal or Kosher) have encouraged the search for alternative milk-clotting sources (Roseiro et al. 2003). Genetically modified organisms (GMO) made it possible to obtain calf chymosin as a fermentation product; nevertheless, it is banned in GMO-free countries. Vegetable coagulants have proven to be suitable rennet substitutes that also offer an alternative to the previous requirements. Aqueous extracts of dried cardoon flowers have been widely used for the manufacture of traditional Mediterranean cheeses.

4.7.1 Proteinases

The stylets and stigmas of *C. cardunculus* flowers (violet part) contain several aspartic proteinases (APs), initially synthesized as zymogens which, after proteolytic processing, yield the mature and active enzyme (Ramalho-Santos et al. 1997). So far, nine different APs have been purified and characterized: cyprosins 1, 2 and 3 (Heimgartner et al. 1990; Cordeiro et al. 1992), and cardosins A, B, E, F, G and H (Sarmento et al. 2009). However, not all the APs are expressed in all the cardoon flowers (Amira et al. 2017a). Even though cardosin and cyprosin genes coexist in the genome of the same plant, they have not yet been copurified. Additionally, there are genetic evidences that other two APs could be expressed by the pistils of *C. cardunculus* (cardosins C and D), although they have never been purified and characterized (Pimentel et al. 2007). It has been observed a high variability of protease content in cardoon flowers, probably due to the natural geographic localization, the flowering stage and the seasonal climatic conditions (Heimgartner et al. 1990).

Cardosins A and B are the best-known proteinases from *C. cardunculus*. Their enzymatic activities have been studied against caseins, insulin B chain and several synthetic substrates (Veríssimo et al. 1995). According to kinetic parameters and specificity, cardosin A was considered to be similar to chymosin, while cardosin B was comparable to pepsin, which shows a broader specificity. Cardosins E, F, G and H share among them most of amino acid sequence and primary structure, resembling cardosin A. Nevertheless, they exhibit different activities and selectivities (Sarmento et al. 2009). Cyprosins, on their part, display a high proteolytic activity and low specificity, being cyprosins 2 and 3 the most active ones (Brodelius et al. 1998).

These enzymes induce milk coagulation like chymosin, through cleavage of the Phe_{105}–

Met_{106} bond in bovine and ovine κ-casein, whereas caprine κ-casein is preferentially cleaved at Lys_{116}–Thr_{117} (Sousa and Malcata 1998). However, the strong proteolytic action and the low specificity of some of them eventually lead to a more extensive breakdown of the milk proteins. In general, proteinases from *C. cardunculus* show higher preference than chymosin for bonds between hydrophobic regions of αs_1-casein and β-casein, (Macedo et al. 1996) releasing hydrophobic peptides, which are sensorially perceived as bitter.

Depending on the milk type (sheep, goat or cow), caseins are hydrolyzed in a different way by the APs from cardoon flowers. The hydrolysis degree of bovine and caprine caseins due to *C. cardunculus* rennet during milk coagulation or cheese ripening is generally higher than that in ovine caseins. β-casein and αs_2-casein from cow and goat contain more peptide bonds susceptible of cleavage than the ovine ones. Instead, ovine or caprine αs_1-casein presented more cleavage sites than bovine αs_1-casein. A more detailed description of the known cleavage sites of *C. cardunculus* proteinases in bovine, ovine and caprine caseins has been reviewed by Sousa and Malcata (2002).

C. cardunculus proteinases have also effect on whey proteins. Cardosin A and B showed proteolytic activity toward bovine α-lactoglobulin, producing several hydrophobic peptides. However, β-lactoglobulin apparently was not hydrolyzed by these APs. In the same way as against caseins, cardosin B displayed a broader specificity than cardosin A (Barros and Malcata 2006).

4.7.2 Proteinases Extraction from Florets

Traditional cardoon rennet is an aqueous extract from *C. cardunculus* flowers. Despite being a simple procedure, it is not standardized, there being significant differences in method and manufacturing conditions among the producers from different Mediterranean regions. It could even be said that each farmhouse or dairy has its own extraction procedure. Nevertheless, different traditional methods have common features. All of them involve soaking the stylets and stigmas of dried flowers in tap water and let it macerate at room temperature. The maceration time may vary from 4 h (Amira et al. 2017b; Ordiales et al. 2016) to 24 h (Roa et al. 1999b; Fernández-Salguero et al. 2002b). The amount of dried flowers used for the coagulant preparation is dependent on the amount of milk and the clotting time desired. As an illustrative example, it would be necessary 50 g of dried flowers to coagulate 100 L of sheep's milk in one hour at 30 °C. After the maceration, the mix is filtered through a cotton piece and the brownish liquor is eventually used to induce the milk coagulation. Cardoon flowers are usually ground in a mortar; however, this process can be carried out dry (Roa et al. 1999b) or wet (Sousa and Malcata 2002), depending on whether it is done before or after mixing with water. If it is done dry, a more intense grinding is possible adding crude kitchen salt, which acts as an abrasive. According to Sousa and Malcata (1996), intense grinding affected positively the proteolytic activity of the extracts, whereas longer homogenization affected it negatively. Although pH has not been considered in traditional preparations, it seems to have a crucial role in the activity of *C. cardunculus* extracts. Their specific action (MCA/PA value) increased considerably with pH drop, exceeding even that of chymosin (Amira et al. 2017b). According to Cavalli et al. (2013), at acid pH ~75–90% of the extracted enzyme activity corresponds to cardosin A, which among APs from *C. cardunculus* is the one that most resembles chymosin. On the contrary, maceration at pH 6 showed the lowest milk-clothing activity and MCA/PA ratio, although the crude extract contained the higher protein concentration and proteolytic activity (Amira et al. 2017b; Sousa and Malcata 1996). Another major factor to consider is the microbiological quality of crude aqueous extracts. The addition of cardoon rennet produces additional contamination of the initial milk and, consequently, of the cheeses (Gómez et al. 2001). Long maceration times, especially at room temperature, may increase the initial microbiological population. Differences in the aforementioned

processing steps involve an extra variability source for cardoon rennet, which may have a significant impact on its specificity and proteolytic activity, and therefore, on the milk-clotting activity and the morpho-organoleptic characteristics of dairy products (Pirisi et al. 2007). Checking the milk-clotting activity of the vegetal rennet may help in standardizing the cheese production. Several methods have been proposed for that purpose, although the Berridge clotting time (Berridge 1945) is used as the indicator of milk flocculation in all the International Dairy Federation Standards for the total milk-clotting activity determination.

C. cardunculus extracts are usually prepared daily, as refrigerated storage (4 °C) decreases significantly its clotting activity (Tavaria et al. 2001). Some authors freeze fresh extracts at −20 °C and store them until further use (Amira et al. 2017b). However, this method is not appropriate for dairy industry, since it involves the freezing of large rennet volumes, which is not cost-effective. Lyophilization might be an alternative option to fresh storage or daily preparation, although it may also affect the proteolytic activity of the vegetal rennet (Sousa and Malcata 2002). With the aim to provide the dairy industry with a more stable and standardized product, several patents for the production of dried powder from vegetal rennet have been published. Most of them dehydrate the original aqueous extract via lyophilization (Roa et al. 1999a; Fernández-Salguero et al. 2002a), air drying or spray drying (Cáceres and Fernández 2000), although previous concentration process (ultrafiltration) may also be considered.

4.7.3 Rennet in Cheesemaking

C. cardunculus extracts are widely used for manufacture of traditional Italian, Spanish and Portuguese cheeses, which are usually characterized as soft creamy texture, and sometimes slightly bitter. Such organoleptic characteristics are the consequence of stronger proteolysis in cheeses made with cardoon rennet than those made with calf rennet. The use of cardoon is especially successful in ewe's milk cheeses. Texture and flavor defects are frequent in cheeses made from cow's milk when cardoon rennet is used, mainly caused by the high proteolytic activity. Barbosa et al. (1976) suggested that different characteristics between caseins from ovine milk and bovine milk might account for differences in production of bitter peptides by proteolysis. However, comparison of cheeses from different species showed that both, primary proteolysis (WSN/TN) and concentration of hydrophobic peptides, during ripening was lower in bovine cheeses than in ovine or caprine cheeses (Sousa and Malcata 2002). Despite this fact, cardoon could satisfactorily replace animal rennet, especially for soft-bodied cheeses (Roseiro et al. 2003). Texture analysis of ewe's milk cheeses coagulated with *C. cardunculus* showed that firmness and consistency decreased along ripening, mostly during the first 30 days, coinciding with the higher proteolysis rate in β-casein (Delgado et al. 2010). Highly significant correlations were found between textural parameters, residual caseins levels and nitrogen fractions during maturation, which show the importance of proteolytic changes for the optimal texture formation in soft cheeses (Delgado et al. 2010; Ordiales et al. 2013). Furthermore, vegetable rennet could be used as a proteinase system in the accelerated ripening of certain types of cheese made from sheep milk (Prados et al. 2007).

Apart from the enzymatic content of the aqueous extract from cardoon flowers, which has proven to be very heterogeneous (Ordiales et al. 2013, 2016), caseins degradation by cardosins could be influenced by other technological parameters. β-caseins are hydrophobic protein, and as such can bind to fat globules, protecting them from enzymatic attack (Tavaria et al. 2001). Thus, cheeses with higher fat content, like ewe's milk cheese, might be less proteolyzed than cow or goat cheeses. Other source of variation in cheese composition, like NaCl content, makes β-casein more resistant to proteolysis via folding or aggregation (Sousa and Malcata 2002). Formation of bitter peptides in cheese during ripening is also influenced by rennet residual activity. Peptides released by residual rennet

activity keep on forming for a longer time in cheese manufactured with cardosin than with chymosin (Gaya et al. 1999). Vegetable rennet appears to be highly stable in cheese during ripening (Roa et al. 1999b); thus, factors that affect coagulant retention in the curd will surely determine bitterness development in cheese. As consequence, the degree of proteolysis in cheese might be modulated adjusting the aforementioned parameters. The concrete details about proteolysis of the different caseins during ripening are still controversial, and further research is needed. According to Delgado et al. (2010), proteolysis in 'Torta del Casar' was slightly more intense during the second month of ripening. In ewe's milk cheese, β-casein proteolysis occurred faster than αs_1-casein (Delgado et al. 2010; Roa et al. 1999b). However, this pattern of degradation of caseins was reversed in other ewe cheeses made with vegetal rennet (Sousa and Malcata 2002). As discussed previously, natural variability, seasonal climatic variations and even enzyme extraction (Amira et al. 2017b) result in a different protease content, which may influence characteristics of the cheese (Heimgartner et al. 1990; Ordiales et al. 2013). Therefore, differences in the results and conclusions obtained might be explained by variances in rennet used for the different studies.

4.8 Use as Ornamental Plant

Various *C. cardunculus* taxa feature a range of plant architecture. The foliage varies in color from green to ash-gray, plant height can reach 3 m, there is a variable amount of branching (Porceddu et al. 1976), and the number of inflorescences (capitula) per plant ranges from 5 to 15 in the globe artichoke and from 30 to 40 in cultivated cardoon. The main capitulum is invariably the largest. The immature capitula are polymorphic with respect to size and shape; some develop spiny outer bracts of various shades of green, which later during development may turn purple (Cravero et al. 2005; Pochard et al. 1969). The mature capitulum can be white or violet. Beyond all the possible cited uses, *C. cardunculus* is also exploited as ornamental, both as garden plants and as cut flowers (Cocker 1967; Sekara et al. 2015); fresh specimens have a long vase life, while dried forms are popular in floral arrangements. The use of the species as ornamental is increasing since, notwithstanding the assortment of ornamental crops is already very large, novelty are constantly in demand by consumers.

The highly heterozygotic nature of the species produces a wide range of phenotypes, especially among progeny of crosses between the different taxa (Cravero et al. 2005; Lopez-Anido et al. 1998; Mauromicale and Ierna 2000). Lanteri et al. (2012) reported on phenotypic diversity released in progenies obtained by crossing a genotype of globe artichoke with both one of cultivated and one of wild cardoon. The two progeny sets showed an astonishing phenotypic variation (Fig. 4.12), with some individuals, the result of specific events of chromosomal segregation and recombination, displaying aspects of morphology not previously observed in either cultivated or wild types. A number of the segregants displayed aspects of morphology which have high potential as ornamentals. On the basis of previous studies (Lanteri et al. 2006), due to the heterozygotic nature of the species, highly segregating progenies may be also obtained following selfing or by crossing spiny with non-spiny globe artichoke varietal types; however, the amount of released phenotypic diversity is by far less pronounced than the one observed in the intertaxa progenies. A particular advantage of this species is that any individual can be readily immortalized by vegetative propagation, either by isolating basal growing shoots or semi-dormant shoots which develop on the underground stem.

The microsatellite (SSR—simple sequence repeat)-based molecular fingerprinting of a wide number of newly generated genotypes was also performed for both sets of progenies, with a particular focus on segregants showing potential for ornamental use. A set of nine SSR loci, evenly dispersed across the genome, was shown to be sufficient to unambiguously identify each segregant obtained by Lanteri et al. (2012). The DNA profiles generated by these SSR assays will have utility in establishing the genetic identity of vegetatively propagated materials.

Fig. 4.12 Examples of the phenotypic variation released in two inter-subspecies hybrid populations. White flower of the cultivated cardoon parent is also included (from Lanteri et al. 2012)

References

Abrantes S, Amaral ME, Costa AP, Duarte AP (2007) *Cynara cardunculus* L. alkaline pulps: alternatives fibres for paper and paperboard production. Bioresour Technol 98:2873–2878

Alexandre AMRC, Dias AMA, Seabra IJ, Portugal AATG, de Sousa HC, Braga MEM (2012) Biodiesel obtained from supercritical carbon dioxide oil of *Cynara cardunculus* L. J Supercrit Fluids 68:52–63

Amira AB, Bauwens J, Pauw ED, Besbes S, Attia H, Francis F, Blecker C (2017a) Identification of proteins from wild cardoon flowers (*Cynara cardunculus* L.) by a proteomic approach. J Chem Biol 10:25–33

Amira AB, Makhlouf I, Flaviu Petrut R, Francis F, Bauwens J, Attia H, Besbes S, Blecker C (2017b) Effect of extraction pH on techno-functional properties of crude extracts from wild cardoon (*Cynara cardunculus* L.) flowers. Food Chem 225:258–266

Angelini LG, Ceccarini L, Nassi o Di Nasso N, Bonari E (2009) Long-term evaluation of biomass production and quality of two cardoon (*Cynara cardunculus* L.) cultivars for energy use. Biomass Bioenergy 33 (5):810–816

Antunes A, Amaral E, Belgacem MN (2000) *Cynara cardunculus* L.: chemical composition and soda-anthraquinone cooking. Ind Crops Prod 12:85–91

Archontoulis SV, Danalatos NG, Struik PC, Batzogiannis D, Savas V (2010a) The effect of nitrogen fertilization and supplemental irrigation on seed and biomass productivity of *Cynara cardunculus* growing in a semi-arid environment in Central Greece. In: Proceedings of the 18th European biomass conference and exhibition, Lyon, France, pp 273–279

Archontoulis SV, Struik PC, Yin X, Bastiaans L, Vos J, Danalatos NG (2010b) Inflorescence characteristics, seed composition, and allometric relationships predicting seed yields in the biomass crop *Cynara cardunculus*. GCB Bioenergy 2:113–129

Ballesteros M, Negro MJ, Manzanares P, Ballesteros I, Sáez F, Oliva JM (2007) Fractionation of *Cynara cardunculus* (Cardoon) biomass by dilute-acid pretreatment. Appl Biochem Biotechnol 137:239–252

Barbosa M, Valles E, Vassal L, Mocquot G (1976) L'utilisation d'extrait de *Cynara cardunculus* L. comme agent coagulant en fabrication de fromages à pâte molle et à pâte cuite. Lait 56:1–17

Barros RM, Malcata FX (2006) Molecular characterization of peptides released from beta-lactoglobulin and alpha-lactalbumin via cardosins A and B. J Dairy Sci 89:483–494

Beck GR (1891) Versuch einer neuen classification der früchte. Verh Zool Bot Ges Wien 41:307–312

Benjelloun-Mlayah B, de Lopez S, Delmas M (1997) Oil and paper pulp from *Cynara cardunculus*: preliminary results. Ind Crops Prod 6:233–236

Bentsen NS, Felby C (2012) Biomass for energy in the European Union—a review of bioenergy resource assessments. Biotechnol Biofuels 5(1):25

Berglund M, Börjesson P (2006) Assessment of energy performance in the life-cycle of biogas production. Biomass Bioenergy 30(3):254–266

Berridge NJ (1945) The purification and crystallization of rennin. Biochem J 39:179–186

Bolohan C, Marin DI, Mihalache M, Ilie L, Oprea AC (2014) Total biomass and grain production of *Cynara cardunculus* L. species grown under the conditions of south easter Romania. Agro Life Sci J 3(1):31–34

Bouriazos A, Ikonomakou E, Papadogianakis G (2014) Aqueous-phase catalytic hydrogenation of methyl esters of *Cynara cardunculus* alternative low-cost non-edible oil: a useful concept to resolve the food, fuel and environment issue of sustainable biodiesel. Ind Crops Prod 52:205–210

Brodelius PE, Cordeiro M, Mercke P, Domingos A, Clemente A, Pais MS (1998) Molecular cloning of aspartic proteinases from flowers of *Cynara cardunculus* subsp. *flavescens* cv. cardoon and *Centaurea calcitrapa*. Adv Exp Med Biol 436:435–439

Cáceres P, Fernández I (2000) Proceso de obtención de extractos en polvo de un coagulante de leche procedente de las flores del cardo *Cynara* sp. Spanish patent ES 2 139 550 A1

Cajarville C, González J, Repetto JL, Rodríguez CA, Martinez A (1999) Nutritive value of green forage and crop by-products of *Cynara cardunculus*. Ann Zootech 8:353–365

Cajarville C, González J, Repetto JL, Alvir MR, Rodríguez CA (2000) Nutritional valuation of cardoon (*Cynara cardunculus*) seed for ruminants. Anim Feed Sci Technol 87:203–214

Carvalho IS, Miranda I, Pereira H (2006) Evaluation of oil composition of some crops suitable for human nutrition. Ind Crops Prod 24:75–78

Castejón D, Herrera A, Heras Á, Cambero I, Mateos-Aparicio I (2017) Oil quality control of culinary oils subjected to deep-fat frying based on NMR and EPR spectroscopy. Food Anal Methods 10(7):2467–2480

Cavalli SV, Lufrano D, Colombo ML, Priolo N (2013) Properties and applications of phytepsins from thistle flowers. Phytochemistry 92:16–32

Cocker H (1967) Il carciofo pianta ornamentale. In: I international congress on artichoke. Edizioni Minerva Medica, Torino, pp 313–317

Cordeiro MC, Jakob E, Puhan Z, Pais MS, Brodelius PE (1992) Milk clotting and proteolytic activities of purified cynarases from *Cynara cardunculus*—a comparison to chymosin. Milchwissenschaft 47:683–700

CORDIS-JOUB0030 (1990) Production and utilization of *Cynara cardunculus* L. biomass for energy, paper-pulp and food industry. Project ID: JOUB0030 (WWW document). http://cordis.europa.eu/project/rcn/237_en.html

Cotana F, Cavalaglio G, Gelosia M, Coccia V, Petrozzi A, Ingles D, Pompili E (2015) A comparison between SHF and SSSF processes from cardoon for ethanol production. Ind Crops Prod 69:424–432

Cravero VP, Picardi LA, Cointry EL (2005) An approach for understanding the heredity of two quality traits

(head color and tightness) in globe artichoke (*Cynara scolymus* L.). Genet Mol Biol 28(3):431–434

Curt MD, Sánchez G, Fernández J (2002) The potential of *Cynara cardunculus* L. for seed oil production in a perennial cultivation system. Biomass Bioenergy 23:33–46

Curt MD, Sánchez G, Fernández J (2005) *Cynara cardunculus* L. as a source of silymarin. Acta Hortic 681:461–467

Curt MD, Sanz M, Aguado PL, Sánchez G, Fernández J (2010) Performance of cardoon oilseed cake in soil substrates. In: Proceedings of the 18th European biomass conference and exhibition, Lyon, France, pp 207–211

Curt MD, Martínez I, Sanz M, Lourenço A, Gominho J, González J, Fernández J (2014) Potential of four selected clones of *Cynara cardunculus* L. for oil production. In: Proceedings of the 22nd European biomass conference and exhibition, Hamburg, Germany, pp 274–279

Dahl J, Obernberger I. (2004) Evaluation of the combustion characteristics of four perennial energy crops (*Arundo donax*, *Cynara cardunculus*, *Miscanthus* × *giganteus* and *Panicum virgatum*). In: 2nd world conference on biomass for energy, industry and climate protection, Rome, Italy, pp 1265–1270

Damartzis T, Vamvuka D, Sfakiotakis S, Zabaniotou A (2011) Thermal degradation studies and kinetic modeling of cardoon (*Cynara cardunculus*) pyrolysis using thermogravimetric analysis (TGA). Bioresour Technol 102(10):6230–6238

De Menna F, Malagnino RA, Vittuari M, Molari G, Seddaiu G, Deligios PA, Solinas S, Ledda L (2016) Potential biogas production from artichoke byproducts in Sardinia, Italy. Energies 9:1–11

De Wit M, Londo M, Faaij A (2011) Productivity developments in European agriculture: relations to and opportunities for biomass production. Renew Sustain Energy Rev 15(5):2397–2412

Del Campo I, Alegria I, Zazpe M, Echeverria M, Echeverría I (2006) Diluted acid hydrolysis pretreatment of agri-food wastes for bioethanol production. Ind Crops Prod 24(3):214–221

Delgado I, Ochoa MJ, Muñoz F, Albiol A (2005) Aprovechamiento invernal para forraje de un cultivo de cardo (*Cynara cardunculus* L.) y valoración de su incidencia sobre la producción de biomasa. In: Producciones agroganaderas: Gestión eficiente y conservación del medio natural. XLV Reunión Científica de la SEEP, Gijón, España, vol II, pp 601–607

Delgado FJ, Rodríguez J, González J, Ramírez R, Roa I (2010) Proteolysis and texture changes of a Spanish soft cheese ('Torta del Casar') manufactured with raw ewe milk and vegetable rennet during ripening. Int J Food Sci Technol 45:512–519

Dressler D, Loewen A, Nelles M (2012) Life cycle assessment of the supply and use of bioenergy: impact of regional factors on biogas production. Int J Life Cycle Assess 17(9):1104–1115

Duca D, Toscano G, Riva G, Mengarelli C, Rossini G, Pizzi A, Del Gatto A, Foppa Pedretti E (2015) Quality of residues of the biodiesel chain in the energy field. Ind Crops Prod 75:91–97

Ek M, Gellerstedt G, Henriksson G (2009) Pulp and paper chemistry and technology. In: Pulping chemistry and technology, vol 2. Walter de Gruyter

Environmental Protection Agency (2012) Global anthropogenic non-CO_2 greenhouse gas emissions: 1990–2030. EPA 430-R-12-006, Dec 2012. Office of Atmospheric Programs Climate Change Division U. S. Environmental Protection Agency, Washington, DC

Fabbri A, Serranti S, Bonifazi G (2014) Biochemical methane potential (BMP) of artichoke waste: the inoculum effect. Waste Manag Res 32:207–214

Fernandes MC, Ferro MD, Paulino AFC, Mendes JAS, Gravitis J, Evtuguin DV, Xavier AMRB (2015) Enzymatic saccharification and bioethanol production from *Cynara cardunculus* pretreated by steam explosion. Bioresour Technol 186:309–315

Fernandes MC, Ferro MD, Paulino AFC, Chaves HT, Evtuguin DV, Xavier AMRB (2018) Comparative study on hydrolysis and bioethanol production from cardoon and rockrose pretreated by dilute acid hydrolysis. Ind Crops Prod 111:633–641

Fernández J, Curt MD (2004) Low-cost biodiesel from cynara oil. In: Van Swaaij WP, Fjällström T, Helm P, Grassi A (eds) Proceedings of the 2nd world conference and exhibition on biomass for energy, industry and climate protection. ETA-Florence & WIP-Munich, Rome, Italy, pp 109–112

Fernández J, Curt MD (2005) State-of-the-art of *Cynara cardunculus* as an energy crop. In: Sjunnesson L, Carrasco JE, Helm P, Grassi A (eds) Proceedings of the 14th European biomass conference, biomass for energy, industry and climate protection. ETA-Florence and WIP-Munich, Paris, France, pp 22–27

Fernández J, Manzanares P (1989) *Cynara cardunculus* L., a new crop for oil, paper pulp and energy. In: 5th European conference on biomass for energy and industry. Elsevier Applied Science, Lisbon, pp 1184–1189

Fernández J, Hidalgo M, del Monte JP, Curt MD (2005) *Cynara cardunculus* L. as a perennial crop for non-irrigated lands: yields and applications. Acta Hortic 681:109–116

Fernández J, Curt MD, Aguado PL (2006) Industrial applications of *Cynara cardunculus* L. for energy and other uses. Ind Crops Prod 24:222–229

Fernández J, Curt MD, Sanz M, Escudero A (2007) Biomass partitioning of cynara as a function of the plant height. In: Maniatis K, Grimm HP, Helm P, Grassi A (eds) Proceedings of the 15th European biomass conference. ETA-Renewable Energies & WIP-Renewable Energies, Berlin, Germany, pp 654–656

Fernández-Salguero J, Gómez R, Tejada L, Vioque M (2002a) Un coagulante vegetal en polvo,

procedimiento para su preparación y sus aplicaciones en la fabricación de quesos. Spanish patent ES 2 166 719 A1

Fernández-Salguero J, Tejada L, Gómez R (2002b) Use of powdered vegetable coagulant in the manufacture of ewe's milk cheeses. J Sci Food Agric 82:464–468

Ferreira-Dias S, Gominho J, Baptista I, Pereira H (2018) Pattern recognition of cardoon oil from different large-scale trials. Ind Crop Prod 118:236–245

Fike JH, Parrish DJ, Wolf DD, Balasko JA, Green JT Jr, Rasnake M, Reynolds JH (2006) Long-term yield potential of switchgrass-for-biofuel systems. Biomass Bioenergy 30(3):198–206

Foti S, Mauromicale G, Raccuia SA, Fallico B, Fanella F, Maccarone E (1999) Possible alternative utilization of *Cynara* spp. I. Biomass, grain yield and chemical composition of grain. Ind Crops Prod 10(3):219–228

Francaviglia R, Bruno A, Falcucci M, Farina R, Renzi G, Russo DE, Sepe L, Neri U (2016) Yields and quality of *Cynara cardunculus* L. wild and cultivated cardoon genotypes. A case study from a marginal land in Central Italy. Eur J Agron 72:10–19

Gaya P, Carrera E, Medina M, Nuñez M (1999) Formation of hydrophobic peptides during the manufacture of ewe's milk Manchego cheese using different milk coagulants. Milchwissenchaft 54:556–558

Genovese C, Platania C, Venticinque M, Calderaro P, Argento S, Scandurra S, Raccuia SA (2016) Evaluation of cardoon seeds presscake for animal feeding. Acta Hortic 1147:323–328

Gherbin P, Monteleone M, Tarantino E (2001) Five year evaluation on cardoon (*Cynara cardunculus* L. var. *altilis*) biomass production in a Mediterranean environment. Ital J Agron 5(1–2):11–19

Gómez R, Sánchez E, Vioque M, Ferreira J, Tejada L, Fernández-Salguero J (2001) Microbiological characteristics of ewes' milk cheese manufactured using aqueous extracts of flowers from various species of cardoon *Cynara* L. Milchwissenschaft 56:16–19

Gominho J, Pereira H (2000) An overview of the research on pulping aptitude of *Cynara cardunculus* L. In: Kyritsis S (ed) 1st world conference on biomass for energy and industry. James & James, Sevilla, pp 1187–1190

Gominho J, Pereira H (2006) Influence of raw-material and process variables in the kraft pulping of *Cynara cardunculus* L. Ind Crops Prod 24:160–165

Gominho J, Fernandez J, Pereira H (2001) *Cynara cardunculus* L.—a new fibre crop for pulp and paper production. Ind Crops Prod 13:1–10

Gominho J, Lourenço A, Curt MD, Fernández J, Pereira H (2009) Characterization of hairs and pappi from *Cynara cardunculus* capitula and their suitability for paper production. Ind Crops Prod 29:116–125

Gominho J, Lourenço A, Palma P, Lourenço ME, Curt MD, Fernández J, Pereira H (2011) Large scale cultivation of *Cynara cardunculus* L. for biomass production—a case study. Ind Crop Prod 33:1–6

Gominho J, Curt MD, Lourenço A, Fernández J, Pereira H (2018) *Cynara cardunculus* L. as a biomass and multi-purpose crop: a review of 30 years of research. Biomass Bioenergy 109:257–275

González JF, González-García CM, Ramiro A, González J, Sabio E, Gañán J, Rodríguez MA (2004) Combustion optimisation of biomass residue pellets for domestic heating with a mural boiler. Biomass Bioenergy 27(2):145–154

González J, Matías L, Royano L, Cabanillas J (2011) Seed production of *Cynara cardunculus* L. cultivars selected in south western Spain. In: Proceedings of the 19th European biomass conference and exhibition, Berlin, Germany, pp 757–760

Grammelis P, Malliopoulou A, Basinas P, Danalatos NG (2008) Cultivation and characterization of *Cynara cardunculus* for solid biofuels production in the Mediterranean region. Int J Mol Sci 9(7):1241–1258

Heimgartner U, Pietrzak M, Geertsen R, Brodelius P, Silva-Figueiredo AC, Pais MSS (1990) Purification and partial characterization of milk clotting proteases from flowers of *Cynara cardunculus*. Phytochemistry 29:1405–1410

Ierna A, Mauromicale G (2010) *Cynara cardunculus* L. genotypes as a crop for energy purposes in a Mediterranean environment. Biomass Bioenergy 34 (5):754–760

Ierna A, Mauro RP, Mauromicale G (2012) Biomass, grain and energy yield in *Cynara cardunculus* L. as affected by fertilization, genotype and harvest time. Biomass Bioenergy 36:404–410

INRA—Institut National de la Recherche Agronomique (2007) Alimentation des bovins, ovins et caprins: besoins des animaux—valeurs des aliments—tables INRA 2007. Editions Quae, INRA, Versailles

Kalamaras SD, Kotsopoulos TA (2014) Anaerobic co-digestion of cattle manure and alternative crops for the substitution of maize in South Europe. Bioresour Technol 172:68–75

Karampinis E, Nikolopoulos N, Nikolopoulos A, Grammelis P, Kakaras E (2012) Numerical investigation Greek lignite/cardoon co-firing in a tangentially fired furnace. Appl Energy 97:514–524

Lag-Brotons A, Gómez I, Navarro-Pedreño J, Mayoral AM, Curt MD (2014) Sewage sludge compost use in bioenergy production—a case study on the effects on *Cynara cardunculus* L. energy crop. J Clean Prod 79:32–40

Lanteri S, Acquadro A, Comino C, Mauro R, Mauromicale G, Portis E (2006) A first linkage map of globe

artichoke (*Cynara cardunculus* var. *scolymus* L.) based on AFLP, S-SAP, M-AFLP and microsatellite markers. Theory Appl Genet 112(8):1532–1542

Lanteri S, Portis E, Acquadro A, Mauro RP, Mauromicale G (2012) Morphology and SSR fingerprinting of newly developed *Cynara cardunculus* genotypes exploitable as ornamentals. Euphytica 184:311–321

Lapuerta M, Armas O, Ballesteros R, Fernández J (2005) Diesel emissions from biofuels derived from Spanish potential vegetable oils. Fuel 84:773–780

Ligero P, Villaverde JJ, Vega A, Bao M (2007) Acetosolv delignification of depithed cardoon (*Cynara cardunculus*) stalks. Ind Crops Prod 25:294–300

Ligero P, Villaverde JJ, Vega A, Bao M (2008) Pulping cardoon (*Cynara cardunculus*) with peroxyformic acid (MILOX) in one single stage. Bioresour Technol 99:5687–5693

Lopez Anido FS, Firpo IT, García SM, Cointry EL (1998) Estimation of genetic parameters for yield traits in globe artichoke (*Cynara scolymus* L.). Euphytica 103:61–66

Lourenço A, Gominho J, Curt MD, Revilla E, Villar JC, Pereira H (2017) Steam explosion as a pretreatment of *Cynara cardunculus* prior to delignification. Ind Eng Chem Res 56:424–433

Maccarone E, Fallico B, Fanella F, Mauromicale G, Raccuia SA, Foti S (1999) Possible alternative utilization of *Cynara* spp. II. Chemical characterization of their grain oil. Ind Crops Prod 10:229–237

Macedo IQ, Faro CJ, Pires EM (1996) Caseinolytic specificity of cardosin, an aspartic protease from the cardoon *Cynara cardunculus* L.: action on bovine αs- and β-casein and comparison with chymosin. J Agric Food Chem 44:42–47

Martínez G, Sánchez N, Encinar JM, González JF (2014) Fuel properties of biodiesel from vegetable oils and oil mixtures. Influence of methyl esters distribution. Biomass Bioenergy 63:22–32

Marzinek J, De-Paula OC, Oliveira DMT (2008) Cypsela or achene? Refining terminology by considering anatomical and historical factors. Rev Bras Bot 31:549–553

Mauro RP, Sortino O, Pesce GR, Agnello M, Lombardo S, Pandino G, Mauromicale G (2015) Exploitability of cultivated and wild cardoon as long-term, low-input energy crops. Ital J Agron 10(1):44–46

Mauromicale G, Ierna A (2000) Characteristics of heads of seed-grown globe artichoke [*Cynara cardunculus* L. var. *scolymus* (L.) Fiori] as affected by harvest period, sowing date and gibberellic acid. Agronomie 20(2):197–204

Mauromicale G, Sortino O, Pesce GR, Agnello M, Mauro RP (2014) Suitability of cultivated and wild cardoon as a sustainable bioenergy crop for low input cultivation in low quality Mediterranean soils. Ind Crops Prod 57:82–89

Mazorra-Manzano MA, Perea-Gutiérrez TC, Lugo-Sánchez ME, Ramirez-Suarez JC, Torres-Llanez MJ, González-Córdova AF, Vallejo-Cordoba B (2013) Comparison of the milk-clotting properties of three plant extracts. Food Chem 141:1902–1907

McSweeney PLH (2007a) Why is the Phe-Met bond of κ-casein so susceptible to rennet action? In: McSweeney PLH (ed) Cheese problems solved. Woodhead Publishing Limited, Abington, pp 52–53

McSweeney PLH (2007b) What enzymes are in rennet. In: McSweeney PLH (ed) Cheese problems solved. Woodhead Publishing Limited, Abington, pp 55–56

Miceli A, De Leo P (1996) Extraction, characterization and utilization of artichoke-seed oil. Bioresour Technol 57:301–302

Mussatto SI, Dragone G, Guimarães PM, Silva JPA, Carneiro LM, Roberto IC, Vicente A, Domingues L, Teixeira JA (2010) Technological trends, global market, and challenges of bio-ethanol production. Biotechnol Adv 28(6):817–830

Neri U, Pennelli B, Simonetti G, Francaviglia R (2017) Biomass partition and productive aptitude of wild and cultivated cardoon genotypes (*Cynara cardunculus* L.) in a marginal land of Central Italy. Ind Crop Prod 95:191–201

Oliet M, Gilarranz M, Domínguez J, Alonso M, Rodríguez F (2005) Ethanol-based pulping from *Cynara cardunculus* L. J Chem Technol Biotechnol 80: 746–753

Oliveira I, Gominho J, Diberardino S, Duarte E (2012) Characterization of *Cynara cardunculus* L. stalks and their suitability for biogas production. Ind Crops Prod 40:318–323

Ordiales E, Benito MJ, Martín A, Casquete R, Serradilla MJ, Córdoba MG (2013) Bacterial communities of the traditional raw ewe's milk cheese "Torta del Casar" made without the addition of a starter. Food Control 33:448–454

Ordiales E, Martín A, Benito MJ, Ruiz-Moyano S, Gallardo G, Córdoba MG (2016) Characterisation of the vegetable rennets used for 'Torta del Casar' cheesemaking by a protein profile method. Int J Dairy Technol 69:272–281

Pari L, Alfano V, Acampora A, Del Giudice A, Scarfone A, Sanzone E (2016) Harvesting and separation of different plant fractions in *Cynara cardunculus* L. In: Barth S, Murphy-Bokern D, Kalinina O, Taylor G, Jones M (eds) Perennial biomass crops for a resource-constrained world. Springer, New York, pp 261–271

Pesce GR (2017) New products from *Cynara cardunculus* L. Doctoral thesis, Università degli Studi di Catania.

Retrieved from http://archivia.unict.it/handle/10761/4000

Pesce GR, Negri M, Bacenetti J, Mauromicale G (2017) The biomethane, silage and biomass yield obtainable from three accessions of *Cynara cardunculus*. Ind Crops Prod 103:233–239

Petropoulos SA, Fernandez A, Calhelha RC, Danalatos N, Barros L, Ferreira ICFR (2018) How extraction method affects yield, fatty acids composition and bioactive properties of cardoon seed oil? Ind Crop Prod 124:459–465

Pimentel C, Van Der Straeten D, Pires E, Faro C, Rodrigues-Pousada C (2007) Characterization and expression analysis of the aspartic protease gene family of *Cynara cardunculus* L. FEBS J 274:2523–2539

Pires E, Faro C, Macedo I, Esteves C, Morgado J, Veríssimo P, Gomes D (1994) Flor do cardo versus quimosina no fabrico de queijos artesanais. Quim Aliment 54(1):66–68

Pirisi A, Pinna G, Addis M, Piredda G, Mauriello R, De Pascale S, Caira S, Mamone G, Ferranti P, Addeo F, Chianese L (2007) Relationship between the enzymatic composition of lamb rennet paste and proteolytic, lipolytic pattern and texture of PDO Fiore Sardo ovine cheese. Int Dairy J 17:143–156

Pirzadah TB, Malik B, Kumar M, Rehman RU (2014) Lignocellulosic biomass: as future alternative for bioethanol production. In: Hakeem K, Jawaid M, Rashid U (eds) Biomass and bioenergy. Springer, Cham, pp 145–163

Piscioneri I, Sharma N, Baviello G, Orlandini S (2000) Promising industrial energy crop. *Cynara cardunculus*: a potential source for biomass production and alternative energy. Energy Convers Manag 41:1091–1105

Pochard E, Foury C, Chambonner D (1969) Il miglioramento genetico del carciofo. In: Atti del 1° Congresso Internazionale di Studi sul Carciofo. Edizioni Minerva Medica, Torino, pp 117–143

Porceddu E, Dellacecca V, Bianco VV (1976) Classificazione numerica di cultivar di carciofo. In: Atti II Congresso Internazionale Carciofo, Bari. Edizioni Minerva Medica, Torino, pp 1105–1119

Portis E, Acquadro A, Tirone M, Pesce GR, Mauromicale G, Lanteri S (2018) Mapping the genomic regions encoding biomass-related traits in *Cynara cardunculus* L. Mol Breed 38(5):64

Poveda P, Tagle JL, Villar JC (1999) Obtencion de pastas al sulfato a partir del cardo (*Cynara cardunculus* L.). Influencia del troceado sobre la calidad de las pastas. Invest Agrar-Sist R 8:305–318

Prados F, Pino A, Fernández-Salguero J (2007) Effect of a powdered vegetable coagulant from cardoon *Cynara cardunculus* in the accelerated ripening of Manchego cheese. Int J Food Sci Technol 42:556–561

Raccuia SA, Melilli MG (2007) Biomass and grain oil yields in *Cynara cardunculus* L. genotypes grown in a Mediterranean environment. Field Crops Res 101:187–197

Raccuia SA, Piscioneri I, Sharma N, Melilli MG (2011) Genetic variability in *Cynara cardunculus* L. domestic and wild types for grain oil production and fatty acids composition. Biomass Bioenergy 35:3167–3173

Ramalho-Santos M, Pissarra J, Verissimo P, Pereira S, Salema R, Pires E, Faro CJ (1997) Cardosin A, an abundant aspartic proteinase, accumulates in protein storage vacuoles in the stigmatic papillae of *Cynara cardunculus* L. Planta 203:204–212

Riahi M (2015) Effet de l´ajout des graines riches en acides gras insaturés ayant des teneurs en fibres et des niveaux de lignification différents sur la réduction des émissions de méthane chez les ruminants in vitro. M. Sc. thesis, Mediterranean Agronomic Institut of Zaragoza, 57 pp

Roa I, González J, Mendiola FJ, Mas M (1999a) Coagulante para leche extraído de *Cynara cardunculus*, procedimiento para su preparación y sus aplicaciones. Spanish patent ES 2 117 958 B1

Roa I, López MB, Mendiola FJ (1999b) Residual clotting activity and ripening properties of vegetable rennet from *Cynara cardunculus* in La Serena cheese. Food Res Int 32:413–419

Romero MJ, Otal J, Lorenzo I, Perez JI (1997) Utilización del cardo (*Cynara cardunculus*) como forraje de invierno en ganado ovino. Composición química y digestibilidad in vivo. Resultados preliminares. ITEA, vol Extra, 18:46–48

Ros M, Franke-Whittle IH, Morales AB, Insam H, Ayuso M, Pascual JA (2013) Archaeal community dynamics and abiotic characteristics in a mesophilic anaerobic co-digestion process treating fruit and vegetable processing waste sludge with chopped fresh artichoke waste. Bioresour Technol 136:1–7

Roseiro LB, Barbosa M, Ames JM, Wilbey RA (2003) Cheesemaking with vegetable coagulants—the use of *Cynara* L. for the production of ovine milk cheeses. Int J Dairy Technol 56:76–85

Sarmento AC, Lopes H, Oliveira CS, Vitorino R, Samyn B, Sergeant K, Debyser G, Van Beeumen J, Domingues P, Amado F, Pires E, Domingues RM, Barros MT (2009) Multiplicity of aspartic proteinases from *Cynara cardunculus* L. Planta 230:429–439

Sękara A, Kalisz A, Gruszecki R, Grabowska A, Kunicki E (2015) Globe artichoke—a vegetable, herb and ornamental of value in central Europe: a review. J Hortic Sci Biotechnol 90(4):365–374

Sengo I, Gominho J, D'Orey L, Martins M, D'Almeida--Duarte E, Pereira H, Ferreira-Dias S (2010) Response surface modeling and optimization of biodiesel production from *Cynara cardunculus* oil. Eur J Lipid Sci Technol 112:310–320

Sharaf-Eldin MA (2016) The effects of electromagnetic treatments on the growth and palmitic acid content of *Cynara cardunculus*. JAPS 26(4):1081–1086

Shatalov AA, Pereira H (2011) Biorefinery of energy crop cardoon (*Cynara cardunculus* L.)—hydrolytic xylose production as entry point to complex fractionation scheme. J Chem Eng Process Technol 2:118

Shatalov AA, Pereira H (2014) Dissolving grade eco-clean cellulose pulps by integrated fractionation of cardoon (*Cynara cardunculus* L.) stalk biomass. Chem Eng Res Des 92:2640–2648

Sousa MJ, Malcata FX (1996) Effects of processing conditions on the caseinolytic activity of proteinases from *Cynara cardunculus* L. Food Sci Technol Int 2:255–263

Sousa MJ, Malcata FX (1998) Proteolysis of ovine and caprine caseins in solution by enzymatic extracts of *Cynara cardunculus*. Enzyme Microb Technol 22:305–314

Sousa MJ, Malcata FX (2002) Advances in the role of a plant coagulant (*Cynara cardunculus*) in vitro and during ripening of cheeses from several milk species. Lait 82:151–170

St-Gelais D, Collet PT (2010) Fromage. In: Vignola CL (ed) Science et technologie du lait. Fondation de Technologie Laitière du Québec Inc. Montréal, Presses Internationales Polytechniques, pp 349–415

Tavaria FK, Sousa MJ, Malcata FX (2001) Storage and lyophilization effects of extracts of *Cynara cardunculus* on the degradation of ovine and caprine caseins. Food Chem 72:79–88

Troch T, Lefébure É, Baeten V, Colinet F, Gengler N, Sindic M (2017) Cow milk coagulation: process description, variation factors and evaluation methodologies. A review. Biotechnol Agron Soc Environ 21 (4):276–287

Vasilakoglou I, Dhima K (2014) Potential of two cardoon varieties to produce biomass and oil under reduced irrigation and weed control inputs. Biomass Bioenergy 63:177–186

Veríssimo P, Esteves C, Faro C, Pires E (1995) The vegetable rennet of *Cynara cardunculus* L. contains two proteinases with chymosin and pepsinlike specificities. Biotechnol Lett 17:621–626

Wellinger A, Murphy JD, Baxter D (2013) Preface. In: Wellinger A, Murphy JD, Baxter D (eds) The biogas handbook: science, production and applications. Woodhead Publishing, pp xxvii–xxviii

Wiesenthal T, Mourelatou A (2006) How much bioenergy can Europe produce without harming the environment? European environmental agency report no 7

Bio-active Compounds and Their Synthetic Pathway

5

Vincenzo Lattanzio, Cinzia Comino, Andrea Moglia and Sergio Lanteri

Abstract

Traditional European medicine has attributed to globe artichoke, as well as to its close relatives cultivated and wild cardoons, many beneficial properties to treat chronic liver diseases, jaundice, hepatitis and arteriosclerosis. Indeed, globe artichoke is a source of bio-active compounds, such as caffeoylquinic acid derivatives (chlorogenic acid and dicaffeoylquinic acid isomers) and sesquiterpene lactones, such as cynaropicrin, grosheimin and its derivatives. Furthermore, globe artichoke roots are rich in inulin, a natural fibre which has been demonstrated to improve the balance of beneficial bacteria in the human gut and it is widely used in food industry to modify texture, replace fat or as low-calorie sweetener. This chapter describes the main bio-active compounds present in globe artichoke as well as their pharmacological properties and potential applications. Furthermore, it reports an updated state of the art on the known steps of their biosynthetic pathway and characterization of genes involved.

5.1 Bio-active Compounds

Globe artichoke immature inflorescences (capitula or heads), with edible fleshy leaves (bracts) and receptacle, as well as other plant organs and tissues, represent a rich source of nutraceuticals and bio-active compounds such as phenolics, terpenoids and polysaccharides (Lattanzio et al. 2009). The edible portion of globe artichoke heads is about 30–40% of its fresh weight, depending on the variety and the harvesting time. This means that the ratio edible fraction/total biomass produced by the plant is very low, being less than 15–20%. This ratio decreases if the contribution to the total biomass represented by offshoots, removed from the field by common cultural procedures, is also considered. Since only the central portion of the immature inflorescences is consumed, the discarded parts of the plant such as leaves, external bracts, stems and offshoot are usually considered a 'waste', although they may represent raw material for the production of food additives and nutraceuticals. Nevertheless, in recent years, some studies have highlighted the possibility to exploit these 'wastes' as a source of phenolics, sesquiterpene lactones, inulin and soluble fibres (Lattanzio

V. Lattanzio
Department of Sciences of Agriculture, Food and Environment, University of Foggia, 71122 Foggia, Italy

C. Comino · A. Moglia · S. Lanteri (✉)
DISAFA, Plant Genetics and Breeding, University of Torino, Largo P. Braccini 2, 10095 Grugliasco, Torino, Italy
e-mail: sergio.lanteri@unito.it

E. Portis et al. (eds.), *The Globe Artichoke Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-030-20012-1_5

et al. 2005; Ruíz-Cano et al. 2014; Fissore et al. 2014).

The use of globe artichoke as well as wild cardoon extracts in folk medicine dates back to the ancient Egyptians, Greeks and Romans, and their hepatoprotective, choleretic, diuretic, inhibiting cholesterol biosynthesis, anti-inflammatory and anti-microbial properties have been widely documented (Adzet et al. 1987; Gebhardt 1998; Bundy et al. 2008).

5.1.1 Phenolics

The phenolics present in globe artichoke extracts, with mono- and dicaffeoylquinic acids being the major components, have shown to possess a wide number of therapeutic properties. In globe artichoke extracts, the presence of caffeic acid depsides, positional isomers of caffeic acid esters of quinic acid, has been well characterized. Starting from the 1950s, several caffeoylquinic derivatives: 1-*O*-caffeoylquinic acid; 3-*O*-caffeoylquinic acid (chlorogenic acid); 4-*O*-caffeoylquinic acid (cryptochlorogenic acid); 5-*O*-caffeoylquinic acid (neochlorogenic acid); 1,3-*O*-dicaffeoylquinic acid; 1,4-*O*-dicaffeoylquinic acid; and 3,5-O-dicaffeoylquinic acid have been identified in globe artichoke extracts (Panizzi and Scarpati 1954; Panizzi et al. 1955; Panizzi and Scarpati 1965; Scarpati et al. 1957; Scarpati and Esposito 1963). The hydroxycinnamic acid moiety of these caffeoylquinic acids is in the *trans* form, although in plant tissues more exposed to UV irradiation, which induces geometric isomerization, *cis* isomers are originated (Towers and Abeysekera 1984).

Panizzi and his research team were the first to isolate and characterize cynarin in globe artichoke leaf extracts. At first, the 1,4-*O*-dicaffeoylquinic acid structure was assigned to cynarin (Panizzi and Scarpati 1954), but later on cynarin structure was identified as 1,5-*O*-dicaffeoylquinic acid (Panizzi and Scarpati 1965). At this regard, it has to be stressed that in the 1950s and 1960s a pre-IUPAC nomenclature for cyclitols (where the positional number of carbon atoms in the quinic acid ring was assigned in an anticlockwise manner) was utilized. In the year 1973 (IUPAC 1976), this nomenclature has been changed and the positional number assigned to the carbon atoms of the quinic acid ring in clockwise sense. As a result, cynarin (formerly 1,5-*O*-dicaffeoylquinic acid) became 1,3-*O*-dicaffeoylquinic acid (Fig. 5.1). Regrettably, in some papers and catalogues a pre-IUPAC numbering is reported or IUPAC numbering is correct but the structure shown is not coherent. To avoid confusion, the IUPAC and pre-IUPAC nomenclatures of caffeoylquinic derivatives identified in globe artichoke extracts are shown in Table 5.1 (see also Clifford et al. 2017).

Lattanzio and his group performed an exhaustive characterization of the caffeoylquinic derivatives present in globe artichoke extracts and identified further caffeoylquinic derivatives, namely 1,3-*O*-dicaffeoylquinic acid; 3,4-*O*-dicaffeoylquinic acid; and 4,5-*O*-dicaffeoylquinic acid (Fig. 5.2). In this regard, it should be stressed that the caffeoylquinic depside level in the plant tissues depends on the physiological stage and ranges from about 8% on dry matter basis in young tissues to less of 1% in senescent tissues. Anyhow, the relative abundances of the different components of the caffeoylquinic derivative fraction are very similar to that shown in Table 5.1 (Lattanzio and Morone 1979; Lattanzio et al. 1989, 1994a).

Caffeoylquinic acids (CQAs) can synergistically or additively provide protection against damage induced by free radicals during oxidative stress and reduce the risk of chronic diseases in humans (Brown and Rice-Evans 1998, Arakawa et al. 2009; Markovic and Tošovic 2016; Puangpraphant et al. 2011). Numerous studies have highlighted that CQAs possess hepatoprotective, diuretic and choleretic activity (Adzet et al. 1987; Gebhardt 1998; Brown and Rice-Evans 1998) and may inhibit HIV integrase, a key enzyme in viral replication and insertion into host DNA (Slanina et al. 2001). However, further studies are required to give adequate scientific support to the growing commercial utilization of globe artichoke active principle as choleretic, hypocholesterolemic and antidyspeptic compounds. The beneficial effects of

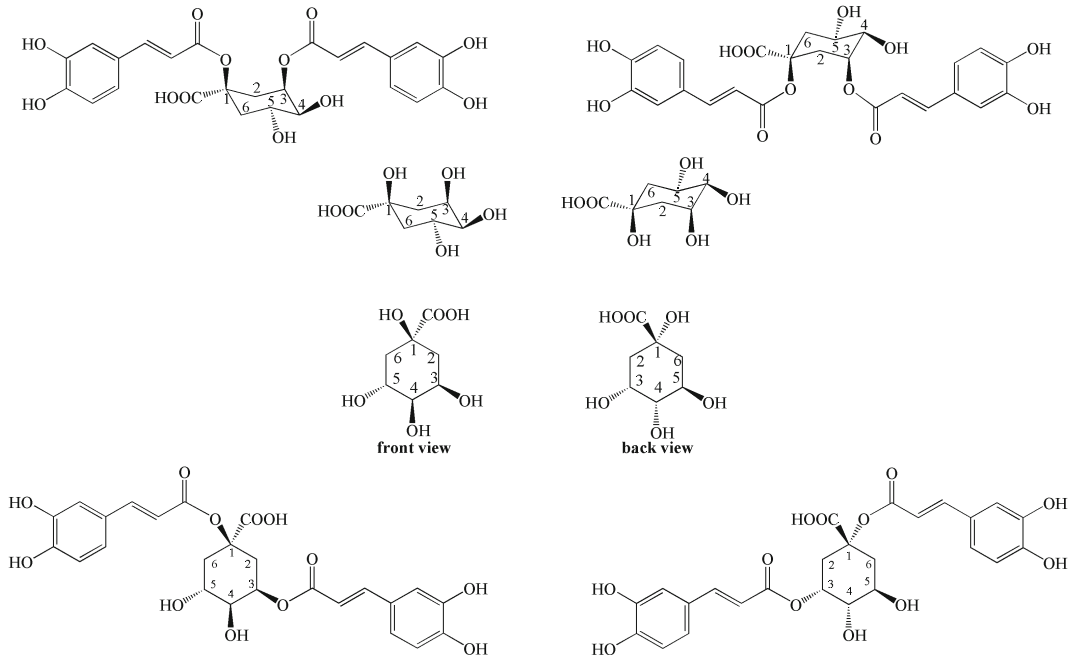


Fig. 5.1 Chemical structure of 1,3-*O*-dicaffeoylquinic acid (cynarin)

Table 5.1 Globe artichoke mono- and dicaffeoylquinic esters

IUPAC recommendations, 1973[a]	Pre-IUPAC numbering	Relative abundance (%)
1-*O*-Caffeoylquinic acid	1-*O*-Caffeoylquinic acid	1.00
3-*O*-Caffeoylquinic acid (neochlorogenic acid)	5-*O*-Caffeoylquinic acid	1.45
4-*O*-Caffeoylquinic acid (cryptochlorogenic acid)	4-*O*-Caffeoylquinic acid	6.76
5-*O*-Caffeoylquinic acid (chlorogenic acid)	3-*O*-Caffeoylquinic acid	38.78
1,3-*O*-Dicaffeoylquinic acid (cynarin)	1,5-*O*-Dicaffeoylquinic acid	1.55
1,4-*O*-Dicaffeoylquinic acid	1,4-*O*-Dicaffeoylquinic acid	3.62
4,5-*O*-Dicaffeoylquinic acid	3,4-*O*-Dicaffeoylquinic acid	5.69
3,5-*O*-Dicaffeoylquinic acid	3,5-*O*-Dicaffeoylquinic acid	8.79
1,5-*O*-Dicaffeoylquinic acid	1,3-*O*-Dicaffeoylquinic acid	21.20
3,4-*O*-Dicaffeoylquinic acid	4,5-*O*-Dicaffeoylquinic acid	10.86

[a]IUPAC (1976)

phenolics in the human body are often attributed to their antioxidant activity, which however is not altogether supported by analytical data highlighting low levels of phenolics found in the blood. The potential beneficial health effects of globe artichoke phenolics are highly dependent upon their uptake from foods, their metabolism and their disposition in target tissues and cells; thus, data concerning in vivo bioaccessibility, bioavailability, absorption, transport and metabolism of ingested phenolics deserve further insights. In this regard, HPLC analyses show that the bioaccessibility of the total caffeoylquinic acid pool is about 58% and that gastrointestinal digestion of globe artichoke extracts induces isomerization in the caffeoylquinic derivative pool as the level of 4-*O*-caffeoylquinic acid and 1,3-*O*-, 1,4-*O*-, 4,5-*O*- and 3,4-*O*-

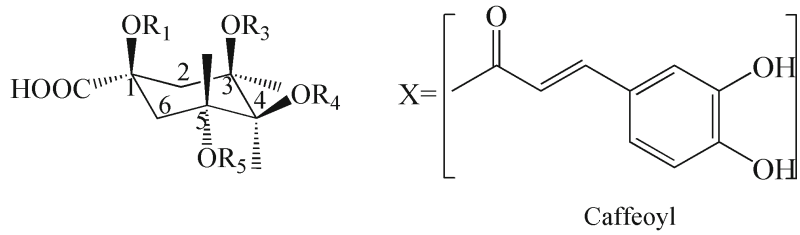


Fig. 5.2 Chemical structure of globe artichoke mono- and dicaffeoylquinic acids

Quinic acid : $R_1 = H$; $R_3 = H$; $R_4 = H$; $R_5 = H$
1-*O*-Caffeoylquinic acid: $R_1 = X$; $R_3 = H$; $R_4 = H$; $R_5 = H$
3-*O*-Caffeoylquinic acid: $R_1 = H$; $R_3 = X$; $R_4 = H$; $R_5 = H$
4-*O*-Caffeoylquinic acid: $R_1 = H$; $R_3 = H$; $R_4 = X$; $R_5 = H$
5-*O*-Caffeoylquinic acid: $R_1 = H$; $R_3 = H$; $R_4 = H$; $R_5 = X$
1,3-*O*-Dicaffeoylquinic acid: $R_1 = X$; $R_3 = X$; $R_4 = H$; $R_5 = H$
1,4-*O*-Dicaffeoylquinic acid: $R_1 = X$; $R_3 = H$; $R_4 = X$; $R_5 = H$
1,5-*O*-Dicaffeoylquinic acid: $R_1 = X$; $R_3 = H$; $R_4 = H$; $R_5 = X$
3,4-*O*-Dicaffeoylquinic acid: $R_1 = H$; $R_3 = X$; $R_4 = X$; $R_5 = H$
3,5-*O*-Dicaffeoylquinic acid: $R_1 = H$; $R_3 = X$; $R_4 = H$; $R_5 = X$
4,5-*O*-Dicaffeoylquinic acid: $R_1 = H$; $R_3 = H$; $R_4 = X$; $R_5 = X$

dicaffeoylquinic acids increases, whereas 3,5-*O*- and 1,5-*O*-dicaffeoylquinic acid content decreases. In addition, it has to be stressed that the majority of these phenolics escape absorption in the small intestine and reach the colon where the microbiota plays a key role in the formation of catabolites (Tomas-Barberan et al. 2014).

In the phenolic fraction of globe artichoke extracts, flavonoid structures have been also identified, although they are considered minor constituents (about 10%) of the total phenolics. Globe artichoke flavonoids belong to flavone (apigenin and luteolin) and anthocyanidins (cyanidin, peonidin and delphinidin) classes (Fig. 5.3). A flavanone structure (naringenin and/or hesperetin) has been also identified by some authors in globe artichoke tissues (Sánchez-Rabaneda et al. 2003; Schütz et al. 2004); however, flavonoids are generally present in plant tissues as glycosides and not as aglycons; thus, these data require additional corroboration.

Flavones are present in both leaves and capitula, while anthocyanidins have been identified in globe artichoke heads only. Again, the flavone content decreases with the age of plant tissues, while the content of anthocyanidins increases in globe artichoke heads during maturation (Lattanzio and van Sumere 1987). Flavone glycosides identified in globe artichoke tissues are luteolin-7-*O*-β-D-glucopyranoside (cynaroside); luteolin-7-*O*-α-L-rhamnosyl(1 → 6)-β-D-glucopyranoside (scolymoside); apigenin-7-*O*-β-D-glucopyranoside and apigenin-7-*O*-α-L-rhamnosyl(1 → 6)-β-D-glucopyranoside. The main anthocyanins identified in globe artichoke heads are: cyanidin 3,5-diglucoside, cyanidin 3-*O*-β-glucoside, cyanidin 3,5-malonyldiglucoside, cyanidin 3-(3″-malonyl)glucoside and cyanidin 3-(6″-malonyl) glucoside. Two peonidin glycosides, namely peonidin 3-*O*-β-glucoside and peonidin 3-(6″-malonyl) glycoside, have been identified as well (Schütz et al. 2006; Lattanzio et al. 2009).

Phenolics in plant tissues play also a key role in plant in response to biotic and abiotic stresses. For example, caffeoylquinic acids are involved in the browning of mechanically damaged (enzymatic browning) and not mechanically injured (formation of iron–chlorogenic acid complexes) tissues. A significant increase of caffeoylquinic acid has been detected when globe artichoke

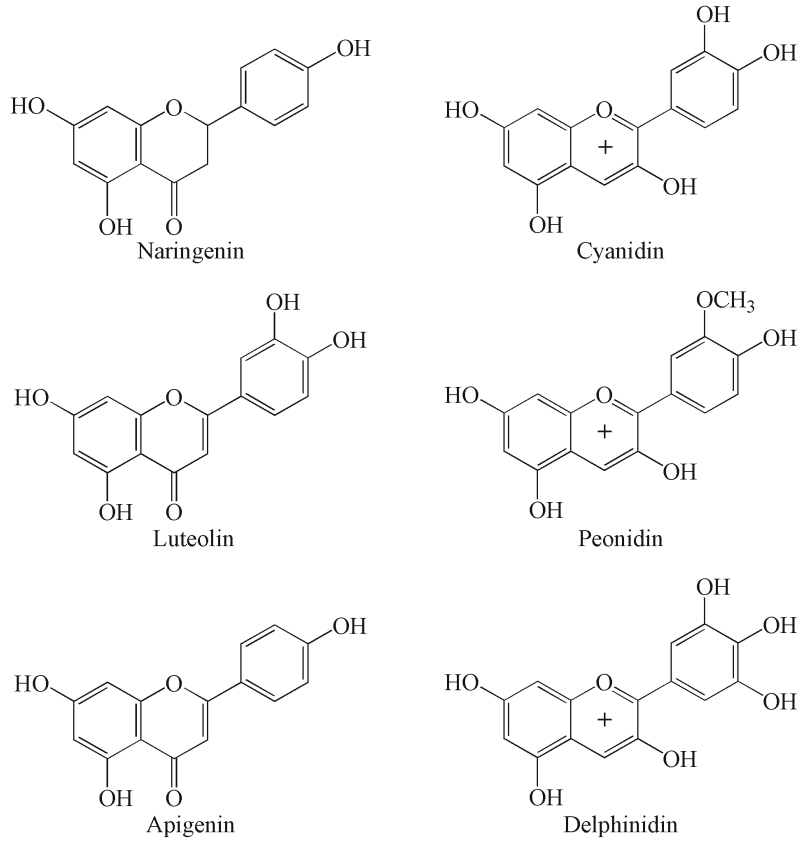


Fig. 5.3 Globe artichoke anthocyanidins

plants are submitted to nutritional or drought stress. In this regard, it has been hypothesized that caffeoylquinic acids and flavonoids may play antioxidant functions in protecting cellular components against the harmful activity of ROS synthesized by stressed tissues. Indeed, endogenous or exogenous applied chlorogenic acid has demonstrated action against *Streptomyces scabies*, *Verticillium albo-atrum*, *Phytophthora infestans* and *Phlyctaena vagabunda* (Lattanzio et al. 1994b; Lattanzio et al. 2001). Anthocyanins are a class of flavonoids providing the red and blue/purple colours common in plant tissues and in flowers of globe artichoke heads, which show different shades of violet. These compounds are synthesized as visual cues, to attract pollinators and other animals for seed dispersal, as well as molecular cues protecting plants from various stress conditions. Flavonoids also play a defensive role against damage by UV-B waveband of solar radiation (280–320 nm). This is due to different mechanisms such as absorption of UV-B radiation by pigments located in the epidermal cells or scavenging of radicals formed by absorption of UV-B photons by antioxidant flavonoids (Rozema et al. 1997; Lattanzio et al. 2012; Lattanzio 2013; Lattanzio et al. 2018).

5.1.2 Sesquiterpene Lactones

The characteristic bitterness of globe artichoke is mainly due to the presence of sesquiterpene lactones (STLs), which represent the terpenoids at higher concentration in its tissues and of which the two major representatives are cynaropicrin and, at a lower concentration, grosheimin and its derivatives (Schneider and Thiele 1974; Cravotto et al. 2005) (Fig. 5.4). Cynaropicrin contributes

to approximately 80% of the characteristic bitter taste of globe artichoke, which is associated with the activation of bitter sensory receptors (Eljounaidi et al. 2015), and can be isolated in gram-scale by employing countercurrent chromatography (Adekenova et al. 2015).

Although the structure of the STLs varies across the *Asteraceae* family, their backbones, derived from the farnesyl diphosphate (FPP), are constrained to a limited set of core skeletons, such as germacranolide, eudesmanolide and guaianolide (Seto et al. 1988; Van Beek et al. 1990; Fischer 1990), of which costunolide is generally considered the common precursor (de Kraker et al. 2002).

In globe artichoke tissues, other terpenoids have been also identified, such as cynaratriol, the grosheimin derivatives: 8-deoxy-11-hydroxy-13-chlorogrosheimin and 8-deoxy-11,13-dihydroxygrosheimin (Fritsche et al. 2002) as well as aguerin B and the sesquiterpene glycosides: cynarascolosides A, B and C (Shimoda et al. 2003).

Cynaropicrin is characterized by a 5-7-5 fused tricyclic skeleton with six carbon stereocenters, four exo-olefins and two hydroxyl groups (Usuki et al. 2014). This compound was at first isolated in globe artichoke by Šorm and co-workers in 1960 (Suchy et al. 1960), and it is considered as a chemotaxonomic marker of the species *Cynara cardunculus* (Chaturvedi 2011), although it has been also identified in other Asteraceae plants such as: *Centaurea solstitialis* L. (Wang et al. 1991), *Hemisteptia lyrata* B., (Ha et al. 2003) and *Saussurea* spp. (Cho et al. 2000, Choi et al. 2005).

In globe artichoke, cynaropicrin has been mainly detected in mature and old leaf tissues, while in heads its content is rather low in the receptacle and undetectable in outer bracts; furthermore, its accumulation is strongly influenced by growth conditions as well as the genotypes in the study (Menin et al. 2012).

Cynaropicrin has been shown to possess various biological activities and pharmacologic as well as antifeedant properties (Elsebai et al. 2016a; Ishida et al. 2010). Because cynaropicrin is a potent antioxidant (Yamada et al. 2015), it can play a supportive role for liver in different hepatic diseases, by acting during the early steps of the hepatitis C virus (HCV) infection (Elsebai et al. 2016b). Cho and co-authors (2000, 2004) found out that cynaropicrin may inhibit the production of inflammatory mediators and the proliferation of lymphocytes, demonstrating its anticancer activity against lymphoma or leukaemia. Tanaka et al. (2012) reported that cynaropicrin extracted from *C. cardunculus* leaves prevents skin photoaging processes and, in addition, it decreases the generation of ROS and the production of inflammatory cytokines in UV-B-irradiated keratinocytes. Hay and colleagues (1994) demonstrated that, due to the presence of α-methylene function, cynaropicrin inhibits the smooth muscle contractility.

STLs are also known for their key roles in plant/environment interactions (Rodriguez et al. 1976; Picman 1986). Cynaropicrin acts as potent feeding deterrent against several species of Lepidoptera (Bhattacharyya et al. 1995) and plays antibacterial activity being a potent, irreversible inhibitor of the bacterial enzyme MurA, which is of vital importance for bacterial cells since it is responsible for the first step in the cytoplasmic biosynthesis of peptidoglycan precursor molecules (Bachelier et al. 2006).

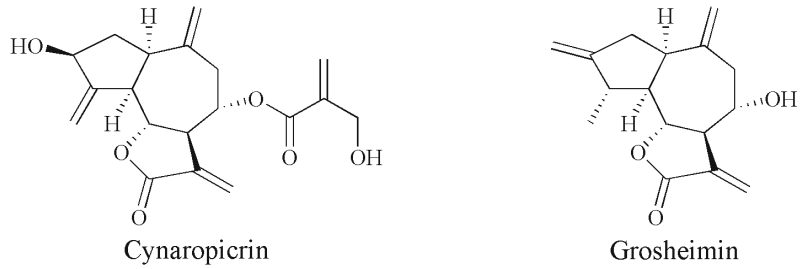


Fig. 5.4 Chemical structures of the two most conspicuous sesquiterpene lactones in globe artichoke

5.1.3 Inulin

Globe artichoke, like other members of the *Asteraceae*, accumulates inulin as a major reserve of carbohydrate in its storage organs such as roots. The edible part of globe artichoke heads is characterized by a high reducing sugar content and a high percentage of inulin, mainly located in the receptacle, which represents about 75% of the total glucidic content. In addition, the inulin content of the edible portion is relatively higher (about 30% more) in globe artichoke heads of marketable quality compared to those at earlier stages of the development (Lattanzio et al. 2009). Inulin (α-D-glucopyranosyl-[β-D-fructofuranosyl] (n-1)-D-fructofuranoside) is a polysaccharide, which possesses a high water solubility and is vacuole-localized. It belongs to the fructose-based polysaccharides called fructans and it is exploited for a wide range of food and non-food applications (Roberfroid and Delzenne 1998). Fructans may have other functions than carbon storage, since they are implicated in protecting plants against water deficit caused by drought or low temperatures, even if a direct correlation between the observed fructan accumulation and the stress tolerance has not uniquely demonstrated. Lastly, it is believed that vacuolar fructans synthesis lowers the sucrose concentration in the cell and prevents sugar-induced feedback inhibition of photosynthesis (Pilon-Smits et al. 1995; Vijn and Smeekens 1999).

Inulin consists of (2 → 1) linked β-D-fructosyl residues, with a terminal (1↔2) α-D-glucose group and a degree of polymerization (DP) up to 150 (Fig. 5.5). The degree of polymerization determines the physicochemical characteristics of inulin and is related to the plant source. Globe artichoke inulin shows an average DP (∼ 50–100) higher than the one present in other Asteraceae such as Jerusalem artichoke, chicory and dahlia. Inulins are not digested in the human upper gastrointestinal tract, which does not contain enzymes (inulinases) able to degrade fructans, and thus are classified as 'non-digestible' carbohydrates. In the colon, they are rapidly fermented to produce short-chain fatty acids (SCFA) that can explain some of the systemic effects of inulins. Inulins stimulate the growth and activity of some desired bacteria (e.g. *Bifidobacterium* and *Lactobacillus*) in the colon, thus positively affecting the composition of the gut microflora and, in turn, improving host health, functioning as a prebiotic, with significantly higher butyrate and propionate production and stimulation of lactic acid-producing bacteria. From a nutritional perspective, inulins have a reduced caloric value and this, combined with its bland to sweet taste, means that they can be used as a low-calorie bulking agent in food, replacing sugar. This property is linked to its particular properties of water solubility. Hydroxyl groups present in the molecular structure of inulin, interacting with water molecules, provide inulin with some surfactant character and the ability to form stable gels with water. These gels provide similar textual characteristics to fat, allowing it to be used to replace fat, resulting in low-fat foods. In addition, research in the field of human nutrition shows that inulins have positive effects on mineral bioavailability and modulate the secretion of gastrointestinal peptides involved in lipid metabolism. Furthermore, they do not lead to a rise in serum glucose or stimulate insulin secretion since when globe artichoke inulin was added to wheat-starch meal of healthy human subjects, the blood glycemic response was low (Rumessen et al. 1990; Roberfroid 2007; Costabile et al. 2010; Mensink et al. 2015). These properties, along with the ability to add texture and improve rheological characteristics and nutritional properties of food, make inulin a functional food ingredient. As far as therapeutic effects of inulin are concerned, it has been observed that dietary inulin, via its fermentation products such as the SCFA butyric and propionic acids, inhibits growth of colon cancer cells. It seems that the beneficial effect of inulin in suppressing tumours may be mediated by an enhancement of gut immune function, by targeting gut-associated lymphoid tissues via increasing secretory immunoglobulin A and interleukin-10 production. Inulin has also shown promising properties in reducing the risk of irritable bowel diseases (ulcerative colitis). Finally, besides their effects on the gastrointestinal tract

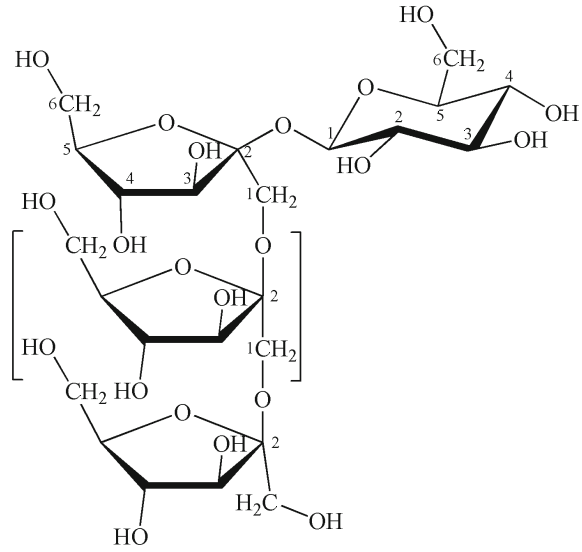


Fig. 5.5 Chemical structure of inulin

diseases, some research showed that inulins contribute to reducing the risk related to systemic diseases (Roller et al. 2004; Roberfroid 2007; Barclay et al. 2010).

5.2 Biosynthetic Pathways

In globe artichoke, the biosynthetic pathways of the main bio-active compounds (i.e. caffeoylquinic and dicaffeoylquinic acids, sesquiterpene lactones and inulin) have been largely studied and the key genes involved have been identified and functionally characterized.

5.2.1 Caffeoylquinic Acids

Caffeoylquinic acids (CQAs) are produced as monoesters (monocaffeoylquinic acids, mono-CQAs, which include chlorogenic acids, CGA) and diesters (dicaffeoylquinic acids, di-CQAs).

In globe artichoke, the biosynthesis of mono-CQAs might occur from: (a) *p*-coumaroyl quinate, synthetized by hydroxycinnamoyl-CoA: shikimate/quinate hydroxycinnamoyltransferase (HCT, EC 2.3.1.133) or hydroxycinnamoyl-CoA: quinate hydroxycinnamoyltransferase (HQT, EC 2.3.1.99), and subsequently hydroxylated by *p*-coumarate-3′-hydroxylase (C3′H, EC 1.14.14.96); and (b) caffeoyl-CoA and quinic acid by means of HQT (Moglia et al. 2013) (Fig. 5.6).

In globe artichoke, the key genes implicated in CQA biosynthesis have been isolated and characterized, in particular, one HCT (NCBI Accession number DQ104740; Comino et al. 2007), three HQTs: HQT1, HQT2 and HQT3 (DQ915589, GU248357, GU248358, respectively; Comino et al. 2009; Menin et al. 2010) and one C3′H (FJ225121; Moglia et al. 2009).

In vitro, characterizations of HCT, HQT1, HQT2 and HQT3 highlighted their involvement in CQA biosynthesis, due to their ability to use either *p*-coumaroyl-CoA or caffeoyl-CoA as an acyl donor and quinic acid as an acceptor to generate, respectively, chlorogenic acid and *p*-coumaroyl quinate (Comino et al. 2007, 2009; Menin et al. 2010).

The in vitro enzymatic activity of C3′H (FJ225121; Moglia et al. 2009) was tested using several substrates at different concentrations, and it was shown that in the presence of *p*-coumaroylshikimate and NADPH, the recombinant protein synthesizes caffeoylshikimate. Moreover, the enzyme was able, although with less efficiency, to convert *p*-coumaroyl quinate into CGA.

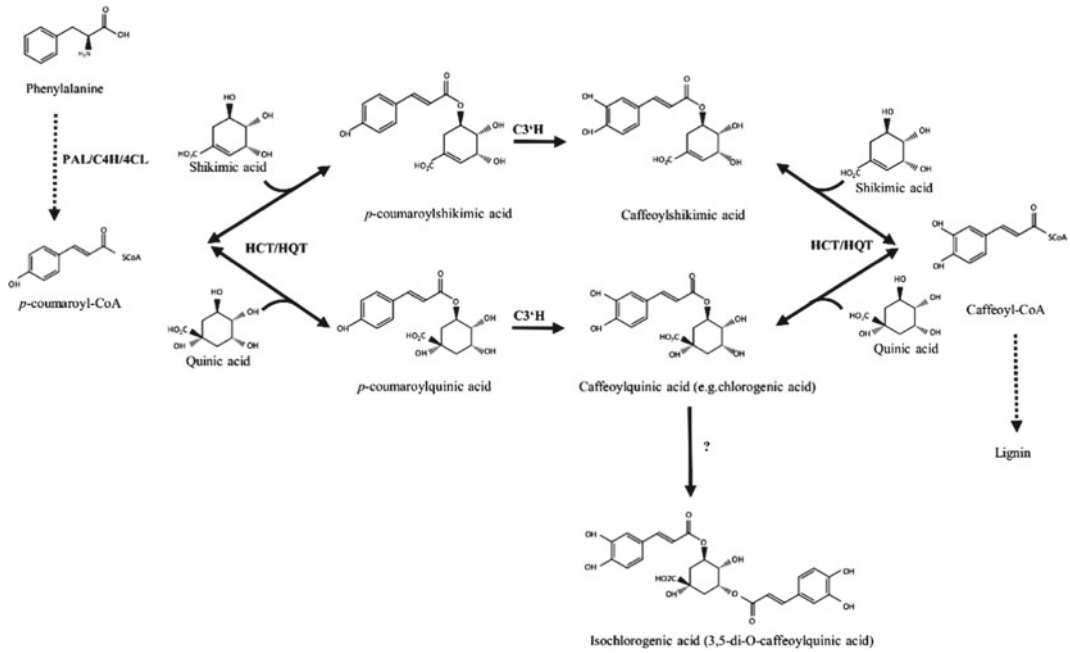


Fig. 5.6 Biosynthetic pathways of CQAs. PAL, phenylalanine ammonia lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-hydroxycinnamoyl-CoA ligase; HCT, hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase; HQT, hydroxycinnamoyl-CoA quinate hydroxycinnamoyl transferase; C3′H, *p*-coumaroyl ester 3′-hydroxylase (from Moglia et al. 2016)

To gain further insight into the relationship between the expression of these genes and the accumulation of CQAs, in vivo functional characterization was carried out through transient and stable over-expression in *Nicotiana benthamiana* and *Nicotiana tabacum*, being globe artichoke recalcitrant to genetic transformation. The *Agrobacterium*-mediated agroinfiltration of *N. benthamiana* leaves with pEAQ expression vector containing HQT1, HQT2 and HQT3 led to a significant increase in mono- (two- to five-fold) and di-CQAs (one- to four-fold) (Moglia et al. 2016). These metabolic results are in accordance with those previously reported (Sonnante et al. 2010) that highlighted an increase in CGA and 1,3-*O*-di-CQAs upon transient and stable over-expression of HQT3 (named HQT1 in that paper) in *N. benthamiana* and *N. tabacum*. Contrary to expectation, *hqt*1-transformed tobacco cell suspensions were found to produce very low levels of CGAs and, as an unintended effect, abscisic acid and benzoic acid derivatives (Mudau et al. 2018).

Virus-induced gene silencing (VIGS) has been routinely used for analysis of gene function in many plant species (Baulcombe 1999; Burch-Smith et al. 2004), due to its simplicity, robustness and circumvention of stable transformation. However, in order to get an efficient single gene silencing, sequences characterized by an identity higher than 22 nt with other genes should be avoided (Gaquerel et al. 2013). Due to the high level of nucleotide identity, this prerequisite was not achievable for HQT2 and HQT3; thus, VIGS was only applied to HQT1. Globe artichoke cotyledons were infiltrated with a suspension of *Agrobacteria* transformed with pTRV2-HQT1. The HQT1 transcript levels in leaves were reduced to 50% as compared to control and its silencing resulted in a significant reduction in content of both CGA and 3,5-di-CQAs (82 and 88% less, respectively).

Over-expression and silencing approaches confirmed the involvement of HQT1, HQT2 and HQT3 in the production of CQAs, supporting the hypothesis that the altered accumulation of dicaffeylquinic acids might be a consequence of enzymatic conversion of CGA acid to di-CQAs.

Recently, Ferro and colleagues (2017) applied the high-resolution melting technology to identify new allelic variants of the C3′H and HQT1 genes in a set of 127 individuals of *C. cardunculus*, including wild cardoon, cultivated cardoon and globe artichoke. The C3′H gene proved to be highly conserved with only 4 haplotypes, while up to 17 haplotypes for HQT1 were de novo identified. Chlorogenic acid content and antioxidant activities were also evaluated in globe artichoke leaf extracts, and an association analysis was performed to assess a putative correlation between these traits and the identified polymorphisms.

5.2.2 Sesquiterpene Lactones

Although the detailed structure of the STLs varies across the *Asteraceae* family, their backbones, derived from the farnesyl diphosphate (FPP), are constrained to a limited set of core skeletons, such as germacranolide, eudesmanolide and guaianolide of which costunolide is generally considered the common precursor.

The biosynthetic pathway of costunolide has been elucidated in globe artichoke and involves three key enzymes (Fig. 5.7): the terpene synthase germacrene A synthase (GAS; Menin et al. 2012) and two cytochrome P450s, the germacrene A oxidase (GAO) and the costunolide synthase (CcCOS; Eljounaidi et al. 2014).

In particular, Menin and colleagues (2012) identified in globe artichoke the GAS gene and demonstrated that the heterologous expression of the protein in *Escherichia coli* converts farnesyl diphosphate (FPP) into β-elemene, which is known to be the Cope rearrangement product of germacrene A (De Kraker et al. 2002), thus confirming its involvement in the initial step for the cynaropicrin biosynthesis.

Later, other two P450s of globe artichoke have been identified, cloned and their sequences compared to the ones previously isolated in other *Asteraceae* species (Eljounaidi et al. 2014). Through their expression in yeast, followed by GC–MS and LC–MS metabolic profiling, their functional characterization was performed and they were identified as GAO and COS genes. In particular, metabolite analyses revealed that the co-expression of globe artichoke GAO together with the chicory GAS (Bouwmeester et al. 2002) resulted in the biosynthesis of germacra-1 (10),4,11(13)-trien-12-oic acid in yeast, an elemene derivative from the corresponding germacrene A derivative. Finally, the co-expression of both globe artichoke GAO and COS with GAS led to biosynthesis of the free costunolide in yeast, demonstrating the involvement of the three genes in biosynthesis of sesquiterpene lactones.

In tobacco agroinfiltration experiments, the enzymatic activity of globe artichoke GAS, GAO and COS was confirmed. Their co-expression led to biosynthesis of the free costunolide in yeast and costunolide conjugates in *N. benthamiana*, demonstrating their involvement in sesquiterpene lactone biosynthesis (Eljounaidi et al. 2014).

Recently, nine haplotypes of GAS have been identified in a collection of *C. cardunculus* from different origins by applying the high-resolution melting technology (Ferro et al. 2018) and the new allelic variants correlated with cynaropicrin content.

5.2.3 Inulin

Globe artichoke like other members of the *Asteraceae* accumulates inulin as a major reserve carbohydrate in their storage organs. The size of the fructosyl polymers deposited in storage organs varies between species, and, as mentioned above, the inulin accumulated in *Cynara scolymus* has a high degree of polymerization.

Inulin is synthesized by the combined action of at least two different fructosyltransferases (Fig. 5.8; Edelman and Jefford 1968). In a first step, the enzyme sucrose: sucrose

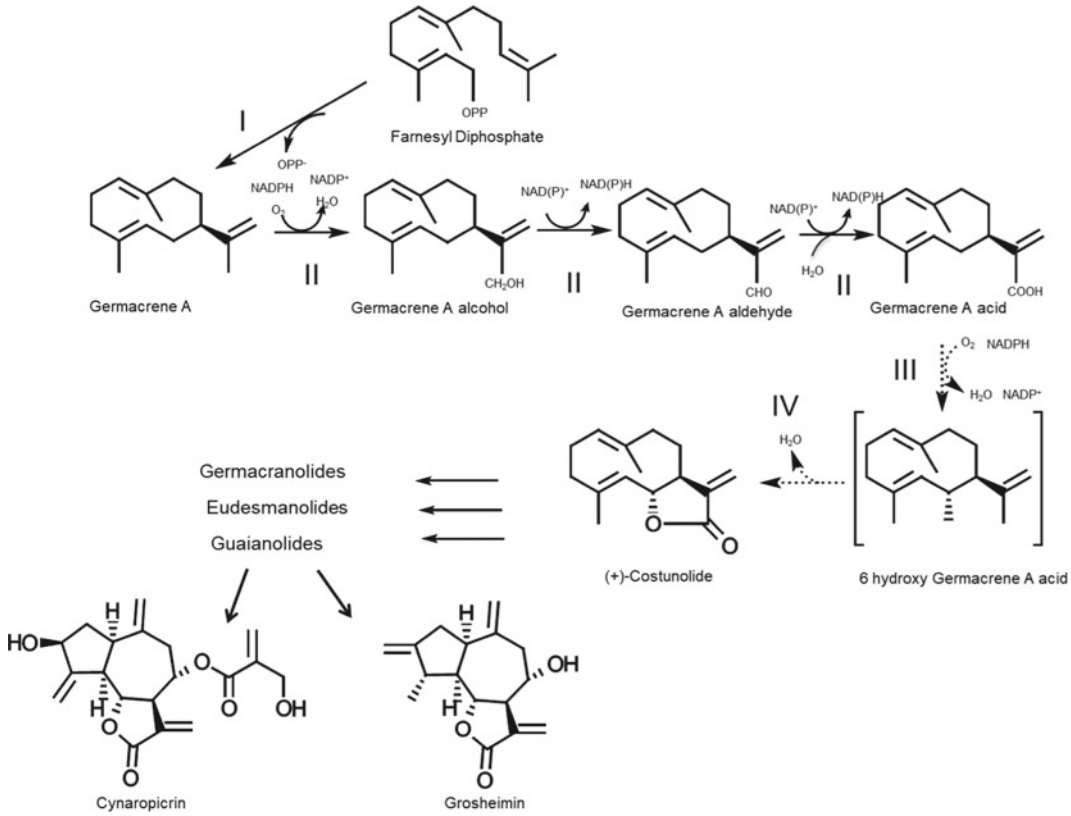


Fig. 5.7 Proposed biosynthetic pathway of STLs. Enzymatic reactions: I cyclization by the germacrene A synthase; II oxidation by the germacrene A oxidase; III hydroxylation by the costunolide synthase; IV subsequent spontaneous lactonization

Fig. 5.8 Biosynthetic pathways for inulin. Enzymes involved are: the sucrose: sucrose 1-fructosyltransferase (1-SST) and the fructan: fructan 1-fructosyltransferase (1-FFT)

1-fructosyltransferase (1-SST) catalyses the formation of the trisaccharide 1-kestose from two molecules of sucrose with the release of glucose; then, the enzyme fructan: fructan 1-fructosyltransferase (1-FFT) mediates the reversible transfer of fructosyl residues between inulins of different chain lengths.

Globe artichokes 1-SST and 1-FFT have been isolated and characterized (Hellwege et al. 1997, 1998). In particular, by screening for fructosyltransferases a cDNA library of globe artichoke (*C. scolymus*) blossom discs, Hellwege and colleagues (1997) isolated a clone designated *Cy21*; its transient expression in *N. tabacum* protoplasts allowed the production of 1-kestose, confirming that *Cy21* codes for a sucrose: sucrose 1-fructosyltransferase (1-SST). Later, the same authors isolated a new cDNA clone, *Cy3*, encoding the fructan: fructan 1-fructosyltransferase (1-FFT) in globe artichoke (Hellwege et al. 1998), and used tobacco protoplasts as expression system. After incubation of protoplast extracts with a mixture of oligofructans, the production of inulins with a high degree

of polymerization was observed, demonstrating that *Cy3* codes for the fructan: fructan 1-fructosyltransferase (1-FFT). In addition, Hellwege and colleagues (2000) demonstrated that expression of globe artichoke 1-SST and 1-FFT genes in transgenic potato plants led to the accumulation of inulin in tubers, with a maximum chain length that is indistinguishable from inulin isolated from globe artichoke roots. These results demonstrated *in planta* that the two enzymes 1-SST and 1-FFT are enough to synthesize inulin molecules of all chain lengths naturally occurring in a given plant species and that the fructan pattern in a given species is mainly defined by the enzymatic characteristics of 1-FFT.

References

Adekenova AS, Sakenova PY, Ivasenko SA, Khabarov IA, Adekenov SM, Berthod A (2015) Gram-scale purification of two sesquiterpene lactones from chartolepsis intermedia boiss. Chromatographia 79:37–43

Adzet T, Camarasa J, Laguna JC (1987) Hepatoprotective activity of polyphenolic compounds from *Cynara scolymus* against CCl_4 toxicity in isolated rat hepatocytes. J Nat Prod 50:612–617

Arakawa T, Yamasaki H, Ikeda K, Ejima D, Naito T, Koyama A (2009) Antiviral and virucidal activities of natural products. Curr Med Chem 16:2485–2497. https://doi.org/10.2174/092986709788682065

Bachelier A, Mayer R, Klein CD (2006) Sesquiterpene lactones are potent and irreversible inhibitors of the antibacterial target enzyme MurA. Bioorganic Med Chem Lett 16:5605–5609

Barclay T, Ginic-Markovic M, Cooper P, Petrovsky N (2010) Inulin—a versatile polysaccharide with multiple pharmaceutical and food chemical uses. J Excip Food Chem 1:27–50

Baulcombe DC (1999) Fast forward genetics based on virus-induced gene silencing. Curr Opin Plant Biol 2 (109–11):3

Bhattacharyya PR, Barua NC, Ghosh AC (1995) Cynaropicrin from *Tricholepis glaberrima*: a potential insect feeding deterrent compound. Ind Crops Prod 4:291–294

Brown J, Rice-Evans C (1998) Luteolin-rich artichoke extract protects low density lipoprotein from oxidation *in vitro*. Free Radic Res 29:247–255. https://doi.org/10.1080/10715769800300281

Bouwmeester HJ, Kodde J, Verstappen FWA, Altug IG, de Kraker JW, Wallaart TE (2002) Isolation and characterization of two germacrene A synthase cDNA clones from chicory. Plant Physiol 129:134–144

Bundy R, Walker AF, Middleton RW, Wallis C, Simpson HCR (2008) Artichoke leaf extract (*Cynara scolymus*) reduces plasma cholesterol in otherwise healthy hypercholesterolemic adults: a randomized, double blind placebo controlled trial. Phytomedicine 15:668–675

Burch-Smith TM, Anderson JC, Martin GB, Dinesh-Kumar SP (2004) Applications and advantages of virus-induced gene silencing for gene function studies in plants. Plant J 39:734–746

Chaturvedi D (2011) Sesquiterpene lactones: structural diversity and their biological activities. In: Opportunity, challenge and scope of natural products in medicinal chemistry. Natural Products Chemistry Division, North-East Institute of Science and Technology (CSIR), Trivandrum, 313–334

Cho JY, Baik KU, Jung JH, Park MH (2000) *In vitro* anti-inflammatory effects of cynaropicrin, a sesquiterpene lactone, from *Saussurea lappa*. Eur J Pharmacol 398:399–407

Cho JY, Kim AR, Joo HG, Kim BH, Rhee MH, Yoo ES et al (2004) Cynaropicrin, a sesquiterpene lactone, as a new strong regulator of CD29 and CD98 functions. Biochem Biophys Res Commun 313:954–961

Choi SZ, Choi SU, Lee KR (2005) Cytotoxic sesquiterpene lactones from *Saussurea calcicola*. Arch Pharmacal Res 28:1142–1146

Clifford MN, Jaganath IB, Ludwig IA, Crozier A (2017) Chlorogenic acids and the acyl-quinic acids: discovery, biosynthesis, bioavailability and bioactivity. Nat Prod Rep 34:1391–1421

Comino C, Hehn A, Moglia A, Menin B, Bourgaud F, Lanteri S, Portis E (2009) The isolation and mapping of a novel hydroxycinnamoyltransferase in the globe artichoke chlorogenic acid pathway. BMC Plant Biol 9:30

Comino C, Lanteri S, Portis E, Acquadro A, Romani A, Hehn A, Larbat R, Bourgaud F (2007) Isolation and functional characterization of a cDNA coding a hydroxycinnamoyltransferase involved in phenylpropanoid biosynthesis in *Cynara cardunculus* L. BMC Plant Biol 7:14

Costabile A, Kolida S, Klinder A, Gietl E, Bauerlein M, Frohburg C, Landschutze V, Gibson GR (2010) A double-blind, placebo-controlled, cross-over study to establish the bifidogenic effect of a very-long chain inulin extracted from globe artichoke (*Cynara scolymus*) in healthy subjects. Br J Nutr 104:1007–1017

Cravotto G, Nano G, Binello A, Spagliardi P, Seu G (2005) Chemical and biological modification of cynaropicrin and grosheimin: a structure–bitterness relationship study. J Sci Food Agric 85:1757–1764

De Kraker JW, Franssen MC, Joerink M, de Groot A, Bouwmeester HJ (2002) Biosynthesis of costunolide, dihydrocostunolide, and leucodin. Demonstration of cytochrome P450-catalyzed formation of the lactone ring present in sesquiterpene lactones of chicory. Plant Physiol 129:257–268

Edelman J, Jefford TG (1968) The mechanism of fructosan metabolism in higher plants as exemplified in *Helianthus tuberosus*. New Phytol 67:517–531

Elsebai MF, Mocan A, Atanasov AG (2016a) Cynaropicrin: a comprehensive research review and therapeutic potential as an anti-hepatitis C virus agent. Front Pharmacol 7:472

Elsebai MF, Koutsoudakis G, Saludes V, Pérez-Vilaró G, Turpeinen A, Mattila S et al (2016b) Pan-genotypic hepatitis C virus inhibition by natural products derived from the wild Egyptian artichoke. J Virol 90:1918–1930

Eljounaidi K, Cankar K, Comino C, Moglia A, Hehn A, Bourgaud F, Bouwmeester H, Menin B, Lanteri S, Beekwilder J (2014) Cytochrome P450s from *Cynara cardunculus* L. CYP71AV9 and CYP71BL5, catalyze distinct hydroxylations in the sesquiterpene lactone biosynthetic pathway. Plant Sci 223:59–68

Eljounaidi K, Comino C, Moglia A, Cankar K, Genre A, Hehn A, Bourgaud F, Beekwilder J, Lanteri S (2015) Accumulation of cynaropicrin in globe artichoke and localization of enzymes involved in its biosynthesis. Plant Sci 239:128–136

Ferro AM, Ramos P, Guerreiro O, Jerónimo E, Pires I, Capel C, Capel J, Lozano R, Duarte MF, Oliveira MM, Gonçalves S (2017) Impact of novel SNPs identified in *Cynara cardunculus* genes on functionality of proteins regulating phenylpropanoid pathway and their association with biological activities. BMC Genom 18:183

Ferro AM, Ramos P, Guerra A, Parreira P, Brás T, Guerreiro O, Jerónimo E, Capel C, Capel J, Yuste-Lisbona FJ, Duarte MF, Lozano R, Oliveira MM, Gonçalves S (2018) Haplotype analysis of the germacrene A synthase gene and association with cynaropicrin content and biological activities in *Cynara cardunculus*. Mol Genet Genomics 293:417–433

Fischer NH (1990) In: Towers GHN, Stafford HA (eds) Biochemistry of the mevalonic acid pathway to terpenoids. Plenum Press, New York

Fissore E, Santo Domingo C, Pujol C, Damonte E, Am Rojas, Gerschenson LN (2014) Upgrading of residues of bracts, stems and hearts of *Cynara cardunculus* L. var. scolymus to functional fractions enriched in soluble fiber. Food Funct 5:463–470

Fritsche J, Beindorff CM, Dachtler M, Zhang H, Lammers JG (2002) Isolation, characterization and determination of minor artichoke (*Cynara scolymus* L.) leaf extract compounds. Eur Food Res Technol 215:149–157

Gaquerel E, Kotkar H, Onkokesung N, Galis I, Baldwin IT (2013) Silencing an N-acyltransferase-like involved in lignin biosynthesis in *Nicotiana attenuata* dramatically alters herbivory-induced phenolamide metabolism. PLoS ONE 8:5

Gebhardt R (1998) Inhibition of cholesterol biosynthesis in primary cultured rat hepatocytes by artichoke (*Cynara scolymus* L.) extracts. J Pharmacol Exp Ther 286:1122–1128

Ha TJ, Jang DS, Lee JR, Lee KD, Lee J, Hwang SW, Jung HJ, Nam SH, Park KH, Yang MS (2003) Cytotoxic effects of sesquiterpene lactones from the flowers of *Hemisteptia lyrata* B. Arch Pharmacal Res 26:925–928

Hay AJ, Hamburger M, Hostettmann K, Hoult JRS (1994) Toxic inhibition of smooth muscle contractility by plant-derived sesquiterpenes caused by their chemically reactive α-methylenebutyrolactone functions. Br J Pharmacol 112:9–12

Hellwege EM, Gritscher D, Willmitzer L, Heyer AG (1997) Transgenic potato tubers accumulate high levels of 1-kestose and nystose: functional identification of a sucrose sucrose 1-fructosyltransferase of artichoke *(Cynara scolymus)* blossom discs. Plant J 12:1057–1065

Hellwege EM, Raap M, Gritscher D, Willmitzer L, Heyer AG (1998) Differences in chain length distribution of inulin from *Cynara scolymus* and *Helianthus tuberosus* are reflected in a transient plant expression system using the respective 1-FFT cDNAs. FEBS Lett 427:25–28

Hellwege EM, Czapla S, Jahnke A, Willmitzer L, Heyer AG (2000) Transgenic potato (*Solanum tuberosum*) tubers synthesize the full spectrum of inulin molecules naturally occurring in globe artichoke (*Cynara scolymus*) roots. PNAS 97:8699–8704

Ishida K, Kojima R, Tsuboi M, Tsuda Y, Ito M (2010) Effects of artichoke leaf extract on acute gastric mucosal injury in rats. Biol Pharm Bull 33:223–229

IUPAC Commission on the Nomenclature of Organic Chemistry (CNOC) and IUPAC-IUB Commission on Biochemical Nomenclature (CBN) (1976) Nomenclature of cyclitols, recommendations. Biochem J 153:23–31

Lattanzio V, Morone I (1979) Variations of the orthodiphenol content in *Cynara scolymus* L. during the plant growing season. Experientia 35:993–994

Lattanzio V, Linsalata V, Palmieri S, Van Sumere CF (1989) The beneficial effect of citric and ascorbic acid on the phenolic browning reaction in stored artichoke (*Cynara scolymus* L.) heads. Food Chem 33:93–106

Lattanzio V, Cardinali A, Di Venere D, Linsalata V, Palmieri S (1994a) Browning phenomena in stored artichoke (*Cynara scolymus* L.) heads: enzymic or chemical reactions? Food Chem 50:1–7

Lattanzio V, De Cicco V, Di Venere D, Lima M, Salerno M (1994b) Antifungal activity of phenolics against different storage fungi. Ital J Food Sci 6(1):23–30

Lattanzio V, Di Venere D, Linsalata V, Bertolini P, Ippolito A, Salerno M (2001) Low temperature metabolism of apple phenolics and quiescence of *Phlyctaena vagabonda*. J Agric Food Chem 49 (12):5817–5821

Lattanzio V, Cicco N, Linsalata V (2005) Antioxidant activities of artichoke phenolics. Acta Hort 681:421–428

Lattanzio V, Kroon PA, Linsalata V, Cardinali A (2009) Globe artichoke: a functional food and source of nutraceutical ingredients. J Funct Foods 1:131–144

Lattanzio V, Cardinali A, Linsalata V (2012) Plant phenolics: a biochemical and physiological perspective. In: Cheynier V, Sarni-Manchado P, Quideau S (eds) Recent advances in polyphenols research, vol 3. Wiley-Blackwell Publishing, Oxford, UK, pp 1–39

Lattanzio V (2013) Phenolic compounds: introduction. In: Ramawat KG, Merillon JM (eds) Handbook of natural products. Springer, Berlin, pp 1543–1580

Lattanzio V, Caretto S, Linsalata V, Colella G, Mita G (2018) Signal transduction in artichoke [*Cynara cardunculus* L. subsp. *scolymus* (L.) Hayek] callus and cell suspension cultures under nutritional stress. Plant Physiol Biochem 127:97–103

Lattanzio V, van Sumere CF (1987) Changes in phenolic compounds during the development and cold storage of artichoke (Cynara scolymus L.) heads. Food Chem 24:37–50

Marković S, Tošović J (2016) Comparative study of the antioxidative activities of caffeoylquinic and caffeic acids. Food Chem 210:585–592. https://doi.org/10.1016/j.foodchem.2016.05.019

Menin B, Comino C, Moglia A, Dolzhenko Y, Portis E, Lanteri S (2010) Identification and mapping of genes related to caffeoylquinic acid synthesis in *Cynara cardunculus* L. Plant Sci 179:338–347

Menin B, Comino C, Portis E, Moglia A, Cankar K, Bouwmeester H, Lanteri S, Beekwilder J (2012) Genetic mapping characterization of the globe artichoke (+)-germacrene A synthase gene, encoding the first dedicated enzyme for biosynthesis of the bitter sesquiterpene lactone cynaropicrin. Plant Sci 190:1–8

Mensink MA, Frijlink HW, van der VoortMaarschalk K, Hinrichs WLJ (2015) Inulin, a flexible oligosaccharide. I: Review of its pharmaceutical applications. Carbohyd Polym 130:405–419

Moglia A, Acquadro A, Eljounaidi K, Milani AM, Cagliero C, Rubiolo P, Genre A, Cankar K, Beekwilder J, Comino C (2016) Genome-wide identification of BAHD acyltransferases and *in vivo* characterization of HQT-like enzymes involved in caffeoylquinic acid synthesis in globe artichoke. Front Plant Sci 7:1424

Moglia A, Comino C, Menin B, Portis E, Acquadro A, Beekwilder J, Hehn A, Bourgaud F, Lanteri S (2013) Caffeoylquinic acids biosynthesis and accumulation in *Cynara cardunculus*: state of the art. Proc. Acta Hort. 983:401–406

Moglia A, Comino C, Portis E, Acquadro A, De Vos RCH, Beekwilder J, Lanteri S (2009) Isolation and mapping of a C3'H gene (CYP98A49) from globe artichoke, and its expression upon UV-C stress. Plant Cell 28(6):963–974

Mudau SP, Steenkamp PA, Piater LA, De Palma M, Tucci M, Madala NE, Dubery IA (2018) Metabolomics-guided investigations of unintended effects of the expression of the hydroxycinnamoyl quinate hydroxycinnamoyltransferase (hqt1) gene from *Cynara cardunculus* var. *scolymus* in *Nicotiana tabacum* cell cultures. Plant Physiol Biochem 127:287–298

Panizzi L, Scarpati ML (1954) Constitution of cynarine, the active principle of the artichoke. Nature 174:1062–1063

Panizzi L, Scarpati ML, Oriente G (1955) Sintesi dell'acido clorogenico. Experientia, XI/10:383–384

Panizzi L, Scarpati ML (1965) Sugli acidi 1, 4- e 1, 5-dicaffeilchinici. Gazz Chim Ital 95:71–82

Picman AK (1986) Biological activities of sesquiterpene lactones. Biochem Syst Ecol 14:255–281

Pilon-Smits EAH, Ebskamp MJM, Paul MJ, Jeuken MJW, Weisbeek PJ, Smeekens SCM (1995) Improved performance of transgenic fructan-accumulating tobacco under drought stress. Plant Physiol 107:125–130

Puangpraphant S, Berhow MA, Vermillion K, Potts G, Gonzalez de Mejia E (2011) Dicaffeoylquinic acids in Yerba mate (*Ilex paraguariensis* St. Hilaire) inhibit NF-κB nucleus translocation in macrophages and induce apoptosis by activating caspases-8 and -3 in human colon cancer cells. Mol Nutr Food Res 55:1509–1522. https://doi.org/10.1002/mnfr.201100128

Roberfroid MB, Delzenne NM (1998) Dietary fructans. Annu Rev Nutr 18:117–143

Roberfroid MB (2007) Inulin-type fructans: functional food ingredients. J Nutr 137:2493S–2502S

Rodriguez E, Towers GHN, Mitchell JC (1976) Biological activities of sesquiterpene lactones. Phytochemistry 15:1573–1580

Roller M, Rechkemmer G, Watzl B (2004) Prebiotic inulin enriched with oligofructose in combination with the probiotics *Lactobacillus rhamnosus* and *Bifidobacterium lactis* modulates intestinal immune functions in rats. J Nutr 134(153–156):2004

Rozema J, van de Staaij J, Bjorn LO, Caldwell MM (1997) UV-B as an environmental factor in plant life: stress and regulation. Trends Ecol Evol 12:22–28

Ruíz-Cano D, Pérez-Llamas F, Frutos MJ, Arnao MB, Espínosa C, López-Jiménez JA, Castillo J, Zamora S (2014) Chemical and functional properties of the different byproducts of artichoke (*Cynara scolymus* L.) from industrial canning process. Food Chem 160:34–140

Rumessen JJ, Bode S, Hamberg O, Gudman-Hoyer E (1990) Fructans of Jerusalem artichokes: intestinal transport, absorption, fermentation and influence on blood glucose, insulin and C-peptide in healthy subjects. Am J Clin Nutr 52:675–681

Sánchez-Rabaneda F, Jáuregui O, Lamuela-Raventós RM, Bastida J, Viladomat F, Codina C (2003) Identification of phenolic compounds in artichoke waste by high-performance liquid chromatography-tandem mass spectrometry. J Chromatogr A 1008:57–72

Scarpati ML, Oriente G, Panizzi L (1957) Sui costituenti caffeici nel carciofo. Anal Chim 47:150–154

Scarpati ML, Esposito P (1963) Neochlorogenic acid and "band 510" structure. Tetrahedron Lett 18:1147–1150

Schneider G, Thiele K (1974) Die Verteilung des Bitter-stoffes Cynaropicrin in der Artischocke. Planta Med 26:174–183

Schütz K, Kammerer D, Carle R, Schieber A (2004) Identification and quantification of caffeoylquinic acids and flavonoids from artichoke (*Cynara scolymus* L.) heads, juice, and pomace by HPLC-DAD-ESI/MSn. J Agric Food Chem 52:4090–4096

Schütz K, Persike M, Carle R, Schieber A (2006) Characterization and quantification of anthocyanins inselected artichoke (*Cynara scolymus* L.) cultivars by HPLC-DAD ESI-MSn. Anal Bioanal Chem 384:511–1517

Seto M, Miyase T, Umehara K, Ueno A, Hirano Y, Otani N (1988) Sesquiterpene lactones from *Cichorium endivia* L. and *C. intybus* L. and cytotoxic activity. Chem Pharm Bull 36:2423–2429

Shimoda H, Ninomiya K, Nishida N, Yoshino T, Morikawa T, Matsuda H et al (2003) Anti-hyperlipidemic sesquiterpenes and new sesquiterpene glycosides from the leaves of artichoke (*Cynara scolymus* L.): structure requirement and mode of action. Bioorganic Med. Chem. Lett. 13:223–228

Slanina J, Taborska E, Bochorakova H, Slaninova I, Humpa O, Robinson W, Schram K (2001) New and facile method of preparation of the anti-HIV-1 agent, 1, 3-dicaffeoylquinic acid. Tetrahedron Lett 42:3383–3385. https://doi.org/10.1016/S0040-4039(01)00448-8

Sonnante G, D'Amore R, Blanco E, Pierri CL, De Palma M, Luo J, Tucci M, Martin C (2010) Novel hydroxycinnamoyl-Coenzyme A quinate transferase genes from artichoke are involved in the synthesis of chlorogenic acid. Plant Physiol 153(3):1224–1238

Suchy M, Herout V, Šorm F (1960) Terpenes. CXVI. Structure of cynaropicrin. Collect Czech Chem Commun 25:2777–2782

Tanaka YT, Tanaka K, Kojima H, Hamada T, Masutani T, Tsuboi M, Akao Y (2012) Cynaropicrin from *Cynara scolymus* L. suppresses photoaging of skin by inhibiting the transcription activity of nuclear factor-kappa B. Bioorg Med Chem Lett 23(2):518–523

Tomas-Barberan F, Garcia-Villalba R, Quartieri A, Raimondi S, Amaretti A, Leonardi A, Rossi M (2014) *In vitro* transformation of chlorogenic acid by human gut microbiota. Mol Nutr Food Res 58:1122–1131

Towers GHN, Abeysekera B (1984) Cell wall hydroxycinnamate esters as UV-A receptors in phototropic responses of higher plants—a new hypothesis. Phytochemistry 23(5):951–952

Usuki T, Sato M, Hara S, Yoshimoto Y, Kondo R, Zimmermann S, Kaiser M, Brun R, Hamburger M, Adams M (2014) Antitrypanosomal structure-activity-relationship study of synthetic cynaropicrin derivatives. Bioorg Med Chem Lett 24:794–798

Van Beek TA, Maas P, King BM, Leclercq E, Voragen AG, De Groot A (1990) Bitter sesquiterpene lactones from chicory roots. J Agric Food Chem 38:1035–1038

Vijn I, Smeekens S (1999) Fructan: more than a reserve carbohydrate? Plant Physiol 20:351–359

Wang Y, Hamburger M, Cheng CH, Costall B, Naylor RJ, Jenner P, Hostettmann K (1991) Neurotoxic sesquiterpenoids from the yellow star thistle *Centaurea solstitialis* L. (Asteraceae). Helv Chim Acta 74:117–123

Yamada K, Ishii Y, Takeda T, Kuroki H, Mitoma C, Uchi H et al (2015) Effect of cynaropicrin on 2,3,4,7,8-pentachlorodibenzofuran-induced wasting syndrome and oxidative stress. Fukuoka Igaku Zasshi 106:169–175

Genetics and Breeding

6

Ezio Portis, Alberto Acquadro and Sergio Lanteri

Abstract

Genetic variation within globe artichoke landraces contributes to the variable yield performance detected in cultivation. It was built up over many generations of vegetative propagation as a consequence of the limited selection applied by farmers as well as the accumulation of mutations. This chapter reports the results obtained by applying clonal selection programmes within local landraces, with the goal to identify genotypes retaining desirable traits and able to ensure a higher and stable yield. The selected high-yielding genotypes while contributing to increase productivity may also implement the conservation 'in situ' of valuable germplasm at risk of extinction. With the goal to move to breeding programme based on crossing strategies, it is pivotal the development of genetic maps with high levels of genome coverage. Genetic maps represent the first step for localizing genes or quantitative trait loci (QTL) that are associated with economically important trait and open the way for marker-assisted selection, comparative mapping between different species, a framework for anchoring physical maps and the basis for map-based cloning of genes. In this chapter, the history of the development of genetic maps in *Cynara cardunculus* is retraced, starting from the first maps based on a limited number of mostly dominant molecular markers, up to the most recently developed maps allowing the identification of genomic region influencing key agronomic traits.

6.1 Clonal Selection and Molecular Fingerprinting

The application of PCR-based markers to assess the genetic variation between and within populations of clonally propagated globe artichoke varietal types (Lanteri et al. 2001; Portis et al. 2005; Ciancolini et al. 2012) has highlighted their multi-clonal composition, which is less evident but also detectable at the phenotypic level. The latter was mainly attributed to the limited selection criteria adopted by farmers since, after some years of cultivation, plants are mown at the end of the growing season and the propagative material used for replanting is collected some months later, without any selection of the mother plants. An additional source of variation is presumably due to spontaneous mutations occurred over time after repeated cycles of cloning. These findings have stimulated the application of clonal selection programmes,

E. Portis · A. Acquadro · S. Lanteri (✉)
DISAFA Plant Genetics and Breeding, University of Torino, Largo P. Braccini 2, 10095 Grugliasco, Torino, Italy
e-mail: sergio.lanteri@unito.it

E. Portis et al. (eds.), *The Globe Artichoke Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-030-20012-1_6

with the goal to identify superior genotypes more productive and/or possessing high qualitative traits of the capitula.

Lanteri et al. (2004) applied a clonal selection programme within the varietal type 'Spinoso sardo' grown in Sardinia Island. A sample of 500 ovoli from as many plants was randomly collected from five populations grown in areas characterized by variable pedo-climatic conditions. The plants obtained were characterized for morphology, production and earliness over two growing seasons. A set of 24 early and more productive plants were identified and, from each of them, five plants were clonally propagated. The resulting 120 plants (24 genotypes each represented by 5 plants) were genotyped by applying 26 inter-simple sequence repeat (ISSR) primers as well as 31 amplified fragment length polymorphism (AFLP) primer combinations. The ISSR markers disclosed an extremely low level of polymorphism, while the UPGMA cluster analysis based on AFLP differentiated each selected clone and demonstrated their suitability for their fingerprinting. In the study, the AFLP markers were also used to assess genetic uniformity within each clone and, from just three plants of one clone, an electrophoretic pattern polymorphic for one or two bands was detected. Since the variant pattern was observed on the same plants of the clone with at least two primer combinations, while the other plants were 'true to clone', the authors excluded possible mis-labelling during clonal propagation or laboratory errors. Otherwise, they hypothesized that the mother plant from which the clone was obtained was a chimera not phenotypically detectable but maintained over time through vegetative propagation.

Mauro et al. (2012) performed a clonal selection within the varietal type 'Violetto di Sicilia', whose cultivation is spread in the eastern area of Sicily Island. Four populations were identified in as many locations representative of the growing area and, in each of them, propagative material was sampled from seven to ten plants. The latter were chosen on the basis of the number of floral stem ramifications (an index of yield potential), earliness as well as proper head pigmentation, shape and thickness. On the whole, 36 genotypes were selected from which 3 to 10 semi-dormant offshoots were collected. The plants obtained were characterized for morphological traits and yield during the two following growing seasons, and this led to discard 26 clones. The remaining 10, including at least one clone for each site, were further on propagated with the goal to obtain at least 20 plants per clone and then characterized for a number of head traits and number of floral stem ramifications during further two growing seasons. All traits were significantly variable and for the most part subjected to a marked seasonal influence. The mean production of the selected clones was about nine heads per plant, corresponding to a fresh weight around 1.3 kg. However, two clones produced an average of 10.6 heads, equivalent to 1.46 kg per plant over both growing season, while two produced the largest heads (on average 165 g), whose receptacle incidence reached on average 19.3 g for 100 g of fresh weight. In the study, a moderate, broad-sense heritability was detected for total yield, number of heads per plant and receptacle incidence, which means these traits represent the primary target for clonal selection within this varietal type. The ten selected clones together with ten genotypes of 'Violetto di Sicilia', which in a previous study were identified by the authors as representatives of the genetic variation within this varietal type, were genotyped by applying seven AFLP primer combinations, and just three were able to fingerprint each clone. Furthermore, three clones were found to possessed unique fragments, which the authors suggested convertible into sequence tagged site assays for their easy identification in nurseries. Interestingly, the principal component analysis based on molecular data did not group separately the ten selected clones and the other ten genotypes included in the study, demonstrating that the former are thus representative of the genetic variation present in cultivation. This result appeared to confirm that the clonal selection within 'Violetto di Sicilia' may provide higher productive genotypes without causing genetic erosion and promoting its preservation in situ. Lastly, the authors highlighted only a limited

correlation between phenotypic and the genotypic data and hypothesized that this was so because most of the AFLP loci were sited in the non-coding portion of the genome and the traits in the study were much affected by environmental conditions.

With an approach analogous to the one previously described, Mauro et al. (2015) performed a clonal selection programme within another Sicilian varietal type, 'Spinoso di Palermo', which is mainly grown in the western area of Sicily Island. Within five populations grown in as many locations, from three to eight plants were selected on the basis of the number of floral stem ramifications as well as earliness in the production of heads and their number, size, colour and firmness. A set of 30 selected plants was clonally propagated and from 8 to 10 plants per clone obtained. Over the following two growing seasons, 25 clones were discarded and 5 maintained. The number of plants per selected clone was then increased to 54 for each clone (for a total of 270 plants) by transplanting their lateral offshoots, with the goal to perform a more reliable morphological characterization during two further growing seasons. A set of 54 plants of the allochthonous and early varietal type 'Violet de Provence' was also included in the study, since being characterized by high yields and long productive cycle is spreading in Southern Italy and endangers native local landraces. The morphological characterization of the whole set of plants was based on 51 traits, according to the guidelines provided by International Union for the Protection of New Varieties of Plants (UPOV). Three of the evaluated traits, e.g. days to first harvest, duration of harvest period and rate of yield (expressed as the ratio yield and harvest duration), were found significantly affected by the clone × season interaction, while the weight of main heads proved to be the most stable. The head yield and number of heads per plant were found associated with a moderate level of broad-sense heritability, suggesting that they represent key traits for primary selection criteria within this landrace. Eighteen out of the 51 scored traits did not differentiate the varietal types 'Spinoso di Palermo' and 'Violet de Provence', while other 18 were distinctive of the 'Spinoso di Palermo' clones. The latter produced an average number of heads of 13.8 per plant which corresponded to a weight of 2 kg/plant, a value higher than the one detected in 'Violet de Provence'. Two clones produced up to 2.4 kg of capitula per plant and three noticeably larger second-order heads (mean of 156 g), thus resulting suitable for the production of commercially valuable heads over a prolonged harvesting period.

Two plants for each selected clone as well as 12 genotypes previously representative of genetic variation within 'Spinoso di Palermo' and two plants of Violetto di Sicilia (used as a reference) were genotyped with seven AFLP primer combinations. The UPGMA dendrogram obtained from AFLP data clustered separately the two varietal types, and just three out of the five primer combinations were able to discriminate each clone. In this study, differently from what reported by the same authors for the varietal type 'Violetto di Sicilia' (Mauro et al. 2012), the grouping of entries generated by AFLP analyses was highly consistent with the grouping based on morphological variation, presumably as a consequence of the greater accuracy in phenotyping the material under study.

More recently, Crinò and Pagnotta (2017) performed the morphological, biochemical and molecular characterization of ten clones, result of a selection programme (Crinò et al. 2008), and representative of the 'Romanesco' varietal type which is mainly grown in Latium region of Central Italy. The ten clones were characterized by early medium and late head production. Twelve plants per clone were phenotyped, over two subsequent growing seasons, on the basis of 47 traits including globe artichoke standard UPOV descriptors and complementary Romanesco-type descriptors reported by Ciancolini et al. (2012). Morphological differences among clones were detected for 15 traits, among which shape of the head (ranging from elongated to elliptical), receptacle shape, head pigmentation (ranging from purple striped to green) and head yield, while 8 traits showed significant clone-per-year interactions highlighting their unsuitability as standard descriptors. The

similarity dendrogram based on hierarchical cluster analysis of morphological data identified three main clusters while principal component analysis highlighted that four traits (i.e. plant height, central flower-head weight, earliness and total flower-head weight) made it possible a clear grouping of the clones.

One plant of each clone was also genotyped with seven AFLP primer combinations, four ISSR and thirteen SSR markers, and 393 molecular polymorphisms were detected. Molecular assessment was allowed to discriminate each clone, and the UPGMA dendrogram based on Nei's genetic distances identified three main clusters which, however, were not consistent with the ones obtained on the basis of morphological variation.

During the first growing season, at least nine primary flower-heads from each clone were collected at the commercial stage and their polyphenol profile obtained on the basis of HPLC. The authors reported that differently from what previously found by Pandino et al. (2015) in leaves of a genotype of 'Romanesco C3', cynarin (1,3-di-O-CQA), apigenin and luteolin were not detected in any of the sample in study. On the other hand, significant differences were detected among clones with respect to the total polyphenol content as well as 1,5-O-dicaffeoylquinic acid, chlorogenic acid and cynaroside (Luteolin 7-O-glucoside).

The result of the study made it possible the identification of two clones, characterized by remarkable earliness and highly productive, which have been included with the names 'Michelangelo' and 'Raffaello' in the National Register of Varieties of the Italian Ministry of Agricultural, Food and Forestry Policies.

The results obtained in these studies highlight that clonal selection is effective in identifying superior genotypes within the population of globe artichoke varietal types which at present are vegetatively propagated. The selected genotypes may represent the starting material for nursery activities, but undoubtedly, a limit is represented by the time needed to produce an adequate quantity of their propagative material for field cultivation. This, together with the fact that the use of the achene as reproductive unit makes it possible to treat the crop as annual and reduce the cost of planting and the diffusion of pathogens, has led to the growing in popularity of seed-propagated cultivars and F_1 hybrids. However in a recent paper, Mauromicale et al. (2018) suggested a selection scheme which combines mass phenotypic selection, self-pollination and marker-assisted selection and suggested its application for the development of open-pollinated varieties from vegetatively propagated landraces. Studies shepherded in this direction are at present very limited; however, the results reported by the authors seem very promising and demonstrate that a practical and cheap means of conducting marker-assisted breeding might be easily adopted also by small seed companies furnished with laboratories equipped for basic molecular analyses and which carry commercial activities at level of local markets. This might provide a significant contribution to their income as well as counteract the ongoing drastic genetic erosion of local germplasm characterized by peculiar qualitative and organoleptic characteristics.

6.2 Genetic Maps

Cynara cardunculus mapping populations were, at first, developed by crossing a clone of the globe artichoke varietal non-spiny-type 'Romanesco C3', used as female parent, with (i) an early-maturing varietal spiny-type 'Spinoso di Palermo', (ii) a clone of cultivated cardoon 'Altilis 41' and (iii) a clone of the wild cardoon 'Creta 4'. The three progenies obtained were highly segregating for both capitula and plant traits; furthermore, high segregation was detected for the content in phenolic esters and sesquiterpene lactones. A number of the segregants displayed aspects of morphology which might also have high potential as ornamentals, taking into consideration that any individual can be readily immortalized by vegetative propagation, as well maintained over time through 'in vitro' micropropagation (Lanteri et al. 2012).

The first genetic maps of globe artichoke were generated following genotyping of a progeny of 94 individuals, obtained by crossing the two globe artichoke clones 'Romanesco C3' and 'Spinoso di Palermo', with AFLP, M-AFLP and S-SAP markers (Lanteri et al. 2006). The female map comprised of 204 loci, spread over 18 linkage groups (LGs) and spanned 1330.5 cM with a mean marker density of 6.5 cM. The equivalent figures for the male parent map were 180 loci, 17 LGs, 1239.4 cM and 6.9 cM. The presence of 78 loci in common to both maps allowed for the alignment of 16 of the linkage groups. The maps were subsequently extended by the inclusion of a number of microsatellite loci, of which 19 were represented in both maps (Acquadro et al. 2009) and three genes involved in the synthesis of caffeoylquinic acid have also been positioned on the maps (Comino et al. 2007, 2009; Moglia et al. 2009).

A further map was developed by Portis et al. (2009a) on the progeny obtained by crossing the clone of 'Romanesco C3' with the cultivated cardoon 'Altilis 41'. The F_1 population was genotyped using a variety of PCR-based marker platforms (AFLP, S-SAP and SSR), resulting in the identification of 708 testcross markers suitable for map construction. The male map consisted of 177 loci arranged in 17 major linkage groups and spanning 1015.5 cM; while the female map was built with 326 loci arranged into 20 major linkage groups, spanning 1486.8 cM. The presence of 84 loci shared between these maps and those previously developed by Lanteri et al. (2006) allowed for map alignment and the definition of 17 homologous linkage groups, corresponding to the haploid number of the species. Furthermore, as 25 mapped markers (8%) corresponded to coding regions, this map had an additional value as functional map and represented the first genetic tool for candidate gene studies in globe artichoke. Later on, the maps have been integrated with the inclusion of all the genes involved in the synthesis of caffeoylquinic acids known in the species (Menin et al. 2010).

Later on, the same authors produced an improvement of the maps (Portis et al. 2012) based on the same progeny, following the integration of 172 microsatellite (SSR) loci derived from expressed sequence tag DNA sequence (Scaglione et al. 2009). Each of the resulting maps detected 17 major linkage groups. Their alignment was followed by the construction of a consensus map which succeeded in capturing 694 loci, 227 (217 SSRs, ten SNPs) of which involved co-dominant markers. The map generated 17 LGs with a total genetic length of 1687.6 cM and a mean inter-marker spacing of 2.5 cM. The newly developed consensus as well as the parental genetic maps was then used to elucidate the pattern of inheritance of key agronomic traits as reported in the next paragraph.

On the basis of the two-way, pseudo-testcross approach, genetic maps were also developed from F_1 progenies obtained by crossing globe artichoke with its wild progenitor, wild cardoon. Sonnante et al. (2011) developed an F_1 progeny of 192 individuals by crossing a single clone of the globe artichoke 'Mola', a non-spiny varietal type which produces green/purple flowers and used as female parent, with a single plant of the wild cardoon 'Tolfa', which produces spines on head bracts and leaves and is white-flowered. The progeny was genotyped with 149 EST-derived and genomic SSRs, AFLPs, ten genes mainly belonging to the chlorogenic acid pathway as well as phenotyped for two morphological traits: presence or absence of spines on the head bracts and flower colour. The map of the globe artichoke parent included 289 markers, ordered on 18 LGs and spanned 1486.4 cM with an average marker distance of 5.2 cM. The map of wild cardoon included 122 markers, ordered into 17 LGs and spanned 865.5 cM, with an average distance of 7.1 cM between markers. From the maternal and paternal maps, an integrated map was obtained which contained 337 molecular markers ordered in 17 LGs and the genes for the traits presence of spines on head bracts and flower colour were located, while one linkage group of globe artichoke remained as a single LG. The integrated map covered 1488.8 cM, with an average distance of 4.4 cM between markers. The integrated map was finally aligned with already existing maps for artichoke

(Lanteri et al. 2006; Acquadro et al. 2009; Portis et al. 2009a), and 17 LGs were linked via 31 bridge markers.

An F_1 mapping population from the cross between an accession of wild cardoon and a single plant belonging to the open-pollinated Argentinian globe artichoke variety 'Estrella del Sur FCA' was bred by Martin et al. (2013) as well. On the basis of a prior study (Cravero et al. 2007), most of the genotypic data needed for linkage map construction were derived from SRAP markers, which constituted the major backbone of the two parental maps; however, other PCR-based markers were included in the map and made it possible to allow cross-referencing with the SSR-based consensus map previously produced by Portis et al. (2009b). The 1465.5 cM map based on the segregation of alleles present in the wild cardoon parent comprised of 214 loci distributed across 16 LGs, while the 910.1 cM globe artichoke-based map featured 141 loci falling into 12 LGs. Three morphological traits (head spininess, leaf spininess and head colour) for which the parents contrasted were found to be inherited monogenetically, and the genes conditioning these traits were also mapped. On the basis of 48 co-dominant markers, an alignment was possible with the Portis et al. (2009b) SSR-based consensus map.

Interestingly, a comparison of heterozygosity between the mapping parents used by Portis et al. (2009a) showed that fewer loci were informative in the seed-propagated variety of cultivated cardoon than in the vegetatively propagated globe artichoke varietal types. This highlighted that clonal propagation allows the maintenance of high heterozygosity while seed propagation introduced an element of purifying selection to stabilize production and reduce the global level of heterozygosity. Martin et al. (2013) reported that the map of the wild cardoon was constituted by a higher number of loci than the one of the globe artichoke (214 vs. 141 markers) and was some 50% longer (1465.5 vs. 910.1 cM). In their work, both progenitors were seed-propagated and it must be kept in mind that the globe artichoke genotype underwent over time some phenotypic selection on the basis of key commercial traits which reduced its global level of heterozygosity, while the cardoon genotype was collected from the wild as a result of open-pollination without any anthropic selection.

With the goal to increase the number of markers available and contribute to the development denser *C. cardunculus* genetic maps, Scaglione et al. (2012) combined the restriction-site-associated DNA (RAD) approach with the Illumina DNA sequencing platform to affect the rapid and mass discovery of SNP markers. RAD tags were sequenced from the genomic DNA of the three parents (globe artichoke 'Romanesco C3', cultivated cardoon 'Altilis 41' and the wild cardoon accession 'Creta 4') of the mapping population previously described. A sample of heterozygous SNP loci was mapped by CAPS assays, and the CAPS-derived genotypic data were incorporated into a pre-existing data set of molecular loci already used to generate the cultivated cardoon genetic map (Portis et al. 2012).

Recently, two high-density genetic maps have been developed through the re-sequencing of 'Romanesco C3' and 'Altilis 41' genotypes at a depth of $\sim 30\times$, together with a $\sim 1\times$ genotyping-by-sequencing of the F_1 mapping population. This enabled linkage analysis to be carried out, facilitated by the use of the SOILoCo pipeline (Scaglione et al. 2016, https://bitbucket.org/dscaglione/soiloco). The maps, each comprised 17 LGs, harbouring 1157 (C3) and 1497 (Altilis 41) independently segregant genetic bins and enclosing respectively 23,366 and 15,227 haplotype markers. The overall length of the C3 map was 1459.6 cM, with an average genome-wide inter-locus distance of 1.33 cM. Assuming the physical size of each chromosome to be as reported by Scaglione et al. (2016), the global mean ratio of physical to genetic distance was 356.4 kbp per cM. The length of the 'Altilis 41' map was 1748.2 cM, and its mean inter-locus distance was 1.23 cM. The global ratio between the physical and genetic distance was 300.7 kbp per cM. The 312 markers on shared scaffolds between the two maps allowed the two sets of LGs to be inter-aligned, and homologous

genomic regions with the LGs to be identified (Portis et al. 2018) (Fig. 6.1). The high-density genetic maps were then used for the mapping of genomic regions encoding biomass-related traits as described in the next paragraph.

6.3 QTL Analysis

The availability of genetic maps is essential for the dissection of quantitative trait loci (QTL), which underline the inheritance of many key

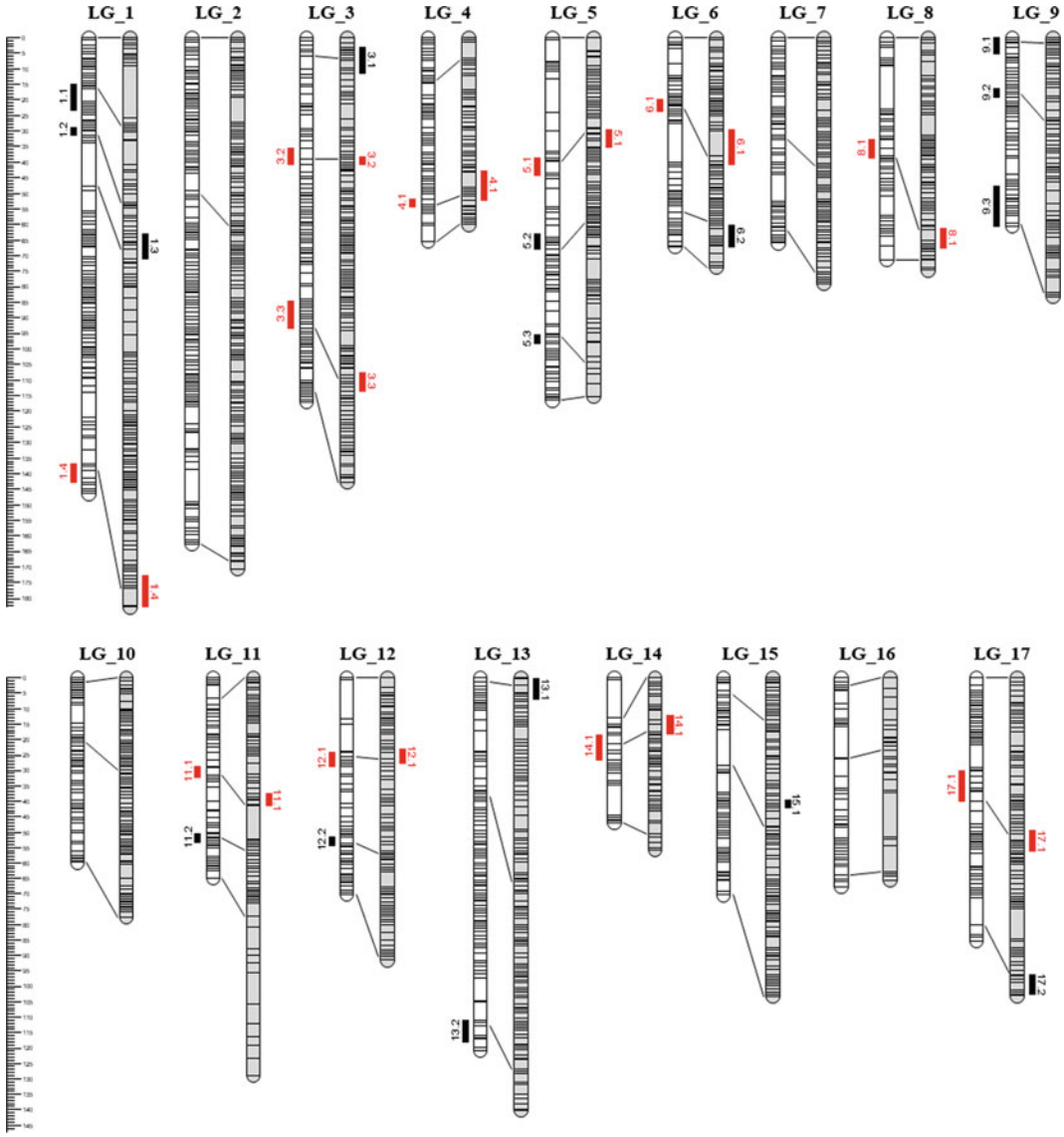


Fig. 6.1 High-density linkage maps derived from the female parent 'Romanesco C3' (white LGs) and the male parent 'Altilis 41' (grey LGs), showing the location of biomass-related QTL. Bars represent merged confidence intervals for each QTL. QTL-harbouring regions are numbered progressively along each LG, and homologous regions are shown highlighted in red. Only the scaffolds shared by the parental maps lying at the top, middle and bottom of each LG are shown, as well as those lying within a QGR. The scale to the left indicates lengths in cM. Adapted from Portis et al. (2018)

agronomic characters. A few studies have been carried out, at present, in *C. cardunculus* on the identification of genomic region influencing QTL, most of them by using the previously described map obtained from the F_1 progeny 'Romanesco C3' × 'Altilis 41'.

Portis et al. (2012) elucidate the pattern of the key commercial trait 'head production earliness'. Indeed, shortening the life cycle is seen as an important breeding goal in terms of both globe artichoke's economic value but can also increase the exploitability of cultivated cardoon as an energy crop. Earliness in both parents and F_1 progeny was assessed over two seasons and scored either as the number of days between transplanting (first season) or awakening (second season) and harvesting of both the main (eMH trait) and first- and second-order heads, obtained from the ramification of the main stem (eFOH and eSOH traits). An evaluation of the variance for the three earliness-related traits established significant genotypic differences between the two mapping parents. All three traits were found to vary continuously among the F_1 progeny, although no progeny was as early flowering as 'Romanesco C3', while a few were later flowering than 'Altilis 41', due to transgressive segregation. Across all three traits, a total of 25 QTLs were detected, of which 19 were stable across both growing seasons, with the other six expressed only during the second season. In the study, a broad-sense heritability for eMH of 0.76 was detected, which indicates the trait to be predominantly under genetic control, but rather lower heritabilities were found for the traits eFOH and eSOH, suggesting that the environment is quite influential in their determination. At any rate, the identified QTL accounted for up to 74% of the overall phenotypic variance and a cluster of large-effect QTL residing on the homologous LG 1 of globe artichoke and LG 2 of cultivated cardoon were claimed to represent a reasonable candidate for marker-assisted breeding and include two SSR loci which might serve as an indirect selection criterion for earliness.

In further work (Portis et al. 2014), the mapping population was phenotyped in order to elucidate the inheritance of seven main and first-order capitulum traits over two growing seasons. This study represented the first report on yield traits' QTL in globe artichoke. Capitula were harvested when they had reached a marketable size (prior to bract divergence), and measured with respect to their number per plant (*nc*), length (*cl*) diameter (*cd*), shape index (*si*) [cl:cd ratio] fresh weight (*fw*), receptacle diameter (*rd*) and receptacle height (*rh*). The cl, cd, si, fw, rd and rh traits were scored for both the main head (MH traits) and the first-order head (FOH traits) developed on the ramifications of the main stem. As previously highlighted for earliness, all the assessed yield-related traits were under polygenic control as were found to be normally distributed among the mapping progeny. A total of 100 QTLs associated with the seven capitulum traits were mapped to 23 chromosomal regions, scattered over 12 of the 17 linkage groups. Among these, 73 were expressed in both growing seasons, while the others were only detected in one season. As reported by the authors, the effect of a QTL has been shown to vary over time in perennial plants, like *C. cardunculus*, with changing biotic and abiotic factors (Brown et al. 2003); furthermore, since the two seasons were not identical in terms of temperature, rainfall, etc., their results seem to suggest that the species is adaptable to yearly variations in climate resulting in a significant degree of genotype × environment interaction.

The capitulum size and weight were found to be genetically determined by multiple QTL, which proved to be located in 18 distinct genomic regions. The number of QTL per trait ranged from five to nine (mean 7.2), and the only traits for which more than one QTL mapped to a single LG were fw and rd. Major QTLs, responsible for some 20% of the phenotypic variation, were detected for capitulum length, diameter, shape index and fresh weight. In the study, both head and receptacle size were found to be positively correlated with their fresh weight, as was the number of capitula with both capitulum length and diameter, and QTL for correlated traits frequently co-localized, presumably due to pleiotropy. In a previous work, Mauro et al. (2009) highlighted that the economic yield of globe

artichoke is positively correlated with the number and weight of the secondary capitula (Mauro et al. 2009), while clonal selection within a specific globe artichoke varietal type has shown that although yield was strongly dependent on the number of capitula, yet there was no significant relationship between yield and mean capitulum weight, (Mauro et al. 2012, 2015). A positive correlation between capitulum diameter and weight was previously been reported by Lopez Anido et al. (1998) and as pointed out by Lanteri et al. (2005), in the spiny elongated capitulum type, the heaviest capitula tend to be those having the greatest length and diameter. The correlations among different traits as well as their co-localization indicate the magnitude and direction of correlated response to selection as well as the relative efficiency of indirect selection (Holland 2006). As suggested by Carter et al. (2011), when traits are highly correlated, plant breeders can select for the trait with higher heritability and simultaneously indirectly select for the other trait, thus maximizing genetic gain for both traits in segregating populations. The results of Portis et al. (2014) suggest that the strong correlation between receptacle height and heads' size and weight might facilitate the indirect selection of the incidence of the edible head fraction, which is desirable for both fresh consumption and processing. Interestingly, the capitulum weight and size of the majority of the mapping progeny individuals were below the mid-parent value, which might be explained from exerting of a degree of semi-dominance of cultivated cardoon alleles over those from globe artichoke.

Following phenotyping of the same progeny, Portis et al. (2015) identified QTL for the traits' bract pigmentation and fleshy thorns. There is considerable variation in cultivated germplasm for the intensity of capitulum pigmentation, a character which is also very sensitive to temperature; however, varieties are conventionally classified as either violet or green. Cravero et al. (2005) have proposed that capitulum colour is genetically determined as follows: *P*—allows anthocyanin production, resulting in purple bracts, while *pp* inhibits anthocyanin production resulting in green bracts; *U*—results in an uneven distribution of anthocyanin pigments encoded by *P*, while *uu* results in an even distribution of pigment in the presence of *P*. This proposed model is presumably sufficient for breeders seeking to enhance colour; however, it was supposed that other modifier genes or multiple alleles might be involved (Pochard et al. 1969). In the work of Portis et al. (2015) by extending the classification of bract pigmentation to seven classes (one green; two green with purple hue; three green-purple; four purple-green; five purple with green hue; six purple; seven dark purple), the authors were able to identify a continuous like distributions appropriate for a QTL analysis. Three QTL regions, each stable across the two seasons, were identified. One was represented in both parents (homologous QTL), one only in globe artichoke and one only in cultivated cardoon. The largest effect QTL (PVE of 66–69%) mapped to globe artichoke LG 5 and cultivated cardoon LG 1 coincident with *P*. The globe artichoke QTL (PVE 16–17%) was located on LG 13 and cultivated cardoon QTL (PVE of 21–23%) on LG 11 confirm the presence of modifier loci involved in the genetic control of anthocyanin distribution, and whose effect might be variable in different progenies. The genetic basis of spine formation relates to the allelic constitution at the *Sp* locus, with the spiny "thistle-like" phenotype determined by the recessive allele *Sp* (Basnitzki and Zohary 1994), and the map location of *Sp* has been reliably established with a few mapping exercises (Lanteri et al. 2006; Sonnante et al. 2011; Martin et al. 2013). In their mapping population, Portis et al. (2015) by scoring the phenotypes on a 1–5 scale, identified a QTL justifying 68–73% of the PVE plus a second locus (PVE of 11–12%) exclusive of cultivated cardoon and possibly related to modifier genes.

Table 6.1 documents the total number of QTL detected using the 'Romanesco C3' and 'Altilis 41' maps by Portis et al. (2012, 2014, 2015), their distribution over years, homologies between maps as well as the maximum values estimated for LOD and the proportion (%) of the total phenotypic variance (PV).

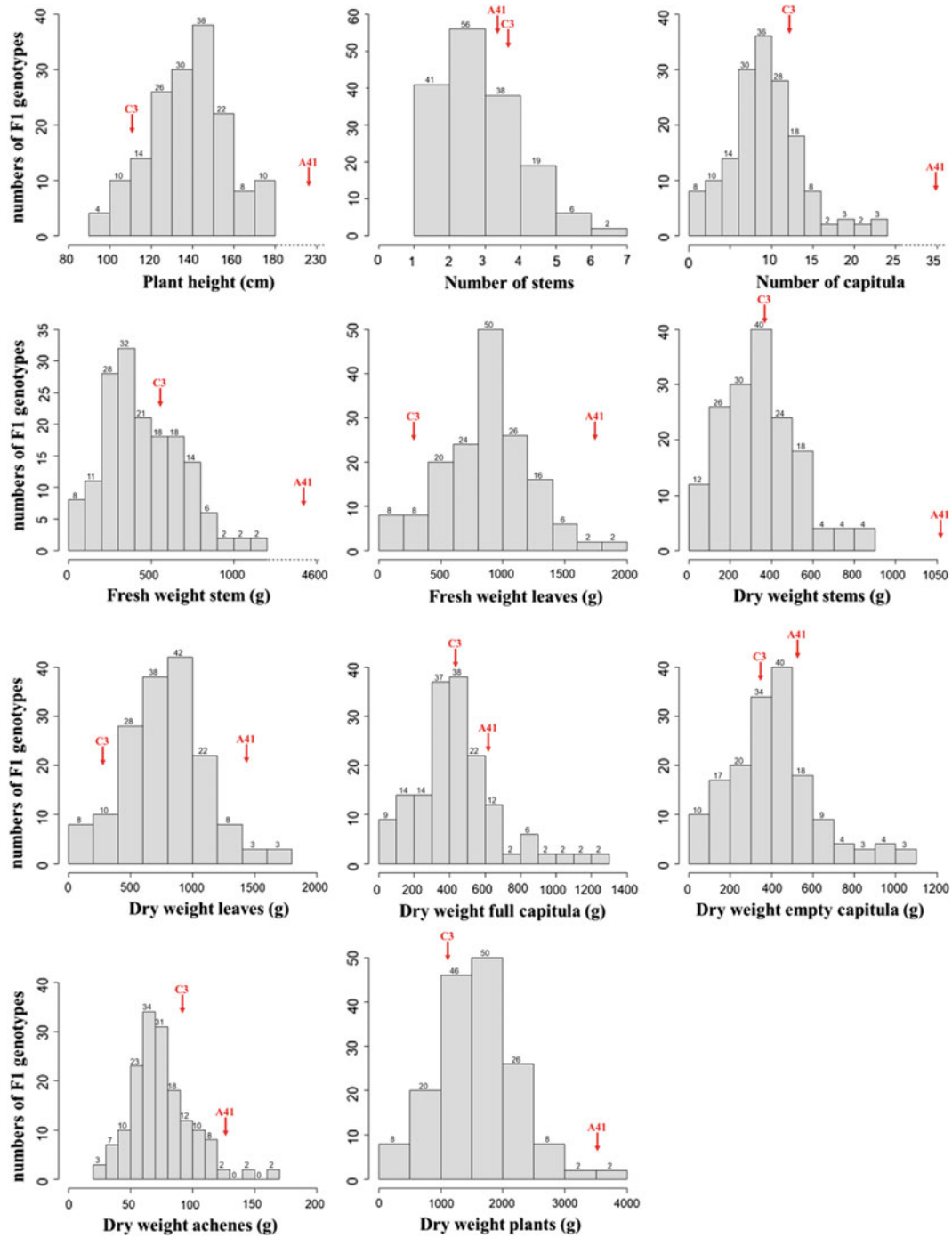


Fig. 6.2 Distribution of phenotypes for the eleven biomass-related traits across the mapping population. Traits segregating for the set of 162 F_1 individuals scored in one season are reported. The parental mean performances are indicated by arrows. Adapted from Portis et al. (2018)

Table 6.1 Total number and characteristics of the main head (MH), first-order head (FOH) and second-order head (SOH) QTL detected using the C3 and Altilis 41 maps by Portis et al. (2012)

Trait	Head	Code	Homolog. QTL	QTL in C3 map			QTL in A41 map			Max LOD	Max PV (%)
				1st season	2nd season	Both seasons	1st season	2nd season	Both seasons		
Earliness	MH	eMH	2	4	5	4	3	4	3	14.7	47.9
	FOH	eFOH	3	3	4	3	4	5	4	8.9	31.0
	SOH	eSOH	3	2	3	2	3	4	3	7.1	32.2
Capitulum length	MH	clMH	2	6	5	5	4	4	4	6.0	21.5
	FOH	clFOH	1	2	1	1	2	2	2	4.4	15.6
Capitulum diameter	MH	cdMH	2	6	5	5	3	2	2	5.9	20.3
	FOH	cdFOH	1	5	4	3	3	3	3	3.7	15.8
Shape index	MH	siMH	0	3	3	3	4	4	4	8.6	29.5
	FOH	siFOH	0	2	2	2	1	1	1	7.9	29.2
Fresh weight	MH	fwMH	3	4	3	3	6	5	5	5.1	19.9
	FOH	fwSOH	1	5	5	5	5	5	5	3.8	13.8
Receptacle diameter	MH	rdMH	2	4	3	2	5	4	3	4.4	18.8
	FOH	rdFOH	1	2	3	1	2	3	1	3.9	16.0
Receptacle height	MH	rhMH	1	5	4	4	1	1	1	4.8	11.3
	FOH	rhFOH	0	4	4	3	0	1	0	3.5	9.6
Number of *capitula*	All	nc	3	4	3	3	3	4	3	3.9	12.6
Bracts colour	All	bc	1	2	2	2	2	2	2	20.6	69.3
Bract thorniness	All	bt	1	1	1	1	2	2	2	14.0	73.2

The 'C3 x Altilis' F_1 population was, finally, phenotyped over two seasons by Portis et al. (2018) with respect to the following biomass-related traits: plant height (PH), number of stems (NS), number of capitula (NC), fresh weight of leaves (FWL), fresh weight of stems (FWS); then the material were dried and evaluated for dry weight plant (DWP), dry weight of leaves (DWL), dry weight of stems (DWS), dry weight of full capitula (DWFC), dry weight of empty capitula (DWEC) after achenes removal, and dry weight of achenes (DWA). In a preliminary study, Martin et al. (2016a) characterized the segregation of five biomass-related traits in a population bred from a cross between var. *sylvestris* and var. *scolymus* taxa, but the scope of the study was limited by its being based on data from a single growing season and its reliance on simple interval mapping for QTL identification. The research of Portis et al. (2018) therefore represents the most exhaustive attempt so far to identify the genetic basis of biomass accumulation in *C. cardunculus.*

The phenotypic data were combined with the high-density linkage maps previously described to identify 81 quantitative trait loci, of which 50 were placed on the C3 map and 31 on the Altilis map. The loci were scattered over 13 linkage groups and were clustered within 27 genomic regions, 22 of which harboured two or more QTL. Ten of these regions were specific to the C3 map and six to the Altilis map, while the other eleven were represented on both maps. The number of QTL per trait ranged from four to twelve (mean 7.4); for the traits FWL, DWL, DWFC, DWEC and DWP, more QTL located in

a single LG. The QTL allele effects were largely consistent with the phenotypic difference between parents, reinforcing the robustness of the mapping. There was a marked degree of transgressive segregation towards C3 for the traits DWS, DWA and NC: non-parental allele combinations are generally invoked to explain this phenomenon (Tanksley and McCouch 1997), but given that the majority of the progeny performed below the mid-parent value for these traits (Fig. 6.2), it is reasonable to suggest instead that a degree of semi-dominance was exerted by the globe artichoke over the cultivated cardoon alleles.

The co-localization of many of the QTL was inferred from the overlap of 1-LOD support intervals, a measure used to provide an approximate confidence interval for a QTL's true location. Only a small number of the QTL-genomic regions harboured QTL influencing the variation for a single trait. In contrast, five regions each harboured at least five QTL influencing a variety of traits—these were the genomic regions named 1.1 (seven traits), 3.3 (nine traits) and 9.1, 17.1 and 17.2 (five traits) in Fig. 6.1. The high inter-trait correlations established between traits for which the controlling QTL were frequently clustered imply that the clusters mostly likely did not comprise a group of linked loci, but rather were the expression of a single pleiotropic locus. The map location of some of the QTL was consistent with the outcomes reported by Portis et al. (2014) and Martin et al. (2016a, b). In particular, the location of NC QTL to LGs 01, 03, 05 and 09 reiterates the conclusions drawn from AFLP/SSR-based maps developed from the same population but phenotyped independently (Portis et al. 2012), as well as those obtained by mapping of a population derived from the var. *sylvestris* and var. *scolymus* cross-analysed by Martin et al. (2016b). The much higher saturation of the high-density maps has allowed further some additional NC QTL (on LGs 06, 08, 13 and 14) to be detected. At least some of the QTL influencing main capitulum fresh weight which have been mapped to LGs 01, 06, 08, 09, 12 and 13 (Portis et al. 2014) and to LGs 03 and 09 (Martin et al. 2016a) may be identical to the DWFC and DWEC QTL reported here.

Based on available genomic information, a total of 80 scaffolds covering each of the 27 QTL-genomic regions were derived and characterized. The length of the scaffolds ranged from 15,521 to 1,019,252 nucleotides (average 326,566 nt), and specified an estimated 1960 genes (24.5 genes per scaffold); 83% of these genes have been functionally annotated. An enrichment for certain gene ontology terms was noted for the gene content of the genomic regions harbouring loci influencing seed yield and the number/weight of stems.

References

Acquadro A, Lanteri S, Scaglione D, Arens P, Vosman B, Portis E (2009) Genetic mapping and annotation of genomic microsatellites isolated from globe artichoke. Theor Appl Genet 118:1573–1587

Basnitzki J, Zohary D (1994) Breeding of seed planted artichoke. Plant Breed Rev 12:253–269

Brown GR, Bassoni DL, Gill GP, Fontana JR, Wheeler NC, Megraw RA, Davis MF, Sewell MM, Tuskan GA, Neale DB (2003) Identification of quantitative trait loci influencing wood property traits in loblolly pine (*Pinus taeda* L.). III. QTL verification and candidate gene mapping. Genetics 164:1537–1546

Carter AH, Garland-Campbell K, Kidwell KK (2011) Genetic mapping of quantitative trait loci associated with important agronomic traits in the spring wheat (*Triticum aestivum* L.) cross 'Louise' × 'Penawawa'. Crop Sci 51:84–95

Ciancolini A, Rey NA, Pagnotta MA, Crinò P (2012) Characterization of Italian spring globe artichoke germplasm: morphological and molecular profiles. Euphytica 186:433–443

Comino C, Hehn A, Moglia A, Menin B, Bourgaud F, Lanteri S, Portis E (2009) The isolation and mapping of a novel hydroxycinnamoyltransferase in the globe artichoke chlorogenic acid pathway. BMC Plant Biol 9:30

Comino C, Lanteri S, Portis E, Acquadro A, Romani A, Hehn A, Larbat R, Bourgaud F (2007) Isolation and functional characterization of a cDNA coding a hydroxycinnamoyltransferase involved in phenylpropanoid biosynthesis in *Cynara cardunculus* L. BMC Plant Biol 7:14

Cravero V, Martin E, Cointry E (2007) Genetic diversity in *Cynara cardunculus* determined by sequence-related amplified polymorphism markers. J Amer Soc Hort Sci 132:1–5

Cravero V, Picardi L, Cointry E (2005) An approach for understanding the heredity of two quality traits (head color and tightness) in globe artichoke (*Cynara scolymus* L.). Genet Mol Biol 28:431–434

Crinò P, Pagnotta MA (2017) Phenotyping, genotyping, and selections within Italian local landraces of Romanesco globe artichoke. Diversity 9:14

Crinò P, Tavazza R, Rey Muñoz NA, Trionfetti Nisini P, Saccardo F, Ancora G, Pagnotta MA (2008) Recovery, morphological and molecular characterisation of globe artichoke 'Romanesco' landraces. Genet Resour Crop Evol 55:823–833

Holland JB (2006) Estimating genotypic correlations and their standard errors using multivariate restricted maximum likelihood estimation with SAS Proc MIXED. Crop Sci 46:642–654

Lanteri S, Acquadro A, Comino C, Mauro R, Mauromicale G, Portis E (2006) A first linkage map of globe artichoke (*Cynara cardunculus* var. *scolymus* L.) based on AFLP, S-SAP, M-AFLP and microsatellite markers. Theor Appl Genet 112:1532–1542

Lanteri S, Ledda L, Di Leo I, Mameli MG, Portis E (2005) Molecular and morphological variation among and within populations of *Cynara scolymus* L. cv. 'Spinoso sardo'. Acta Hortic 681:333–340

Lanteri S, Portis E, Acquadro A, Mauro RP, Mauromicale G (2012) Morphology and SSR fingerprinting of newly developed *Cynara cardunculus* genotypes exploitable as ornamentals. Euphytica 184:311–321

Lanteri S, Acquadro A, Saba E, Portis E (2004) Molecular fingerprinting and evaluation of genetic distances among selected clones of globe artichoke (*Cynara cardunculus* L. var. *scolymus* L.) 'Spinoso sardo'. J Hortic Sci Biotech 79:863–870

Lanteri S, Di Leo I, Ledda M, Mameli M, Portis E (2001) RAPD, variation within and among populations of globe artichoke (*Cynara scolymus* L.), cv 'Spinoso sardo'. Plant Breed 120:243–246

Lopez Anido F, Firpo I, Garcia S, Cointry E (1998) Estimation of genetic parameters for yield traits in globe artichoke (*Cynara scolymus* L.). Euphytica 103:61–66

Martin E, Cravero V, Portis E, Scaglione D, Acquaviva E, Cointry E (2013) New genetic maps for globe artichoke and wild cardoon and their alignment with an SSR-based consensus map. Mol Breed 32:177–187

Martin EA, Cravero VP, Cointry EL (2016a) Quantitative trait loci (QTLs) related to biomass production in *Cynara cardunculus* L. Acta Hortic 1147:189–196

Martin EA, Cravero VP, López Anido FS, Cointry EL (2016b) QTLs detection and mapping for yield-related traits in globe artichoke. Sci Hortic 202:156–164

Mauro R, Portis E, Acquadro A, Lombardo S, Mauromicale G, Lanteri S (2009) Genetic diversity of globe artichoke landraces from Sicilian small-holdings: implications for evolution and domestication of the species. Conserv Genet 10:431–440

Mauro R, Portis E, Lanteri S, Lo Monaco A, Mauromicale G (2015) Clonal selection in a globe artichoke landrace: characterization of superior germplasm to improve cultivation in Mediterranean environment. J Agric Sci 153:102–113

Mauro R, Portis E, Lanteri S, Mauromicale G (2012) Genotypic and bio-agronomical characterization of an early Sicilian landraces of globe artichoke. Euphytica 186:357–366

Mauromicale G, Portis E, Acquadro A, Lo Monaco A, Pesce GR, Lanteri S (2018) An integrated model to accelerate the development of seed-propagated varieties of globe artichoke. Crop Breed Appl Biotechnol 18:72–80

Menin B, Comino C, Moglia A, Dolzhenko Y, Portis E, Lanteri S (2010) Identification and mapping of genes related to caffeoylquinic acid synthesis in *Cynara cardunculus* L. Plant Sci 179:338–347

Moglia A, Comino C, Portis E, Acquadro A, De Vos R, Beekwilder J, Lanteri S (2009) Isolation and mapping of a C3'H gene (CYP98A49) from globe artichoke, and its expression upon UV-C stress. Plant Cell Rep 28:963–974

Pandino G, Lombardo S, Moglia A, Portis E, Lanteri S, Mauromicale G (2015) Leaf polyphenol profile and SSR-based fingerprinting of new segregant *Cynara cardunculus* genotypes. Front Plant Sci 5:800

Pochard E, Foury C, Chambonet D (1969) Il miglioramento genetico del carciofo. Proceedings 1° Congresso Internazionale sul carciofo. Bari, Italy, pp 117–155

Portis E, Acquadro A, Scaglione D, Mauromicale G, Mauro R, Taylor CA, Knapp SJ, Lanteri S (2009b) Construction of an SSR-based linkage map for *Cynara cardunculus*. In: Proceeding of the 8th Plant Genomics European Meeting, p 142

Portis E, Mauro RP, Acquadro A, Moglia A, Mauromicale G, Lanteri S (2015) The inheritance of bract pigmentation and fleshy thorns on the globe artichoke capitulum. Euphytica 206:523–531

Portis E, Mauro RP, Barchi L, Acquadro A, Mauromicale G, Lanteri S (2014) Mapping yield-associated trait QTL in globe artichoke. Mol Breed 34:615–630

Portis E, Mauromicale G, Barchi L, Mauro R, Lanteri S (2005) Population structure and genetic variation in autochthonous globe artichoke germplasm from Sicily Island. Plant Sci 168:1591–1598

Portis E, Mauromicale G, Mauro R, Acquadro A, Scaglione D, Lanteri S (2009a) Construction of a reference molecular linkage map of globe artichoke (*Cynara cardunculus* var. *scolymus*). Theor Appl Genet 120:59–70

Portis E, Scaglione D, Acquadro A, Mauromicale G, Mauro R, Knapp S, Lanteri S (2012) Genetic mapping and identification of QTL for earliness in the globe artichoke/cultivated cardoon complex. BMC Res Notes 5:252

Portis E, Acquadro A, Tirone M, Pesce GR, Mauromicale G, Lanteri S (2018) Mapping the genomic regions encoding biomass-related traits in *Cynara cardunculus* L. Mol Breed 38:64

Scaglione D, Acquadro A, Portis E, Taylor CA, Lanteri S, Knapp SJ (2009) Ontology and diversity of

transcript-associated microsatellites mined from globe artichoke EST database. BMC Genom 10:454

Scaglione D, Acquadro A, Portis E, Tirone M, Knapp SL, Lanteri S (2012) RAD tag sequencing as a source of SNP markers in *Cynara cardunculus* L. BMC Genom 13:3

Scaglione D, Reyes Chin Wo S, Acquadro A, Froenicke L, Portis E, Beitel C, Tirone M, Mauro R, Lo Monaco A, Mauromicale G, Faccioli P, Cattivelli L, Rieseberg L, Michelmore R, Lanteri S (2016) The genome sequence of the outbreeding globe artichoke constructed *de novo* incorporating a phase-aware low-pass sequencing strategy of F1 progeny. Sci Rep 6:19427

Sonnante G, Gatto A, Morgese A, Montemurro F, Sarli G, Blanco E, Pignone D (2011) Genetic map of artichoke × wild cardoon: toward a consensus map for *Cynara cardunculus*. Theor Appl Genet 123:1215–1229

Tanksley SD, McCouch SR (1997) Seed banks and molecular maps: unlocking genetic potential from the wild. Science 277:1063–1066

7 Insights into the Population Structure and Association Mapping in Globe Artichoke

Domenico Rau, Monica Rodriguez, Giovanna Attene, Limbo Baghino, Anna Barbara Pisanu, Davide Sanna, Alberto Acquadro, Ezio Portis and Cinzia Comino

Abstract

Understanding the population genetic structure of crop gene pool is a fundamental prerequisite in plant evolutionary biology and plant breeding. Based on this consideration, in this chapter we pursued two main aims. The first was to review the status of the research pointed at deciphering the population structure of the gene pools of globe artichoke. The second was to show how the use of a well-characterized germplasm collection together with deep genotyping could serve the identification of QTLs for traits of agronomic and economic relevance in globe artichoke.

7.1 Introduction

To 'unlock' the potential of crop germplasm collections by their effective use, researchers need detailed information about the amount and distribution of genetic diversity present within collections (Schreiber et al. 2018). For example, the knowledge of population structure allows the identification of accessions chosen to represent as much as possible of the diversity in the collection (Brown 1989). This is useful (i) to better plan phenotypic characterization, (ii) to look for genetic resources adapted to different environmental conditions, (iii) to select materials for hybridization experiments and (iv) to guide the choice of accessions for genome sequencing and comparative studies. Moreover, with the advent of next-generation sequencing (NGS) technologies, genome-wide association (GWA) studies have been conducted routinely to unravel genotype–phenotype associations in many species (Liu and Yan 2019). A serious problem with association mapping is that population structure can lead to false positives (spurious associations) between a candidate marker and a phenotype. Such associations can occur when the trait mean (or frequency) varies across subpopulations (Pritchard et al. 2000a). However, if population structure of the collections is inferred, it is possible to curtail the risk of false positives (Pritchard et al. 2000b; Yu and Buckler 2006; Kang et al. 2008).

D. Rau · M. Rodriguez · G. Attene
Dipartimento di Agraria, Sezione di Agronomia, Coltivazioni Erbacee e Genetica (SACEG), Università degli Studi di Sassari, Sassari, Italy

L. Baghino · A. B. Pisanu · D. Sanna
Agenzia Agris Sardegna, Servizio Ricerca sui Sistemi Colturali Erbacei, Settore Innovazione dei Modelli Gestionali e Studio Della Biodiversità Nelle Colture Intensive, Sassari, Italy

A. Acquadro · E. Portis · C. Comino (✉)
DISAFA Plant Genetics and Breeding, University of Torino, Largo P. Braccini 2, 10095 Grugliasco, Torino, Italy
e-mail: cinzia.comino@unito.it

E. Portis et al. (eds.), *The Globe Artichoke Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-030-20012-1_7

A review on the current knowledge of the population structure of *Cynara cardunculus* var. *scolymus*, the cultivated globe artichoke, is here reported. The survey is divided considering two hierarchical levels of genetic organization: *within* and *among* varietal types.

7.2 Variation Within Varietal Types

The study of genetic variation within varietal types and landraces has been driven mainly by two needs. The first was to valorize local germplasm. The second was to understand the feasibility of and to guide the clonal selection. The main aims have been to improve the agronomic performances of traditional types at risk of genetic erosion or to gain more varietal uniformity.

In globe artichoke, commercial production is mainly based on the cultivation of clones. These are obtained by vegetative propagation using 'ovoli' and 'carducci'. The first are semi-dormant shoots with a limited root system, while the latter are basal shoots. Clones are highly heterozygous and segregate widely when propagated by seed (Basnizki and Zohary 1987). They are relatively uniform as farmers apply purging selection by eliminating off-types that often arise in the field because of accidental seed dissemination (see Chap. 2 '*Cynara cardunculus* Propagation').

Notwithstanding, several studies reported genetic variability within the fields of the same clonal varietal type. This was observed, for example, within 'Spinoso Sardo' the main variety cultivated in Sardinia Island (Lanteri et al. 2001; Cadinu et al. 2007); within 'Romanesco' mainly cultivated in centre Italy (Crinò et al. 2008; Crinò and Pagnotta 2017); and within 'Spinoso di Palermo' and 'Violetto di Sicilia' which are typical of Sicilian Island (Portis et al. 2005; Mauro et al. 2012, 2015).

The type 'Spinoso Sardo' showed phenotypic diversity for several morpho-phenological and yield traits (Deidda 1967). A programme of clonal selection started more than fifty years ago (Deidda 1967) and allowed the researchers of the Agenzia AGRIS Sardegna to identify and select clones with different pigmentations, morphological features of head, yield performance (a number of heads per plant), earliness and low susceptibility to head atrophy (Cadinu et al. 2007). Some of the selected clones (Fig. 7.1) are currently subjected to the procedures for variety release. Molecular markers such as random amplified polymorphic DNA (RAPD; Lanteri et al. 2001), amplified fragment length polymorphisms (AFLPs) and inter simple sequence repeats (ISSRs) (Lanteri et al. 2004) confirmed a significant degree of genetic diversity within 'Spinoso Sardo'. The evaluation of five populations (= fields) showed well-defined differences between populations (28.1% of the total genetic diversity) and revealed substantial diversity within populations (71.9% of total genetic diversity). Results overall showed the need to apply clonal selection to diminish the genetic variability within the varietal type and to gain more uniformity, as described in Chap. 6.

Using AFLPs, Portis and colleagues (2005) analysed three populations (= fields) of 'Spinoso di Palermo' (a spiny type) and four populations (= fields) of 'Violetto di Sicilia' (a non-spiny type). Within both types, populations showed nontrivial genetic differences (G_{ST} = 0.20 and 0.12, respectively) likely because of fragmentation and adaptation to local pedoclimatic conditions. However, within each type, most of the genetic variation was allocated, again, within populations. Mauro and colleagues (2012, 2015) further showed that genetic variation useful for breading exists either within 'Violetto di Sicilia' or within 'Spinoso di Palermo' (Chap. 6). Mauro and colleagues (2009) investigated 24 landraces of globe artichoke collected from smallholdings in Sicily. Local farmers appreciated and cultivated these landraces for centuries because of their culinary uniqueness. Landraces were characterized using phenotypic traits and molecular markers and were compared with wild cardoon accessions collected from different sites in Sicily. The comparison also included accessions of 'Spinoso di Palermo' and 'Violetto di Sicilia'. High genetic diversity was detected, and landraces showed a variable level of admixture either with the domesticated spiny type ('Spinoso di Palermo') or with the non-spiny type

Fig. 7.1 Example of variation within varietal type. Top: The picture represents three clones selected within the varietal type 'Spinoso Sardo' by the Agricultural Research Agency of Sardinia Region. AGRIS 2 has more oblong heads compared to AGRIS 1 and 3. AGRIS 1 is the most early, and AGRIS 3 is the most productive (as a number of heads per plant). *Photograph* L. Baghino (AGRIS)

('Violetto di Sicilia'). These results outlined the pivotal role of farmers' practice in the maintenance of '*on farm*' genetic diversity.

Several studies were conducted on the 'Romanesco' varietal type (Crinò et al. 2008; Ciancolini et al. 2012; Crinò and Pagnotta 2017). Crinò and colleagues (2008) characterized 19 clones derived from plants cultivated in different areas of Latium region, the major world producer of this varietal type. They integrated morphological traits (UPOV descriptors) and DNA markers (AFLP and ISSR). A very strong correlation structure among the 39 recorded traits was observed. Such observation was interpreted by the authors as a signature of human selection: different genotypes might have been selected based on the multiple traits that appeared to be correlated. Moreover, molecular markers clearly distinguished most of the clones and identified two major genetic groups; however, there was no correlation between genetic and molecular distances. Interestingly, some 'Romanesco' clones also differed for productive traits (particularly for the number of heads) and for metabolomic profiles of heads (Crinò and Pagnotta 2017).

Thus, overall these studies emphasize that varietal types are not uniform and can be regarded as composed by several genotypes, as affirmed by De Vos (1992).

As suggested for other clonally propagated domesticated plants, somatic mutations are likely the main cause of this variability (Whitham and Slobodchikoff 1981). Somatic mutations accumulate over time, and different parts of the plant can inherit different mutations. Thus, the plant can become a mosaic of tissues with different genotypes, a phenomenon known as chimerism (Gill et al. 1995). If the plant is used for vegetative propagation, descendants can be genetically heterogeneous. Thus, the strict genetic identity of 'clone mates' is highly improbable (Lushai and Loxdale 2002).

Moreover, different somatic mutations can have different consequences on the cells. Thus, within a meristem, cell lineages bearing different mutations can have a different fitness and can compete among them (diplontic selection). Cell lineages with higher fitness can replace other lineages. However, in plants, soma and germ-line are not separated. Thus, it is more likely that gametes will carry the favourable somatic mutations of the 'winner' cell lineages from which they descend (McKey et al. 2010). Mathematical models show that this process is quantitatively comparable to mutations occurring during meiosis (Orive 2001).

All this renders the evolution of clonal species, likely including globe artichoke, more dynamic than is usually thought (McKey et al. 2010).

As discussed above, the genetic diversity present within varietal types gives an opportunity to the clonal selection. From another perspective, this means that GWA studies should point to exploit not only the genetic variation *among* but also that *within* varietal types. For example, seven accessions of Bayrampasa and ten accessions of 'Spinoso Sardo' differed on average by 15 and 11% of the AFLP loci examined, respectively. Moreover, genetic differences detected *within* the same varietal type were sometimes higher than that found *between* different varietal types (Lanteri et al. 2004).

Thus, it might be useful that collections for association mapping include, as far as possible, a high number of accessions per varietal type. The number of accessions sampled should be proportional to the genetic diversity observed within varietal types.

7.3 Variation *Among* Varietal Types

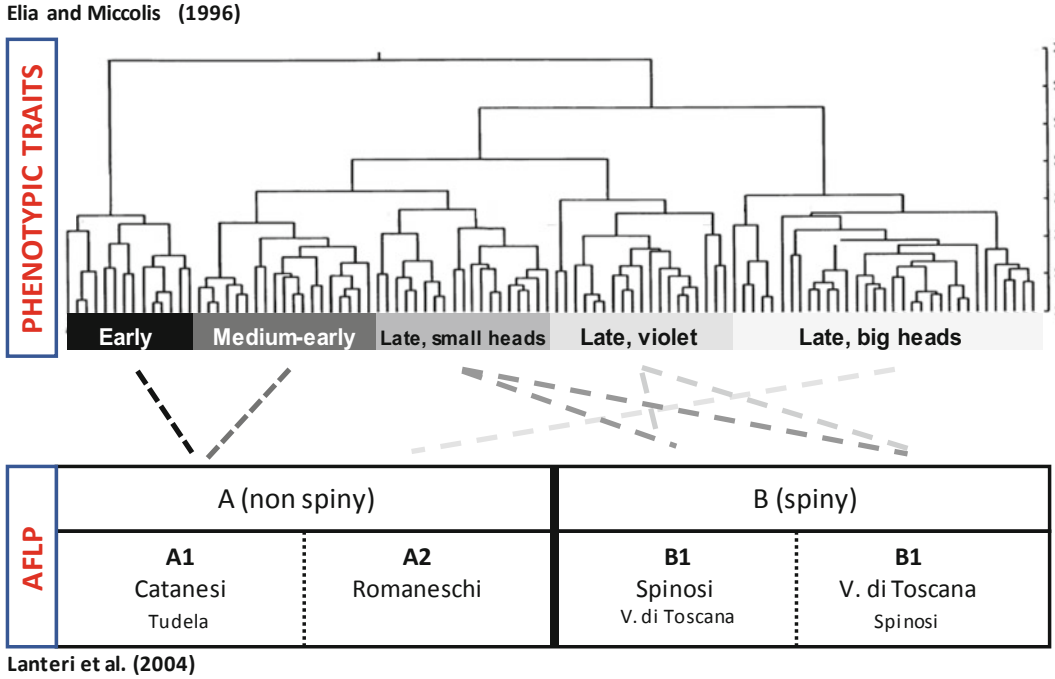
Many distinct clonal varietal types of globe artichoke exist. These were mainly 'local' as often they have a favoured area or territory of cultivation. This is further marked by their names that often refer to the location of origin. For example, 'Violetto di Sicilia', 'Spinoso di Palermo' and 'Brindisino' are common in South Italy; 'Romanesco' type is more usual in centre Italy, and 'Spinoso Sardo' in Sardinia (Calabrese 2016). In France, the most cultivated types are 'Camus de Bretagne', 'Gros Vert de Laon', 'Blanc Hyerois', 'Violet du Gapeau', 'Castel' and 'Petit Violet de Provence' (Calabrese 2016). In Spain, 'Blanca de Tudela' ecotype represents 90% of the whole production (Calabrese 2016). However, accession names are far from being a good indicator of genetic differences. Indeed, many cases of homonymy as also of synonymy have been recognized (Pagnotta et al. 2017).

Clonal varietal types are likely the results of clonal selection for adaption to different environmental conditions. However, it is possible that, such as in several other horticultural crops, the selection was mainly 'consumer-driven'. Likely, market requirements concerning the appearance, taste and properties of the edible parts dictated selection *criteria* no less strictly than environmental conditions.

The early attempts to classify the cultivated germplasm of globe artichoke date about 40 years ago (Porceddu et al. 1976; Dellacecca et al. 1976; Vanella et al. 1981). Four major groups were distinguished mainly based on capitulum traits: (1) 'Spinosi', comprising types with long spines on bracts and leaves; (2) 'Violetti', with medium-sized, violet-coloured and less spiny heads; (3) 'Romaneschi', with spherical or subspherical non-spiny heads; and (4) 'Catanesi', with rather small, elongated and non-spiny heads. Another classification is based on harvest time (Mauromicale and Ierna 2000). Two types can be distinguished. Indeed, if ovoli are planted and irrigated during summer (July) certain varieties can be 'forced' to grow 'out of season'. These produce capitula very early, starting in autumn (October) and continuing until the spring. Differently, varieties that cannot be 'forced' produced capitula only during spring and early summer. With some exception, 'Catanesi', 'Violetti di Provenza', 'Spinoso Sardo' and 'Spinoso di Palermo' fit into the first class while 'Romaneschi' and 'Violetti di Toscana' in the second one.

Elia and Miccolis (1996) conducted a morpho-phenological characterization of 104 accessions grown in a common garden. Using

multivariate analysis and considering eight quantitative traits, they distinguished five main clusters (Fig. 7.2). These were 'early', 'medium–early', 'late with a small head', 'late violet' and 'late with a large head'. The 'early' and 'medium–early' clusters were mainly made up by 'Catanesi' and 'Violetti di Provenza', characterized by long-shaped, small main and secondary



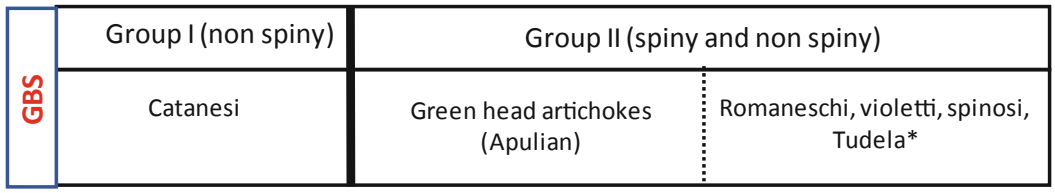


Fig. 7.2 Summary of the results on the population structure of cultivated globe artichoke. *Notes* The intention of the figure is to sketch the main results of each study. Phenotypic analysis: dendrogram has been re-adapted from Elia and Miccolis (1996). Molecular analyses—black continuous vertical line: main genetic subdivision. Black vertical dotted line: secondary genetic subdivisions. Subdivisions are not proportional to sample size. Within each subdivision, the more frequent typology (as discussed by the authors) was reported with big font while less frequent with small font. In this figure, the 'Catanesi' type also includes 'Violetti di Provenza'

heads, low-medium yield, with an early harvest and low vigorous plant (short, and with a low number of heads). The 'late with a small head' cluster mainly included 'Spinosi' types characterized by long-shaped and medium–small heads and late harvest. Within the cluster, 'late violet' prevailed 'Violetti di Toscana' and other varieties characterized by vigorous plants (tall and high yielding) and less long-shaped than the previous groups. Finally, the 'late with a large head' cluster mainly comprised 'Romaneschi' and 'Macau', characterized by late harvest, big principal and secondary heads with a spherical shape.

Lanteri and colleagues (2004) reported the first study on the molecular population structure of the cultivated gene pool of globe artichoke. They studied a large collection of cultivated globe artichoke by using AFLPs. Their multivariate analysis evidenced two major clusters named 'A' and 'B' (Fig. 7.2). The first comprised non-spiny types ('Catanesi' and 'Romaneschi'), while the second, 'B', included all the spiny types ('Spinosi' and 'Violetti') with few exceptions. Differently, there was no association between capitulum pigmentation and genetic clustering. Cluster 'A' divided into two sub-clusters: 'A1' and 'A2'. Cluster 'A1' contained 'Catanesi' and 'Violetti di Provenza' types. 'A2' encompassed 'Romaneschi' and 'Macau'. Cluster 'B' also split into two sub-clusters: 'B1' and 'B2'. In this case, the association between phenotypic typologies and AFLP clusters did not hold as both B1 and B2 included 'Spinosi' and 'Violetti' types.

The AFLP classification of Lanteri and colleagues (2004) was partially congruent with that proposed by Elia and Miccolis (1996) on phenotypic data. Accessions included in cluster 'A1' ('Catanesi') corresponded to the 'early' and 'medium–early types' of Elia and Miccolis (1996); sub-cluster 'A2' (mainly 'Romaneschi') comprised the 'late types with large heads'. However, the accessions included in the group 'late with small heads' and 'late violet' of Elia and Miccolis (1996) were distributed randomly across the clusters 'B1' and 'B2' of Lanteri et al. (2004) comprising 'Violetti' and 'Spinosi'.

More recently, Pagnotta and colleagues (2017) analysed 174 leafy cardoon and globe artichoke accessions using SSR and ISSR markers. Based on the Bayesian clustering method implemented in Structure software (Pritchard et al. 2000a) and based on Evanno's method (2005), they inferred that the uppermost level of population structure was represented by three genetic groups. Leafy cardoon accessions clustered separately from cultivated globe artichoke; however, this latter further split into two groups (Fig. 7.2). The first group comprised almost all 'Catanesi' types together with 'Tudela' types and two 'Romanesco' accessions. The second group included the remaining accessions. However, the authors also considered secondary structures at $K = 5$. Here, the first group split into two subgroups, named SG2 and SG6. SG2 included most 'Catanesi' and two accessions of 'Tudela', while SG6 included most of 'Tudela' accessions, two 'Romaneschi' and two 'Catanesi'. The second group split into three subgroups named SG3, SG4 and SG5. SG3 included all the spiny types and some accessions of 'Romanesco' and 'Violetti'; SG4 grouped most of the 'Romaneschi' types with two 'Violetti' and one 'Catanese'; SG5 contained twelve accessions of 'Romanesco' and two 'Violetti'.

Thus, based on the analysis of Pagnotta and colleagues (2017), the majority of 'Catanesi' clustered together, 'Spinosi' types were grouped with some 'Romaneschi' and 'Violetti' types, and 'Romaneschi' appeared less uniform being distributed across several genetic subgroups. Consequently, except for the group of 'Catanesi', the association between phenotypic types and molecular groups was not clear-cut. Moreover, as noted by the authors, accessions clustered not only based on their typology but also based on the gene bank of origin. They argued that the environmental conditions of the different field gene banks or the method used to conserve plant materials might have affected germplasm characteristics. The authors suggested the use of accessions from different gene banks to conduct association mapping studies (Pagnotta et al. 2017).

Recently, Pavan et al. (2018) used genotyping by sequencing (GBS) to analyse 65 globe artichoke, 9 leafy cardoon and 21 wild cardoon

samples. The collection showed a strong population structure at $K = 2$, separating the globe artichoke from the leafy and wild cardoon. At higher K values, wild cardoon was separated from the leafy cardoon. However, genetic subdivisions within the globe artichoke were also discussed (Fig. 7.2). When the analysis was repeated considering globe artichoke and assuming two populations, structure analysis pointed out the genetic differentiation between the 'Catanesi' typology and all other samples. When a higher number of populations were assumed, the separation of the 'Catanesi' group still held true, and green-headed landraces from Apulia region, Italy ('Green Apulian'), formed a distinct group. 'Romaneschi' fell in a variable group with admixed samples, showing that they should not be considered as a genetically uniform typology (Pavan et al. 2018) as also observed by Pagnotta et al. (2017).

Comparing different studies, some differences emerged. Lanteri et al. (2004) find that 'Catanesi' and 'Romaneschi' clustered together and were separated from 'Violetti' and 'Spinosi'. Differently, both Pagnotta et al. (2017) and Pavan et al. (2018) found that the main genetic subdivision is between 'Catanesi' and the remaining accessions.

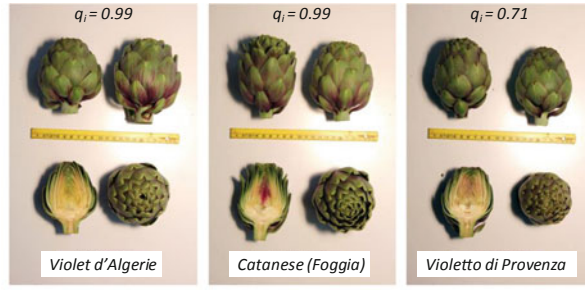
There are several examples in the literature of partially non-overlapping or different results when using different marker types to decipher population structure (e.g. Schlötterer 2004). Such differences can stem because the various marker types have distinct mutational mechanisms and different mutational rates. Specifically, it was estimated that 10^{-3}–10^{-4} mutations per locus per generation occur for SSR markers whereas the mutation rate for AFLP is about 10^{-6} per locus per generation (Mariette et al. 2001). This means that SSRs and AFLPs could depict scenarios reflecting evolutionary processes that operated at different timescales. Indeed, markers with low mutation rate have a stronger phylogenetic signal than those with high mutation rate. Thus, the divergence between spiny type and non-spiny types (AFLP; Lanteri et al. 2004) might reflect a process more ancient than the separation between the 'Catanesi' and the other globe artichokes (SSR; Pagnotta et al. 2017).

Moreover, AFLPs are dominant markers while SNPs and SSRs are codominant markers. Thus, contrary to AFLP, SNPs and SSRs allow (i) to distinguish heterozygotes from homozygotes, and (ii) to recognize different heterozygotes. This can affect the estimation of the genetic distances among individuals (and populations) in an outbreeding species such as globe artichoke (Gaudeul et al. 2004). Finally, the collections analysed in these studies were not the same and varied in sample sizes; thus, the sampling of different plant materials could also account for the differences observed among studies.

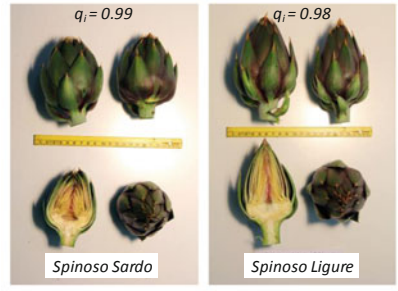
A wide germplasm collection comprising about 120 globe artichoke accessions and most of the globe artichoke varieties cultivated worldwide was characterized integrating SSR analysis with the phenotypic characterization in a common garden (Rau et al. 2015; Rau et al. submitted). This living collection is maintained by AGRIS at Oristano, Sardinia, Italy. Based on morpho-phenotypic traits primarily related to head characteristics, the accessions were classified into seven groups: 'Catanesi' (CAT), 'Violetti di Provenza' (VPR), 'Romaneschi' (ROM), 'Green et alia (GEA)'; 'Violetti di Toscana' (VTO), 'Macau' (MAC) and 'Spinosi' (SPI). The study showed that the cultivated gene pool of globe artichoke includes five distinct genetic groups (Fig. 7.3). They were associated with the major phenotypic typologies. The most relevant structure was between the 'Catanesi' and the remaining accessions like also observed by Pagnotta et al. (2017) and Pavan et al. (2018). Differently, in the study of Rau and colleagues (submitted) the secondary genetic subdivisions were more precisely associated with the other phenotypic typologies (Fig. 7.3). The first group attracted 94.1% of 'Catanesi' and 'Violetti di Provenza'. However, the second group separated into four subgroups. Group 2a attracted 71.1–100% of 'Spinosi', Group 2b 100% of 'Violetti di Toscana', Group 2c 62.5% of 'Romaneschi' and Group 2d 57.2% of 'Macau'. Thus, these two last phenotypic typologies were those less genetically defined as also noted by Pagnotta et al. (2017) and Pavan et al. (2018).

Rau and colleagues determined that 17 and 20% of phenotypic and molecular variances,

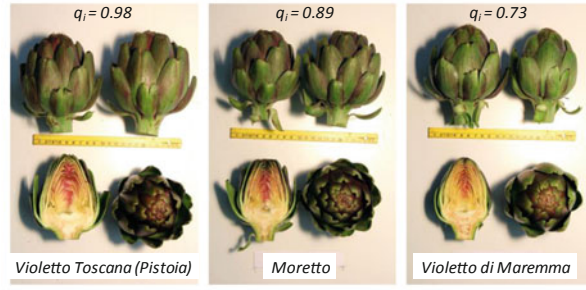
CAT-VPR-related		
Group 1	**N**	**%**
V. Provenza	17	94.4
Catanesi	16	94.1
Green et al.	3	13.0
Romaneschi	2	12.5
Macau	1	14.3
Spinosi AGRIS	0	0
Spinosi	0	0
V. Toscana	0	0



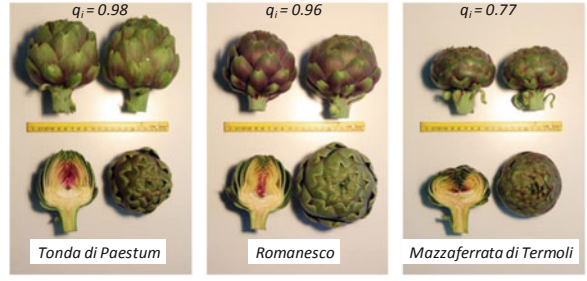
SPI-related		
Group 2a	**N**	**%**
Spinosi-AGRIS	9	100
Spinosi	5	71.4
Green et al.	1	4.3
Catanesi	0	0
Macau	0	0
Romaneschi	0	0
V. Provenza	0	0
V. Toscana	0	0



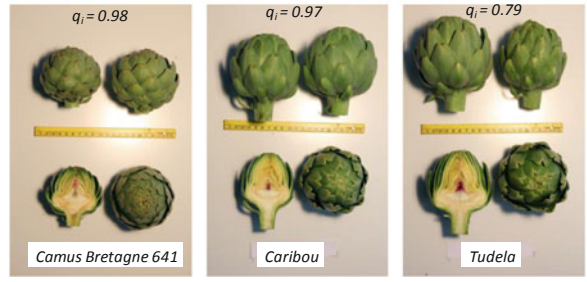
VTO-GEA-related		
Group 2b	**N**	**%**
V. Toscana	12	100.0
Green et al.	14	60.9
Macau	2	28.6
Romaneschi	4	25.0
Spinosi	1	14.3
Catanesi	1	5.9
V. Provenza	1	5.6
Spinosi-AGRIS	0	0



ROM-related		
Group 2c	**N**	**%**
Romanesco	10	62.5
Green et al.	1	4.3
Spinosi-AGRIS	0	0
Catanesi	0	0
Macau	0	0
Spinosi	0	0
V. Provenza	0	0
V. Toscana	0	0



MAC-related		
Group 2d	**N**	**%**
Macau	4	57.1
Green et al.	4	17.4
Spinosi	1	14.3
Spinosi-AGRIS	0	0
Catanesi	0	0
Romaneschi	0	0
V. Provenza	0	0
V. Toscana	0	0



◀ **Fig. 7.3** Results of model-based cluster analysis (structure) for 109 accessions of cultivated globe artichoke. Each of the five lines exemplifies the composition of each genetic cluster. On the left, the table reports the composition of each genetic cluster in each of the main phenotypic typologies recognized in globe artichoke. CAT-VPR = 'Catanesi' and 'Violetti di Provenza', SPI ='Spinosi', GEA = Green *et alia*, MAC = 'Macau', ROM = 'Romaneschi'. Tables are compiled considering a coefficient of membership (q_i) higher than 0.50. For 'Spinosi' types, the clones of 'Spinoso Sardo' selected by AGRIS were considered separately from the other 'Spinosi' accessions. *N* = number of accessions. % = percentage referred to the total number of accessions of that type. On the right: for each genetic groups, the head characteristics of three representative ($q_i > 0.70$) accessions are shown. Photographs by L. Baghino (AGRIS)

respectively, were among SSR groups. This was also the degree of differentiation between phenotypic typologies; i.e. the typologies are far from being genetically uniform. Moreover, several accessions showed an SSR profile admixed between different genetic groups (Fig. 7.4). These have phenotypes that fit none of the major typological categories and represented a remarkable part of the variability observed within the collection.

The phenotypic characterization of the collection was achieved under field condition at the AGRIS experimental station, Oristano, Sardinia, Italy, during the growing season 2015–2016. The analysis considered primary, secondary and tertiary heads and thirty-one quantitative traits and 19 UPOV descriptors as qualitative traits. This comprised phenological traits (harvest time and duration); key capitulum traits (e.g. number per plant, length, diameter, fresh weight, receptacle diameter and height); plants (e.g. height, number of branches, growth habit); and leaf traits (e.g. size, shape). Overall, the collection showed good levels of phenotypic variability (e.g. see Figs. 7.3, 7.4 and 7.5). Within the SSR-derived groups, i.e. after solving for the population structure, 82% of the total phenotypic variance was even observed, i.e. a considerable variation for most of the recorded phenotypic traits. This renders this collection an interesting material to perform future GWA studies.

Several evidences indicated a long history of clonal propagation in globe artichoke (Rau et al. submitted): (i) molecular divergence among SSR groups was strongly correlated with differences for phenotypic traits, (ii) within SSR groups, strong heterozygote excess was present, and (iii) linkage disequilibrium persisted also after solving for population structure. All these suggest that any association mapping exercise in globe artichoke must accurately consider the population structure. The aim would be to find the balance between the needs to curtail the risk of false positives and that to keep sufficient power to detect true associations. To this regard, it would be of relevance to compare the results of gene mapping obtained considering the uppermost level of population structure with those obtained considering also secondary structures.

7.4 Towards an Association Mapping Approach in Globe Artichoke

Advances in 'next-generation sequencing' (NGS), through multiplexed sequencing of barcoded samples in a single sequencing run, have decreased experimental costs. This allowed the re-sequencing of entire germplasm collections and to obtain a comprehensive picture of existing variation in a crop species. However, so far, the globe artichoke takes little advantages of this high-throughput technology.

A GWA analysis in globe artichoke was reported by Comino and colleagues (2016) and Acquadro and colleagues (2017a). This study was based on a panel of 111 accessions belonging to the living collection maintained at AGRIS (Sardinia, Italy). This collection was the same previously investigated by means of microsatellite (SSRs) genotyping (Rau et al. 2015; Rau et al. submitted; see also the first section of this chapter).

The collection was characterized with a genotyping by sequencing approach, using a two-enzyme restriction site-associated DNA sequencing (RADseq; Fig. 7.6) technology

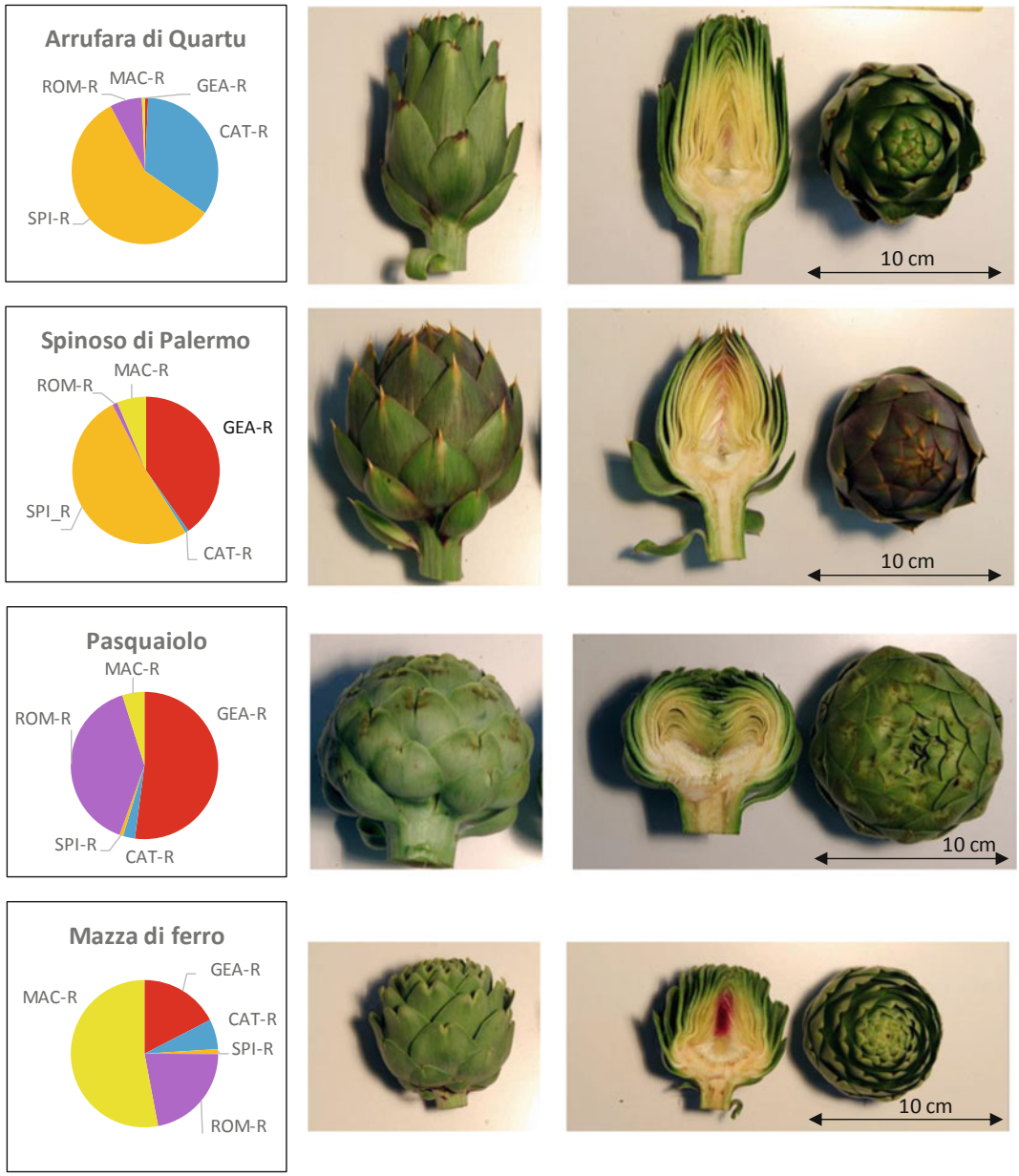


Fig. 7.4 Example of admixed accessions. The pies on the left represent the contribution of each SSR group to the genetic profile of the accessions. The colours represent the different SSR groups and are the same used in Fig. 7.1. The suffix '-R' means 'related' and indicates the correspondence between genetic group and phenotypic typology as shown in Fig. 7.1: CAT = 'Catanesi' and 'Violetti di Provenza', SPI = 'Spinosi', GEA = Green *et alia*, MAC = 'Macau', ROM = 'Romaneschi'. For example, the first line shows 'Arrufara di Quartu' that resulted admixed between Group 2a (that was 'Spinosi-related', SPI-R, as it captured most of 'Spinosi' types) and Group 1 (that was typical of 'Catanesi' and 'Violetti di Provenza', CAT-R; see also Fig. 7.3). Photographs by L. Baghino (AGRIS)

(Peterson et al. 2012) with minor modifications (Acquadro et al. 2016). Peterson and colleagues (2012) proposed the use of two restriction enzymes (a frequent and a rare cutter) followed by ligation of two adapters (P1 and P2), coupled with a size selection procedure (range: 300–

Fig. 7.5 Degree of variation for pigmentation (top line) and spininess (bottom line). Top line, from the left to the right: 'Camus Bretagne', 'Bianco di Ostuni', 'Locale di Mola', 'Violetto di Provenza', 'Violetto di Teramo', 'Violetto di Margot'. Bottom line, from the left to the right: 'Macau' (non-spiny), 'Bianco Tarantino' (with mucro), 'Kiss of Durgum' (GEA), 'Spinoso Sardo', 'Spinoso di Sciacca', 'Spinoso Violetto di Liguria' (all spiny types). Photographs by L. Baghino (AGRIS)

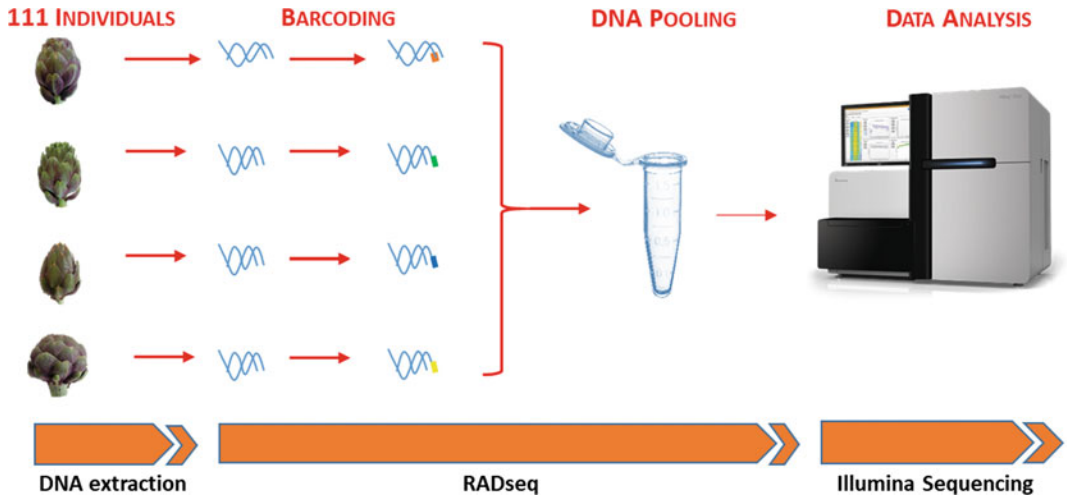


Fig. 7.6 Schematic workflow of the RADseq approach used to characterize 111 globe artichoke accessions

450 bp) to remove undesired homo-adapter fragments (P1/P1 and P2/P2). This procedure makes it possible to produce only a subset of genomic restriction digest fragments; however, the number of homo-adapter fragments P2/P2 and P1/P1 remains high. Acquadro and colleagues (2016) reported on a two-enzyme RADseq protocol, with an alternative pooling procedure and fragment selection step, guided by a biotin-/streptavidin-coated beads-based step, before the final PCR enrichment, which consistently reduces the presence of short P2/P2 and large P1/P1 fragments resulting in enrichment of P1/P2 fragments.

RAD tag libraries for each genotype were produced, pooled and sequenced on a NextSeq 500 Illumina platform (SE 1 × 75). Raw reads obtained from sequencing were aligned to the globe artichoke reference genome (www.artichokegenome.unito.it), and GWA analysis was conducted using loci having, on average, a minimum of 15 reads for sample (DP > 15).

Overall, 9694 robust SNPs (DP > 15) have been identified, of which about 1400 belonged to the gene space (Comino et al. 2016; Acquadro et al. 2017a). Population structure was inferred, using the whole SNP data set, by applying the unpaired group method with arithmetic average (UPGMA), the Principal Coordinate Analysis (PCoA) and the Bayesian clustering implemented in *fastSTRUCTURE v1.0* software. Accessions were classified into three main groups (based on a coefficient of membership, q_i, higher than 0.60) that were associated with varietal types. The highest divergence was seen between the groups comprising most of 'Spinosi' (SPI) and that comprising 'Catanesi' and 'Violetti di Provenza' (CAT-VPR). The third group that included most of 'Violetti di Toscana' (VTO), 'Romaneschi' (ROM), 'Macau' (MAC) and 'Verdi' (GEA) types was in an intermediate position and showed the highest variability. The 'Catanesi' was clearly distinct from the other groups and appeared genetically quite uniform. Eleven accessions had ambiguous membership (q_i for the prevalent group < 0.60) and were classified as admixed (Fig. 7.7).

Moreover, in Acquadro et al. (2017a), association analyses between SNP and phenotypic traits (previously collected and reported in Rau et al. submitted) were performed by applying the software *TASSEL* (Bradbury et al. 2007). As a result, the high-throughput sequencing of RAD tags allowed the establishment of many significant associations between SNP alleles and different key breeding traits including head traits (Fig. 7.8) and several other morpho-phenological features.

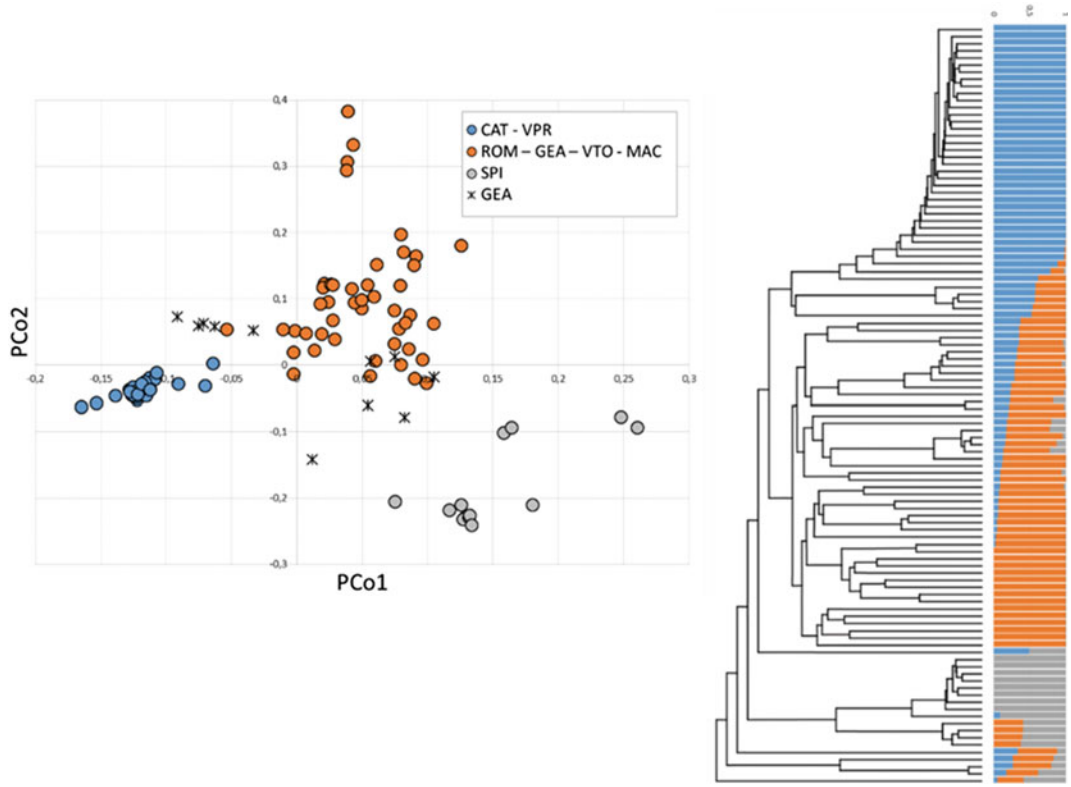


Fig. 7.7 Genetic structure of the globe artichoke collection analysed using 9k SNPs. Left: result of Principal Coordinate Analysis (PCoA). Right: results of multivariate cluster analysis (UPGMA method) fastSTRUCTURE analysis. Four groups are highlighted: (I) 'Catanesi' and 'Violetti di Provenza' (CAT-VPR), (II) 'Violetti di Toscana' (VTO), 'Romaneschi' (ROM), 'Macau' (MAC), (III) 'Spinosi' (SPI) and (IV) Green types plus other undefined types (GEA) types

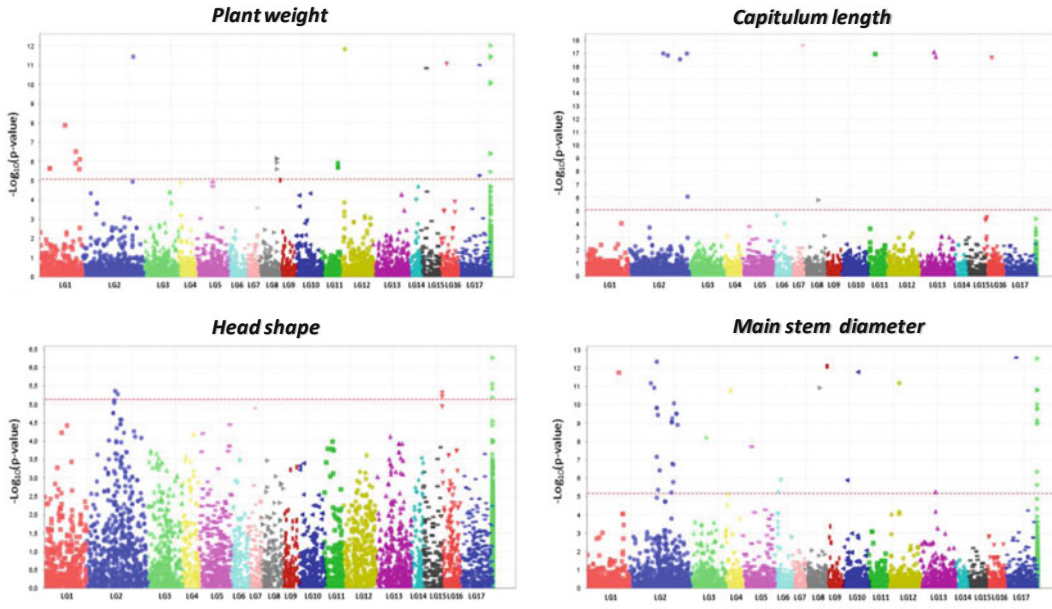


Fig. 7.8 Manhattan plots showing the results of association mapping for four traits. Dashed red line: threshold *P* value (loci that showed peaks above the red lines can be considered significantly associated with the trait)

7.5 Perspectives

Globe artichoke is a plant species with extreme plasticity (e.g. Crinò et al. 2008). This implies that the same collection should be better characterized under different environmental conditions to disentangle genotype mean effects and genotype by environment interaction effects. Moreover, even if strong efforts are done to remove the confounding effect of population structure, false-positive associations still occur due to the high number of statistical tests and to other factors such as erroneous genotype calling at some loci, small population size and synthetic associations (but see Liu and Yan 2019 for a review). This implies that the identified QTLs must be validated by independent experiments. At least two methodologies can be pursued in future. The first consists in validating candidate QTLs or genes in different genetic backgrounds (populations). The second uses knockout, over-expression, genetic complementation, transgenesis and genome editing of candidate genes to collect functional proofs. Four globe artichoke genotypes, representatives of the core varietal types, were recently re-sequenced (Acquadro et al. 2017b). Moreover, a set of SNPs regularly spaced along the chromosomes was identified and annotated. This resource will be useful to pinpoint the regions responsible for the QTLs and will support marker-assisted selection after that reliable markers for genes of interest will be obtained (Acquadro et al. 2017b).

References

Acquadro A, Barchi L, Portis E, Carrasquilla-Garcia N, Tirone M, Lanteri S, Comino C (2016) RAD2seq: an efficient protocol for plant genotyping by sequencing. Acta Hortic 1147:1–8. https://doi.org/10.17660/ActaHortic.2016.1147.1

Acquadro A, Cerruti E, Comino C, Rodriguez M, Pisanu AB, Baghino L, Sanna D, Attene G, Portis E, Rau D (2017a) Rad-tag sequencing and association mapping in a globe artichoke germplasm collection. In: Proceedings of the joint congress SIBV-SIGA, Pisa, Italy, 19/22 Sep 2017. ISBN 978-88-904570-7-4. Poster Communication Abstract—2.19

Acquadro A, Barchi L, Portis E, Mangino G, Valentino D, Mauromicale G, Lanteri S (2017b) Genome reconstruction in *Cynara cardunculus* taxa gains access to chromosome-scale DNA variation. Sci Rep 7:5617

Basnizki J, Zohary D (1987) A seed planted cultivar of globe artichoke. HortScience 22:678–679

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 2633–2635

Brown AHD (1989) Core collections: a practical approach to genetic resources management. Genome 31:818–824

Cadinu M, Mallica GM, Lanteri S, Repetto A, Frau A, Portis E, Baghino L, Pisanu AB (2007) Possibilità di rilancio per lo Spinoso Sardo. Speciale carciofo. L'informatore Agrario 22:44–46

Calabrese N (2016) From Mediterranean basin to the Andean mountains, the long journey of the artichoke that unites different people and cultures. In: IX International symposium on artichoke, cardoon and their wild relatives, pp 23–34

Ciancolini A, Rey NA, Pagnotta MA, Crinò P (2012) Characterization of Italian spring globe artichoke germplasm: morphological and molecular profiles. Euphytica 186:433–443

Comino C, Portis E, Gianoglio S, Acquadro A, Pisanu AB, Baghino L, Sanna D, Rau D, Rodriguez M, Attene G (2016) Towards a genome-wide association (GWA) mapping approach in globe artichoke. Acta Hortic 1147. ISHS 2016. https://doi.org/10.17660/actahortic.2016.1147.7

Crinò P, Pagnotta MA (2017) Phenotyping, genotyping, and selections within Italian local landraces of Romanesco globe artichoke. Diversity 9:14

Crinò P, Tavazza R, Munoz NAR, Nisini PT, Saccardo F, Ancora G, Pagnotta MA (2008) Recovery, morphological and molecular characterization of globe artichoke 'Romanesco' landraces. Genet Resour Crop Evol 55:823–833

De Vos NE (1992) Artichoke production in California. HortTechnology 2:438–444

Deidda M (1967) Contributo al miglioramento genetico del carciofo. In: Atti 1° Congresso Internazionale Di Studi sul Carciofo. Minerva Medica, Torino, pp 157–174

Dellacecca VV, Magnifico V, Marzi V, Porceddu E, Scarascia Mugnozza GT (1976) Contributo alla conoscenza delle varietà di carciofo coltivate nel mondo. In: Proceeding of international congress on artichoke. Minerva Medica, Torino, pp 119–316

Elia A, Miccolis V (1996) Relationship among 104 artichoke (*Cynara scolymus* L.) accessions using cluster analysis. Adv Horticult Sci 10:158–162

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

Gaudeul M, Till-Bottraud I, Barjon F, Manel S (2004) Genetic diversity and differentiation in *Eryngium alpinum* L. (Apiaceae): comparison of AFLP and microsatellite markers. Heredity 92:508

Gill DE, Chao L, Perkins SL, Wolf JB (1995) Genetic mosaicism in plants and clonal animals. Annu Rev Ecol Syst 26:423–444

Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ, Eskin E (2008) Efficient control of population structure in model organism association mapping. Genetics 178:1709–1723

Lanteri S, Di Leo I, Ledda L, Mameli MG, Portis E (2001) RAPD variation within and among populations of globe artichoke cultivar 'Spinoso sardo'. Plant Breeding 120:243–246

Lanteri S, Acquadro A, Saba E, Portis E (2004) Molecular fingerprinting and evaluation of genetic distances among selected clones of globe artichoke (*Cynara cardunculus* L. var. *scolymus* L.). J Hortic Sci Biotechnol 79:863–870

Liu HJ, Yan J (2019) Crop genome-wide association study: a harvest of biological relevance. Plant J 97:8–18

Lushai G, Loxdale HD (2002) The biological improbability of a clone. Genet Res 79:1–9

Mariette S, Chagné D, Lézier C, Pastuszka P, Raffin A, Plomion C, Kremer A (2001) Genetic diversity within and among *Pinus pinaster* populations: comparison between AFLP and microsatellite markers. Heredity 86:469–479

Mauro RP, Portis E, Acquadro A, Lombardo S, Mauromicale G, Lanteri S (2009) Genetic diversity of globe artichoke landraces from Sicilian small-holdings: implications for evolution and domestication of the species. Conserv Genet 10:431–440

Mauro RP, Portis E, Lanteri S, Mauromicale G (2012) Genotypic and bio-agronomical characterization of an early Sicilian landrace of globe artichoke. Euphytica 186:357–366

Mauro RP, Portis E, Lanteri S, Monaco AL, Mauromicale G (2015) Clonal selection in a globe artichoke landrace: characterization of superior germplasm to improve cultivation in Mediterranean environments. J Agric Sci 153:102–113

Mauromicale G, Ierna A (2000) Panorama varietale e miglioramento genetico del carciofo. L'Informatore Agrario 26:39–45

McKey D, Elias M, Pujol B, Duputié A (2010) The evolutionary ecology of clonally propagated domesticated plants. New Phytol 186:318–332

Orive ME (2001) Somatic mutations in organisms with complex life histories. Theor Popul Biol 59:235–249

Pagnotta MA, Fernández JA, Sonnante G, Egea-Gilabert C (2017) Genetic diversity and accession structure in European *Cynara cardunculus* collections. PLoS One 12:e0178770

Pavan S, Curci PL, Zuluaga DL, Blanco E, Sonnante G (2018) Genotyping-by-sequencing highlights patterns of genetic structure and domestication in artichoke and cardoon. PLoS One 13:e0205988

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(5): e37135. PubMed http://dx.doi.org/10.1371/journal.pone.0037135

Porceddu E, Dellacecca V, Bianco VV (1976) Classificazione numerica di cultivar di carciofo. In: Atti 2nd Congresso Internazionale Di Studi sul Carciofo. Minerva Medica, Torino, pp 1105–1119

Portis E, Mauromicale G, Barchi L, Mauro R, Lanteri S (2005) Population structure and genetic variation in autochthonous globe artichoke germplasm from Sicily Island. Plant Sci 168:1591–1598

Pritchard JK, Stephens M, Donnelly P (2000a) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Pritchard JK, Stephens M, Rosenberg NA, Donnelly P (2000b) Association mapping in structured populations. Am J Hum Genet 67:170–181

Rau D, Comino C, Acquadro A, Pisanu AB, Baghino L, Sanna D, Rodriguez M, Attene G, Portis E (2015) First insights into the population structure, linkage disequilibrium & association mapping in a globe artichoke germplasm collection. In: Proceedings of the joint congress SIBV-SIGA, Milano, Italy, 8/11 Sep 2015. ISBN 978-88-904570-5-0. Poster Communication Abstract—D.34

Schlötterer C (2004) The evolution of molecular markers—just a matter of fashion? Nat Rev Genet 5:63

Schreiber M, Stein N, Mascher M (2018) Genomic approaches for studying crop evolution. Genome Biol 19:140

Vanella B, Porceddu E, De Pace C (1981) Applicazioni di metodi di analisi numerica per il miglioramento genetico del carciofo. In: Atti 3° Congresso Internazionale Di Studi sul Carciofo. Laterza, Bari, pp 797–807

Whitham TG, Slobodchikoff CN (1981) Evolution by individuals, plant–herbivore interactions, and mosaics of genetic variability: the adaptive significance of somatic mutations in plants. Oecologia 49:287–292

Yu J, Buckler ES (2006) Genetic association mapping and genome organization of maize. Curr Opin Biotechnol 17:155–160

8 Cytological and Molecular Cytogenetic Insights into the *Cynara cardunculus* Genome

Debora Giorgi, Gianmarco Pandozy, Anna Farina, Valentina Grosso, Paola Crinò and Sergio Lucretti

Abstract

Before the recent advances in molecular biology and sequencing techniques enabled a deep knowledge of the complete nucleotide sequence of many plant genomes, cytological and karyological studies have provided the first information concerning the genome organization of living beings. As stated by Figueroa and Bass (2010), plant cytogeneticists were among the earliest researchers who started to visualize genomes nearly a century before the first plant genome was sequenced (The *Arabidopsis* genome initiative 2000). In spite of the high development of next-generation sequencing approach and of the dramatic increase in available data regarding sequenced genomes, the cytological analysis of cell nucleus content remains a valuable tool for evolutionary studies and for structural and functional genomic research. Since Wilhelm Gottfried von Waldeyer-Hartz coined the term chromosome in 1888, many cytological techniques have been developed to disclose nucleus content and to allow a detailed description of the chromosome complement of plant species. Recently, some of these techniques have been applied to the cytological analysis and the molecular cytogenetic characterization of *Cynara cardunculus* complement, a traditional vegetable crop of the Mediterranean basin. The DNA content and a detailed karyotype of the two cultivated botanical varieties *C. cardunculus* var. *altilis* DC (cultivated cardoon) and *C. cardunculus* L. var. *scolymus* L. (globe artichoke) have been reported (Khaldi et al. 2014; Falistocco 2016; Giorgi et al. 2016). In this chapter, some of the used methodological approaches and the main results obtained by different authors will be discussed.

8.1 Introduction

Considering the general definition of cytology as a branch of biology that deals with the study of cells, particularly of chromosomes, and other cell organelles (Singh 2018), the present chapter will focus on studies related to *Cynara cardunculus* nucleus, in terms of both DNA content and its organization in chromosome structures.

Since Robert Hooke in 1665 coined the term "cell" after observing a slide of cork on microscope, great progress has occurred in cytology. This happened also thanks to the technological achievements as microscope refinement and to the development of a number of methodologies to investigate nuclear DNA content and

D. Giorgi (✉) · G. Pandozy · A. Farina · V. Grosso · P. Crinò · S. Lucretti
Italian National Agency for New Technologies, Energy and Sustainable Economic (ENEA), via Anguillarese 301, 00123 Rome, Italy
e-mail: debora.giorgi@enea.it

E. Portis et al. (eds.), *The Globe Artichoke Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-030-20012-1_8

chromosome structure through classical cytological techniques and molecular cytogenetic approach.

The nuclear DNA content was considered as a constant value (C-value; Swift 1950) within an organism, and the un-replicated haploid chromosome complement (n) uniquely represents the basic DNA content of a species. During the mitotic cell cycle, the nuclear DNA amount dynamically changes through the S phase from 2C in phase G_0/G_1 to 4C in phase G_2, where two copies of the genome are present. The knowledge of the C-value of an organism is of relevance either for several disciplines such as systematics, ecology and molecular biology or for evolutionary and phylogenetic studies (Bennett and Leitch 2011).

The estimation of nuclear DNA content can be performed mainly by two approaches: analysis of DNA extracted from a large number of cell, as in chemical analysis (Schdmidt and Thannhauser 1945), and measurements of individual nuclei (see Greilhuber 2008 for a review). Considering that cells in tissues and organs may contain different amount of DNA depending on the cell cycle phase (Van'T Hof 1965), the single nuclei procedure could represent a more precise and reliable approach (Doležel and Bartoš 2005). One of the earliest and best established procedures for quantifying nuclear DNA was based on the Feulgen staining (Feulgen and Rossenbeck 1924) of root tips and squash preparation, followed by the absorbance measurements of nuclear DNA in micro-densitometric analysis (Swift 1950). However, as stated by Bennet and Smith (1976), this procedure "has several time-consuming steps which can be subjected to error". An alternative and more effective approach is represented by flow cytometry (FCM) (Doležel and Greilhuber 2010), a technique that allows a very fast and statistically accurate analysis of thousands of nuclei stained with DNA-selective fluorophores. FCM requires the preparation of nuclei in suspension and has been greatly enhanced by the development of the fast and easy method of nuclei isolation based on fresh leaf tissue chopping (Galbraith et al.1983).

Despite some still existing issues, as the possible interference of cytosolic compound with DNA staining, or the need of some internationally agreed DNA reference standards (Doležel and Bartoš 2005), FCM is now the prevailing method for the measurement of nuclear DNA content in plants, and a number of data concerning plant C-value have been produced and collected in electronic databases (Bennett and Leitch 2005). However, plant DNA content archive is far from being completed, as it is known for a fraction of plant species (e.g. only 1.4% of angiosperms C-value has been reported; Hanson et al. 2003; Greilhuber 2008).

This ascertainment also applies to the largest angiosperm family such as that of *Asteraceae* that comprises about 24,000 species, 1600–1700 genera and 29 tribes. However, most genome size data (96.74%) only refer to five *Asteraceae* tribes: *Anthemideae*, *Cichorieae*, *Senecioneae*, *Heliantheae* and *Cardueae* to which the genus *Cynara* belongs. Such tribes contribute to around 50% of the species richness of the family (Funk et al. 2009; Vallès et al. 2013). Although the coverage of the family is still incomplete, a public database of *Asteraceae* genome sizes (www.asteraceaegenomesize.com) was built (Garnatje et al. 2011) and updated in 2013, providing an easy access to the available information on nuclear DNA content in the family. Genus *Cynara* accounts for only *C. cardunculus* var. *scolymus* nuclear DNA content and genome size (Marie and Brown 1993), but further studies have been carried out by Kaldhi et al. (2014) and Giorgi et al. (2016) which will be discussed in the following Sect. 8.2.2.

One of the main targets of plant cytogenetics is to count and discriminate the chromosomes that constitute the complement of a species, with the final aim of describing its karyotype. As stated by Battaglia first (1994), and recently by Singh (2018), "*Karyotype is the number, size, and morphology of the chromosome set of a cell in an individual or species*". Firstly, the analysis and classification of chromosomes are based on physical characteristics as chromosome size, position of the primary constriction

(or centromere), relative length of whole chromosomes and chromosome arms, presence of a secondary constriction and of a satellite body. Somatic chromosomes in mitotic metaphase are the most used for the analysis of the morphological markers described above, because they are well condensed and defined. The standard approach to handle the mitotic chromosomes of all plant species consists in the collection of fast growing specimens (e.g. root tips), tissue pre-treatments, fixation and DNA staining. Explants' pre-treatments are necessary for several purposes: (i) to increase the number of cells in metaphase and of chromosomes at the metaphase plate, by blocking the formation of spindles during mitosis; (ii) to enhance chromosome contraction and clarify the chromosome morphology; (iii) to clear the cytoplasm. To this aim, a number of pre-treatment agents are available and should be tested depending upon plant species. The same principle would also apply to tissue fixation and DNA staining: fixation is an essential step to get good-quality chromosomes, avoiding their distortion, swelling or shrinkage, while DNA staining is necessary to visualize at microscope the chromosomes and their structures. The different fixatives, dyes and the main techniques used in plant cytogenetics can be found in the comprehensive manual compiled by Singh (2018).

In some species as rice, tomato, rose, cocoa (Andras et al. 1999) and even globe artichoke and cardoon (Giorgi et al. 2013a, 2016; Falistocco 2016), chromosome discrimination can be hampered by the small size and the similar shape of mitotic chromosomes. The identification of the morphological markers previously described becomes a hard task. In such cases, the use of additional techniques, such as chromosome banding (Giemsa C- and N-banding) or molecular labelling by fluorescent in situ hybridization (FISH), can be of help for improving cytogenetic resolution and karyotyping (see Singh 2018). FISH allows the direct localization of properly labelled specific DNA sequences (e.g. probes) on chromosomes. To enhance plant chromosome discrimination, different kinds of probes can be used, but the most common ones consist of repetitive sequences, as the tandemly repeated DNA sequences encoding ribosomal RNAs (rDNA and rRNA, respectively), and satellite and microsatellite DNA, which are abundant and widely spread in most plant genomes.

Ribosomal DNA has been extensively used at both molecular and cytogenetic levels to infer phylogenetic relationships (Gatto et al. 2013) and to allow chromosome identification in a number of plant species. Garcia et al. (2010) studied the dynamic evolution of 5S and 35S ribosomal RNA genes in about 200 species belonging to the *Asteraceae* family, and several other cytogenetic studies on *Asteraceae* species used rDNA (Fregonezi et al. 2004; Weiss-Schneeweiss et al. 2008) to help in chromosome characterization.

Microsatellite DNA or simple sequence repeats (SSRs) consist in very short tandem repeats (mainly di- or tri-nucleotides) differently spread into the plant genome. They may differ in terms of abundance, organization and copy number and can be easily transformed in synthetically labelled oligonucleotides which have been extensively used in several plant karyotyping as probes (Cuadrado et al. 2008; Cuadrado and Jouve 2010; Fu and Tang 2015). Both rDNA and SSR oligonucleotides have been employed for the characterization of the chromosome complement of *C. cardunculus* varieties. In addition, original techniques as fluorescent in situ hybridization in suspension (FISHIS, Giorgi et al 2013b) and a fast FISH method have been used and will be shortly described.

8.2 Cytology

8.2.1 An Overview

The cytological study of *C. cardunculus* has taken advantage of the improvements and methodologies mentioned above, considering that all cytological studies related to the complement of *C. cardunculus* species are quite recent, with the exception of the first chromosome number

identification performed by Ancora et al. (1981) and of the analysis on globe artichoke DNA content by Marie and Brown in 1993.

The DNA content of globe artichoke and cardoon has been recently estimated by Khaldi et al. (2014) and by Giorgi et al. (2016) who both used the flow cytometric approach, considered as a reliable method to this aim (Doležel and Greilhuber 2010).

8.2.2 DNA Content Estimation in *Cynara cardunculus* Varieties

In 2012, Bennett and Leitch listed genome sizes for more than 7500 plant species; since then, there have been a regular increasing of papers dealing with plant DNA content (more than 80 papers/year till 2017) which have brought the gross inventory up to about 8000 species, as of now. The wide diffusion of flow cytometry (FCM) techniques in several laboratories has increased and supported the interest of knowing plant genome size for enforcing cellular, molecular, genomic, taxonomy and evolutionary genetic studies (Ochatt 2008; Sliwinska 2018).

During years, FCM got a predominant role in DNA content evaluation thanks to the easy and fast procedure developed for plant nuclei isolation and DNA labelling (Galbraith et al. 1983; Doležel et al. 1989), and quickness of analysis. Few years ago, a DNA content list has been released focused on DNA content estimation in the *Asteraceae* family for 1279 species, and more than 75% were FCM measurements (*Asteraceae* DNA content data bank, GSAD: http://www.etnobiofic.cat/gsad_v2/index.php). This database presents two papers dealing with *C. cardunculus* DNA content, referred as Khaldi et al. (2014) and Marie and Brown (1993). The first one lists data from twelve accessions of *C. cardunculus* var. *altilis* (ten wild cardoon accessions and two cultivated ones) and one accession of *C. cardunculus* var. *scolymus* (artichoke) showing a range of variability in DNA content, from 2C = 1.98 to 3.03 pg in wild cardoon and 2C = 2.10–2.11 pg for two accessions of cultivated cardoon, and 2C = 2.05 pg for a single variety of globe artichoke. Marie and Brown (1993) stated a DNA content of 2C = 2.22 pg for one accession of globe artichoke. Even if FCM is doubtless the preferred method for genome size determination, yet it deserves a careful experimental design to prevent problematic measurements and misleading data evaluation (Greilhuber 1988; Wang et al. 2015). The main issues affecting DNA content measurements in plants are related to (a) a proper selection of the isolation tissue (Galbraith et al. 1983); (b) the choice of an effective extraction buffer (Loureiro et al. 2006); (c) the nuclear DNA labelling with an intercalating dye (Doležel et al. 1992); (d) the usage of an available and well-described internal standard (Doležel and Greilhuber, 2010). In Table 8.1, a short list of reference papers which can be linked to *Cynara* ssp. DNA content is presented, which rises from the first plant standard DNA content chemically estimated, up to the most recent ones.

Most of the listed references consider the first colorimetric determination of DNA content in onion (*A. cepa* Ailsa Craig, Van'T Hof 1965) as a primary standard and the DNA content of pea (*P. sativum* L. cv Ctirad) as "golden standards" for local calibration and FCM analysis in plant. The availability and the use of a set of internationally agreed DNA reference standards are crucial for data comparison among laboratories (Doležel et al. 1998; Doležel and Bartoš 2005; Doležel and Greilhuber 2010). In agreement with these authors, we used the available plant species of *Pisum sativum* cv Ctirad as an internal standard for *Cynara* ssp. DNA content measurements (DNA reference standards freely available at http://olomouc.ueb.cas.cz/plant-dna-cytometry-standards).

The flow cytometry analysis was performed on nuclei isolated from fresh and fixed tissue of globe artichoke and cardoon. When using the most common isolation buffers (Loureiro et al. 2006) and the classical method of chopping fresh leaf tissues (Galbraith et al. 1983), a low yield of *Cynara* nuclei in suspensions has been obtained, which almost prevented a significant FCM analysis. A large increase of nuclei yield (about 5X) was achieved after leaf fixation in 2%

Table 8.1 DNA reference standards "involved" in *Cynara* ssp. genome size measurement

References papers	Species	Methods	Stain	2C DNA pg	DNA standard pg	DNA reference standard
Van't Hof (1965)	*A. cepa* Ailsa Craig	D[a]	C[b]	33.5 (34.52)		Van't Hof (1965)
Bennett and Smith (1976)	*P. sativum*	D[a]	Feulgen	9.73		
Galbraith et al. (1983)	Chicken	C[b]	MI	2.33		Galbraith et al. (1983)
	P. sativum	FCM		7.72	Chicken 2.33	
Tiersch et al. (1989)	Human male leucocytes	FCM	PI	7.0	Chicken 2.5	Rasch et al. (1971)
Arumuganathan and Earle (1991)	*P. sativum*	FCM	PI	8.18	Chicken 2.33	Galbraith et al. (1983)
	A. cepa	FCM	PI	31.69	Chicken 2.33	
Marie and Brown (1993)	*P. sativum*	FCM	PI or EB	8.37	Petunia	Marie and Brown (1993)
	A. cepa	FCM	PI or EB	32.74	Petunia	Marie and Brown (1993)
	C. cardunculus scolymus	FCM	PI or EB	2.2	Petunia 2.85	Marie and Brown (1993)
	Petunia hybrida	FCM	PI or EB	2.85	Chicken 2.33	Galbraith et al. (1983)
Doležel et al. (1998)	*P. sativum* cv Ctirad	FCM	PI	9.09	Human leucocytes 7.0	Tiersch et al. (1989)
	A cepa cv Alice	FCM	PI	34.89		
Johnston et al. (1999)	*H. vulgare cv.* Sultan	FCM	PI	11.12	*A. cepa* 33.5	Vant' Hof (1965)
	Gallus domesticus	FCM	PI	2.49–3.01		Johnston et al. (1999)
	P. sativum Minerva Maple	FCM	PI	9.39–9.21	*Hordeum*/ chicken	Idem
	A. cepa Ailsa Craig	FCM	PI	32.97	*Hordeum*	Idem
Khaldi et al. (2014)	*C. cardunculus scolymus* Violet d'Hyeres	FCM	PI	2.05	*P. sativum* 8.37	
	C. cardunculus altilis W[c]	FCM	PI	1.98–3.03	*P. sativum* 8.37	
	C. cardunculus altilis c[d]	FCM	PI	2.10	*P. sativum* 8.37	
Giorgi et al. (2016)	*C. cardunculus scolymus* c[d]	FCM	PI	2.4	*P. sativum* 9.09	Doležel et al. (1998)
	C. cardunculus altilis c[d]	FCM	PI	2.2	*P. sativum* 9.09	Doležel et al. (1998)

[a]*D* Densimetry
[b]*C* Colorimetric
[c]*W* Wild
[d]*c* cultivated

formaldehyde and standard chopping of leaves in LB isolation buffer. The addition of a higher amount of detergent can be beneficial to nuclei isolation (0.7% Triton X-100 instead of 0.1%, named buffer LB07). This way, the leaf tissues can also be mechanically homogenized in LB07 (Ultraturrax Mini—IKA, with a GN5 generator at 9500 rpm for 10 s) to increase further nuclei yield. Fixation was also helpful to reduce interactions among DNA and antioxidant molecules, of which *Cynara* species are rich of, therefore stabilizing DNA labelling (Greilhuber 1988). These modified isolation procedures resulted to an average amount of more than 600 nuclei per milligram of leaf tissue, which showed as adequate for a real flow cytometric analysis in a difficult plant such as *C. cardunculus*, generating a good DNA fluorescence histogram, with a low noise background (Fig. 8.1).

In our hands, the DNA content for cardoon and globe artichoke was estimated to be 2C = 2.20 pg ± 0.04 and 2C = 2.40 pg ± 0.04 corresponding to a C genome size of 1.07×10^9 bp and 1.17×10^9 bp, respectively (Giorgi et al. 2016).

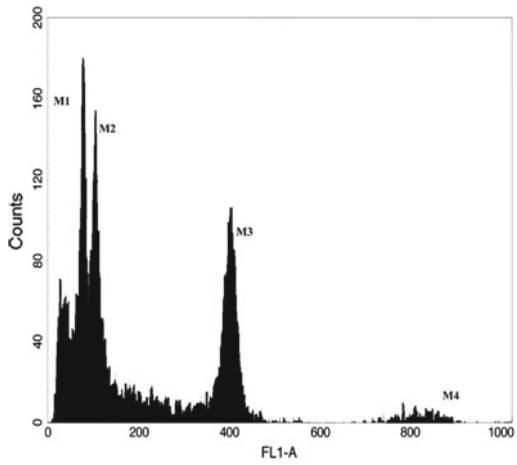


Fig. 8.1 DNA content histogram. Flow cytometry analysis of DNA fluorescence peaks (FL1A) from G0/G1 (2C) PI stained 2C leaf nuclei from cultivated cardoon (M1), globe artichoke (M2) and *P. sativum* (pea) (M3) used as an internal standard. Pea was set at channel 400; M4 shows pea G2 nuclei (4C DNA content). From Giorgi et al. (2016)

The genome size measured for leafy cardoon and globe artichoke is slightly different to those previously reported by Marie and Brown (1993) and by Khaldi et al. (2014). The disparities between our estimates in DNA content and those of other authors could be ascribed to the DNA content considered for the applied internal standard (*P. sativum*). Both papers counted a pea DNA content of 2C = 8.37 pg, while we employed *P. sativum* cv "Ctirad" with an agreed DNA content as 2C = 9.09 pg (Doležel et al. 1998). The exploit of a diverse pea 2C value, the use of fixed tissues and a possible intra-specific genome size variability, which were reported by Khaldi et al. (2014) among ten accessions of wild cardoon, may well account for such little differences in 2C DNA content estimates, assessed on different *C. cardunculus* samples. A well-tuned FCM analysis might effectively contribute to a better knowledge on plant DNA content, and keeps up with the availability of always faster and less expensive procedures for genome sequencing.

8.3 Karyotyping

8.3.1 Overview

Preliminary studies on globe artichoke karyotype were first performed by Giorgi et al. (2013a), using the classical cytological approach based on chromosome morphology analysis, and were further investigated through molecular cytogenetic characterization by Falistocco (2016) and Giorgi et al. (2016), in both globe artichoke and cardoon varieties. Fluorescent in situ hybridization was applied to nuclei in suspension (FISHIS) by Giorgi et al. (2016) and to chromosomes on slide (FISH) (Falistocco 2016; Giorgi et al. 2016), with the aim of selecting reliable cytogenetic probes and producing a detailed karyotype of *C. cardunculus* varieties, respectively. Here, we present some methodological approaches and the main results achieved in globe artichoke and cardoon karyotyping.

8.3.2 Setting Conditions for *C. Cardunculus* Chromosomes Analysis

As previously stated, the detailed morphological or molecular analysis of the chromosome complement of a plant species requires the availability of complete metaphases and of visible, good-quality chromosomes. To this purpose, several steps and tissue pre-treatments are necessary and they must be optimized depending on the plant species (Fig. 8.2).

In preliminary cytogenetic studies of globe artichoke, Giorgi et al. (2013a) evaluated the reliability of the two spindle inhibitors such as amiprophos-methyl (APM) and 8-hydroxyquinoline (8HQ) as well as that of the classical method with ice-cold water, to increase the mitotic index in root tip meristems, together with two different procedures to obtain good chromosome spread and DNA staining.

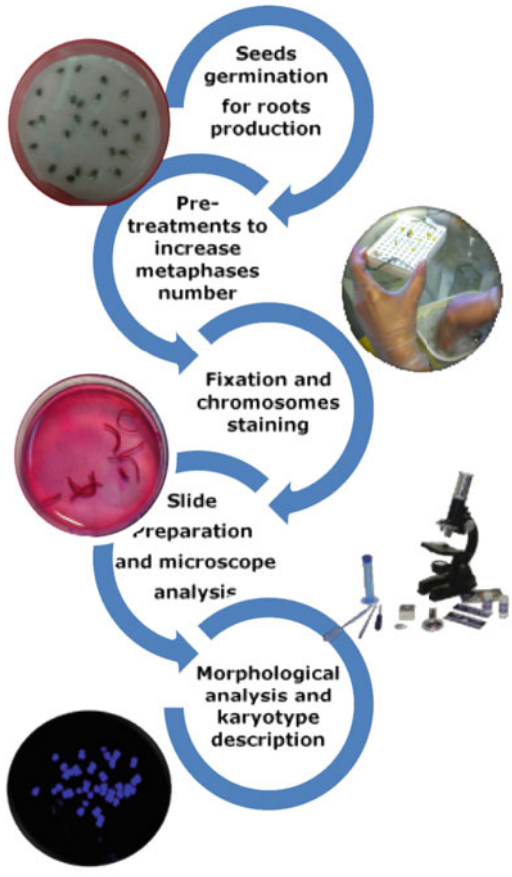


Fig. 8.2 Preliminary steps to cytogenetic analysis

In the APM treatment, germinated seeds were immersed for 5 h in an aerated Hoagland's solution (Gamborg 1975) containing different concentrations of APM (2.5, 12, 32 and 64 μM). The APM 2.5 μM treatment was two times performed: the first for 5 h and the second for 18 h. After the treatment, seeds were rinsed out with water and overnight incubated on ice.

In the 8HQ processing, actively growing roots (0.5–1 cm in length) were immersed in 2 mM 8HQ at room temperature for 3 and for 5 h.

Finally in ice-cold water treatment, germinated seeds or seedlings were kept on ice for 18–24 h. In a further experiment, a saturated aqueous solution of α-bromonaphthalene for 2–4 h at room temperature, a 30 μM solution of oryzalin for 20 h at 4 °C or a 5% colchicine solution for 1–2 h were also tested. After the mentioned treatments, roots were excised, fixed in Carnoy's solution (ethanol: glacial acetic acid 3:1) and stored (−20 °C) for at least 16 h. A sample of untreated roots, fixed in Carnoy's solution, was made and used as control. None of the pre-treatments gave relevant results in terms of a higher mitotic index: ice-cold water for 24 h showed a slight increase in the metaphase number compared to APM or 8HQ, but produced shortening and loss of chromosome morphology definition. Only the 8HQ treatment for 3 h maintained a suitable chromosome morphology, allowing a better definition of chromosome constrictions, and therefore, it became the treatment of choice for further cytological analysis (Giorgi et al. 2013a, 2016; Falistocco 2016).

The preparation of mitotic chromosome spreads usually requires the enzymatic digestion of the cell wall of root tips, followed by the cells spreading through the application of coverslip (Schwarzacher and Heslop-Harrison 2000). In addition to this standard procedure, Andras et al. (1999) developed a "Drop-spreading" technique particularly suited for plants with small chromosome. This method employs a hydrochloric acid treatment to hydrolyse the cytoplasm. This facilitates the subsequent removal of cytoplasmic

debris as well as the evaporative force of a methanol-based fixative to disperse the cytoplasm and to deposit the chromosomes in a single optical plane by dropping cells on slide. Both the approaches, the standard chromosome spread and the drop-spreading technique, were used and compared in *Cynara* (Giorgi et al. 2013a), confirming the latter procedure as the most efficient one for the small globe artichoke chromosomes and resulting in good metaphase spreads with low cytoplasm debris.

As regards chromosome staining, it was performed both by the Feulgen reaction, according to Kiernan (1999), and with 4,6-diamidino-2-phenylindole (DAPI) at the final concentration of 2 μg ml^{-1} (Schwarzacher and Heslop-Harrison 2000). The DAPI staining showed a slight increase in chromosome morphology definition (Giorgi et al. 2013a) and was used for *Cynara* karyotyping as well as in further cytogenetic experiments for chromosome counterstaining after FISH by Falistocco (2016) and Giorgi et al. (2016).

In both studies, the morphological analysis of globe artichoke and cardoon complements has confirmed their chromosome number to be 2n = 34 (Fig. 8.3). According to their size, chromosomes have been roughly divided for both varieties in three main groups composed by seven large (2.1 μm < tl < 3.2 μm), six medium (1.6 μm < tl < 2.0 μm) and four small (tl < 1.6 m) chromosome pairs; within each group, single chromosome discrimination and even chromosome pairing become quite hard if only based on morphological features.

8.3.3 Molecular Cytogenetic Characterization

Molecular cytogenetic characterization is based on fluorescence in situ hybridization techniques. By potentially producing chromosome-specific labelling pattern, it helps in chromosome discrimination and has been performed on the two botanical varieties of *C. cardunculus*, using rRNA gene sequences (Falistocco 2016; Giorgi et al. 2016) and SSR synthetic oligonucleotides as probes (Giorgi et al. 2016).

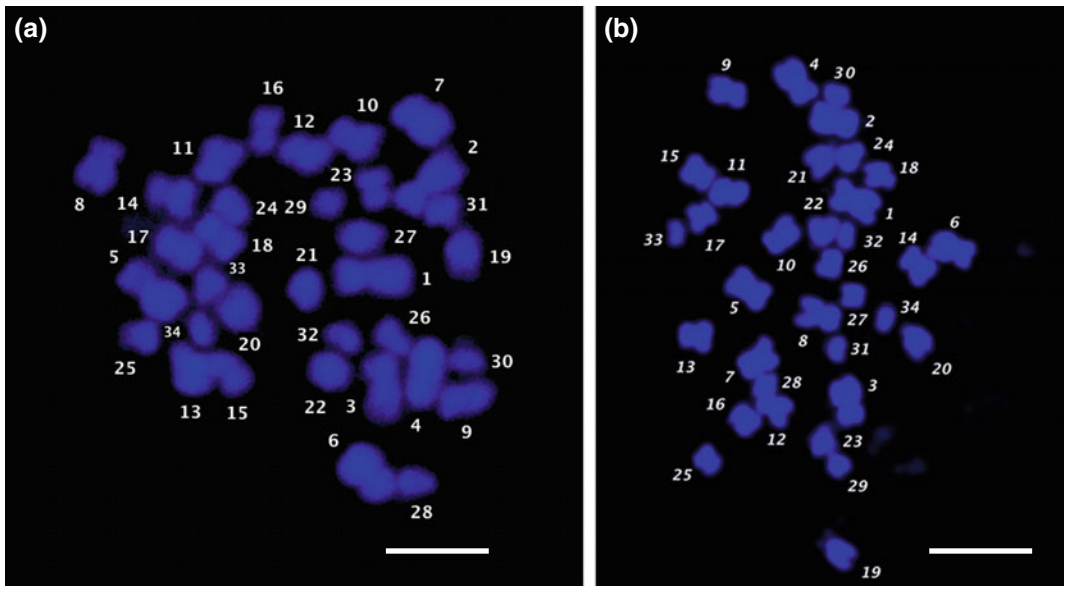


Fig. 8.3 Chromosome numbering. Cardoon (**a**) and globe artichoke (**b**) metaphase spreads DAPI counterstained. From Giorgi et al. (2016). Scale bar: 5 μm

8.3.3.1 FISH by rDNA Sequence Repeats

The analysis of chromosome distribution of ribosomal DNA in the genome of the *scolymus* and *altilis* varieties has been performed using the same pTa71 clone as FISH probe containing the 18S-5.8S-26S rRNA genes and the non-transcribed spacers of *T. aestivum* (Gerlach and Bedbrook 1979). The probe has been labelled by using the "nick translation" procedure and has produced eight hybridization sites (Fig. 8.4) localized at the very distal part of two medium and two small chromosomes, in both crops. The latter have been classified as acrocentric by Giorgi et al. (2016) and as metacentric by Falistocco (2016).

According to Giorgi et al. (2016), a detailed morphological analysis witnessed very small satellite bodies on some acrocentric chromosomes (Fig. 8.4), which were not consistently present, because they were probably damaged during slide preparation and FISH labelling. This ascertainment is in agreement with Roa and Guerra (2012), who observed a very high frequency of rDNA sites on acrocentric chromosomes in several genera.

8.3.3.2 FISHIS by SSR

Before FISH investigation with synthetic SSR oligonucleotides and in order to quickly select the ones best performing as probes, a fluorescent in situ hybridization in suspension (FISHIS) technique was carried on by Giorgi et al. (2016) on nuclei suspension of globe artichoke and cardoon. FISHIS is a fast, reliable method particularly effective for the selection of probes which produce strong, discrete and localized hybridization signals. This technique, developed by Giorgi et al. (2013b), enables a fast labelling of nuclei or chromosomes not spread and immobilized on a slide but freely floating in a liquid buffer. It is based on DNA alkaline denaturation and can be performed in less than 2 h at room temperature. A brief comparison among standard FISH and FISHIS is presented in Fig. 8.5. The following oligonucleotides $(AT)_{12}$, $(AAC)_5$, $(AGG)_5$, $(ATC)_5$, $(GAA)_7$, $(CAT)_5$, $(CA)_{10}$, $(GA)_{10}$, $(CAG)_5$, $(GACA)_4$ have been

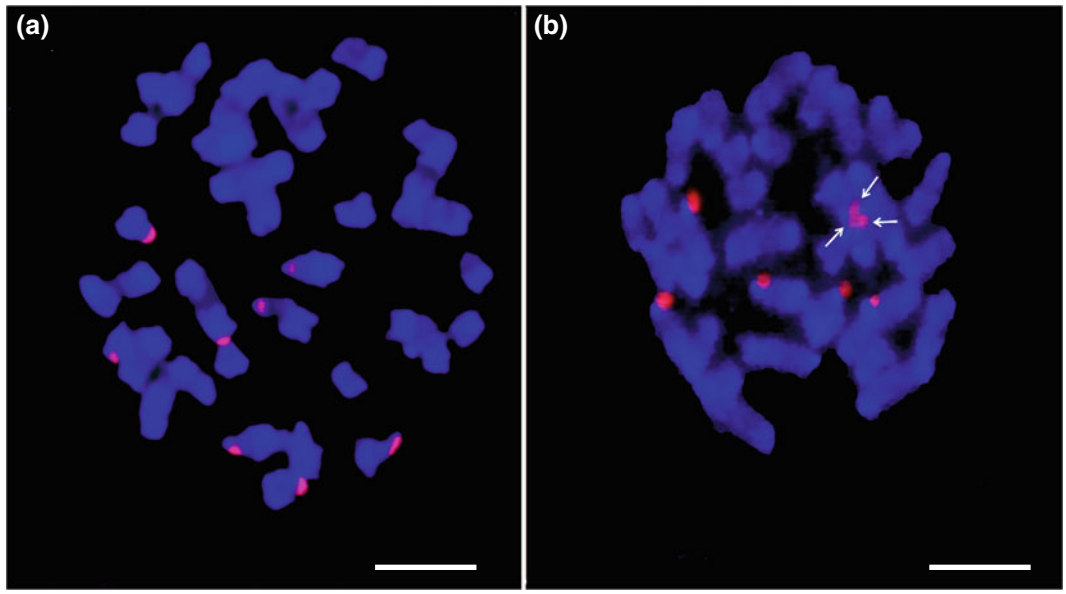


Fig. 8.4 From Giorgi et al. (2016). FISH on metaphase spreads of cultivated cardoon (**a**) and globe artichoke (**b**) using the rDNA probe pTa71-Cy3 (red fluorescence). In **b**, arrows point to the hybridization spots localized on the three small acrocentric chromosomes, placed head to head. Scale bar: 5 μm

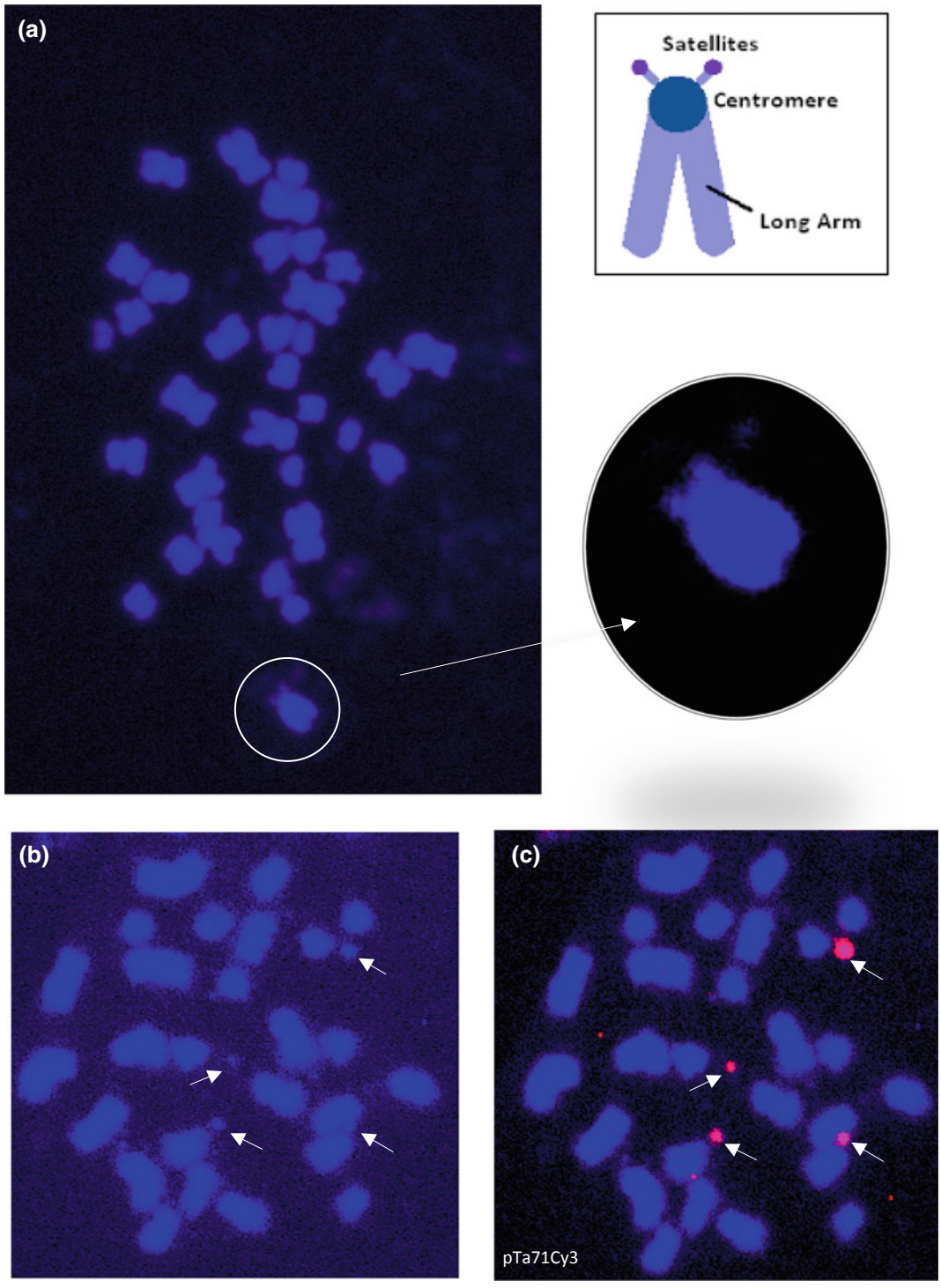


Fig. 8.5 Globe artichoke metaphases stained with DAPI. In **a**, an acrocentric chromosome showing very small satellite body is circled (enlarged to the right). In **b** and **c** (which are incomplete metaphases), arrows indicate the small satellite bodies detached from chromosome (brightness has been exasperated to make visible the satellite bodies) but still hybridizing with the pTa71 rDNA red probe (**c**)

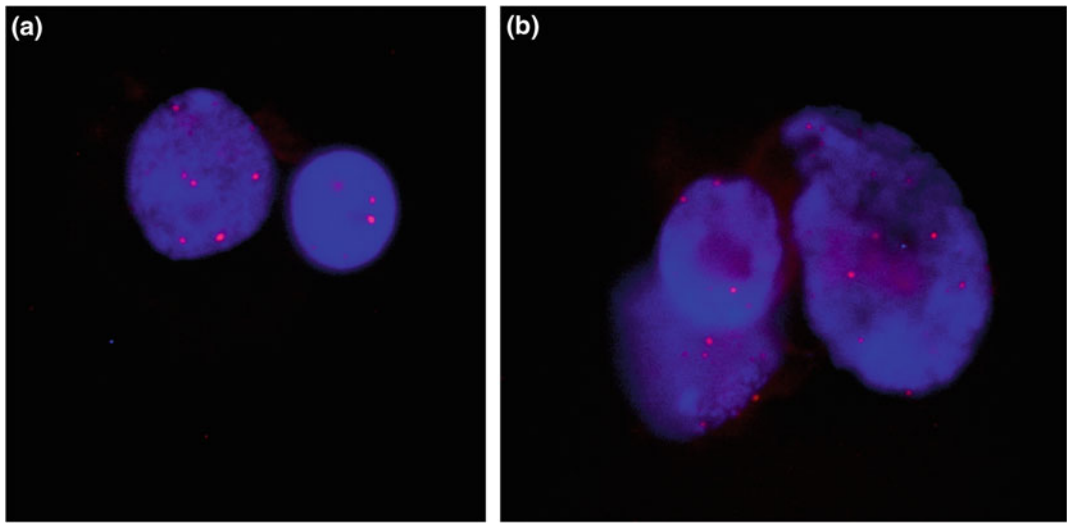


Fig. 8.6 FISHIS hybridization on cardoon (**a**) and globe artichoke (**b**) nuclei in suspension, counterstained with DAPI, with the telomeric probe (red) $(TTTAGGG)_5$

synthesized and labelled by fluorescein-5-isothiocyanate (FITC) or by Cyanine 3 (Cy3) by Eurofins MWG Operon (Ebersberg, Germany) and tested by FISHIS, together with the telomeric sequence $(TTTAGGG)_5$. In both crops, only the telomeric sequence $(TTTAGGG)_5$ (Fig. 8.6) and the $(GAA)_7$ SSR enabled the detection of discrete hybridization signals, while the two di-nucleotide $(CA)_{10}$ and $(GA)_{10}$ showed diffuse and faint hybridization (data not shown). Based on these results, four oligonucleotides $(TTTAGGG)_5$, $(GAA)_7$, $(GA)_{10}$ and $(CA)_{10}$ were selected out of 11 for further FISH characterization of the chromosome complement of *C. cardunculus* on microscope slide.

8.3.3.3 Fast FISH Labelling by Oligonucleotides

FISH labelling by synthetic oligonucleotides has drawn benefits from a FISH method based on DNA alkaline denaturation on slide, first described by Andras et al. (1999) and then modified by Giorgi et al. (2016) for a fast labelling of artichoke and cardoon chromosome complements. Metaphases were denatured in an alkaline 70% ethanol solution (pH 13 using 4 N NaOH) at room temperature for 5 min. After dehydration through an ice-cold ethanol series (70, 85 and 100%), 1.5–3 ng μL^{-1} of labelled oligonucleotide in 2X SSC (300 mM sodium chloride, 0.3 mM trisodium citrate) was applied on the slide (final volume 60 µL per slide) and hybridization was carried on at room temperature for 1 h. Afterwards, slides were washed in 4X SSC, 0.2% TWEEN 20 for 10 min, and DNA was counterstained with DAPI and mounted in antifade solution. A short description of the fast FISH procedure is presented in Fig. 8.7, together with the standard FISH and FISHIS.

The $(TTTAGGG)_5$ FISH hybridization enables the identification of chromosome ends and allows a more accurate measurement of chromosome length of both varieties. For this reason, it has been used also in some double FISH experiments (Fig. 8.8).

A variable and sometimes asymmetric distribution on sister chromatids of the two crops was observed using the oligonucleotide $(GAA)_7$, even if a reproducible hybridization signal in all the metaphases analysed was detected on two large, one medium and one small chromosome pairs (data not shown). In order to get more reliable results on all chromosomes, further analyses of $(GAA)_7$ distribution are required.

The two di-nucleotides $(CA)_{10}$ and $(GA)_{10}$ revealed a strong similarity in chromosome

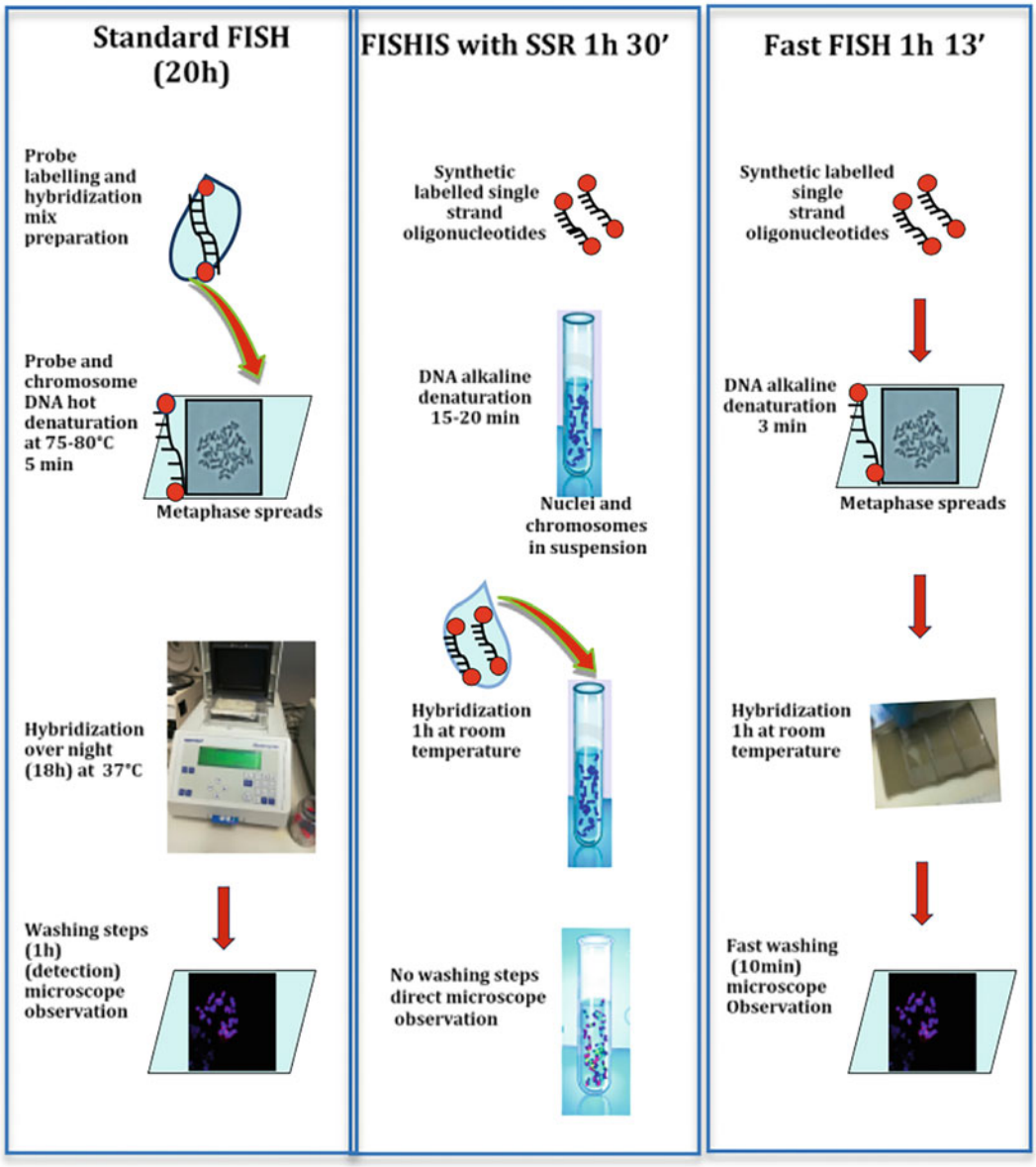


Fig. 8.7 Comparison of the main steps of the three FISH techniques: standard FISH, FISHIS and fast FISH, mentioned in the text

distribution and hybridization pattern among the two crops (Fig. 8.9). Wide hybridization bands have been detected at the telomeric and sub-telomeric regions of most chromosome arms, with some differences in band sizes among chromosomes. In chromosome pairs 3, 5 and 8, the extent of the hybridization signal was bigger on the long arm compared to the short one and to the other chromosome pairs. Chromosome pairs 10, 11, 14, 15 16 and 17 in cardoon and 10, 11, 15, 16 and 17 in globe artichoke appear to be acrocentric and can be discriminated from all the

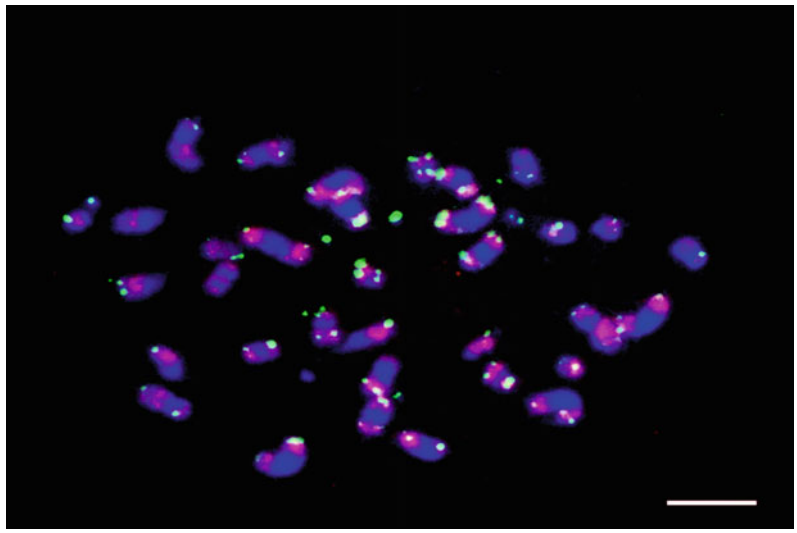


Fig. 8.8 Double target FISH hybridization with $(TTTAGGG)_5$ telomeric probe (green) and $(GA)_{10}$ oligonucleotides (red) to globe artichoke chromosomes. Scale bar: 5 μm

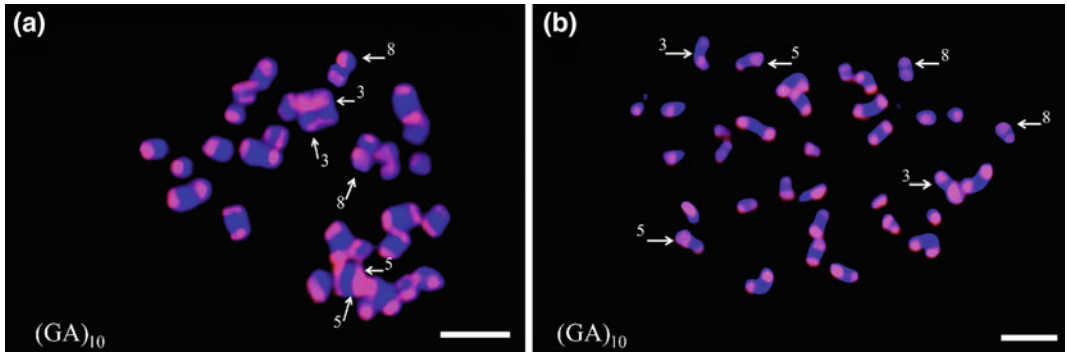


Fig. 8.9 FISH hybridization with $(GA)_{10}$ oligonucleotide on cardoon (A) and globe artichoke (B) spreads; chromosomes 3, 5 and 8 are indicated by arrows and can be discriminated by the widespread probe distribution on the long arms. Scale bar: 5 μm

other chromosomes by the distribution of $(GA)_{10}$ and $(CA)_{10}$ along only one side of the chromosome arms.

Giorgi et al. (2016) stated that the FISH-based molecular characterization of the chromosome complement of cardoon and globe artichoke enabled the identification of some chromosome pairs and the execution of a more detailed karyomorphological analysis, which were performed by measuring the total length (tl), the arm length and the arm ratio (AR) long/short arm length) of all 34 individual chromosomes (Table 8.2 from Giorgi et al. 2016).

According to the Levan et al. (1964) system, Falistocco (2016) classified all chromosomes of both varieties as metacentric (M) with an arm ratio (AR) from 1.0 to 1.5, but without reporting single chromosome size. Otherwise, Giorgi et al. (2016) described the following karyotypic formula: 2n = 14M + 10SM + 10A and 2n = 12M + 10 SM + 12A for, respectively, *C. cardunculus* L. var. *scolymus* (L.) Fiori and *C. cardunculus*

Table 8.2 Morphometric analysis of *C. cardunculus* chromosomes

C.p.	T.l. (μm)		*p* (μm)		*q* (μm)		AR		Class	
	A	B	A	B	A	B	A	B	A	B
1	3.16	3.19	1.23	1.27	1.93	1.92	1.57	1.51	SM	SM
2	2.97	3.16	1.07	1.24	1.78	1.93	1.78	1.55	SM	SM
3	2.58	2.81	1.05	1.21	1.55	1.60	1.46	1.32	M	M
4	2.35	2.65	0.97	1.16	1.38	1.49	1.42	1.28	M	M
5	2.35	2.45	0.78	0.95	1.57	1.50	2.01	1.57	SM	SM
6	2.12	2.27	0.85	0.95	1.27	1.32	1.49	1.39	M	M
7	2.10	2.25	0.87	0.93	1.23	1.32	1.41	1.42	M	M
8	2.03	2.00	0.79	0.78	1.24	1.22	1.56	1.56	SM	SM
9	2.00	1.98	0.92	0.92	1.08	1.06	1.17	1.15	M	M
10	2.00	1.98	–	–	2.00	1.98	>3	>3	A	A
11	1.95	1.95	–	–	1.95	1.95	>3	>3	A	A
12	1.79	1.95	0.81	0.71	0.98	1.24	1.20	1.74	M	SM
13	1.68	1.70	0.62	0.79	1.06	0.91	1.70	1.15	SM	M
14	1.40	1.41	–	0.65	1.40	0.76	>3	1.17	A	M
15	1.30	1.38	–	–	1.30	1.38	>3	>3	A	A
16	1.25	1.27	–	–	1.25	1.27	>3	>3	A	A
17	1.22	1.22	–	–	1.22	1.22	>3	>3	A	A

A cardoon; *B* globe artichoke;[†]*C. p.* chromosome pairs; *T.l* total length; *p* average length of short arm; *q* average length of long arm; *AR* arm ratio (p/q); *Class* classification: *M* metacentric; *SM* submetacentric; *A* acrocentric

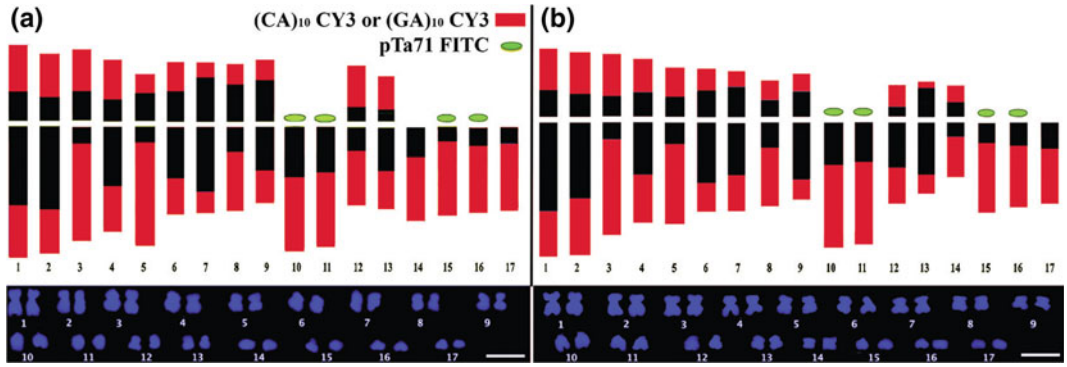


Fig. 8.10 Ideogram of cardoon (**a**) and globe artichoke (**b**) complements, showing chromosome morphology (in black) and the $(CA)_{10}$, $(GA)_{10}$ di-nucleotides (red) and pTa71 sequence (green) distribution. The two di-nucleotides localize at similar chromosome regions. Scale bar: 5 μm

L. var. *altilis* DC (see Table 8.2) and according to Guerra's nomenclature (Guerra 1986) as follows: AR = 1.00–1.49 metacentric (M), AR = 1.50–2.99 submetacentric (SM), AR $\geq$ 3 acrocentric (A). Figure 8.10 reports the ideogram of the crops. The different chromosome morphological classification in metacentric or in submetacentric described by the authors could be ascribed to

the different classification systems used as reference, Levan et al. (1964) for Falistocco while Guerra (1986) for Giorgi. Anyway, this observation does not apply to chromosome classified as acrocentric by Giorgi which shows an AR >3 since the small size of *Cynara* chromosomes makes it difficult to clearly identify the position of centromere and secondary constrictions, and it could explain the different results independently obtained by the two laboratories.

8.4 Outlook

The close phylogenetic relationship between cardoon and globe artichoke stated by several authors (Fiori and Beguinot 1904; Wiklund 1992; Gatto et al. 2013) and described in detail in Chap. 1 (this book) has been confirmed by the cytological and molecular cytogenetic studies. The new findings have revealed a very similar karyotype for the two varieties, and an analogous distribution of rRNA genes and of the SSR used for the analysis. Only a few differences have been detected, as those related to DNA content; namely, the variety *altilis* contains a slightly lower amount of DNA than *scolymus*, and the chromosome morphology of pairs 14 is acrocentric in the cultivated cardoon and metacentric in globe artichoke. Considering the proposed common origin of the two cultivated varieties from wild cardoon (Rottenberg and Zohary 1996), such structural chromosome differences could have occurred during the process of domestication (probably two distinct events, see Chap. 1, this book) with either the deletion of the small arm of cardoon chromosome or a translocation or insertion on chromosome 14 in globe artichoke.

For a more comprehensive view of the evolutionary pathway of *C. cardunculus*, the cytogenetic analysis should be extended to the proposed wild progenitor *C. cardunculus* var. *sylvestris* (Falistocco 2016). In addition, the use of centromeric probes and the selection of chromosome-specific cytogenetic markers would help in further characterizing all the *C. cardunculus* varieties.

References

Ancora G, Belli Donini ML, Cuozzo L (1981) Globe artichoke plants obtained from shoot apices through rapid in vitro micropropagation. Sci Hortic 14:207–213

Andras SC, Hartman TPV, Marshall JA, Marchant R, Power JB, Cocking EC, Davey MR (1999) A drop-spreading technique to produce cytoplasm-free mitotic preparations from plants with small chromosomes. Chromosome Res 7(8):641–647

Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. Plant Mol Biol Rep 9(3):208–218

Battaglia E (1994) Nucleosome and nucleotype: a terminological criticism. Caryologia 47:193–197

Bennett MD, Leitch IJ (2005) Nuclear DNA amounts in angiosperms: progress, problems and prospects. Ann Bot 95(1):45–90

Bennett MD, Leitch IJ (2011) Nuclear DNA amounts in angiosperms: targets, trends and tomorrow. Ann Bot 107(3):467–590

Bennett MD, Smith JB (1976) Nuclear DNA amounts in angiosperms. Philos Trans R Soc Lond B Biol Sci 274(933):227–274

Cuadrado Á, Jouve N (2010) Chromosomal detection of simple sequence repeats (SSRs) using nondenaturing FISH (ND-FISH). Chromosoma 119(5):495–503

Cuadrado Á, Cardoso M, Jouve N (2008) Increasing the physical markers of wheat chromosomes using SSRs as FISH probes. Genome 51:809–815

Doležel J, Bartoš J (2005) Plant DNA flow cytometry and estimation of nuclear genome size. Ann Bot 95(1): 99–110

Doležel J, Greilhuber J (2010) Nuclear genome size: are we getting closer? Cytom Part A J Int Soc Anal Cytol 77(7):635–642

Doležel J, Binarova P, Lucretti S (1989) Analysis of nuclear-DNA content in plant-cells by flow-cytometry. Biol Plant 31(2):113–120

Doležel J, Sgorbati S, Lucretti S (1992) Comparison of 3 DNA fluorochromes for flow cytometric estimation of nuclear-DNA content in plants. Physiol Plant 85(4): 625–631

Doležel J, Greilhuber J, Lucretti S, Meister A, Lysak MA, Nardi L, Obermaier R (1998) Plant genome size estimation by flow cytometry: inter-laboratory comparison. Ann Botany 82(Suppl. A):17–26

Falistocco E (2016) Cytogenetic characterization of cultivated globe artichoke (*Cynara cardunculus* var. *scolymus*) and cardoon (*C. altilis*). Caryologia 69:1–4

Feulgen R, Rossenbeck H (1924) Mikroskopisch-chemischer Nachweis einer Nucleinsaure vom Typus der Thymonucleinsaure und die darauf beruhende elektive Farbung vom Zellkernen in mikroskopischen Praparaten. Zts Phys Chem 135:203–248

Figueroa DM, Bass HW (2010) A historical and modern perspective on plant cytogenetics. Brief Funct Genomics 9(2):95–102

Fiori A, Beguinot A (1904) Flora analitica d'Italia (Vol. III). Tipografia del Seminario, Padova, p 527

Fregonezi JN, Torezan JMD, Vanzela ALL (2004) A karyotypic study of three southern Brazilian *Asteraceae* species using fluorescence in situ hybridization with a 45S rDNA probe and C-CMA_3 banding. Genet Mol Biol 27(2):223–227

Fu S, Tang Z (2015) Oligonucleotide probes for ND-FISH analysis to identify rye and wheat chromosomes. Sci Rep 5:10552

Funk VA, Susanna A, Stuessy TF, Robinson H (2009) Classification of Compositae. In: Funk VA, Susanna A, Stuessy T, Bayer R (eds) Systematics, evolution and biogeography of the Compositae. IAPT, Vienna, pp 171–189

Galbraith DW, Harkins KR, Maddox JM, Ayres NM, Sharma DP, Firoozabady E (1983) Rapid flow cytometric analysis of the cell cycle in intact plant tissues. Science 3:220(4601):1049–1051

Gamborg OL (1975) Callus and cell culture. In: Gamborg OL, Wettwr LR (eds) Plant tissue culture methods. National Research Council Canada, Prairie Regional Laboratory, Saskaton Sask, pp 1–10

Garcia S, Panero JL, Siroky J, Kovarik A (2010) Repeated reunions and splits feature the highly dynamic evolution of 5S and 35S ribosomal RNA genes (rDNA) in the *Asteraceae* family. BMC Plant Biol 10:176

Garnatje T, Canela MÁ, Garcia S, Hidalgo O, Pellicer J, Sánchez I, Siljak JS, Vitales JD, Vallès J (2011) GSAD: A genome size in the *Asteraceae* database. Cytom Part A 79A(6):401–404

Gatto A, De Paola D, Bagnoli F, Vendramin GG, Sonnante G (2013) Population structure of *Cynara cardunculus* complex and the origin of the conspecific crops artichoke and cardoon. Ann Bot 112:855–865

Gerlach WL, Bedbrook JR (1979) Cloning and characterization of ribosomal RNA genes from wheat and barley. Nucleic Acids Res 7(7):1869–1885

Giorgi D, Pandozy G, Farina A, Grosso V, Lucretti S, Crinò P, Saccardo F (2013a) Karyotype of globe artichoke (*Cynara cardunculus* var. *scolymus*): preliminary studies to define its chromosome morphology. Acta Hort 983:133–138

Giorgi D, Farina A, Grosso V, Gennaro A, Ceoloni C, Lucretti S (2013b) FISHIS: Fluorescence In Situ Hybridization in Suspension and Chromosome Flow Sorting Made Easy. PLoS ONE 8(2):e57994

Giorgi D, Pandozy G, Farina A, Grosso V, Lucretti S, Gennaro A, Crinò P, Saccardo F (2016) First detailed karyo-morphological analysis and molecular cytological study of leafy cardoon and globe artichoke, two multi-use Asteraceae crops. Comp Cytogenet 10 (3):447–463

Greilhuber J (1988) "Self-tanning"—a new and important source of stoichiometric error in cytophotometric determination of nuclear DNA content in plants. Plant Syst Evol 158(2):87–96

Greilhuber J (2008) Cytochemistry and C-values: the less-well-known world of nuclear DNA amounts. Ann Bot 101(6):791–804

Guerra M (1986) Reviewing chromosome nomenclature of Levan et al. Rev Brasil Genet 4:741–743

Hanson L, Brown RL, Boyd A, Johnson MAT, Bennett MD (2003) First nuclear DNA C-values for 28 angiosperm genera. Ann Bot 91:1–8

Johnston JS, Bennett MD, Rayburn AL, Galbraith DW, Price HJ (1999) Reference standards for determination of DNA content of plant nuclei. Am J Bot 86(5):609–613

Khaldi S, Hidalgo O, Garnatje T, El Gazzah M (2014) Karyological and genome size insights into cardoon (*Cynara cardunculus* L., *Asteraceae*) in Tunisia. Caryologia 67(1):57–62

Kiernan JA (1999) Histological and histochemical methods: theory and practice, 3rd edn. Oxford, Butterworth-Heinemann, Boston

Levan A, Fredga K, Sandberg A (1964) A nomenclature for centromeric position of chromosomes. Hereditas 52(2):201–220

Loureiro J, Rodriguez E, Doležel J, Santos C (2006) Comparison of four nuclear isolation buffers for plant DNA flow cytometry. Ann Bot 98(3):679–689

Marie D, Brown SC (1993) A cytometric exercise in plant DNA histograms, with 2C values for 70 species. Biol Cell 78(2):41–51

Ochatt SJ (2008) Flow cytometry in plant breeding. Cytom Part A Jo Int Soc Anal Cytol 73(7):581–598

Rasch EM, Barr HJ, Rasch RW (1971) The DNA content of sperm of Drosophila melanogaster. Chromosoma 33(1):1–18

Roa F, Guerra M (2012) Distribution of 45S rDNA sites in chromosomes of plants: structural and evolutionary implications. BMC Evol Biol 12:225

Rottenberg A, Zohary D (1996) The wild ancestry of the cultivated artichoke. Genet Resour Crop Evol 43:53–58

Schdmidt G, Thannhauser SJ (1945) A method for the determination of desoxyribonucleic acid, ribonucleic acid, and phosphoproteins in animal tissues. J Biol Chem 161:83

Schwarzacher T, Heslop-Harrison P (2000) Practical in situ hybridization. BIOS Scientific Publishers

Singh RJ (2018) Practical manual on plant cytogenetics. CRC Press

Sliwinska E (2018) Flow cytometry—a modern method for exploring genome size and nuclear DNA synthesis in horticultural and medicinal plant species. Folia Horticulturae 30(1):103

Swift H (1950) The constancy of deoxyribose nucleic acid in plant nuclei. Proc Natl Acad Sci 36:643–654

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 408:796–815

Tiersch TR, Chandler RW, Wachtel SS, Elias S (1989) Reference standards for flow cytometry and application in comparative studies of nuclear DNA content. Cytometry 10:706–710

Vallès J, Canela MÁ, Garcia S, Hidalgo O, Pellicer J, Sánchez-Jiménez I, Siljak-Yakovlev S, Vitales D, Garnatje T (2013) Genome size variation and evolution in the family Asteraceae. Caryologia 66(3):221–235

Van'T Hof J (1965) Relationships between mitotic cycle duration, S period duration and the average rate of DNA synthesis in the root meristem cells of several plants. Exp Cell Res 39(1):48–58

Wang J, Liu J, Kang M (2015) Quantitative testing of the methodology for genome size estimation in plants using flow cytometry: a case study of the Primulina genus. Front Plant Sci 6:354

Weiss-Schneeweiss H, Tremetsberger K, Schneeweiss GM, Parker JS, Stuessy TF (2008) Karyotype diversification and evolution in diploid and polyploid south American *Hypochaeris* (*Asteraceae*) inferred from rDNA localization and genetic fingerprint data. Ann Bot 101(7):909–918

Wiklund A (1992) The genus *Cynara L.* (Asteraceae-Cardueae). Bot J Linn Soc 109:75–123

Genomics

9

Davide Scaglione, Alberto Acquadro, Ezio Portis and Sergio Lanteri

Abstract

A first assembly of globe artichoke genome was released in 2016 and the project required the adoption of a set of strategies to overcome the lack of a homozygous reference genotype. Moreover, the inability to exploit standard genotyping methods, normally adopted in inbred species where homozygous segregating progenies can be more easily obtained, posed some challenges. A low-pass genotyping-by-sequencing strategy for heterozygous individuals was implemented for the first time and required the development of dedicated tools. The method allowed to leverage very low coverage to impute genotypic states of F_1 segregating individuals and use this information to arrange scaffolds in pseudomolecules. The availability of a chromosome-scale reference sequence allowed to depict, for the first time, the genomic landscape of globe artichoke. The work shed lights on its genes, their structure and function, orthology and synteny with other species, repeats and the timing of their expansion, age of speciation and whole-genome duplications. All the features revealed in this study expanded the knowledge base for comparative genomics across all plants. More importantly, the genome sequence of globe artichoke will facilitate transferability of genomic and genetic findings in the *Compositae* family, hence accelerating the discovery of the molecular mechanism at the base of economically important traits and their exploitation for selection, genetic engineering or utilization in biotechnological processes, in the light of the richness of nutraceutical properties of this species and its overall economic importance for the Mediterranean horticulture.

9.1 The Globe Artichoke Sequencing Initiative

The world production of globe artichoke is led by Italy. The genome of this commercial important species (*Cynara cardunculus* var. *scolymus* L.) has been recently sequenced by an international Consortium which included the University of Torino (DISAFA, Italy), the University of California (The Genome Centre, Davis, CA, USA) and the Università di Catania (Di3A, Italy). The project started in 2011 in the framework of the Compositae Genome Project (CGP) and later on was joined by the University of British Columbia

D. Scaglione
IGA Technology Services, Via J. Linussio, 51 ZIU, Udine, Italy

A. Acquadro (✉) · E. Portis · S. Lanteri
DISAFA Plant Genetics and Breeding, University of Torino, Largo P. Braccini 2, 10095 Grugliasco, Torino, Italy
e-mail: alberto.acquadro@unito.it

E. Portis et al. (eds.), *The Globe Artichoke Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-030-20012-1_9

(Canada) and CREA (Genomics Research Centre, Italy). The sequence has been recently published in Scientific Reports, and data is available at www.artichokegenome.unito.it and on NCBI.

The globe artichoke (*C. cardunculus* var. *scolymus* L.), together with cultivated cardoon (*C. cardunculus* var. *altilis*), was domesticated in the Mediterranean region from the wild cardoon (*C. cardunculus* var. *sylvestris*). These species are members of the *Compositae* (aka Asteraceae), which is the largest botanical family among dicots, thus one of the most successful botanical clade, including species of relevant economic scale such as sunflower, lettuce and chicory. The family most likely originated in South America (30–100 Mya, Raven and Axelrod 1974). Rapid diversification of the Compositae in the mid-Eocene (40 Mya) is associated with a polyploidization event near the base of the family (Barker et al. 2008), and it has been hypothesized that the base chromosome number for the family is nine (Kesseli and Michelmore 1997). Italy, which harbours the richest primary gene pool of cultivated globe artichoke, is the top producing country, but globe artichoke cultivation has also emerged in other Mediterranean countries and more recently to the Americas and China. The primary product of globe artichoke is the immature inflorescence (head or *capitulum*), whose inner bracts and fleshy receptacle are consumed as a fresh, preserved or frozen delicacy. Noteworthy, similar to other members of the *C. cardunculus* complex, it is also a rich source of bio-active compounds, such as antioxidant phenolics, sesquiterpene lactones and inulin. The rich biomass of cardoon ecotype can be exploited for the production of lignocellulosic biomass while seed oil is valuable for both edible and biofuel purposes.

C. cardunculus is a diploid species ($2n = 2x = 34$) with a genome size estimated by flow cytometry to be 1.07 Gb (Marie and Brown 1993). The few genomics data currently available for this crop consists of a diversity survey through restriction site-associated DNA sequencing (RADseq, Scaglione et al. 2012a) and RNAseq data with the release of reference transcriptome of 38 K unigenes (Scaglione et al. 2012b). In the last decade, several genetic maps of *C. cardunculus* have been developed relying on a two-way pseudo-test-cross approach and were exploited for the study of key breeding traits via QTL analysis with the scope to develop molecular tools for marker-assisted selection (Lanteri et al. 2006; Portis et al. 2009, 2012, 2014). Also, the sequence of the chloroplast genome has recently been reported (Curci et al. 2015).

Some *Compositae* crop genomes have been already assembly at a good quality level: the 2.7-Gbp lettuce genome (http://lgr.genomecenter.ucdavis.edu; GenomeProject ID: PRJNA68025) and the 3.5-Gbp sunflower genome (http://www.sunflowergenome.org; GenomeProject ID: PRJNA64989) have been already obtained. The 335-Mbp genome of horseweed (*Conyza canadensis*) has been also released, but the sequence scaffolds have not been placed into linkage groups or integrated into pseudomolecules. The comparative genomics with other asterid species which have been recently sequenced (potato, Xu et al. 2011; tomato, Sato et al. 2012; pepper, Kim et al. 2014; Qin et al. 2014) and eggplant (Hirakawa et al. 2014; Rotino et al. 2014) will provide unprecedented insights into the molecular mechanisms at the basis of the successful evolution of *Compositae*. The availability of the globe artichoke genome will contribute to this research and provide understanding of ancestral genome duplications, rearrangements and burst of repetitive elements which shaped its structure.

On a more applicative outcome, the understanding of the genomic features of globe artichoke will serve the foundation for the study of complex traits (Portis et al. 2018), where the exploitation of comparative genomics and the study of synteny within the family will facilitate precision and effectiveness in the transfer of genetic findings across species while fostering community research and collaboration. Short-read sequencing technologies are now offering a democratized opportunity for non-model species to gain access to the genomic era. Still, chromosome-scale assembly of genomes remains an endeavour that is constrained

by genome complexity, technologies, costs and analytical tools (Xie et al. 2010; Mascher et al. 2013). Indeed, to facilitate this task a globe artichoke clone with low residual heterozygosity was chosen. To assign scaffolds to genetic bins, ordered along with chromosomal linkage groups, a novel pipeline named SOILoCo—Scaffold Ordering by Imputation with Low Coverage—was developed. In particular, it provides genotyping-by-sequencing capability in a low-coverage (<1×) whole-genome shotgun, even if segregation of heterozygous loci is the only source of informative markers and none of the segregating haplotypes is currently represented in the reference sequence.

The tools developed for the genome reconstruction of globe artichoke are not exhausted with the genome completion itself, but they will be valued by adoption in further studies of genetic traits and, possibly, to complement other genome assembly works for this and other species towards a pan-genomic era of *Compositae*.

9.2 Genome Sequencing and Assembly

The genome assembly for a highly heterozygous, outbreeding species like globe artichoke is known to be a more challenging exercise than for a comparable genome in terms of size and complexity but with a nearly homozygous genotype. The latter would ease the reconstruction of single reference sequence without redundancies caused by the presence of two alleles, which may harbour structural variation other than single-nucleotide polymorphisms. A globe artichoke clone designated "2C" was obtained by three cycles of self-cross (S_3) and analysis of 115 SSR markers assessed its residual heterozygosity to approximately 10% (Acquadro et al. 2009; Scaglione et al. 2009). DNA was extracted from young leaves and one paired-end and two mate-pair libraries were generated. The 100 bp paired reads sequencing, with an average insert size of 170 bp, produced 67 Gbp of raw data, on average covering 62 times the estimated genome; 96% of pairs successfully generated 3'-overlaps (with mean of 26 bp), thus satisfying the ALLPATH-LG algorithm requirements. The two mate-pair libraries, with effective insert sizes of 2.5 Kbp and 5.5 Kb, produced 56 Gbp and 7 Gbp of raw data, respectively (corresponding to 52× and 6.5× genome coverages). After cleaning the reads from ligation adapters and low-quality bases, a total of 133.7 Gbp was available for the assembly exercise. The limited residual heterozygosity of the 2C breeding line was confirmed by the analysis of k-mer spectrum (Fig. 9.1) as compared to those of the parental lines (globe artichoke and cultivated cardoon) of the mapping population used in this study (see below). Based on k-mer statistics and sequencing depth (Li et al. 2010), the globe artichoke genome size was estimated to be 1084 Mb, being in close agreement with the previous estimates by flow cytofluorometry (Marie and Brown 1993). Sequencing errors were corrected by a function implemented in ALLPATHS-LG. After error correction, 87.5% of paired-end reads were retained to be used in the assembly, while the remainder were discarded as containing sequence k-mer neither confirmed nor corrected by those found in the raw data pool. The largest 50% set of sequences from the de novo assembly result (N_{50}) was included in 10,596 contigs of 17.5 Kbp or larger, with 32% G + C content, while all contigs, once assembled into 13,588 scaffolds ($\geq$ 1000 bp), returned a N_{50} of 1408 and a L_{50} equal to 125.9 Kb (Table 9.1). The final validated assembly spanned approximately 724.7 Mb, with the longest scaffold of 1.5 Mb (Table 9.1). Chloroplast sequences that were found to populate the final assembly were removed via a BLAST analysis.

The completeness of the assembly was assessed in two ways. First, the available unigene set (Scaglione et al. 2012b) of the *C. cardunculus* gene space was aligned to the genome sequence and 37,018 of the 38,726 unigenes (95.6%) found a positive hit, while unaligned unigenes were analysed by blastx against Viridiplantae proteins (Refseq) and the 69% of these had hits. This data converged to a hypothetical total of 27,564 gene models, which would estimate the proportion of non-assembled genes up to 2.5%. Second, a survey based on 248 conserved eukaryotic genes (CEGs, Parra et al. 2007)

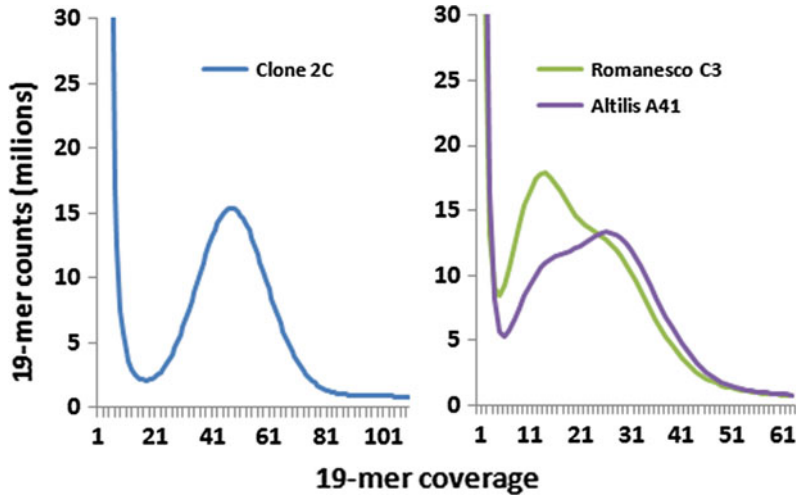


Fig. 9.1 K-mer spectra analysis using 19-mer counts. Clone "2C" (left) showed significantly lower heterozygosity compared to that of cultivated "Romanesco C3" artichoke and "Altilis A41" cardoon genotypes (right). Among cultivated genotypes, the difference between globe artichoke and cardoon was also evident and in agreement with the previous reports (Portis et al. 2005) (from Scaglione et al. (2016), used under CC BY 4.0)

Table 9.1 Genome assembly statistics (from Scaglione et al. (2016), used under CC BY 4.0)

Features	Contigs	Scaffolds
Number of sequences	79,681	13,588
Sequences/Mb	121.6	18.9
Total length (Mb)	654.6	724.7
L_{50} (Kb)	17.5	125.8
N_{50}	10,596	1411
L_{90} (Kb)	3.4	28.7
N_{90}	41,711	6109
Number of sequences > 10 Kb	20,561	9187
Median size of gaps in scaffolds	–	660 ± 86 bp

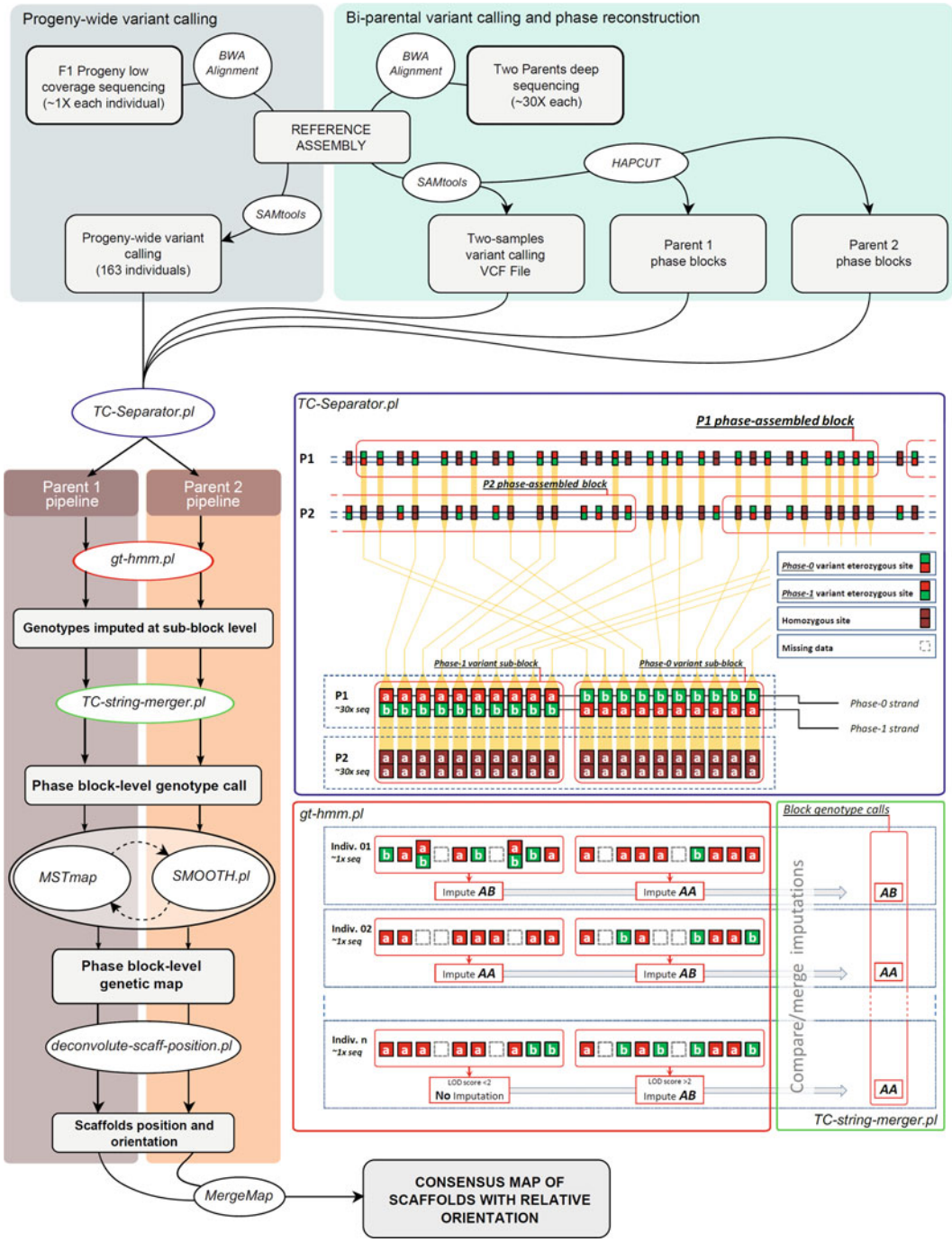
showed that their complete sequence was present for more than 85% of them, while admission of partial matches increased the percentage of mapped CEGs to 96%. In particular, the segmentation of CEGs based on their level of sequence conservation (across CEGs' species) highlighted that the most divergent set "Group1" report an average level of similarity of 80.3%, which increased to 93.85% of "Group4", the most conserved. However, the level of similarity with only partial alignments was similar among the groups (95.45–98.46%). This suggested that partial alignments were possibly due to sequence divergence, and thus, at least 95% of the gene space was represented in the draft genome assembly.

9.3 Production of the Pseudomolecules

The ordering and orientation of scaffolds in chromosome-representative pseudomolecules were achieved via linkage analysis. For this purpose, a F_1 population of 163 individuals was obtained by crossing globe artichoke "C3"

genotype as the female parent and cultivated cardoon "Alt41" genotype as the male parent. The goal was to leverage segregation data obtained by a low-pass sequencing of each progeny line to generate a consensus map of both maternal and paternal segregating alleles. Low-pass sequencing has the benefit to provide information across the whole genome, but it requires to accommodate sparsity of data and lack of coverage to call heterozygous genotypes on each single base for all individuals. To achieve this, we developed and implemented a novel analysis concept for low-pass genotyping-by-sequencing sequencing data combined with conventional deep sequencing of the heterozygous parents, whose chromosomal phases were unknown. The method relies on the de novo identification of parental allelic phases (haplotype blocks), with no reliance on haplotypes being shared with the reference sequence assembly. The pipeline, whose workflow is depicted in Fig. 9.2, was named SOILoCo (Scaffold Ordering by Imputation with Low Coverage). Its recipe requires: (i) a draft assembly to a good state of contiguity (scaffold L50 > 50 kb); (ii) paired-end whole-genome sequencing data of each parent to a coverage of at least 20× (used as input data for HAPCUT); and (iii) whole-genome sequencing of their segregating progeny to a coverage as low as 0.5× each. In our experiment, short reads were generated on HiSeq2000 platform. Those of parental lines were mapped to the draft assembly, populating the reference set of heterozygous SNPs for each parent. Then, haplotype blocks, as a contiguous set of heterozygous SNPs defined above, were deconvoluted using the HAPCUT algorithm (Bansal and Bafna 2008). For each parental genome, pairs of single-nucleotide variants, residing either on the same read or in sister reads, were used to build a graph of phased single-nucleotide variants, interrupted only when the lack of heterozygous sites or repetitive regions made unable to maintain a robust allelic path. SNP polymorphisms for each progeny individual were then collected with a shallow coverage; most of the sites were called based on single-read coverage, being representative of only one of the two alleles, randomly. The SOILoCo pipeline, made of a set of custom Perl scripts, was then applied in consecutive steps to: (i) extract SNP sites expected to segregate in a test-cross fashion; (ii) infer the two segregating haplotype block for each parent (phase 0 and phase 1 in Fig. 9.2; (iii) analyse base calls for each individual and given the haplotype block/segregating phase; (iv) invoke an Hidden Markov Model (HMM) engine to infer a block-based haplotype call for each individual (Fig. 9.3); and finally, (v) extract a confident set of block calls filtered by marginal probability estimates, obtained during the decoding function of the Hidden Markov Model, and the rate of successfully genotyped samples.

A total of 520 M reads (52.0 Gb, 48× coverage) and 467 M reads (46.8 Gb, 43× coverage) were obtained by the sequencing of the "C3" and the "Alt41" genotypes, respectively. Once completed the quality cleaning of reads and their mapping on the draft assembly, 5.8 M and 3.1 M heterozygous SNP sites were collected in the two parental genotypes. The de novo reconstruction of haplotype blocks in the "C3" genotype of globe artichoke yielded 192,938 segments, corresponding to 364 Mbp (L_{50}: 4772 bp, N_{50}: 19,921) with an average of 1 SNP per 59 bp within such blocks. In cultivated cardoon "Alt41", the resulting haplotype blocks were limited to 122,570 and totalled only 187 Mbp but with similar contiguity as compared to "C3" (L_{50}: 4104 bp, N_{50}: 12,001). Also, the rate of polymorphic sites within blocks was equal to that of "C3" (1 SNP every 59 bp). This statistic indicates that SNP-dense regions with similar polymorphism content contributed in extending haplotype blocks in both parents, but similar contiguities and shorter overall size of "Alt41" blocks confirmed the graphical observation that average lower heterozygosity of "Alt41" is due to long homozygous segments in the genome (Fig. 9.4) and not to an average genome-wide reduction of allelic diversity. The catalogue of SNP sites which were expected to desegregate according to the cross model (TC-SNPs) was made of 2,128,486 sites from globe artichoke "C3" alleles, which were


Progeny-wide variant calling
F1 Progeny low coverage sequencing (~1X each individual)
BWA Alignment
REFERENCE ASSEMBLY
SAMtools
Progeny-wide variant calling (163 individuals)
Bi-parental variant calling and phase reconstruction
BWA Alignment
Two Parents deep sequencing (~30X each)
SAMtools
HAPCUT
Two-samples variant calling VCF File
Parent 1 phase blocks
Parent 2 phase blocks
TC-Separator.pl
Parent 1 pipeline
Parent 2 pipeline
gt-hmm.pl
Genotypes imputed at sub-block level
TC-string-merger.pl
Phase block-level genotype call
MSTmap
SMOOTH.pl
Phase block-level genetic map
deconvolute-scaff-position.pl
Scaffolds position and orientation
MergeMap
CONSENSUS MAP OF SCAFFOLDS WITH RELATIVE ORIENTATION
TC-Separator.pl
P1 phase-assembled block
P1
P2 phase-assembled block
P2
Phase-0 variant eterozygous site
Phase-1 variant eterozygous site
Homozygous site
Missing data
Phase-1 variant sub-block
Phase-0 variant sub-block
P1 ~30x seq
P2 ~30x seq
Phase-0 strand
Phase-1 strand
gt-hmm.pl
Indiv. 01 ~1x seq
Impute AB
Impute AA
Indiv. 02 ~1x seq
Impute AA
Impute AB
Indiv. n ~1x seq
LOD score <2
No Imputation
LOD score >2
Impute AB
Block genotype calls
Compare/merge imputations
AB
AA
AA
TC-string-merger.pl

◀ **Fig. 9.2** Workflow schema for the SOILoCo pipeline used to anchor the *C. cardunculus* scaffolds in chromosomal pseudomolecules. Alignments of parental reads to the draft scaffolds were used to (i) identify potential heterozygous test-cross sites and (ii) to compute haplotype phases in both parents (P1 and P2). A multi-sample VCF file of all the progeny was then processed to identify informative heterozygous sites based on parental SNPs and the phase of haplotype blocks (TC-separator.pl, blue box). This assigned the sites according to which phase (i.e. homologous chromosomes) they are expected to segregate in. Subsequently, an HMM-based algorithm was used to impute the most likely genotypes of each haplotype block segregating in the progeny (gt-hmm.pl, red box). A LOD score was also calculated to permit filtering of ambiguous imputations. Genotype imputation from the two alternative segregating phases was then summarized; when there was a discordant call between phases, a majority rule was applied and the highest LOD score for each segregating haplotype used to impute the most likely genotype (TC-string-merger.pl, green box). After grouping markers, linkage maps were generated for each parent using reiterative ordering with the MSTmap software (http://alumni.cs.ucr.edu/~yonghui/mstmap.html) and error correction using a Perl implementation of SMOOTH algorithms (van Os et al. 2005). Maps were finally merged to generate a consensus map and to maximize the resolution of the order and orientation of scaffolds in chromosomal pseudomolecules (from Scaglione et al. (2016), used under CC BY 4.0)

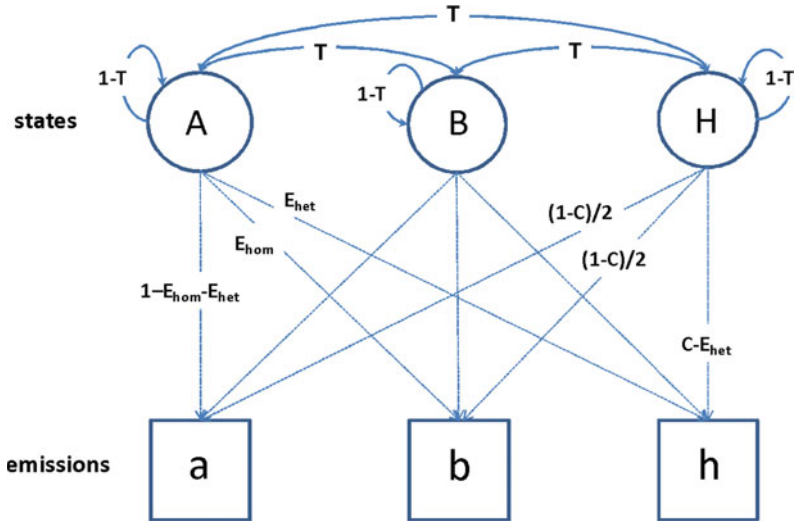


Fig. 9.3 Graphical representation of the HMM model developed for the decoding of most probable genotype paths given the observation of SNP calls at low coverage in an expected heterozygosity scenario. A, B, H are possible states of underlying sequence, while *a*, *b* and *h* represent low-coverage calls obtained from Illumina reads. E_{het} and E_{hom} are the probability of false heterozygous calling for a given site. *C* is the probability of obtaining a heterozygous call at a given site by the presence of reads from both alleles. *T* is the probability of transition from one genotype to another (i.e. crossing-over events) (from Scaglione et al. (2016), used under CC BY 4.0)

organized in 138,608 haplotype blocks, while in cultivated cardoon "Alt41" 1,356,998 variant sites defined 78,458 haplotype blocks. For each parent, and for each block, the invocation of the HMM engine was performed on either phase 0 or phase 1 segregating sites. When available, the haplotype calls from both phase 0 and phase 1 of the same block were used to cross-validate a consensus call, while in case of contradictory calls, those with a significantly higher probability were chosen. A minimum calling rate of 70% was adopted to retain a given block as informative maker for the segregating parent. The F_1 mapping population was previously analysed for linkage mapping purposes and did not showed specific segregation (Portis et al. 2009); therefore, in order to clean up from spurious block (false heterozygous stretches), X^2 test threshold was applied (p-value < 0.03). "C3" parental line obtained successful calls for 23,381 blocks,

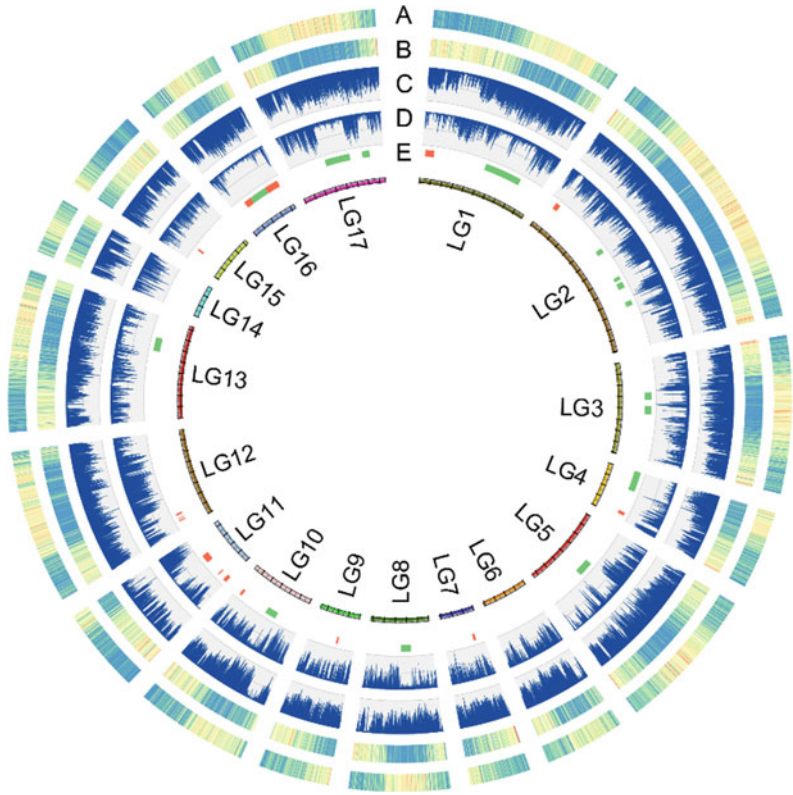


Fig. 9.4 Characteristics of the globe artichoke genome. (a) Heat map of repeat density in the reference genome generated from globe artichoke 2C (blue, low; orange, high density); (b) heat map of gene density in the reference genome; (c) density of heterozygous SNPs in the globe artichoke parental genotype, C3; (d) density of heterozygous SNPs in the cultivated cardoon parental genotype, Alt41; (e) extended homozygous regions in cultivated cardoon: the colour of the boxes indicates whether they occur in repeat-rich regions (green) or gene-rich regions (red) (from Scaglione et al. (2016), used under CC BY 4.0)

distributed in 4234 scaffolds, while "Alt41" male parent collected 15,262 genotyped blocks for 2779 scaffolds; both "C3" and "Alt41" had 1677 mapped scaffolds in common. The two-way pseudo-test cross design allowed to obtain two independent parental linkage analysis which defined 1157 and 1497 genetic bins for the globe artichoke and the cultivated cardoon, respectively. Overall, this corresponded to 445 Mbp of the assembly, or 23,366 markers for "C3" and 358 Mbp or 15,227 markers for "Alt41". A Perl implementation of the SMOOTH algorithm (van Os et al. 2005) was used to impute missing calls in the matrix, and further iterative steps (two) of markers re-ordering were carried to re-locate corrected makers and generated the final maps. In "C3", the linkage map counted a total of 1460 cM of segregation and 1748 cM were calculated for "Alt41"; the average density of markers was 1.28 cM and 1.18 cM, respectively. The two maps resulted in an extraordinary high level of co-linearity easing the generation of a consensus map which worked as a framework to separate co-segregating makers within the same bin in a given map. The final map of 17 chromosomes contained 2178 distinct genetic bins, anchoring 5322 scaffolds and thus resulting in 526 Mbp arranged as pseudomolecules (Table 9.2). Of these scaffolds, 936 (210 Mbp) were oriented by correlating the genetic position of each marker within a given scaffold along with their physical coordinates. Pseudomolecules

Table 9.2 Pseudomolecule statistics

Chromosome	Length (bp)	Average scaffold length	Average GC content (%)	Non-AGCT bases	N° scaffolds	Oriented scaffolds	Oriented bp	Oriented bp (%)	C3 (cM)	Alt41 (cM)	'C3' parental map
1	49,707,439	104,647	34.40	4,211,747	475	104	20,760,124	41.8	146.35	182.62	C1
2	70,339,030	68,623	35.80	6,880,431	1025	174	26,305,745	37.4	162.69	170.83	C2
3	40,261,365	111,527	34.70	3,216,535	361	68	17,965,994	44.6	117.13	142.67	C3
4	20,147,618	119,926	34.60	1,493,262	168	23	7,754,853	38.5	65.27	60.02	C4
5	37,162,917	110,276	34.40	3,124,327	337	50	11,700,443	31.5	116.42	115.36	C5
6	20,611,051	89,225	34.60	2,027,785	231	36	7,554,461	36.7	67.31	73.75	C6
7	15,559,887	170,988	34.10	1,081,502	91	27	8,988,908	57.8	66.03	79.05	C7
8	25,924,184	112,714	34.80	2,005,747	230	38	8,648,700	33.4	71.41	74.74	C8
9	18,326,914	106,552	34.40	1,576,491	172	43	9,416,779	51.4	60.65	83.17	C9
10	29,104,343	100,707	34.70	2,400,099	289	52	12,098,280	41.6	59.93	77.68	C10
11	22,005,725	196,480	33.70	1,381,011	112	34	11,778,224	53.5	64.93	128.76	C11 + C20
12	39,646,955	85,816	35.10	3,702,100	462	51	11,008,553	27.8	70.28	91.45	C12
13	41,505,999	91,222	34.70	3,530,600	455	90	17,212,358	41.5	120.78	140.35	C13 + C18
14	14,475,048	113,086	34.00	1,073,474	128	25	5,746,158	39.7	47.03	55.68	C14
15	21,262,825	172,868	33.50	1,556,777	123	45	14,447,054	67.9	70.39	103.32	C15 + C19
16	21,914,710	115,951	34.60	1,799,640	189	21	6,168,798	28.1	67.82	65.64	C16
17	37,690,487	79,516	35.90	3,591,453	474	55	12,083,853	32.1	85.16	103.06	C17
Chromosome total	525,646,497	114,713	34.59	44,652,981	5322	936	2,1E+08	40	–	–	–
Unmapped	199,020,768	24,077	37	25,635,496	8266	–	–	–	–	–	–
Grand total	724,667,265	53,091	36	70,288,477	13,588	–	–	–	–	–	–

In the last column, the pseudomolecules and linkage group, as depicted in Portis et al. (2009), relationships are reported; the nomenclature is from the C3 map (from Scaglione et al. (2016), used under CC BY 4.0)

were named to be consistent with previously published linkage groups on the same progeny (Table 9.2).

Since no other reference was available to validate the current maps, the robustness of the genetic anchoring of scaffolds was assessed by means of comparing the two parental maps. Briefly, all possible pairs of scaffolds separated by 5 cM or less were analysed. Concordant pairs were defined if two scaffolds maintained the same reciprocal position across the two maps. The former exercise indicated that 23,558 pairs of scaffolds (87%) were collinear each other across the two maps, while only 3480 had discordant positions. However, for scaffolds of 100 Kbp and more the proportion of valid pairs raised to 94%. This indicated that the genotype imputation method via low-coverage sequencing was effective and errors where only introduced by too small scaffolds (unable to provide a robust path of allelic sites). Moreover, small structural variation across the two genotypes cannot be excluded.

High-density genetic maps have proved to be a powerful tool for final ordering of scaffolds, nowadays competing only with Hi-C data and optical mapping. In the last years, increase in the throughput of next-generation sequencing democratized the production of massive amounts of genotyping data in a cost-effective manner. Different methods are based on a complexity reduction of template genome operated by restriction enzymes (GBS, Elshire et al. 2011; RADseq, Etter et al. 2011; ddRAD, Peterson et al. 2012; MSG, Andolfatto et al. 2011) which showed to be widely adopted tools for NGS-based genotyping. Still, restriction enzymes recognition site, diversity of parental genotypes, size of sequenced fragments will define the pool of informative loci and its final size, also determining the rate of recall across the population cohort. On the other side, several publications on the ability of exploiting low-coverage whole-genome sequencing of segregating individuals, requiring sequencing yields comparable to those of restriction-based techniques (Mascher et al. 2013). However, previous works were based on availability of at least one parental haplotype as represented in the reference sequence or in homozygous parental lines. More specifically, robust determination of genotype states from shallow sequencing data (<1×) has been so far explored only in the context of two alternative haplotypes and a single read is virtually enough for genotype calling, thus restricting their application to RILs. In other populations, such as F_2 and BC, higher local depth of coverage would be required to obtain a single-base heterozygous genotype state and then proceed to aggregate more sites to validate the genotype of a given locus. For these reasons, to make it possible a linkage-based chromosomal anchoring in F1 populations from heterozygous parents, as in the outbreeding *C. cardunculus*, we proposed a novel methodology, which employs low-coverage genotyping-by-sequencing of an F_1 progeny based on haplotype phase information computed at parental level. The pipeline proved to allow a correct positioning and orienting of large fraction of sequence from de novo draft assembly of globe artichoke. Although the pipeline was developed to analyse F_1 progeny data, F_2, RILs and BC populations can be analysed similarly.

9.4 Gene Prediction

The coding fraction of the genome was identified in 26,889 gene models (extended to 27,121 predicted transcripts), and 23,895 of them were placed in the chromosomal pseudomolecules (Table 9.3) with only 2994 residing in unmapped scaffolds. The sum of gene models covered a total of 146.5 Mbp, where 38 Mbp were made by exons having a median length of 141 bp, while introns had a median of 334 bp; coding sequences represented 32.8 Mbp. Gene-rich regions along with repeat-dense regions were analysed and suggested the locations of the centromeres.

Among 27,121 predicted protein sequences, the molecular function of 19,554 (72%) was found. Sequence clustering resulted in 3724 groups of proteins, accounting for 21,850 sequences (80.5% of the total number of

Table 9.3 Statistics on the predicted genes for each chromosomal pseudomolecule (from Scaglione et al. (2016), used under CC BY 4.0)

Chromosome	Gene models	Transcripts	Exons	Average exons per gene	# syntenic genes with *L. sativa*	
1	2630	2647	15,479	5.89	1942	(73.8%)
2	2351	2370	13,909	5.92	1639	(69.7%)
3	1868	1888	11,068	7.45	1187	(63.5%)
4	962	972	5809	6.04	703	(73.1%)
5	1640	1653	8632	5.26	1036	(63.2%)
6	903	915	5133	5.68	530	(58.7%)
7	907	914	5190	5.72	604	(66.6%)
8	1196	1209	7126	5.96	861	(72.0%)
9	1006	1012	5550	5.52	785	(78.0%)
10	1436	1444	8692	6.05	872	(60.7%)
11	1453	1462	9001	6.19	1238	(85.2%)
12	1404	1415	7814	5.57	898	(64.0%)
13	1801	1815	10,157	5.64	1140	(63.3%)
14	646	651	3515	5.44	396	(61.3%)
15	1466	1478	8288	5.65	1024	(69.8%)
16	949	960	5830	6.14	556	(58.6%)
17	1277	1290	7435	5.82	821	(64.3%)
Unmapped	2994	3026	15,576	5.19		0
Total	26,889	27,121	154,207	5.84	16,232	(67.9%)

Table 9.4 Most common functional annotations in the artichoke predicted proteins (from Scaglione et al. (2016), used under CC BY 4.0)

Description	Number of proteins
Protein kinase domain	1357
Zinc finger	1041
Serine/threonine protein kinases	898
Leucine-rich repeat	607
PPR repeat family	598
Protein tyrosine kinase	498
EF-hand domain	424
Cytochrome P450	306
ATPase family associated with various cellular activities (AAA)	283
Ring finger domain	267

predicted proteins). A functional enrichment analysis (Table 9.4) of such groups indicated that protein kinase domain is the most common molecular function (primarily serine/threonine kinases and tyrosine kinases) and followed by the zinc finger domain and the leucine-rich repeats.

The same clustering was also performed including proteins from annotated species *Arabidopsis thaliana*, *Brassica rapa*, *Fragaria vesca*, *Solanum lycopersicum* and *Lactuca sativa* (Fig. 9.5). This resulted in 25,981 clusters, considered as gene families, made from 153,126

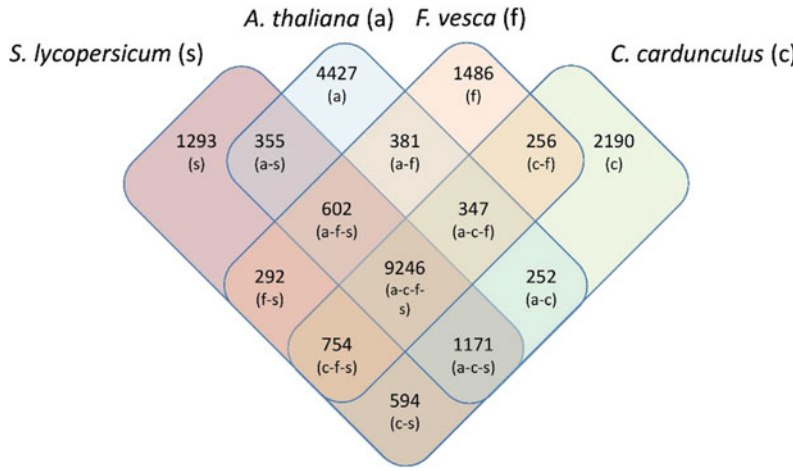


Fig. 9.5 Venn diagram of orthologous gene clusters among *Arabidopsis thaliana* (a), *Cynara cardunculus* (c), *Fragaria vesca* (f) and *Solanum lycopersicum*(s), showing a total of 9246 common gene clusters (from Scaglione et al. (2016), used under CC BY 4.0)

Table 9.5 Statistical analysis of gene families per species (from Scaglione et al. (2016), used under CC BY 4.0)

Species	Gene number	Genes in families	Unclustered genes	Number of families	Unique families	Genes in unique families	Average number of genes per family
A. thaliana	27,416	23,172	4244	16,781	422	1254	1.38
B. rapa	40,905	31,625	9280	17,070	876	2916	1.85
C. cardunculus	27,196	20,509	6687	14,810	418	1170	1.38
F. vesca	32,831	23,729	9102	13,364	1394	6603	1.77
L. sativa	38,919	29,184	9735	16,258	1435	5950	1.79
S. lycopersicon	34,727	24,907	9820	14,307	1118	4743	1.74

proteins from the pool of 201,990 across the six taxa. 20,509 proteins of globe artichoke were distributed into 14,810 gene families, of which, 6005 were shared by all the six species; only 2911 of these were detected as single copy in all the species. Globe artichoke and tomato shared the largest number of gene families, presumably due to their closer phylogenetic relationship. The shared families represented the 79.4% (for globe artichoke) and 82.2% (for tomato) of their individual total gene clusters (Table 9.5).

It was also investigated for the presence of significantly larger gene families within species (Table 9.6). Forty-six families returned a significant deviation, mostly driven by larger counts in either *Arabidopsis*, tomato, strawberry or lettuce. Only in four clusters globe artichoke showed expanded number of genes. For three groups, namely (i) cystatin-like proteins with the cysteine proteases inhibitor signature (33 members), (ii) replication factor A, C-terminal domain (15 members) and (iii) glycosyltransferase family 10 (fucosyltransferase, 8 members), it was found no local clustering on chromosomes. Only a group (27 members), encoding "p450 class E enzyme group I proteins" was physically clustered in chromosome 13 (15 out of 27 genes), in a locus of 800 Kbp which harboured other 44 interspersed genes (Fig. 9.6). A phylogenetic analysis suggested that duplication events caused the amplification of this gene family within the locus (Fig. 9.6).

Moreover, globe artichoke and lettuce shared five gene families with significant higher gene

Table 9.6 Statistical analysis of expanded gene families in globe artichoke/lettuce and in globe artichoke in a six-species comparison

Prevalent expansion in globe artichoke														
Clustering	Annotation	Group name	Atha	Brap	Ccar	Fves	Lsat	Slyc	No. of species	Total count	Mean count	SD	CV	Adj X^2
–	Cystatin domain	Comp10193	0	0	33	1	1	1	4	36	9.00	13.24	1.47	1.17E −24
Chromosome 13	E-class P450 group I signature	Comp10249	0	0	27	0	1	4	3	32	5.33	10.73	2.01	2.11E −3
–	Replication fact-A C-term domain (DUF23)	Comp10806	0	0	15	0	0	2	2	17	2.83	6.01	2.12	1.68E −3
–	Glycosyltransferase family 10	Comp13550	0	0	8	0	1	0	2	9	1.50	3.21	2.14	NA
Prevalent expansion in globe artichoke/lettuce														
Clustering	Annotation	Group name	Atha	Brap	Ccar	Fves	Lsat	Slyc	No. of species	Total count	Mean count	SD	CV	Adj X^2
2—Chromosome 10	LRR	Comp10022	1	1	20	10	55	9	6	96	16.00	20.36	1.27	2.50E −20
3—Chromosomes 5,6,9	Bulb-type mannose-specific lectin	Comp10125	0	0	19	8	18	2	4	47	11.75	8.77	0.75	1.04E −7
1—Chromosome 10	LRR	Comp10027	0	0	19	3	71	0	3	93	15.50	28.18	1.82	6.14E −3
4—Chromosomes 2,13,3,6	Pentatricopeptide (PPR)	Comp10088	0	0	36	0	21	0	2	57	9.50	15.46	1.63	5.66E −3
1—Chromosome 15	Pathogenesis-related protein Bet v I family	Comp10405	0	0	13	0	11	0	2	24	4.00	6.23	1.56	2.38E −3

Atha = arabidopsis; Brap = brassica; Ccar = globe artichoke; Fves = strawberry; Slyc = tomato
From Scaglione et al. (2016), used under CC BY 4.0

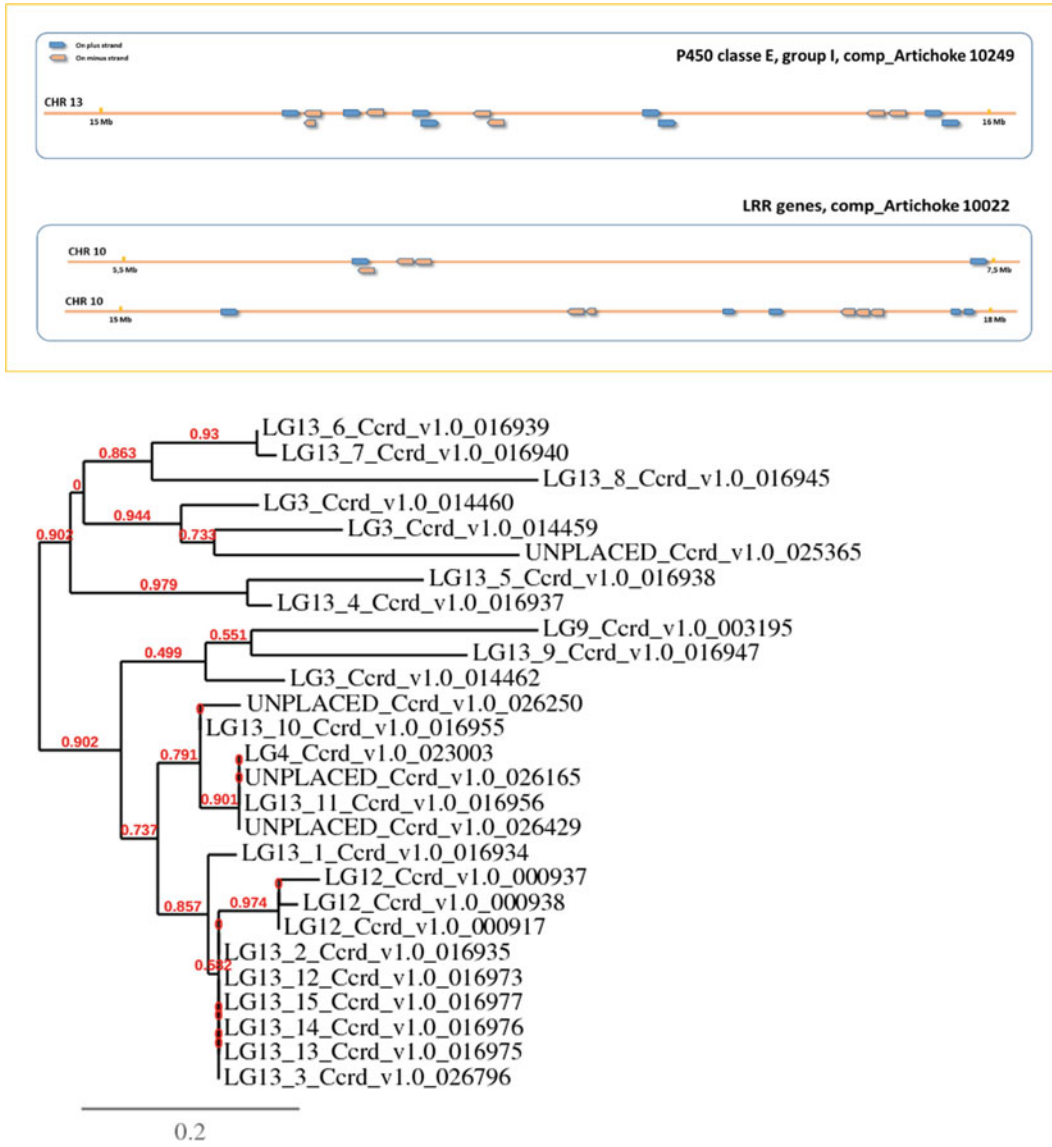


Fig. 9.6 Top: Gene clustering of some globe artichoke-specific OrthoMCL gene groups. Bottom: Phylogeny of globe artichoke expanded P450 class E, group I, comp_Artichoke 10249 (from Scaglione et al. (2016), used under CC BY 4.0)

members in both. Two of them encoded NB-LRR disease resistance-related proteins (with 20 and 19 members) and one having most of its genes located on chromosome 10 but in two distinct loci of 2 Mbp each. Another gene family representing bulb-type mannose-specific lectin associated protein kinases (19 members) was mapped on three chromosomes (chromosome 5, 5 members; chromosome 6, 5 members and chromosome 9, 4 members). Pentatricopeptide proteins (36 members) were the largest group while allergen-related protein family (13 members) was mostly localized on chromosome 15 within a 1 Mbp region.

Gene families found to be exclusive of globe artichoke were 418, counting 1170 genes

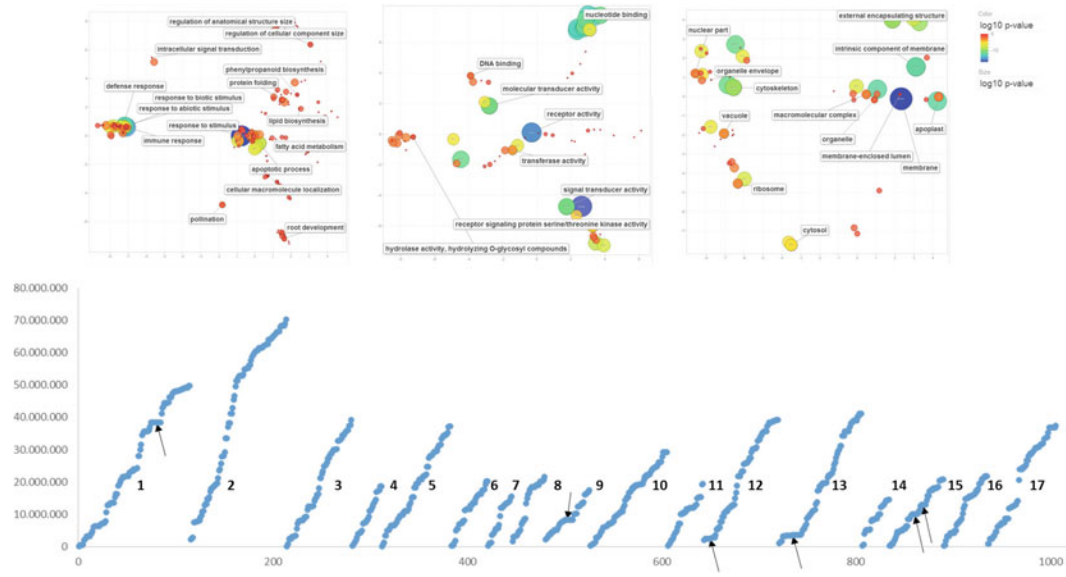


Fig. 9.7 Top: GO enrichment diagram for the 1170 globe artichoke-specific genes in the process, function and component. Enrichment was calculated with AgriGo (http://bioinfo.cau.edu.cn/agriGO) and visualized with the REVIGO suite (http://revigo.irb.hr). Bottom: Genomic distribution of the 1170 globe artichoke-specific genes along the 17 chromosomes. Black arrows show gene clusters (from Scaglione et al. (2016), used under CC BY 4.0)

(Fig. 9.7). Functions such as "response to stress", "phenylpropanoid biosynthesis" and "lipid biosynthesis" were found as over-represented classes in these genes. The specific genes are widely distributed over the 17 pseudomolecules and 18 regions showed clusters of genes likely related to the same biological process (Fig. 9.8).

First insights into the coding gene space of *C. cardunculus* were possible thanks to the availability of genome sequence. Gene duplication has shown, like in other species, to be an important factor contributing to gene families' evolution and the neo-functionalization or sub-functionalization of the same (Ohta 1991). Also, the analysis of non-coding regions is a focal point for understanding regulatory pathways in higher organisms, since their level of complexity has been associated with the length of non-coding sequences (Barrett et al. 2012).

Gene families such as the genes encoding phytocystatins (PhyCys), responsible for pest defence function (Martinez and Diaz 2008), and P450 proteins which are known to contribute in the metabolism of secondary metabolites were found to contain many members as local clusters (Fig. 9.6). Other expansions of gene families such as NBS-LRR proteins were also shown to be common in lettuce. All these genes can be the subject of further studies to better understand their role in biotic stress resistance and their contribution to secondary metabolite pathways. In general, physically clustered and/or species-specific genes families are shown to be frequent in globe artichoke. A comparative analysis at genome scale with *Cynara cardunculus* wild relatives will be a powerful tool to dissect the domestication footprint in the landscape of large gene families.

9.5 Repetitive Content of the Genome

The genome sequence was assessed to be repetitive in 58.4% (424 Mbp) of its current assembled length. The 41.73% of the repeats were not present in current database and thus were not classified. Retroelements were the most

gene	strand	LG	LG coordinates	PHMMER (uniprotKB db)	e-value (phmmer)	PROCESS
Ccrd_v1.0_013577	-	1	38.330.947	Protein_ENHANCED_DOWNY	9,8E-39	Defense response to fungus
Ccrd_v1.0_013578	-	1	38.332.679	Protein_ENHANCED_DOWNY	1,3E-09	
Ccrd_v1.0_013579	-	1	38.333.957	Protein_ENHANCED_DOWNY	2,1E-17	
Ccrd_v1.0_013580	-	1	38.335.514	Protein_ENHANCED_DOWNY	4,8E-13	
Ccrd_v1.0_013581	-	1	38.335.793	Protein_ENHANCED_DOWNY	5E-23	
Ccrd_v1.0_013582	-	1	38.337.064	Protein_ENHANCED_DOWNY	2,6E-28	
Ccrd_v1.0_013583	-	1	38.342.744	Protein_ENHANCED_DOWNY	6,8E-14	
Ccrd_v1.0_013584	-	1	38.350.535	Protein_ENHANCED_DOWNY	4,5E-30	
Ccrd_v1.0_013585	-	1	38.352.862	Protein_ENHANCED_DOWNY	9,3E-57	
Ccrd_v1.0_013586	-	1	38.356.901	Protein_ENHANCED_DOWNY	1,1E-14	
Ccrd_v1.0_013587	-	1	38.357.345	Protein_ENHANCED_DOWNY	5,2E-28	
Ccrd_v1.0_013882	+	1	47.958.758	LRR_receptor-like_serine/threonine-protein	1,2E-103	Defense response signaling pathway
Ccrd_v1.0_013882	+	1	47.962.409	LRR_receptor-like_serine/threonine-protein	1,2E-103	
Ccrd_v1.0_013882	+	1	47.966.732	LRR_receptor-like_serine/threonine-protein	1,2E-103	
Ccrd_v1.0_013892	-	1	48.128.505	Receptor-like_protein_12	4,6E-107	
Ccrd_v1.0_013892	-	1	48.131.550	Receptor-like_protein_12	4,6E-107	
Ccrd_v1.0_013906	+	1	48.313.838	Myb-related_protein	1,2E-17	
Ccrd_v1.0_013907	+	1	48.325.308	Transcription_factor_MYB51	1,1E-12	
Ccrd_v1.0_013918	-	1	48.499.373	#N/D	#N/D	
Ccrd_v1.0_013927	-	1	48.653.548	Probable_acyl-activating_enzyme_peroxisomal	3,5E-22	
Ccrd_v1.0_013928	+	1	48.667.122	Probable_acyl-activating_enzyme_peroxisomal	1,2E-20	
Ccrd_v1.0_013947	-	1	48.983.694	Disease_resistance-like_protein	3E-13	
Ccrd_v1.0_013947	-	1	48.994.352	TMV_resistance_protein	3E-20	
Ccrd_v1.0_013952	+	1	49.085.948	Disease_resistance_protein	2,4E-16	
Ccrd_v1.0_013953	-	1	49.095.145	Putative_Myb_family_factor	8,6E-15	
Ccrd_v1.0_013954	+	1	49.100.100	Disease_resistance_protein	5,9E-26	
Ccrd_v1.0_013998	+	1	49.601.344	Syntaxin-61	1E-17	
Ccrd_v1.0_013998	+	1	49.606.095	Syntaxin-61	1E-17	
Ccrd_v1.0_009654	+	2	7.461.604	Transcription_factor_BIM1	0,0035	Transcription regulation
Ccrd_v1.0_009654	+	2	7.466.862	Transcription_factor_BIM1	0,0035	
Ccrd_v1.0_009658	+	2	7.547.756	unknown	0,42	
Ccrd_v1.0_009659	-	2	7.549.840	Tubulin_beta-2_chain	7,4E-172	
Ccrd_v1.0_009676	-	2	7.776.546	Transcriptional_activator	8,9E-48	
Ccrd_v1.0_009678	-	2	7.804.691	Myb-related_protein	1E-47	
Ccrd_v1.0_010683	+	2	61.492.858	Stearoyl-[acyl-carrier-protein]_9-desaturase_6	7,7E-159	Fatty Acids synthesis
Ccrd_v1.0_010683	+	2	61.497.880	Stearoyl-[acyl-carrier-protein]_9-desaturase_6	7,7E-159	
Ccrd_v1.0_010691	+	2	61.661.811	Stearoyl-[acyl-carrier-protein]_9-desaturase_6	8,6E-165	
Ccrd_v1.0_018157	+	5	6.441.006	Cyclin-dependent_kinase	5,2E-113	Cell cycle regulation
Ccrd_v1.0_018158	+	5	6.446.869	Cyclin-dependent_kinase	2,5E-114	
Ccrd_v1.0_018159	-	5	6.449.546	Cyclin-dependent_kinase	1,4E-129	
Ccrd_v1.0_001281	-	5	6.451.678	Cyclin-dependent_kinase	2,3E-100	
Ccrd_v1.0_018176	-	5	6.708.574	Cyclin-dependent_kinase	4,5E-114	
Ccrd_v1.0_018581	+	5	20.768.511	GTP-binding_nuclear_protein	5,2E-18	Protein transport
Ccrd_v1.0_018581	+	5	20.769.926	GTP-binding_nuclear_protein	5,2E-18	
Ccrd_v1.0_018581	+	5	20.771.685	GTP-binding_nuclear_protein	5,2E-18	
Ccrd_v1.0_018581	+	5	20.772.331	GTP-binding_nuclear_protein	5,2E-18	
Ccrd_v1.0_022377	-	6	11.215.709	Calreticulin	1,5E-227	Signal transduction
Ccrd_v1.0_022382	-	6	11.427.186	B2_protein	0,0018	
Ccrd_v1.0_022383	+	6	11.441.485	Putative_calmodulin-like_protein	0,000000063	
Ccrd_v1.0_022383	+	6	11.442.672	Putative_calmodulin-like_protein	0,000000063	
Ccrd_v1.0_022384	+	6	11.465.132	Calmodulin-like_protein	0,000000025	
Ccrd_v1.0_022580	+	6	16.109.439	Two-pore_potassium_channel	1,2E-30	ion TM transport
Ccrd_v1.0_022583	-	6	16.135.474	Two-pore_potassium_channel	1,4E-29	
Ccrd_v1.0_022582	+	6	16.138.823	Two-pore_potassium_channel	2,3E-23	
Ccrd_v1.0_022584	+	6	16.168.013	Two-pore_potassium_channel	5,2E-30	
Ccrd_v1.0_003123	+	9	7.792.930	Fibrillin-1	1,2E-79	Photooxidative stress
Ccrd_v1.0_003124	+	9	7.827.169	Fibrillin-3	4,6E-48	
Ccrd_v1.0_003138	-	9	8.112.291	Wall-associated_receptor_kinase	5E-14	
Ccrd_v1.0_003138	-	9	8.122.522	Wall-associated_receptor_kinase	5E-14	
Ccrd_v1.0_003138	-	9	8.125.063	Wall-associated_receptor_kinase	5E-14	
Ccrd_v1.0_003138	-	9	8.126.938	Wall-associated_receptor_kinase	5E-14	
Ccrd_v1.0_003139	-	9	8.158.833	Fibrillin-1	2,4E-75	
Ccrd_v1.0_003140	-	9	8.180.750	Wall-associated_receptor_kinase	4,6E-12	
Ccrd_v1.0_003141	-	9	8.203.319	Fibrillin-3	3,1E-85	
Ccrd_v1.0_003142	-	9	8.224.128	Fibrillin-3	3,3E-86	

Fig. 9.8 Globe artichoke-specific genes and their clustering coordinates. The main biological process for each cluster is reported (from Scaglione et al. (2016), used under CC BY 4.0)

gene	strand	LG	LG coordinates	PHMMER (uniprotKB db)	e-value (phmmer)	PROCESS
Ccrd_v1.0_021222	+	10	19.090.822	Protein_SUPPRESSOR_OF_CONSTITUTIVE_1	5,5E-36	Defence response
Ccrd_v1.0_023603	+	10	19.179.725	Probable_serine/threonine-protein_kinase	1,4E-99	
Ccrd_v1.0_021229	+	10	19.252.203	Receptor-like_protein_kinase	2,7E-53	
Ccrd_v1.0_021232	-	10	19.311.453	Probable_serine/threonine-protein_kinase	4,1E-26	
Ccrd_v1.0_021232	-	10	19.317.722	Probable_serine/threonine-protein_kinase	4,1E-26	
Ccrd_v1.0_016160	+	13	3.282.715	Putative_receptor-like_protein	4,8E-71	Defence response
Ccrd_v1.0_016161	+	13	3.286.163	Putative_receptor-like_protein	4,8E-59	
Ccrd_v1.0_016162	+	13	3.299.609	Receptor-like_protein_kinase	4E-67	
Ccrd_v1.0_016164	+	13	3.315.293	Receptor-like_protein_kinase_1	1,9E-68	
Ccrd_v1.0_016165	-	13	3.322.034	TMV_resistance_protein	1,4E-46	
Ccrd_v1.0_016170	+	13	3.370.130	Protein_SUPPRESSOR_OF_CONSTITUTIVE_1	1,7E-47	
Ccrd_v1.0_016173	+	13	3.412.192	unknown	0,09	
Ccrd_v1.0_016174	+	13	3.427.346	Protein_SUPPRESSOR_OF_CONSTITUTIVE_1	1,7E-47	
Ccrd_v1.0_016175	+	13	3.440.527	TMV_resistance_protein	3,1E-16	
Ccrd_v1.0_016175	+	13	3.440.527	TMV_resistance_protein	1,7E-16	
Ccrd_v1.0_016175	+	13	3.447.201	TMV_resistance_protein	1,7E-16	
Ccrd_v1.0_016175	+	13	3.447.202	TMV_resistance_protein	3,1E-16	
Ccrd_v1.0_016176	+	13	3.462.836	Protein_SUPPRESSOR_OF_CONSTITUTIVE_1	3,4E-43	
Ccrd_v1.0_016177	+	13	3.476.304	TMV_resistance_protein	8,4E-20	
Ccrd_v1.0_016178	+	13	3.482.883	TMV_resistance_protein	2,5E-15	
Ccrd_v1.0_016179	+	13	3.492.917	TMV_resistance_protein	1,4E-18	
Ccrd_v1.0_016196	-	13	3.799.430	RNA_polymerase_II_5-mediating_protein	0,0000022	
Ccrd_v1.0_016267	-	13	4.833.186	Serine/threonine-protein_kinase-like_protein	0,0000082	
Ccrd_v1.0_016291	-	13	5.184.329	Probable_receptor-like_protein	2,4E-69	
Ccrd_v1.0_008635	+	14	4.964.500	Chlorophyll_a-b_binding_16_chloroplastic	1,6E-162	response to light stimulus
Ccrd_v1.0_008642	-	14	5.121.295	Chlorophyll_a-b_binding_16_chloroplastic	6,6E-163	
Ccrd_v1.0_008644	+	14	5.146.900	ATP-dependent_Clp_protease_subunit_ClpA	0,0015	
Ccrd_v1.0_008645	-	14	5.155.501	Chlorophyll_a-b_binding_16_chloroplastic	2,1E-160	
Ccrd_v1.0_008646	-	14	5.168.538	Chlorophyll_a-b_binding_16_chloroplastic	2,8E-161	
Ccrd_v1.0_008647	-	14	5.220.185	Chlorophyll_a-b_binding_16_chloroplastic	3,4E-164	
Ccrd_v1.0_008736	-	14	8.288.381	Transcription_factor_MYB48	6,3E-14	Transcr. regulation
Ccrd_v1.0_008736	-	14	8.292.935	Transcription_factor_MYB48	6,3E-14	
Ccrd_v1.0_008737	-	14	8.293.175	Myb-related_protein_306_majus	1,3E-16	
Ccrd_v1.0_008997	-	14	14.153.985	Auxin_response_factor	2,7E-71	Auxin-activ. Signal. pathway
Ccrd_v1.0_008999	+	14	14.177.331	Auxin_response_factor	1,9E-83	
Ccrd_v1.0_009000	+	14	14.188.710	Auxin_response_factor	1,2E-57	
Ccrd_v1.0_001911	-	15	9.910.627	Auxin-responsive_protein_SAUR22	4,3E-30	Auxin-activated signaling pathway
Ccrd_v1.0_001912	-	15	9.919.876	Auxin-responsive_protein_SAUR22	4E-30	
Ccrd_v1.0_001914	-	15	9.929.794	Auxin-responsive_protein_SAUR21	5E-27	
Ccrd_v1.0_001921	-	15	10.026.525	Auxin-responsive_protein_SAUR19	5,4E-26	
Ccrd_v1.0_001923	+	15	10.029.056	Auxin-responsive_protein_SAUR22	3E-21	
Ccrd_v1.0_001922	-	15	10.029.639	Auxin-responsive_protein_SAUR21	1,5E-24	
Ccrd_v1.0_006281	-	16	4.067.666	Early_light-induced_protein_chloroplastic	1,6E-51	response to high light intensity
Ccrd_v1.0_006282	-	16	4.074.913	Early_light-induced_protein_chloroplastic	1,6E-48	
Ccrd_v1.0_006282	-	16	4.075.565	Early_light-induced_protein_chloroplastic	1,6E-48	
Ccrd_v1.0_006282	-	16	4.079.627	Early_light-induced_protein_chloroplastic	1,6E-48	
Ccrd_v1.0_006285	-	16	4.116.754	Early_light-induced_protein_chloroplastic	1E-50	
Ccrd_v1.0_006864	+	16	21.154.883	Protein_SUPPRESSOR_OF_CONSTITUTIVE_1	3,4E-22	Defence response
Ccrd_v1.0_006905	-	16	21.626.285	Protein_SUPPRESSOR_OF_CONSTITUTIVE_1	2,3E-37	
Ccrd_v1.0_006906	-	16	21.677.103	Protein_SUPPRESSOR_OF_CONSTITUTIVE_1	1,3E-33	
Ccrd_v1.0_006908	-	16	21.698.804	Protein_SUPPRESSOR_OF_CONSTITUTIVE_1	2,3E-36	
Ccrd_v1.0_019921	+	17	7.249.743	WAT1-related_protein	8,9E-51	Auxin-activated signaling pathway
Ccrd_v1.0_019923	+	17	7.281.844	WAT1-related_protein	1,8E-28	
Ccrd_v1.0_019925	+	17	7.311.521	WAT1-related_protein	5,4E-51	
Ccrd_v1.0_019926	+	17	7.327.575	WAT1-related_protein	5,9E-36	

Fig. 9.8 (continued)

abundant, with *Ty/1 Copia* (27.84%) and *Gypsy* elements (16.48%) as the most represented super-families, followed by class II DNA elements (7.66%). Of the 14,449 complete LTR retrotransposon-like elements predicted, 610 clustered families were obtained and 354 were successfully classified into families and sub-families, while 154 were assigned to main families (RLG—retrotransposon LTR Gypsy, RLC—retrotransposon LTR Copia and RIX—retrotransposon unclassified). Of the remainder 12,107 unclustered elements, a BLAST search against the TREP database (http://wheat.pw.usda.gov/ggpages/Repeats/) aided to the assignment of further 7909 repeats to their type/family. Among annotated repeats, *Gypsy* and *Copia* families showed a rate of 57 and 37%, respectively. Age of elements was estimated with the analysis of LTRs sequences, using as a proxy the rate of divergence within a cluster. The data provided evidence of a large expansion across all elements that happened some 2.5 Mya ago. *Gypsy* elements showed a lower variability in the rate of mutations across sub-type, suggesting a more stable family across time. This was not true for the *Fatima* sub-family and for most of sub-types of the *Copia* family were some sub-families showed a reduction of expansion over the time, while others had a higher proportion of young elements (Fig. 9.9). Timing of repeats expansion in *C. cardunculus* indicates it took place quite recently and after divergence from the *Heliantheae–Cichorieae* lineage (see next section). Recent increase in retrotransposon activity has also been reported in sunflower (Staton et al. 2012), and this suggested this common trend is related to evolutionary events since they all occurred after the separation of the tribes. The sunflower repeats' expansion is younger, less than 1 Mya, than the expansion that occurred 2.5 Mya in the globe artichoke lineage. This might find an explanation in the second WGD event which occurred only in the *Heliantheae* lineage and justify the significant genome size differences between the two tribes.

Satellite sequences were also searched in the assembly, and the three most abundant were monomers of 96, 94 and 103 bp, found 493, 477 and 191 times, respectively. Less frequent satellites were also found; however, the overall landscape of satellite motifs across the pseudomolecules showed a general localization within presumptive centromeres along with other repeats. It was not possible to determine a specific satellite motif involved in centromeres.

Pseudomolecules and all unmapped scaffolds were surveyed for candidate SSR markers and populated the first microsatellite database *CyMSatDB* (*Cynara cardunculus* MicroSatellite DataBase), available at www.artichokegenome.unito.it containing nearly hundred eighty thousands of perfect SSR motifs with some 15% being compound SSRs and some two hundred forty-four thousands of imperfect repeats. The SSR catalogue (namely *CyMSat*) was classified on the basis of the repeat motif, the number of units and their localization on the pseudomolecules (Portis et al. 2016). Di-nucleotides are the most frequent (73.0%), followed by tri- (11.2%), tetra-nucleotides (6.1%) and mono- (4.7%); penta- and hexa-nucleotides are rare (2.4 and 2.5%, respectively). The database is serving the community as a repository of candidate markers of known genomic positions to be exploited for rapid marker development of traits of interest in varying genetic background given the high level of polymorphisms provided by simple repeats.

9.6 Age of Speciation

Ages of speciation for *C. cardunculus* were estimated through the analysis of synonymous nucleotide substitutions (Ks). Whole-genome duplication preceding the divergence of the *Compositae* family was placed at Ks $\sim$ 0.78 in *C. cardunculus* and Ks $\sim$ 0.95 in *L. sativa*. After correcting those values for nuclear rate heterogeneity across *Compositae* with factors already estimated elsewhere (Barker et al. 2008), the *C. cardunculus* and *L. sativa* WGD were narrowed at Ks = 0.84 and Ks = 0.87, respectively, providing a solid evidence for a common

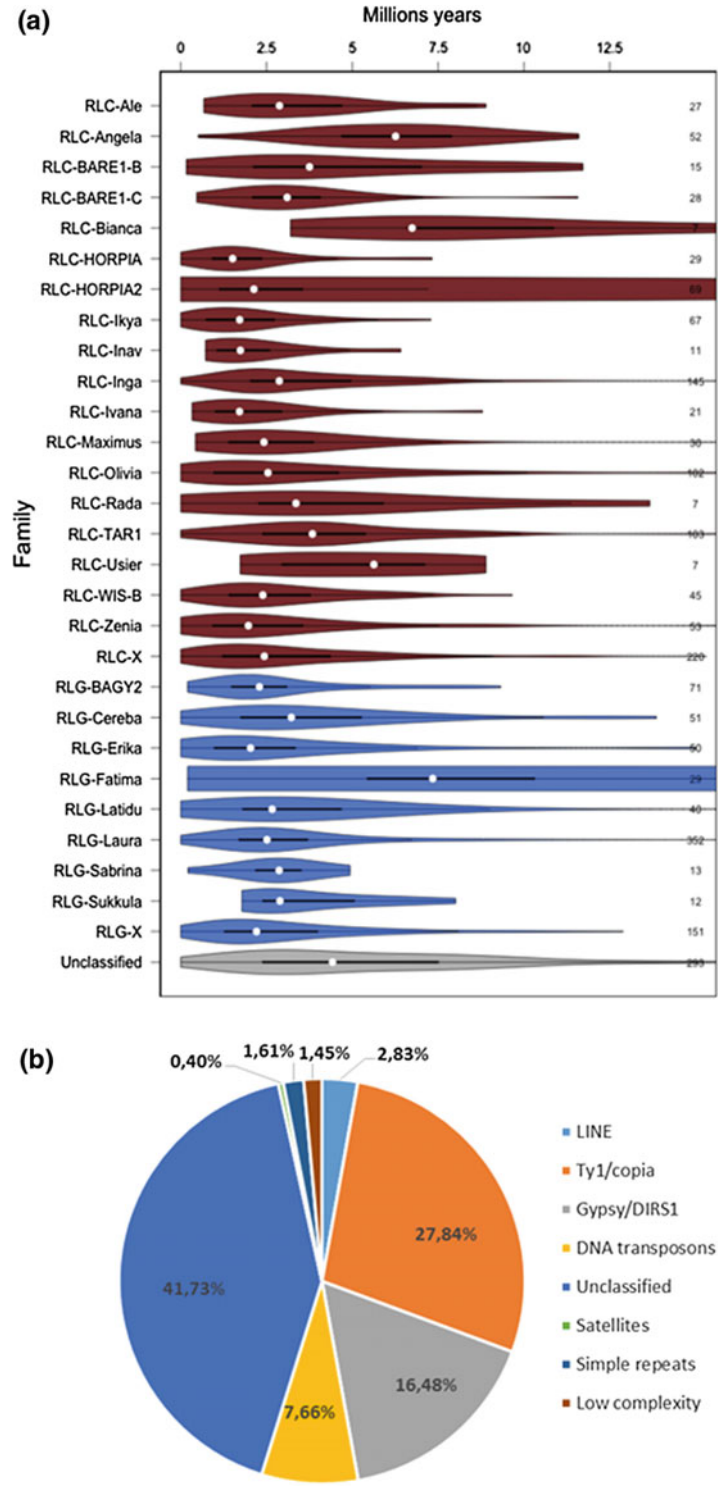


Fig. 9.9 Repeats' identification and dating in the genome of *C. cardunculus*. **a** Overall distribution of the repetitive fraction. **b** Distributions of insertion ages in different families of LTR elements (from Scaglione et al. (2016), used under CC BY 4.0)

Compositae WGD. The diversification of *Cardueae* and *Cichorieae* was placed at Ks ~ 0.55. Since the same age estimation pipeline was adopted for pepper and tomato which reported Ks = 0.27, it agreed with the previous estimation (Ks = 0.30, Kim et al. 2014). Therefore, pepper and tomato divergence, previously reported at 19.1 Mya, places the time of divergence of the *Cardueae*, *Cichorieae* and *Heliatheae* between 35 and 40 Mya.

References

Acquadro A, Lanteri S, Scaglione D, Arens P, Vosman B, Portis E (2009) Genetic mapping and annotation of genomic microsatellites isolated from globe artichoke. Theor Appl Genet, 1573–1587

Andolfatto P, Davison D, Erezyilmaz D, Hu T, Mast J, Sunayama-Morita T, Stern D (2011) Multiplexed shotgun genotyping for rapid and efficient genetic mapping. Genome Res 21:610–617

Bansal V, Bafna V (2008) HapCUT: an efficient and accurate algorithm for the haplotype assembly problem. Bioinformatics 24:I153–I159

Barker M, Kane N, Matvienko M, Kozik A, Michelmore W, Knapp S, Rieseberg L (2008) Multiple paleopolyploidizations during the evolution of the compositae reveal parallel patterns of duplicate gene retention after millions of years. Mol Biol Evol 25:2445–2455

Barrett L, Fletcher S, Wilton S (2012) Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements. Cell Mol Life Sci 69:3613–3634

Curci P, De Paola D, Danzi D, Vendramin G, Sonnante G (2015) Complete chloroplast genome of the multifunctional crop globe artichoke and comparison with other Asteraceae. PLoS ONE 10

Elshire R, Glaubitz J, Sun Q, Poland J, Kawamoto K, Buckler E, Mitchell S (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE 6

Etter P, Preston J, Bassham S, Cresko W, Johnson E (2011) Local de novo assembly of RAD paired-end contigs using short sequencing reads. PLoS ONE 6

Hirakawa H, Shirasawa K, Miyatake K, Nunome T, Negoro S, Ohyama A, Yamaguchi H, Sato S, Isobe S, Tabata S, Fukuoka H (2014) Draft genome sequence of eggplant (*Solanum melongena* L.): the representative *Solanum* species indigenous to the old world. DNA Res

Kesseli RV, Michelmore RW (1997) The Compositae: systematically fascinating but specifically neglected. In: Paterson AH (ed) Genome mapping in plants. R.G. Landes Co. Georgetown, TX, pp 179–191

Kim S, Park M, Yeom S-I, Kim Y-M, Lee JM, Lee H-A, Seo E, Choi J, Cheong K, Kim K-T, Jung K, Lee G-W, Oh S-K, Bae C, Kim S-B, Lee H-Y, Kim S-Y, Kim M-S, Kang B-C, Jo YD, Yang H-B, Jeong H-J, Kang W-H, Kwon J-K, Shin C, Lim JY, Park JH, Huh JH, Kim J-S, Kim B-D, Cohen O, Paran I, Suh MC, Lee SB, Kim Y-K, Shin Y, Noh S-J, Park J, Seo YS, Kwon S-Y, Kim HA, Park JM, Kim H-J, Choi S-B, Bosland PW, Reeves G, Jo S-H, Lee B-W, Cho H-T, Choi H-S, Lee M-S, Yu Y, Do Choi Y, Park B-S, van Deynze A, Ashrafi H, Hill T, Kim WT, Pai H-S, Ahn HK, Yeam I, Giovannoni JJ, Rose JKC, Sorensen I, Lee S-J, Kim RW, Choi I-Y, Choi B-S, Lim J-S, Lee Y-H, Choi D (2014) Genome sequence of the hot pepper provides insights into the evolution of pungency in *Capsicum* species. Nat Genet 46:270–278

Lanteri S, Acquadro A, Comino C, Mauro R, Mauromicale G, Portis E (2006) A first linkage map of globe artichoke (*Cynara cardunculus* var. *scolymus* L.) based on AFLP, S-SAP, M-AFLP and microsatellite markers. Theor Appl Genet 112:1532–1542

Li R, Fan W, Tian G, Zhu H, He L, Cai J, Huang Q, Cai Q, Li B, Bai Y, Zhang Z, Zhang Y, Wang W, Li J, Wei F, Li H, Jian M, Li J, Zhang Z, Nielsen R, Li D, Gu W, Yang Z, Xuan Z, Ryder O, Leung F, Zhou Y, Cao J, Sun X, Fu Y, Fang X, Guo X, Wang B, Hou R, Shen F, Mu B, Ni P, Lin R, Qian W, Wang G, Yu C, Nie W, Wang J, Wu Z, Liang H, Min J, Wu Q, Cheng S, Ruan J, Wang M, Shi Z, Wen M, Liu B, Ren X, Zheng H, Dong D, Cook K, Shan G, Zhang H, Kosiol C, Xie X, Lu Z, Zheng H, Li Y, Steiner C, Lam T, Lin S, Zhang Q, Li G, Tian J, Gong T, Liu H, Zhang D, Fang L, Ye C, Zhang J, Hu W, Xu A, Ren Y, Zhang G, Bruford M, Li Q, Ma L, Guo Y, An N, Hu Y, Zheng Y, Shi Y, Li Z, Liu Q, Chen Y, Zhao J, Qu N, Zhao S, Tian F, Wang X, Wang H, Xu L, Liu X, Vinar T, Wang Y, Lam T, Yiu S, Liu S, Zhang H, Li D, Huang Y, Wang X, Yang G, Jiang Z, Wang J, Qin N, Li L, Li J, Bolund L, Kristiansen K, Wong G, Olson M, Zhang X, Li S, Yang H, Wang J (2010) The sequence and de novo assembly of the giant panda genome. Nature 463:311–317

Marie D, Brown S (1993) A cytometric exercise in plant DNA histograms, with 2C-values for 70 species. Biol Cell 78:41–51

Martinez M, Diaz I (2008) The origin and evolution of plant cystatins and their target cysteine proteinases indicate a complex functional relationship. BMC Evol Biol 8

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(10):e76925

Ohta T (1991) Multigene families and the evolution of complexity. J Mol Evol 33:34–41

Parra G, Bradnam K, Korf I (2007) CEGMA: a pipeline to accurately annotate core genes in eukaryotic genornes. Bioinformatics 23:1061–1067

Peterson B, Weber J, Kay E, Fisher H, Hoekstra H (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS ONE 7(5):e37135

Portis E, Barchi L, Acquadro A, Macua JI, Lanteri S (2005) Genetic diversity assessment in cultivated cardoon by AFLP (amplified fragment length polymorphism) and microsatellite markers. Plant Breed 124(3):299–304

Portis E, Mauro RP, Barchi L, Acquadro A, Mauromicale G, Lanteri S (2014) Mapping yield-associated QTL in globe artichoke. Mol Breed 34:615–630

Portis E, Mauromicale G, Mauro R, Acquadro A, Scaglione D, Lanteri S (2009) Construction of a reference molecular linkage map of globe artichoke (*Cynara cardunculus* var. *scolymus*). Theor Appl Genet, 59–70

Portis E, Portis F, Valente L, Moglia A, Barchi L, Lanteri S, Acquadro A (2016) A genome-wide survey of the microsatellite content of the globe artichoke genome and the development of a web-based database. PLoS ONE 11(9):e016284

Portis E, Scaglione D, Acquadro A, Mauromicale G, Mauro R, Knapp SJ, Lanteri S (2012) Genetic mapping and identification of QTL for earliness in the globe artichoke/cultivated cardoon complex. BMC Res Notes 5:252

Portis E, Acquadro A, Tirone M, Pesce GR, Mauromicale G, Lanteri S (2018) Mapping the genomic regions encoding biomass-related traits in *Cynara cardunculus* L. Mol Breed 38(5):64

Qin C, Yu C, Shen Y, Fang X, Chen L, Min 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, Luo 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 111:5135–5140

Raven PH, Axelrod DI (1974) Angiosperm biogeography and past continental movements. Ann Mo Bot Gard 61:539–637

Rotino GL, Lanteri S, Sala T, Toppino L, Acquadro A, Barchi L, Portis E, Rinaldi R, Scaglione D, Dal Molin A, Minio A, Ferrarini A, Tononi P, Zamperin G, Fantini E, Pietrella M, Giuliano G, Delledonne M (2014) An eggplant (*Solanum melongena* L.) high quality genome draft. In: Plant and animal genome XXIInd conference, January 11–15, 2014, San Diego, CA, USA

Sato S, Tabata S, Hirakawa H, Asamizu E, Shirasawa K, Isobe S, Kaneko T, Nakamura Y, Shibata D, Aoki K, Egholm M, Knight J, Bogden R, Li C, Shuang Y, Xu X, Pan S, Cheng S, Liu X, Ren Y, Wang J, Albiero A, Dal Pero F, Todesco S, Van Eck J, Buels R, Bombarely A, Gosselin J, Huang M, Leto J, Menda N, Strickler S, Mao L, Gao S, Tecle I, York T, Zheng Y, Vrebalov J, Lee J, Zhong S, Mueller L, Stiekema W, Ribeca P, Alioto T, Yang W, Huang S, Du Y, Zhang Z, Gao J, Guo Y, Wang X, Li Y, He J, Li C, Cheng Z, Zuo J, Ren J, Zhao J, Yan L, Jiang H, Wang B, Li H, Li Z, Fu F, Chen B, Han B, Feng Q, Fan D, Wang Y, Ling H, Xue Y, Ware D, McCombie W, Lippman Z, Chia J, Jiang K, Pasternak S, Gelley L, Kramer M, Anderson L, Chang S, Royer S, Shearer L, Stack S, Rose J, Xu Y, Eannetta N, Matas A, McQuinn R, Tanksley S, Camara F, Guigo R, Rombauts S, Fawcett J, Van de Peer Y, Zamir D, Liang C, Spannagl M, Gundlach H, Bruggmann R, Mayer K, Jia Z, Zhang J, Ye Z, Bishop G, Butcher S, Lopez-Cobollo R, Buchan D, Filippis I, Abbott J, Dixit R, Singh M, Singh A, Pal J, Pandit A, Singh P, Mahato A, Dogra V, Gaikwad K, Sharma T, Mohapatra T, Singh N, Causse M, Rothan C, Schiex T, Noirot C, Bellec A, Klopp C, Delalande C, Berges H, Mariette J, Frasse P, Vautrin S, Zouine M, Latche A, Rousseau C, Regad F, Pech J, Philippot M, Bouzayen M, Pericard P, Osorio S, del Carmen A, Monforte A, Granell A, Fernandez-Munoz R, Conte M, Lichtenstein G, Carrari F, De Bellis G, Fuligni F, Peano C, Grandillo S, Termolino P, Pietrella M, Fantini E, Falcone G, Fiore A, Giuliano G, Lopez L, Facella P, Perrotta G, Daddiego L, Bryan G, Orozco M, Pastor X, Torrents D, van Schriek K, Feron R, van Oeveren J, de Heer P, daPonte L, Jacobs-Oomen S, Cariaso M, Prins M, van Eijk M, Janssen A, van Haaren M, Jo S, Kim J, Kwon S, Kim S, Koo D, Lee S, Hur C, Clouser C, Rico A, Hallab A, Gebhardt C, Klee K, Jocker A, Warfsmann J, Gobel U, Kawamura S, Yano K, Sherman J, Fukuoka H, Negoro S, Bhutty S, Chowdhury P, Chattopadhyay D, Datema E, Smit S, Schijlen E, van de Belt J, van Haarst J, Peters S, van Staveren M, Henkens M, Mooyman P, Hesselink T, van Ham R, Jiang G, Droege M, Choi D, Kang B, Kim B, Park M, Yeom S, Lee Y, Choi Y, Li G, Gao J, Liu Y, Huang S, Fernandez-Pedrosa V, Collado C, Zuniga S, Wang G, Cade R, Dietrich R, Rogers J, Knapp S, Fei Z, White R, Thannhauser T, Giovannoni J, Botella M, Gilbert L, Gonzalez R, Goicoechea J, Yu Y, Kudrna D, Collura K, Wissotski M, Wing R, Schoof H, Meyers B, Gurazada A, Green P, Mathur S, Vyas S, Solanke A, Kumar R, Gupta V, Sharma A, Khurana P, Khurana J, Tyagi A, Dalmay T, Mohorianu I, Walts B, Chamala S, Barbazuk W, Li J, Guo H, Lee T, Wang Y, Zhang D, Paterson A, Wang X, Tang H, Barone A, Chiusano M, Ercolano M, D'Agostino N, Di Filippo M, Traini A, Sanseverino W, Frusciante L, Seymour G, Elharam M, Fu Y, Hua A, Kenton S, Lewis J, Lin S, Najar F, Lai H, Qin B, Qu C, Shi R, White D, White J, Xing Y, Yang K, Yi J, Yao Z, Zhou L, Roe B, Vezzi A,

D'Angelo M, Zimbello R, Schiavon R, Caniato E, Rigobello C, Campagna D, Vitulo N, Valle G, Nelson D, De Paoli E, Szinay D, de Jong H, Bai Y, Visser R, Lankhorst R, Beasley H, McLaren K, Nicholson C, Riddle C, Gianese G, Consortium TG, Consortium TG (2012) The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485:635–641

Scaglione D, Acquadro A, Portis E, Taylor C, Lanteri S, Knapp S (2009) Ontology and diversity of transcript-associated microsatellites mined from a globe artichoke EST database. BMC Genom 10:454

Scaglione D, Acquadro A, Portis E, Tirone M, Knapp S, Lanteri S (2012a) RAD tag sequencing as a source of SNP markers in *Cynara cardunculus* L. BMC Genom 13:3

Scaglione D, Lanteri S, Acquadro A, Lai Z, Knapp SJ, Rieseberg L, Portis E (2012b) Large-scale transcriptome characterization and mass discovery of SNPs in globe artichoke and its related taxa. Plant Biotechnol J 10:956–969

Scaglione D, Reyes-Chin-Wo S, Acquadro A, Froenicke L, Portis E, Beitel C, Tirone M, Mauro R, Lo Monaco A, Mauromicale G, Faccioli P, Cattivelli L, Rieseberg L, Michelmore R, Lanteri S (2016) The genome sequence of the outbreeding globe artichoke constructed de novo incorporating a phase-aware low-pass sequencing strategy of F_1 progeny. Sci Rep 6:19427

Staton S, Bakken B, Blackman B, Chapman M, Kane N, Tang S, Ungerer M, Knapp S, Rieseberg L, Burke J (2012) The sunflower (*Helianthus annuus* L.) genome reflects a recent history of biased accumulation of transposable elements. Plant J 72:142–153

van Os H, Stam P, Visser R, van Eck H (2005) SMOOTH: a statistical method for successful removal of genotyping errors from high-density genetic linkage data. Theor Appl Genet 112:187–194

Xie W, Feng Q, Yu H, Huang X, Zhao Q, Xing Y, Yu S, Han B, Zhang Q (2010) Parent-independent genotyping for constructing an ultrahigh-density linkage map based on population sequencing. Proc Natl Acad Sci USA 107:10578–10583

Xu X, Pan S, Cheng S, Zhang B, Mu D, Ni P, Zhang G, Yang S, Li R, Wang J, Orjeda G, Guzman F, Torres M, Lozano R, Ponce O, Martinez D, De la Cruz G, Chakrabarti S, Patil V, Skryabin K, Kuznetsov B, Ravin N, Kolganova T, Beletsky A, Mardanov A, Di Genova A, Bolser D, Martin D, Li G, Yang Y, Kuang H, Hu Q, Xiong X, Bishop G, Sagredo B, Mejia N, Zagorski W, Gromadka R, Gawor J, Szczesny P, Huang S, Zhang Z, Liang C, He J, Li Y, He Y, Xu J, Zhang Y, Xie B, Du Y, Qu D, Bonierbale M, Ghislain M, Herrera M, Giuliano G, Pietrella M, Perrotta G, Facella P, O'Brien K, Feingold S, Barreiro L, Massa G, Diambra L, Whitty B, Vaillancourt B, Lin H, Massa A, Geoffroy M, Lundback S, DellaPenna D, Buell C, Sharma S, Marshall D, Waugh R, Bryan G, Destefanis M, Nagy I, Milbourne D, Thomson S, Fiers M, Jacobs J, Nielsen K, Sonderkaer M, Iovene M, Torres G, Jiang J, Veilleux R, Bachem C, de Boer J, Borm T, Kloosterman B, van Eck H, Datema E, Hekkert B, Goverse A, van Ham R, Visser R, Consortium PGS (2011) Genome sequence and analysis of the tuber crop potato. Nature 475:189–U194

The Chloroplast Genome

10

Gabriella Sonnante

Abstract

In the last years, the availability of chloroplast genomes for plants and algae is increasing. These data can be useful both in the biotechnological area and in the evolutionary field, for the analysis of phylogenetic relationships among groups of organisms. Thanks to the advance of high-throughput sequencing, it is now possible to obtain the chloroplast genome sequence as a contaminant of whole-genome sequencing projects. Otherwise, the entire organelle genome can be recovered by means of long-range PCR amplification, using universal primers designed on highly conserved regions. The fragments obtained can be massively sequenced and the reads mapped on a reference genome. The chloroplast genomes of the *Cynara cardunculus* taxa, namely the globe artichoke, wild and cultivated cardoon and of other *Cynara* wild species are now available in the public databases. The structure of the *Cynara* chloroplast genome shows the classical quadripartite organization, with a long- and a short-single copy region and two inverted repeats. Within the globe artichoke, a marked variation is observed between the Romaneschi group and the other artichoke varietal types, due to some deletion events present in the Romaneschi. The highly variable regions found in the chloroplast genome can be used for barcode applications and phylogenetic relationships, not only in the genus *Cynara*, but also in the Asteraceae family.

10.1 Introduction

Chloroplast organelles are found exclusively in plant and algal cells and are specialized in performing photosynthesis, which converts solar energy to carbohydrates releasing oxygen (Jarvis and López-Juez 2013), thus allowing life on our earth. The only prokaryotic organisms evolving this metabolism were the cyanobacteria (Sánchez-Baracaldo et al. 2017), which were hypothesized, already at the beginning of the twentieth century, to have given origin to the chloroplasts about 1.5 billion years, when a cyanobacterial cell was englobed into a heterotrophic eukaryote (Mereschkowsky 1905). Besides carrying out photosynthesis, the chloroplast (cp) plays other crucial roles in plant physiology and development, being involved in the synthesis of many compounds such as amino acids and fatty acids (Pogson et al. 2015).

Chloroplasts possess multiple copies of a small circular genome ranging from 110 to 200 kb about, which is maternally inherited in angiosperms, and is highly conserved in structure and function (Olmstead and Palmer 1994;

G. Sonnante (✉)
Institute of Biosciences and Bioresources—National Research Council (CNR), Via G. Amendola, 165/a, 70126 Bari, Italy
e-mail: gabriella.sonnante@ibbr.cnr.it

E. Portis et al. (eds.), *The Globe Artichoke Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-030-20012-1_10

Daniell et al. 2016). The cp genome contains 30–50 RNA genes and about 100 protein-coding genes in land plants and green algae; the protein-coding genes can be involved in the expression and translational machinery of the chloroplast or in photosynthesis and other metabolic processes (De Las Riva et al. 2002). The typical cp genome is composed of two single copy regions (a long one and a short one), separated by two inverted repeat regions. Although conserved in its structure, the plastid genome may present striking features in some groups of plants. In legumes, for instance, the inverted repeat regions are missing (Palmer and Thompson 1982; Wu et al. 2011), while in grass cp genome, some important genes have been lost, and other plant species show large rearrangements (Guisinger et al. 2010). The plastid genome stable structure, but at the same time, the presence of variable sites makes chloroplast an informative and important resource for phylogenetic analysis at various taxonomic levels and for barcode applications (Curci et al. 2016; Bi et al. 2018). Knowledge of the chloroplast genome can be also useful for biotechnological purposes. Transplastomic plants, in fact, offer some advantages compared to classical nuclear transgenic plants. A high number of transgene copies can be engineered into the chloroplast genome, ensuring very high levels of expression; moreover, since cp genomes are generally maternally inherited, transgene escape via pollen is eliminated or reduced (Daniell 2007; Daniell et al. 2016).

Cynara L. is a Mediterranean genus composed of nine perennial species, which are characterized by spiny leaves and heads (Wiklund 1992; Rottenberg 2015). The interesting feature of these species is that they typically grow either in the eastern (*C. cornigera* and *C. syriaca*) or in the western (*C. humilis* and *C. baetica*) part of the Mediterranean basin and share molecular features according to their geographical distribution (Sonnante et al. 2008). Within the *Cynara* genus, the most plastic entity appears to be the wild *C. cardunculus*, distributed over the south Mediterranean area, from Greece to Spain and North Africa.

The globe artichoke, *C. cardunculus* var. *scolymus* (L.) Fiori, is a diploid outcrossing plant ($2n = 2x = 34$), which constitutes a single biological species together with the cultivated leafy cardoon (*C. cardunculus* var. *altilis* DC) and the wild ancestor of both crops, *C. cardunculus* var. *sylvestris* Lam. Based on molecular and phenotypic differences between wild *C. cardunculus* from the eastern (Greece/Italy) and the western (Iberian Peninsula) Mediterranean basin, it has been suggested that wild cardoon from Spain and Portugal might be a feral form of the cultivated cardoon (Gatto et al. 2013; Pavan et al. 2018). Within the globe artichoke, four morpho-agronomic groups have been recognized: Catanesi (CAT), Spinosi (SPI), Violetti (VIO) and Romaneschi (ROM), which can be distinguished according to flower head shape and colour, and for flowering time, CAT artichokes being early varieties, while most of the others are late varieties. Besides these recognized types, other artichoke varieties have different or intermediate flower head traits and therefore do not fall in any of the groups above (Porceddu et al. 1976; Sonnante et al. 2003).

In recent years, the availability of next-generation sequencing (NGS) has boosted the obtainment of whole plant genome sequences, due to the continuous improvement of technologies and the reduction of costs. Although during the preparation of DNA for whole-genome sequencing chloroplast DNA is generally removed, the high abundance of this molecule in the cells prevents its complete elimination. Therefore, when analysing sequencing outputs, reads from cp DNA can be found, and the cp genome can be reconstructed. The first complete cp genome sequence was obtained for *Nicotiana tabacum* using a strategy based on the sequencing of plasmid and cosmid libraries (Shinozaki et al. 1986), and since then hundreds of plant and algae cp sequences have been produced and deposited in the Organelle Genome Resources database of NCBI.

10.2 Chloroplast Genome Sequencing of the Globe Artichoke and Other *Cynara* Genotypes

Before the advent of next-generation sequencing (NGS) technologies, the complete cp genome sequencing was quite a laborious task and could be obtained either by screening bacterial artificial chromosome (BAC) libraries (e.g. Sato et al. 1999; Jansen et al. 2006), or by means of rolling circle amplification (e.g. Lee et al. 2006; Saski et al. 2007). High-throughput sequencing has now become the preferred approach for the obtainment of the cp genome, with Illumina being the major platform (Daniell et al. 2016). The first globe artichoke cp genome was sequenced by using a combined methodology: Illumina and BAC clone sequencing. Illumina short reads deriving from whole-genome sequencing of the globe artichoke, Brindisino variety, were filtered against the closest cp genome available by that time in NCBI database, *Lactuca sativa* (Curci et al. 2015). However, by this approach, the artichoke cp genome sequence was not entirely covered, and about 10% was still missing. Thanks to the availability of a BAC clone library from the same Brindisino genotype, a PCR screening of the library with primers designed on cp sequence allowed to identify a clone containing the whole cp genome. The BAC clone was thus isolated and sequenced (Curci et al. 2015). The assembly of NGS short reads can be performed using a reference-based strategy when a cp sequence is already available for a closely related organism, or a de novo assembly approach, when such a sequence is not available. Reference assemblies usually require less computational effort, but can provide a limited number of variants between the two cp genomes compared; in most cases, a de novo strategy is the preferred approach for a precise assembly (Dierckxsens et al. 2016). In the case of the globe artichoke, both assemblies were carried out, the former using the cp genome of *L. sativa* as a reference. The de novo reconstruction of the globe artichoke cp genome sequence provided almost the same results as the reference-based assembly, with a few incongruences, which could be solved by Sanger sequencing. Mostly, when a lack of correspondence between the two assembly methods was observed for artichoke cp genome, the de novo prediction was correct (Curci et al. 2015).

The cp genome sequence was also obtained for the wild progenitor of the globe artichoke, "Sylv Calabria" from Calabria, Italy. In this case, short reads from whole-genome sequencing were mapped on the Brindisino cp genome (Curci et al. 2016).

Long-range PCR of cp fragments, coupled with Illumina sequencing, can provide a cheap method to obtain complete cp sequence (Yang et al. 2014b) . In order to investigate cp sequence variation within *C. cardunculus* and in the genus *Cynara*, a long-range PCR approach was used on 19 genotypes representing the globe artichoke, cultivated and wild cardoon and other wild *Cynara* species. Primers were designed on conserved cp sequences shared among Brindisino, "Sylv Calabria" and *L. sativa*. Long PCR fragments were subsequently subjected to high-throughput sequencing using Illumina technology (Curci and Sonnante 2016; Curci et al. 2016).

10.3 Genome Organization of *Cynara* Chloroplast

The cp genome of the globe artichoke, as well as of the other *C. cardunculus* organisms and wild *Cynara* spp., possesses the typical quadripartite structure, comprising one long single copy (83,578 bp in Brindisino), one short single copy (18,641 bp in Brindisino) region and two inverted repeats (IR), each of 25,155 bp in Brindisino (Fig. 10.1).

Inverted repeats are generally conserved in terms of gene content and sequence, primarily due to the occurrence of rRNA genes (Li et al. 2013); however, expansion or reduction of the IR region and the SSC or LSC boundary regions are the main causes for size differences in angiosperm cp genomes (Chung et al. 2006; Yang et al. 2014a) . Due to these variations, the structure of these regions should be confirmed by classical

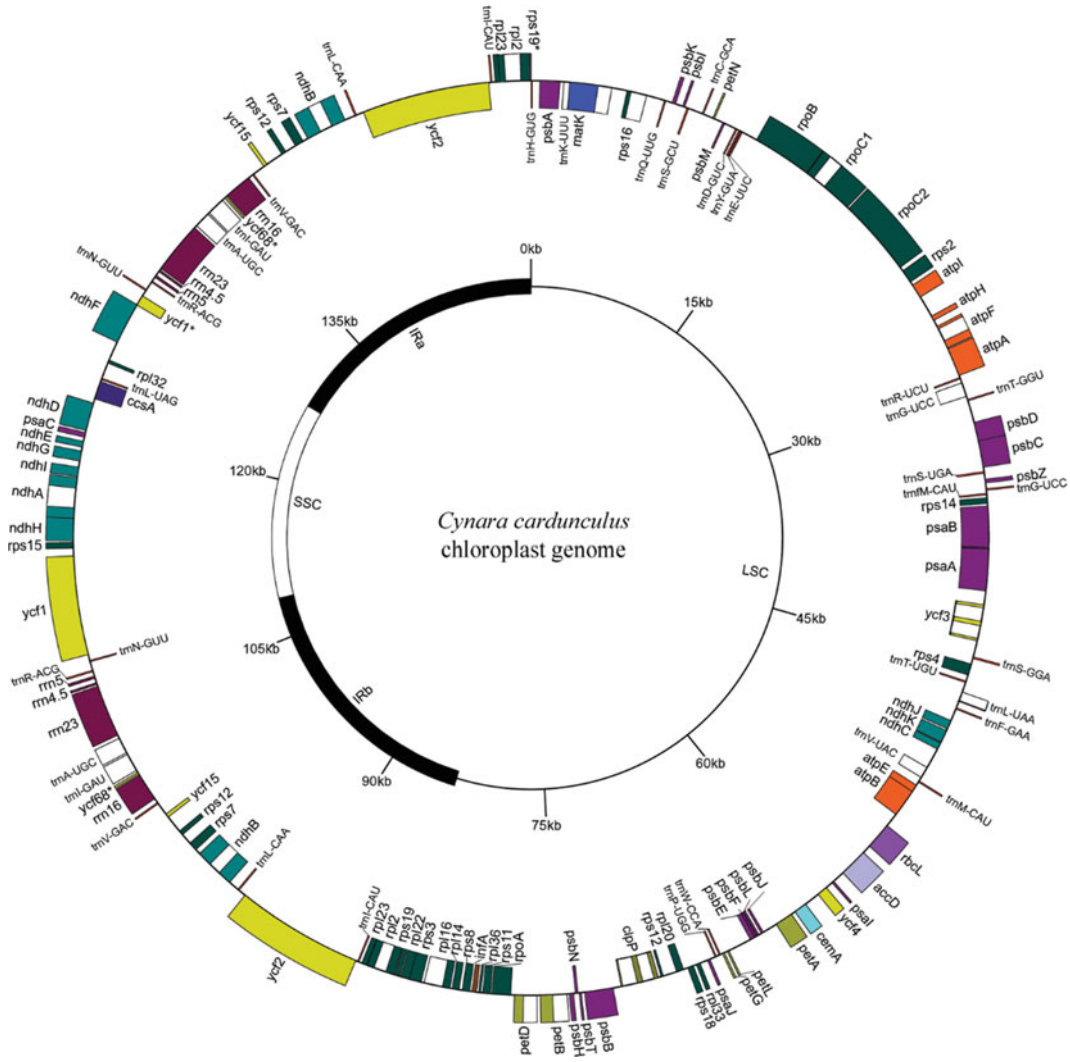


Fig. 10.1 Map of the *Cynara cardunculus* chloroplast genome. In the larger circle, genes are placed outside or inside, if they are transcribed clockwise or counterclockwise, respectively. Adapted from Curci et al. (2015, 2016)

Sanger sequence when obtaining cp genome sequences from NGS short reads. The use of high-throughput sequencing producing long reads, such as PacBio technology, may overcome this problem and facilitate the assembly especially in the junction regions (Ferrarini et al. 2013; Li et al. 2014). In *C. cardunculus* taxa and even in wild *Cynara* species, the borders between single copy regions and IRs are organized in the same manner, highlighting that no IR expansion has occurred within this genus (Curci and Sonnante 2016; Curci et al. 2016). On the other hand, when aligning these regions with the corresponding sequences of other Asteraceae cp genomes, various degrees of expansion have been noticed in *C. cardunculus* compared to the other species belonging to the same family (Curci et al. 2015).

The total length of the Brindisino cp genome is 152,529 bp, and a range between 152,462 and 152,585 bp has been observed within the *Cynara* genus (Table 10.1).

Variation in the whole cp genome length has been detected also within the globe artichoke

Table 10.1 List of the complete *Cynara* chloroplast genomes available in NCBI database and their size

Variety (morpho-group)	Origin	Acc. number/s	Length (bp)
Brindisino (CAT)	Italy	KM035764	152,529
Violetto di Toscana (VIO)	Italy	KP842713	152,529
Spinoso Sardo (SPI)	Italy	KP299293	152,585
Romano-Tondo Paestum/Camus Bretagne (ROM)	Italy/France	KP842709/10/11	152,462
Blanca de Tudela (OUT)	Spain	KP842708	152,529
Bianco Avorio/Verde Peralta (CC)	Italy/Spain	KP842704/05	152,510/11
Wild cardoon	Italy	KP842714/17/18/19	152,501/02/10
Wild cardoon	Greece	KP842715	152,509
Wild cardoon	Malta	KP842716	152,503
Wild cardoon	Tunisia	KP842721	152,528
Wild cardoon	Spain	KP842720	152,527
Wild *C. baetica*	Spain	KP842706	152,548
Wild *C. cornigera*	Egypt	KP842707	152,550
Wild *C. humilis*	Spain	KP299292	152,585

CAT Catanesi; *VIO* Violetti; *SPI* Spinosi; *ROM* Romaneschi; *OUT* no group; and *CC* cultivated cardoon

genotypes sequenced so far. The "Spinoso Sardo" shows a longer cp genome sequence compared to "Brindisino", while the three ROM-type artichokes possess a shorter sequence (152,462 bp). The reduced length of the ROM artichoke cp sequence is due to the presence of some deletion events, the more relevant one (30 bp) occurring in the *ycf1* gene (Curci et al. 2016). Therefore, it seems that the ROM-type artichokes are very similar, if not identical, to each other, at the level of cp genome, while a higher variation has been observed at the nuclear genomic level among the different varieties belonging to the ROM type (Gatto et al. 2013; Pagnotta et al. 2017; Pavan et al. 2018). *Ycf1* is one of the most variable genes in several cp genomes, showing SNPs and insertion/deletion variations higher than average (Wu et al. 2009; Doorduin et al. 2011). The two cultivated cardoon genotypes from Italy and Spain have only one nucleotide difference, while the cp genome length for the wild cardoon ranges between 152,501 and 152,510 for the material from Italy, Greece and Malta, and between 152,527 and 152,528 for the Spanish and the Tunisian genotypes, respectively (Table 10.1). Based on SSR and SNP analyses, the wild cardoons from the Iberian Peninsula have proven to resemble more the cultivated than the wild cardoon, and it has been hypothesized that they could represent feral forms (Gatto et al. 2013; Pavan et al. 2018).

The cp genomes of all the *Cynara* taxa and genotypes analysed so far are composed of 114 unique genes: 30 tRNA, 4 rRNA and 80 deduced protein-coding genes; moreover, the IR regions contain 17 duplicated genes, and thus, the total number of genes adds up to 131 (Curci et al. 2015; Curci and Sonnante 2016; Curci et al. 2016). About 42% of the whole sequence is composed of non-coding introns, intergenic spacers or pseudogenes. Although gene number and structure are well conserved in the cp genome of most plants (Tiller and Bock 2014), some genes may have moved to nuclear or mitochondrial genomes in a few species, providing useful evidence for evolutionary studies (Jansen et al. 2007).

In general, IR regions contain a lower number of ATs compared to the single copy regions, due to the reduced content of AT nucleotides in rRNA genes. It has been hypothesized that the IR regions play a role in the stabilization of the cp genome, and in fact, some legumes that have lost the IR regions are more exposed to

rearrangements than plants that have not (Palmer and Thompson 1982).

In the Brindisino cp genome, alternative start codons have been observed for some genes: for instance, in *nadhD* and *psbL* genes, ACG is used instead of ATG as a start codon; this type of conversion has also been noticed in tobacco (Sasaki et al. 2003). Another case for alternative start codon usage (GUG) in Brindisino is present in *rps19* gene (Curci et al. 2015). As in the majority of angiosperms (Qian et al. 2013; Cheng et al. 2017), also in *Cynara* leucine is the most abundant amino acid (about 10%), while cysteine is the most infrequent, with about 1% (Curci et al. 2015).

10.4 Repeat Composition and INDELs

Repeats are usually divided into two groups: microsatellites (SSRs) and other repeats. While SSRs are tandem repeats of one to six nucleotides, the other class consists of longer repeat units. Although less variable than nuclear microsatellite regions, cp SSRs can vary as well and can be therefore used to assess population diversity and relationships (George et al. 2015). The Brindisino cp genome contains 127 SSRs, most of which are mononucleotides, as usually observed in most plant species (Powell et al. 1995). Moreover, in higher plants, there is a tendency of mononucleotide repeats to occur in the non-coding region of the genome (George et al. 2015). The preferred location for artichoke cp SSRs is in the LSC and in spacer regions, followed by SSC and coding regions; fewer SSRs are found in the IRs, in intronic and finally in pseudogene regions (Curci et al. 2015). Thirty-seven primer pairs designed on SSR flanking regions shared within *Cynara* genus are available and have been used to validate these markers in *Cynara* genotypes (Curci et al. 2016).

Excluding SSRs, 23 direct and 16 palindromic repeats are present in the Brindisino cp genome. Most of these repeats are organized in tandem, and half of them are located in intergenic regions, while about one-third are in coding regions and 13% in intronic regions. The coding region of *ycf1* has the highest number of repeats in cds and possesses the longest repeat (45 bp) present in two copies and organized in tandem (Curci et al. 2015).

Insertion/deletion (INDEL) events different from microsatellites are useful markers for biodiversity and phylogenetic analyses (Yamane et al. 2006). Within the genus *Cynara*, a total of 34 of such cp INDELs are present, from 1 to 33 bp in length, with a prevalent distribution in non-coding (spacer) regions, as for other plant species (Chen et al. 2016; Daniell et al. 2016). Some of these INDELS are specific to either one of the wild *Cynara* species (*C. baetica*, *C. humilis*, *C. syriaca* or *C. cornigera*). *Cynara cardunculus* taxa share a deletion of 33 bp in the spacer of *ycf3-trnS* genes, which is not present in all the wild *Cynara* species listed above. Within *C. cardunculus*, a long deletion of 30 bp (in *ycf1*) and a deletion of 6 bp (in *psbC-trnS* spacer) are specific to the Romaneschi artichokes and are not found in the other artichoke varieties or in wild and cultivated cardoon (Curci et al. 2016). Moreover, Romaneschi genotypes are the only artichoke samples sharing a deletion of 24 bp (in *petD* intron) and an insertion of 5 bp (in *petA-psbJ* spacer) with most of the wild genotypes and with the wild cardoons (Curci et al. 2016).

10.5 Phylogenetic Analysis and DNA Barcode

In the past years, the sequences of a few cp regions were used to infer phylogenetic relationships among related species (Downie and Palmer 1992; Stefanović et al. 2009). In *Cynara*, the *psbA-trnH* spacer was amplified and sequenced in genotypes belonging to *C. cardunculus* and to other *Cynara* species, but no variation was observed in that portion of the cp genome (Sonnante et al. 2007). Recently, entire plastid genomes have proved to be a valuable resource for phylogenetic analyses and barcode applications, since the whole cp genome can be considered as a "super barcode" (Li et al. 2015).

In this study, the whole cp sequence available on NCBI database was retrieved for a subset of

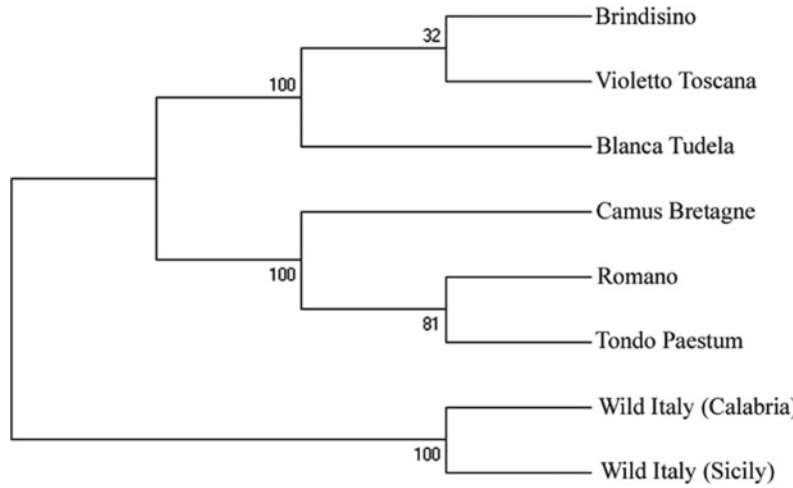


Fig. 10.2 Neighbour-joining phylogenetic tree based on whole cp genome sequence of six artichoke varieties and two wild cardoon samples. Numbers on branch nodes indicate bootstrap values

the genotypes used in Curci et al. (2016), in order to represent some of the main types of the artichoke morpho-types (Catanesi, Violetti, Romaneschi), together with Blanca de Tudela, and two wild cardoons. As for the Romaneschi type artichokes, all the three genotypes for which a cp genome sequence is available were considered. A phylogenetic analysis was carried out based on the Neighbour-Joining method (Saitou and Nei 1987), considering all the variable sites in the cp genome, with 1000 bootstrap replications. The result of this analysis (Fig. 10.2) evidences a clear separation, within the globe artichoke genotypes, of the Romaneschi-type artichokes from the others. This is mainly due to the long (30 bp) and other shorter INDELs characterizing particularly the three ROM genotypes. A similar result was obtained by Curci et al. (2016) when comparing a higher number of genotypes: in their phylogenetic tree, ROM genotypes occupied a more basal position, compared to the other artichoke types, probably because these artichokes, characterized by a big flower head, are the most ancient, being head size one of the early traits selected for domestication.

It has been observed that, for phylogenetic analysis within the genus *Cynara*, the use of all variable sites in the chloroplast genome provides a better resolution compared to the use of just parsimony-informative characters (Curci et al. 2016).

The sequence of six cp intronless genes was used to perform a phylogenetic analysis among 69 Asteraceae species, belonging to the main Asteraceae subfamilies, providing evidence on the position of *C. cardunculus* in the Asteraceae metatree (Curci et al. 2015).

Moreover, the comparison of fully sequenced cp genomes of nine Asteraceae species has allowed the identification of 19 highly polymorphic regions. In particular, eight of these cp DNA sequences with a parsimony-informative rate above 4% (*ycf1*, *rps16*, *ccsA*, *rbcL*, *ndhA*, *matK*, *clpP* and *accD*) have been proposed as possible barcode regions in the Asteraceae family (Curci et al. 2015).

In conclusion, cp whole-genome sequencing is an important tool for exploring molecular diversity, for phylogenetic and evolutionary investigations and for the development of barcode regions in *C. cardunculus* and in plants in general.

References

Bi Y, Zhang MF, Xue J, Dong R, Du YP, Zhang XH (2018) Chloroplast genomic resources for phylogeny and DNA barcoding: A case study on *Fritillaria*. Sci Rep 8:1184

Chen Z, Feng K, Grover CE, Li P, Liu F, Wang Y, Xu Q, Shang M, Zhou Z, Cai X, Wang X, Wendel JF, Wang K, Hua J (2016) Chloroplast DNA structural

variation, phylogeny, and age of divergence among diploid cotton species. PLoS ONE 11:e0157183

Cheng H, Li JF, Zhang H, Cai BH, Gao ZH, Qiao YS, Mi L (2017) The complete chloroplast genome sequence of strawberry (*Fragaria* × *ananassa* Duch.) and comparison with related species of Rosaceae. Peer J 5:e3919

Chung HJ, Jong DJ, Park HW (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

Curci PL, De Paola D, Danzi D, Vendramin GG, Sonnante G (2015) Complete chloroplast genome of the multifunctional crop globe artichoke and comparison with other Asteraceae. PLoS ONE 10:e0120589

Curci PL, De Paola D, Sonnante G (2016) Development of chloroplast genomic resources for *Cynara*. Mol Ecol Resour 16:562–573

Curci PL, Sonnante G (2016) The complete chloroplast genome of *Cynara humilis*. Mitochondrial DNA Part A 27:2345–2346

Daniell H (2007) Transgene containment by maternal inheritance: effective or elusive? Proc Natl Acad Sci USA 104:6879–6880

Daniell H, Lin CS, Yu M, Chang WJ (2016) Chloroplast genomes: diversity, evolution and applications in genetic engineering. Genome Biol 17:134

De Las Riva J, Lozano JJ, Ortiz AR (2002) Comparative analysis of chloroplast genomes: functional annotation, genome-based phylogeny, and deduced evolutionary patterns. Genome Res 12:567–583

Dierckxsens N, Mardulyn P, Smits G (2016) NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucl Acids Res 45:e18

Doorduin L, Gravendeel B, Lammers Y, Ariyurek Y, Chin-A-Woeng T, Vrieling K (2011) The complete chloroplast genome of 17 individuals of pest species *Jacobaea vulgaris*: SNPs, microsatellites and barcoding markers for population and phylogenetic studies. DNA Res 18:93–105

Downie SR, Palmer JD (1992) Use of chloroplast DNA rearrangements in reconstructing plant phylogeny. In: Soltis PS, Soltis DE, Doyle JE (eds) Molecular systematics of plants. Chapman & Hall, New York, USA, pp 14–35

Ferrarini M, Moretto M, Ward JA, Šurbanovski N, Stevanović V, Giongo L, Viola R, Cavalieri D, Velasco R, Cestaro A, Sargent DJ (2013) An evaluation of the PacBio RS platform for sequencing and de novo assembly of a chloroplast genome. BMC Genom 14:670

Gatto A, De Paola D, De Bagnoli F, Vendramin GG, Sonnante G (2013) Population structure of *Cynara cardunculus* complex and the origin of the conspecific crops artichoke and cardoon. Ann Bot 112:855–865

George B, Bhatt BS, Awasthi M, George B, Singh AK (2015) Comparative analysis of microsatellites in chloroplast genomes of lower and higher plants. Curr Genet 61:665–667

Guisinger MM, Chumley TW, Kuehl JV, Boore JL, Jansen RK (2010) Implications of the plastid genome sequence of *Typha* (Typhaceae, Poales) for understanding genome evolution in poaceae. J Mol Evol 70:149–166

Jansen RK, Cai Z, Raubeson LA, Daniell H, Leebens-Mack J, Müller KF, Guisinger-Bellian M, Haberle RC, Hansen AK, Chumley TW, Lee S-B, Peery R, McNeal JR, Kuehl JV, Boore JL (2007) Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. Proc Natl Acad Sci USA 104:19369–19374

Jansen RK, Kaittanis C, Saski C, Lee SB, Tomkins J, Alverson AJ, Daniell H (2006) Phylogenetic analyses of *Vitis* (Vitaceae) based on complete chloroplast genome sequences: effects of taxon sampling and phylogenetic methods on resolving relationships among rosids. BMC Evol Biol 6:32

Jarvis P, López-Juez E (2013) Biogenesis and homeostasis of chloroplasts and other plastids. Nat Rev Mol Cell Biol 14:787–802

Lee SB, Kaittanis C, Jansen RK, Hostetler JB, Tallon LJ, Town CD, Daniell H (2006) The complete chloroplast genome sequence of *Gossypium hirsutum*: organization and phylogenetic relationships to other angiosperms. BMC Genom 7:61

Li Q, Li Y, Song J, Xu H, Xu J, Zhu Y, Li X, Gao H, Dong L, Qian J, Sun C, Chen S (2014) High-accuracy de novo assembly and SNP detection of chloroplast genomes using a SMRT circular consensus sequencing strategy. New Phytol 204:1041–1049

Li R, Ma PF, Wen J, Yi TS (2013) Complete sequencing of five Araliaceae chloroplast genomes and the phylogenetic implications. PLoS ONE 8:e78568

Li X, Yang Y, Henry RJ, Rossetto M, Wang Y, Chen S (2015) Plant DNA barcoding: from gene to genome. Biol Rev 90:157–166

Mereschkowsky C (1905) Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. Biol Centralbl 25:593–604

Olmstead RG, Palmer JD (1994) Chloroplast DNA systematic: a review of methods and data analysis. Am J Bot 81:1205–1224

Pagnotta MA, Fernández JA, Sonnante G, Egea-Gilabert C (2017) Genetic diversity and accession structure in European *Cynara cardunculus* collections. PLoS ONE 12:e0178770

Palmer JD, Thompson WF (1982) Chloroplast DNA rearrangements are more frequent when a large inverted repeat sequence is lost. Cell 29:537–550

Pavan S, Curci PL, Zuluaga DL, Blanco E, Sonnante G (2018) Genotyping-by-sequencing highlights patterns of genetic structure and domestication in artichoke and cardoon. PLoS ONE 13(10):e0205988

Pogson BJ, Ganguly D, Albrecht-Borth V (2015) Insights into chloroplast biogenesis and development. Biochim Biophys Acta 847:1017–1024

Porceddu E, Dellacecca V, Bianco VV (1976) Classificazione numerica di cultivar di carciofo. In: Proceedings of the II International Congress on Artichoke, Bari. Minerva Medica (in Italian), Torino, pp 1105–1119

Powell W, Morgante M, Andre C, McNicol JW, Machray GC, Doyle JJ, Tingey SV, Rafalski JA (1995) Hypervariable microsatellites provide a general source of polymorphic DNA markers for the chloroplast genome. Curr Biol 5:1023–1029

Qian J, Song J, Gao H, Zhu Y, Xu J, Pang X, Yao H, Sun C, Li X, Li C, Liu J, Xu H, Chen S (2013) The complete chloroplast genome sequence of the medicinal plant *Salvia miltiorrhiza*. PLoS ONE 8:e57607

Rottenberg A (2015) The wild gene pool of globe artichoke. Israel J Plant Sci 62:1–2

Saitou N, Nei M (1987) The neighbor-joining method—a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Sánchez-Baracaldo P, Raven JA, Pisani D, Knoll AH (2017) Early photosynthetic eukaryotes inhabited salinity habitats. Proc Natl Acad Sci USA 114:E7737–E7745

Sasaki T, Yukawa Y, Miyamoto T, Obokata J, Sugiura M (2003) Identification of RNA editing sites in chloroplast transcripts from the maternal and paternal progenitors of tobacco (*Nicotiana tabacum*): comparative analysis shows the involvement of distinct trans-factors for ndhB editing. Mol Biol Evol 20:1028–1035

Saski C, Lee SB, Fjellheim S, Guda C, Jansen RK, Luo H, Tomkins J, Rognli OA, Daniell H, Clarke JL (2007) Complete chloroplast genome sequences of *Hordeum vulgare*, *Sorghum bicolor* and *Agrostis stolonifera*, and comparative analyses with other grass genomes. Theor Appl Genet 115:571–590

Sato S, Nakamura Y, Kaneko T, Asamizu E, Tabata S (1999) Complete structure of the chloroplast genome of *Arabidopsis thaliana*. DNA Res 6:283–290

Sonnante G, Carluccio AV, De Paolis A, Pignone D (2008) Identification of artichoke SSR markers: molecular variation and patterns of diversity in genetically cohesive taxa and wild allies. Genet Resour Crop Evol 55:1029–1046

Sonnante G, Carluccio AV, Vilatersana R, Pignone D (2007) On the origin of artichoke and cardoon from *Cynara* gene pool as revealed by rDNA sequence variation. Genet Resour Crop Evol 54:483–495

Sonnante G, De Paolis A, Pignone D (2003) Relationships among artichoke cultivars and some related wild taxa based on AFLP markers. Plant Genet Resour: Characteriz Utiliz 1:125–133

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

Stefanović S, Pfeil BE, Palmer JD, Doyle JJ (2009) Relationships among phaseoloid legumes based on sequences from eight chloroplast regions. Syst Bot 34:115–128

Tiller N, Bock R (2014) The translational apparatus of plastids and its role in plant development. Mol Plant 7:1105–1120

Wiklund A (1992) The genus *Cynara* L. (Asteraceae-Cardueae). Bot J Linn Soc 109:75–123

Wu FH, Kan DP, Lee SB, Daniell H, Lee YW, Lin CC, Lin NS, Lin CS (2009) Complete nucleotide sequence of *Dendrocalamus latiflorus* and *Bambusa oldhamii* chloroplast genomes. Tree Physiol 29:847–856

Wu CS, Wang YN, Hsu CY, Lin CP, Chaw SM (2011) Loss of different inverted repeat copies from the chloroplast genomes of Pinaceae and cupressophytes and influence of heterotachy on the evaluation of gymnosperm phylogeny. Genome Biol Evol 3:1284–1295

Yamane K, Yano K, Kawahara T (2006) Pattern and rate of indel evolution from whole chloroplast intergenic regions in sugarcane, maize and rice. DNA Res 13:197–204

Yang Y, Dang YY, Li Q, Lu JJ, Li XW, Wang YT (2014a) Complete chloroplast genome sequence of poisonous and medicinal plant *Datura stramonium*: organizations and implications for genetic engineering. PLoS ONE 9:e110656

Yang JB, Li DZ, Li HT (2014b) Highly effective sequencing whole chloroplast genomes of angiosperms by nine novel universal primer pairs. Molec Ecol Resour 14:1024–1031

miRNome

11

Lorenzo Barchi and Alberto Acquadro

Abstract

MicroRNAs in plants are endogenous small non-coding RNAs involved in many biological processes related to growth, development and stress response, through post-transcriptional silencing of key target genes. Only a few studies have described the miRNome of globe artichoke. In a very first attempt, 122 new miRNAs were mined through miRNA transcriptomic analysis in leaves and roots of young plants in a NaCl treatment experiment. Computational approach based on known miRNAs sequences was adopted to predict candidate miRNAs in the reference genome sequence, while the recent resequencing of four globe artichoke and one cultivated cardoon genotypes has allowed to expand the prediction of high confidence miRNA-coding sequences. Although many plant microRNAs are evolutionary conserved, polymorphisms exist in many loci affecting miRNA biogenesis, function and target recognition, with a possible effect on phenotype. The existing variation observed in the globe artichoke germplasm could explain the variable number of predicted miRNAs and related target genes, across the tested genotypes and could be exploited in the future as a source of genetic variation for future crop improvement.

11.1 Introduction to miRNAs

MicroRNAs (miRNAs) are members of the endogenous small non-coding RNA group (with a typical length of 20–24 nt) which are involved in post-transcriptional repression of gene expression both in animals and plants (Axtell 2013). They are coded by MIR genes (usually found in intergenic regions but also within genes, Colaiacovo et al. 2010), whose transcription, mainly by RNA polymerase II, produces the primary miRNA (pri-miRNA). Pri-miRNA is then capped at 5′ end and poly-adenylated at the 3′ end and introns are spliced out. Pri-miRNA is then further processed by the dicing complex (which includes as core components the RNase III enzyme DICER-LIKE (DCL), the HYPONASTIC LEAVES1 protein (HYL1) and the SERRATE protein (SE), into a stem loop-containing intermediate (the precursor miRNA or pre-miRNA) and finally into the miRNA/miRNA* duplex which, after stabilization through 2′-*O*-methylation of the 3′ termini by the methyltransferase HUA ENHANCER 1 (HEN1), moves to the cytoplasm thanks to the protein HASTY. The duplex is then unwound allowing only one strand (the one whose 5′-end is less stably, i.e. the guide strand) to guide

L. Barchi · A. Acquadro (✉)
DISAFA Plant Genetics and Breeding, University of Torino, Largo P. Braccini 2, 10095 Grugliasco, Torino, Italy
e-mail: alberto.acquadro@unito.it

E. Portis et al. (eds.), *The Globe Artichoke Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-030-20012-1_11

mature microRNA for a sequence-based recognition of the specific target mRNA, while the other strand (passenger strand) is generally degraded, even if recent studies found that some minor miRNAs persist and have a functional role in gene regulation (Pashkovskiy and Ryazansky 2013). Methylation of the 3′ termini by HEN1 is necessary to protect the unwound mature miRNA from degradation by exonucleases, such as small RNA-degrading nuclease (SDN) protein. The protein responsible for loading microRNA and taking it into the RNA-induced silencing complex (RISC complex) is the ARGONAUTE (AGO). At this point, there are two major ways for miRNA action: one is based on the cleavage of the recognized mRNA, while the other one leads to its translational inhibition, resulting in the post-transcriptional silencing of the target gene expression. Generally, plants miRNAs and their target mRNAs have nearly perfect complementarity (Yu et al. 2017) and was thought that transcript cleavage was the predominant mode of action of plant miRNAs (Voinnet 2009; Rogers and Chen 2013). However, a high degree of sequence complementarity may result in miRNA-mediated translation inhibition (Yang et al. 2012; Li et al. 2013), including some key processes such as developmental switches (Budak and Akpinar 2015). MicroRNA targets in plants are involved in many biological processes related to growth, development and stress response. Many transcription factors have been identified as targets of microRNAs (Samad et al. 2017): this finding highlights the role of these tiny regulators as central hubs in complex regulatory networks. Consequently, a growing body of interest has been recently devoted to the exploration of possible applications of miRNAs and their target genes to crop improvement (Djami-Tchatchou et al. 2017). Besides, conserved miRNAs (Colaiacovo and Faccioli 2013) tend to have also conserved targets suggesting the possibility of applying them to the agronomical improvement of several crops (Tang and Chu 2017). Another very interesting group of microRNA targets includes the major class of R genes which are involved in plant immune response. Besides protein-coding genes, long non-coding RNAs (lncRNAs) have been identified that act as a target mimic for microRNAs leading to the inhibition of physical interaction between a miRNA and its target (Liu et al. 2015).

Plant microRNAs are part of gene families with variable size which in some cases can be very large and can be found as independent loci or in clusters coding for homologous miRNAs (such that co-transcription has a consequent dosage effect) or non-homologous clusters whose members target different proteins belonging to a specific family, thus leading to their co-regulation (Budak and Akpinar 2015). As above reported, miRNAs are classified into two groups: conserved and non-conserved (Baldrich et al. 2018). Conventionally, conserved miRNAs are those present in most or all of the land plant lineages analysed, are abundantly expressed, show low sequence variation across diverse plant families and target mainly genes involved in plant development and stress response. On the contrary, non-conserved, young miRNAs are lineage-specific, and they are expressed at low levels and are often involved in the control of species-restricted processes. Recently Axtell and Meyers (2018) defined a conserved miRNA land plant miRNA family as a family "which has been annotated in miRBase21 in at least two out of the eight major taxonomic divisions" with at least one of the annotations designed as "high confidence" in miRBase21.

11.2 miRNAs Identification in *Cynara cardunculus* Taxa

The first description of the globe artichoke miRNome was reported by De Paola et al. (2012). The study was based on small RNA (sRNA) sequencing in leaf and root tissues of young globe artichoke plants, untreated or treated with NaCl. The work yielded 98 known *C. cardunculus* miRNAs (referred as cca-miRNA) identified by comparing sRNA sequences to miRBase Sequence Database (release 17.0, Kozomara and Griffiths-Jones 2011) and assigned to 25 families, five of them with only one member. About half of them, 46 out of 98,

were shared with other plant species, and some were previously reported to be involved in the plant response to salt stress. In Arabidopsis, miR159 and 319 were up-regulated following saline treatment, as observed in both leaves and root of *C. cardunculus*. In other cases, as for miR156 and 167, the regulation after salt stress was opposite in Arabidopsis compared to globe artichoke (Covarriubas and Reyes 2010). Furthermore, De Paola et al. (2012) have also reported the identification of 122 novel artichoke miRNAs.

The sequencing of the globe artichoke genome (accession 2C, Scaglione et al. 2016) has allowed the genome-wide identification of candidate miRNAs by computational analysis using a dedicated software (MIReNA, Mathelier and Carbone 2010) and known miRNAs sequences (miRBase release 20, Kozomara and Griffiths-Jones 2011) from 11 plants and algae. The predicted miRNAs were searched for transcriptional evidences on pre-existing RNA-seq data (Scaglione et al. 2012) re-analysed using HTSeq-count (Anders et al. 2015). The putative miRNA positions in the reference genome were considered as genomic features, and subsequently used for reads alignment, position with more than 10 mapped reads/miRNA loci was considered as expressed genomic features. Finally, when the putative miRNA-coding sequences were subjected to BLASTn analysis against *C. cardunculus* ESTs (NCBI dbEST) and *C. cardunculus* transcript sequences (Scaglione et al. 2012), a total of 73 different miRNAs have been identified. Some known and conserved miRNAs (i.e. miR162 and miR482) were not retrieved by this analysis, probably due to the absence of these genomic loci in the assembly. Notwithstanding, the computational pipeline allowed to identify all but one of the 19 miRNAs previously verified by qPCR by De Paola et al. (2012). Fifty-eight miRNAs were represented by less than 10 loci, nine between 10 and 99 loci and four ranged between 100 and 999; finally,

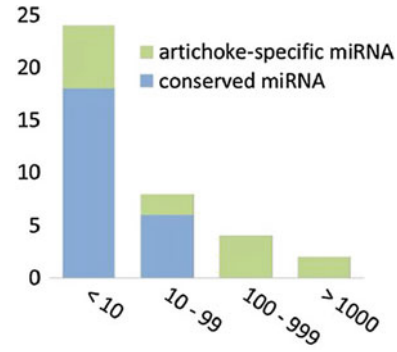


Fig. 11.1 Number of loci identified per miRNA grouped in four frequency classes: (1) <10 loci; (2) 10–99 loci; (3) 100–999 loci; (4) >1,000 loci (from Scaglione et al. 2016, used under CC BY 4.0)

two were found to be present in the 2C genome with more than 1000 loci (Fig. 11.1). All the miRNAs with more than 100 loci were *C. cardunculus*-specific (considering miRNAs present in miRBase release 21), although some low-copy species-specific miRNAs were also detected.

The recent resequencing of four globe artichoke genotypes (Violetto di Sicilia, Violetto di Toscana, Romanesco C3 and Spinoso di Palermo) and of one cultivated cardoon (Altilis 41) has expanded the prediction of high confidence miRNA-coding sequences (Acquadro et al. 2017). MIReNA software (Mathelier and Carbone 2010) was run on each pseudomolecule of the resequenced genotypes, together with the reference artichoke genome (accession "2C"), and known high confidence miRNAs belonging to 13 species (plants and algae) from the miRBase database (release 21, Kozomara and Griffiths-Jones 2011) were used to carry out a homology search. The predicted non-redundant miRNAs ranged between 51 (mapped to 74 regions of Violetto di Toscana genome) and 143 (mapped to 241 regions of the reference 2C genome). The identified miRNA belonged from 32 (Spinoso di Palermo) to 45 (reference 2C) miRNA families (Fig. 11.2).

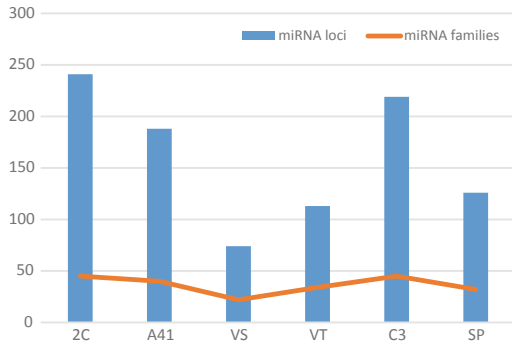


Fig. 11.2 Relative number of predicted miRNAs loci (data from Acquadro et al. 2017, used under CC BY 4.0)

11.3 miRNAs and Their Target Identification in *Cynara cardunculus* Taxa

As reported above, three main studies were dedicated to the analysis of the artichoke miRNome, and each of them contains some attempts to identify miRNA targets. The first report (De Paola et al. 2012) aligned conserved and novel miRNAs against artichoke ESTs in NCBI and Illumina genomic sequences (Sonnante et al. 2010). After removing redundant sequences and retaining only those having a match in the Arabidopsis proteome, 43 sequences were retained, 20 of which matching conserved and 23 novel artichoke miRNAs. Several artichoke miRNAs were predicted to target known transcription factors related to flowering time, plant development and morphology, for example, NAC and AP2 domain-containing protein, auxin response factor involved in developmental growth. Some putative targets were annotated as laccase gene family, a well-known component of the salt response (Cai et al. 2006) and of pathways leading to lignin (Gavnholt et al. 2002), proanthocyanidin and tannin biosynthesis (Pourcel et al. 2005). Among the predicted novel artichoke miRNAs, some target genes were found to be involved in several stress responses (NaCl, cadmium, fungi infection, oxidative conditions).

After the sequencing of the artichoke reference genome (Scaglione et al. 2016), the miRNAs and the predicted transcripts were analysed using the psRNATarget tool (http://plantgrn.noble.org/psRNATarget) to search for putative miRNA:mRNA interactions. Putative miRNA targets GO were then compared against the whole genome GO annotation data set of the reference artichoke genome by mean of the AgriGo suite (Du et al. 2010) at p values $<e^{-5}$ and with a false discovery rate <0.01. Finally, DAVID tool (https://david.ncifcrf.gov/) was used for clustering the miRNA target transcripts according to miRNA families. A total of 256 transcripts were found to be targeted by 57 miRNAs at 297 different sites. As examples, the conserved miR156, miR395 and miR396 have as putative targets 23, 26 and 39 genes, respectively, while, the non-conserved miRNAs, miR6108, miR6111 and miR6114 were found to putatively interact with 28, 24 and 15 genes, respectively. The GO enrichment analysis of the miRNA target transcripts identified several categories showing an enrichment compared to the whole GO annotation (Fig. 11.3). For processes (P), enrichments were observed for the regulation of transcription (GO:0006355, GO:0045449), protein localization in the mitochondrion (GO:0070585) and protein import/transport (GO:0017038, GO:0015031). With respect to functions (F), enrichments were observed in nucleic acid binding (GO:0003676), protein transporter activity (GO:0008565) and ATP binding (GO:0005524), while for components (C) enrichments were present for: nucleus (GO:0005634) and mitochondrial membrane outer membrane translocase complex (GO:0005742).

As expected several conserved miRNAs were found to target known transcription factors related to plant development, as well as to flowering time and plant morphology. Examples include the miR156, which was found to target SQUAMOSA promoter-binding-like proteins, miR160 targeting the auxin response factor (ARF), the miR164 and NAC-like proteins, miR171 with GRAS-like proteins and miR172 targeting AP2-like proteins. On the other side, putative artichoke-specific miRNAs were found to preferentially target genes coding for enzymes and proteins involved in adaptation mechanisms. Among these, miR6108 was found to target

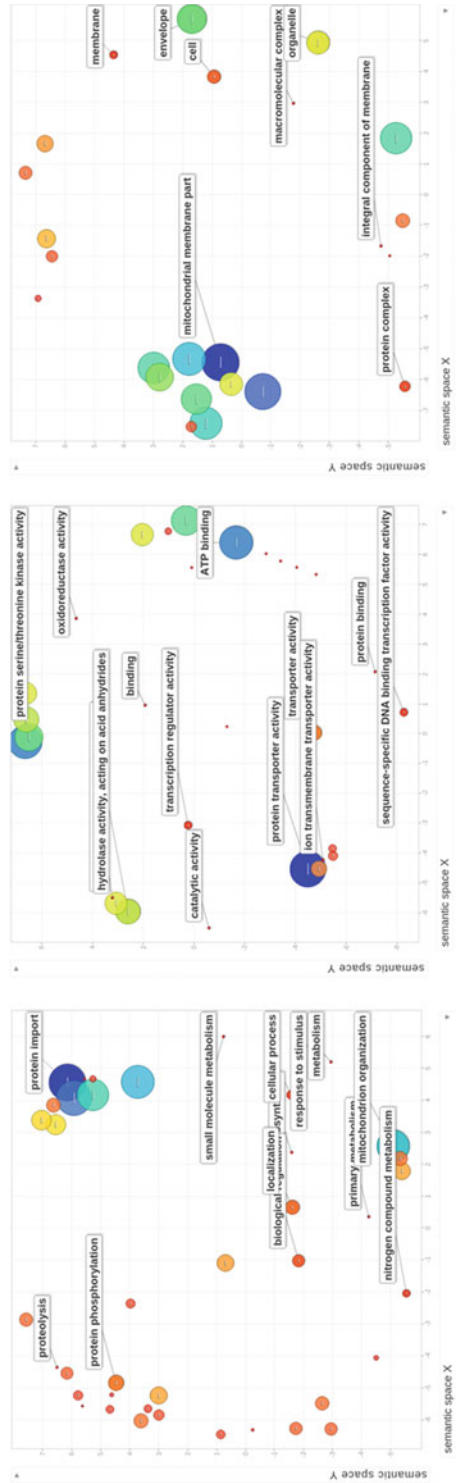


Fig. 11.3 Representation of enriched GO categories in the miRNA target transcriptome (from Scaglione et al. 2016, used under CC BY 4.0). Left: biological process; centre: molecular function; right: cellular component. *X* axis is expressed in log10 *p*-value, *Y* axis is expressed in log10 *p*-value

metal ion-binding proteins involved in ion transport in plasma membrane, miR6110 on nucleotide binding, miR6111 was predicted to act on serine/threonine protein kinase, miR6114 target metal ion-binding proteins and ATP binding proteins, while miR6118 putatively control ATP binding proteins and leucine-rich repeat (LRR), as well as protein with serine/threonine kinase activity.

More recently, Acquadro et al. (2017) identified the miRNA targets with Tapir software (Bonnet et al. 2010) and corresponding GO enrichments using the information generated through the resequencing of the five artichoke genotypes. Overall, between 307 (Violetto di Sicilia) and 1,167 (accession 2C), putative miRNA:mRNA duplexes were identified by the Tapir analyses. Almost 90% of genes encoding predicted target transcripts were found to have an InterPro annotation using InterProScan pipeline (Jones et al. 2014). The miRNA families involved in miRNA:mRNA interactions varied from 20 in Violetto di Sicilia to 45 in the reference genome and in the cultivated cardoon genotype Altilis 41. With the exception of Spinoso di Palermo, miRNA172 was found to be the top-ranked family involved miRNA:mRNA duplex formation in the six artichoke genotypes considered with a number of duplex between 83 (Violetto di Sicilia) and 283 (cultivated cardoon Romanesco C3). Besides miRNA172, the other main families involved in miRNA:mRNA duplex formation showed a genotype-specific behaviour. AGRIGO (Du et al. 2010) cross-comparison of SEA (SEACOMPARE) was applied to spot common and different enrichment GO terms between the genotypes. Significant GO enrichments in the miRNA target genes were found for some GO terms in all genotypes except in Violetto di Toscana (Fig. 11.4).

Artichoke miRNAs were found to putatively target transcription factors related to plant development, flowering time and plant morphology as described in other species (Li and Zhang 2016; Acquadro et al. 2017). For instance, SQUAMOSA promoter-binding-like proteins (miR156), auxin response factor (miR160), NAC-like protein-coding genes (miR164), GRAS-like-protein-coding genes (miR171) and AP2-like-protein-coding genes (miRNA172). Furthermore, several NBS-LRR genes involved in the resistance to pathogens (Zhang et al. 2016) were found among the transcript putatively targeted by globe artichoke miRNAs.

11.4 miRNA Relationship with Repeated Sequences

Several evidences have indicated that variants altering gene expression, rather than variants altering protein sequences, act as the primary cause of natural variation in complex traits (Chen et al. 2016). Although, as above reported, many plant miRNAs are evolutionary conserved, polymorphism exists in the corresponding loci that can affect miRNA biogenesis, function and target recognition with a possible phenotypic variation outcome (Chen et al. 2016). Natural miRNA variants can thus be valuable sources for genome-wide association studies (GWAS) and crop improvement (Tang and Chu 2017).

Simple sequence repeats (SSRs) are ubiquitous, highly polymorphic, usually co-dominant and readily assayed by conventional PCR so very suitable for plant breeding applications. The presence of SSRs within miRNA-coding loci is of great significance for gene expression, thus taking it beyond their conventional use as simple genetic markers. Several putative miR-SSR have been found throughout the genome of globe artichoke, among them the most frequent are di-nucleotides (mainly TA/AT), followed by tri- and tetra-nucleotides, while mono-, penta- and hexa-nucleotides are rarest (Scaglione et al. 2016). Recently a web-based database (CyMSatDB) was developed to provide a searchable interface for artichoke SSR data (Portis et al. 2016), which can facilitate the future identification of miR-SSR markers as well as suggesting the design of the best primers to be used for the corresponding analysis.

miR-SNPs (single-nucleotide polymorphisms) are another interesting class of markers with putative application to crop improvement. The resequencing of six globe artichoke genotypes

Go information			CM				
GOTerm	Onto	Description	2C	A41	VS	C3	SP
GO:0016070	P	RNA metabolic process					
GO:0080090	P	regulation of primary metabolic process					
GO:0019222	P	regulation of metabolic process					
GO:0009698	P	phenylpropanoid metabolic process					
GO:0060255	P	regulation of macromolecule metabolic process					
GO:0019439	P	aromatic compound catabolic process					
GO:0009889	P	regulation of biosynthetic process					
GO:0006355	P	regulation of transcription, DNA-dependent					
GO:0006350	P	transcription					
GO:0006351	P	transcription, DNA-dependent					
GO:0032774	P	RNA biosynthetic process					
GO:0009808	P	lignin metabolic process					
GO:0046274	P	lignin catabolic process					
GO:0046271	P	phenylpropanoid catabolic process					
GO:0051252	P	regulation of RNA metabolic process					
GO:0031326	P	regulation of cellular biosynthetic process					
GO:0045449	P	regulation of transcription					
GO:0042219	P	cellular amino acid derivative catabolic process					
GO:0010467	P	gene expression					
GO:0010556	P	regulation of macromolecule biosynthetic process					
GO:0010468	P	regulation of gene expression					
GO:0019219	P	regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process					
GO:0051171	P	regulation of nitrogen compound metabolic process					
GO:0006139	P	nucleobase, nucleoside, nucleotide and nucleic acid metabolic process					
GO:0031323	P	regulation of cellular metabolic process					
GO:0007275	P	multicellular organismal development					
GO:0003676	F	nucleic acid binding					
GO:0043227	C	membrane-bounded organelle					
GO:0043231	C	intracellular membrane-bounded organelle					
GO:0005634	C	nucleus					
GO:0019748	P	secondary metabolic process					
GO:0006725	P	cellular aromatic compound metabolic process					
GO:0005507	F	copper ion binding					
GO:0005576	C	extracellular region					
GO:0048046	C	apoplast					
GO:0032502	P	developmental process					
GO:0044249	P	cellular biosynthetic process					
GO:0003677	F	DNA binding					
GO:0009719	P	response to endogenous stimulus					
GO:0010033	P	response to organic substance					
GO:0009725	P	response to hormone stimulus					

Fig. 11.4 SEACOMPARE from AGRIGO of the GO terms enriched for genotypic exclusive genes data (from Acquadro et al. 2017, used under CC BY 4.0)

has led to the identification of more than 23 million SNPs/indels among them (Acquadro et al. 2017). Interestingly, the number of predicted miRNAs, as well as the number of predicted target genes, varied across the different tested genotypes and this variable number of identified miRNAs might arise from SNPs present in some loci that hamper their identification in some genotypes.

SSRs and indels have been identified among chloroplast genomes of seven globe artichokes, two cultivated cardoons, eight wild artichokes and three other wild *Cynara species* (*C. baetica*, *C. cornigera* and *C. syriaca*) by Curci et al. (2016). This resource has become very interesting also for miRNAs studies in *Cynara* genus since it has been hypothesized the existence of organellar miRNA-like molecules residing in

mitochondria (mitomiRs) and chloroplast (cipomiRs), opening the possibility that some small non-coding RNAs could be encoded and processed by the genomes of these organelles (Budak et al. 2015; Budak and Akpinar 2015).

11.5 Conclusions

Many conserved miRNAs detected in the globe artichoke genome are predicted to target well-known biological processes. The number of identified miRNAs varied across the different genotypes (Acquadro et al. 2017). These results can be due to incomplete sequence and/or from SNPs present in some miRNA loci that might hamper their identification in some genotypes. The analysis of GO term enrichments of miRNA target transcripts revealed that GO terms related to binding and transcription were common supporting the involvement of miRNAs in transcription factors regulation. Furthermore, artichoke miRNAs have been shown to be involved in the response of the species to stress response, as shown for salinity (De Paola et al. 2012), suggesting their possible role in environmental adaptation (Qin et al. 2014). Altogether, the miRNA features, as well as the genetic variation, highlighted represent a source of information which could assist the future implementation of crop improvement programs.

Acknowledgements We thank Luigi Cattivelli (Consiglio per la ricerca in agricoltura e l'analisi dell'economia Agraria, Centro di Ricerca Genomica e Bioinformatica) for critically revising the chapter and for its helpful suggestions.

References

Acquadro A, Barchi L, Portis E, Mangino G, Valentino D, Mauromicale G, Lanteri S (2017) Genome reconstruction in *Cynara cardunculus* taxa gains access to chromosome-scale DNA variation. Sci Rep 7:5617. https://doi.org/10.1038/s41598-017-05085-7

Anders S, Pyl PT, Huber W (2015) HTSeq–a Python framework to work with high-throughput sequencing data. Bioinformatics 31:166–169. https://doi.org/10.1093/bioinformatics/btu638

Axtell MJ (2013) Classification and comparison of small RNAs from plants. Annu Rev Plant Biol 64:137–159. https://doi.org/10.1146/annurev-arplant-050312-120043

Axtell MJ, Meyers BC (2018) Revisiting criteria for plant microrna annotation in the era of big data. Plant Cell 30:272–284. https://doi.org/10.1105/tpc.17.00851

Baldrich P, Beric A, Meyers BC (2018) Despacito: the slow evolutionary changes in plant microRNAs. Curr Opin Plant Biol 42:16–22. https://doi.org/10.1016/J.PBI.2018.01.007

Bonnet E, He Y, Billiau K, Van de Peer Y (2010) TAPIR, a web server for the prediction of plant microRNA targets, including target mimics. Bioinformatics 26:1566–1568

Budak H, Akpinar BA (2015) Plant miRNAs: biogenesis, organization and origins. Funct Integr Genomics 15:523–531. https://doi.org/10.1007/s10142-015-0451-2

Budak H, Bulut R, Kantar M, Alptekin B (2015) MicroRNA nomenclature and the need for a revised naming prescription. Brief Funct Genomics 15:elv026. https://doi.org/10.1093/bfgp/elv026

Cai X, Davis EJ, Ballif J, Liang M, Bushman E, Haroldsen V, Torabinejad J, Wu Y (2006) Mutant identification and characterization of the laccase gene family in Arabidopsis. J Exp Bot 57:2563–2569. https://doi.org/10.1093/jxb/erl022

Chen J, Xie J, Chen B, Quan M, Li Y, Li B, Zhang D (2016) Genetic variations and miRNA-target interactions contribute to natural phenotypic variations in populus. New Phytol 212:150–160. https://doi.org/10.1111/nph.14040

Colaiacovo M, Faccioli P (2013) Plant microRNAs. In: From plant genomics to plant biotechnol, pp 15–30. https://doi.org/10.1533/9781908818478.15

Colaiacovo M, Subacchi A, Bagnaresi P, Lamontanara A, Cattivelli L, Faccioli P (2010) A computational-based update on microRNAs and their targets in barley (*Hordeum vulgare* L.). BMC Genomics 11:595. https://doi.org/10.1186/1471-2164-11-595

Covarriubas AA, Reyes JL (2010) Post-transcriptional gene regulation of salinity and drought responses by plant microRNAs. Plant, Cell Environ 33:481–489. https://doi.org/10.1111/j.1365-3040.2009.02048.x

Curci PL, De Paola D, Sonnante G (2016) Development of chloroplast genomic resources for *Cynara*. Mol Ecol Resour 16:562–573. https://doi.org/10.1111/1755-0998.12457

De Paola D, Cattonaro F, Pignone D, Sonnante G (2012) The miRNAome of globe artichoke: conserved and novel micro RNAs and target analysis. BMC Genomics 13:41. https://doi.org/10.1186/1471-2164-13-41

Djami-Tchatchou AT, Sanan-Mishra N, Ntushelo K, Dubery IA (2017) Functional roles of microRNAs in agronomically important plants-potential as targets for crop improvement and protection. Front Plant Sci 8:378. https://doi.org/10.3389/fpls.2017.00378

Du Z, Zhou X, Ling Y, Zhang Z, Su Z (2010) agriGO: a GO analysis toolkit for the agricultural community. Nucleic Acids Res 38:W64–W70

Gavnholt B, Larsen K, Rasmussen SK (2002) Isolation and characterisation of laccase cDNAs from meristematic and stem tissues of ryegrass (*Lolium perenne*). Plant Sci 162:873–885. https://doi.org/10.1016/S0168-9452(02)00035-3

Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong S-Y, Lopez R, Hunter S (2014) InterProScan 5: genome-scale protein function classification. Bioinformatics 30:1236–1240

Kozomara A, Griffiths-Jones S (2011) miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res 39:D152–D157

Li C, Zhang B (2016) MicroRNAs in control of plant development. J Cell Physiol 231:303–313. https://doi.org/10.1002/jcp.25125

Li S, Liu L, Zhuang X, Yu Y, Liu X, Cui X, Ji L, Pan Z, Cao X, Mo B, Zhang F, Raikhel N, Jiang L, Chen X (2013) MicroRNAs inhibit the translation of target mRNAs on the endoplasmic reticulum in Arabidopsis. Cell 153:562–574. https://doi.org/10.1016/j.cell.2013.04.005

Liu H, Searle IR, Watson-Haigh NS, Baumann U, Mather DE, Able AJ, Able JA (2015) Genome-wide identification of microRNAs in leaves and the developing head of four durum genotypes during water deficit stress. PLoS ONE 10:e0142799. https://doi.org/10.1371/journal.pone.0142799

Mathelier A, Carbone A (2010) MIReNA: finding microRNAs with high accuracy and no learning at genome scale and from deep sequencing data. Bioinformatics 26:2226–2234

Pashkovskiy PP, Ryazansky SS (2013) Biogenesis, evolution, and functions of plant microRNAs. Biochem 78:627–637. https://doi.org/10.1134/S0006297913060084

Portis E, Portis F, Valente L, Moglia A, Barchi L, Lanteri S, Acquadro A (2016) A Genome-wide survey of the microsatellite content of the globe artichoke genome and the development of a web-based database. PLoS ONE 11:e0162841. https://doi.org/10.1371/journal.pone.0162841

Pourcel L, Routaboul J-M, Kerhoas L, Caboche M, Lepiniec L, Debeaujon I (2005) TRANSPARENT TESTA10 Encodes a laccase-like enzyme involved in oxidative polymerization of flavonoids in arabidopsis seed coat. Plant Cell 17:2966. https://doi.org/10.1105/TPC.105.035154

Qin Z, Li C, Mao L, Wu L (2014) Novel insights from non-conserved microRNAs in plants. Front Plant Sci 5:586. https://doi.org/10.3389/fpls.2014.00586

Rogers K, Chen X (2013) Biogenesis, turnover, and mode of action of plant microRNAs. Plant Cell 25:2383–2399. https://doi.org/10.1105/tpc.113.113159

Samad AFA, Sajad M, Nazaruddin N, Fauzi IA, Murad AMA, Zainal Z, Ismail I (2017) MicroRNA and transcription factor: key players in plant regulatory network. Front Plant Sci 8:565. https://doi.org/10.3389/fpls.2017.00565

Scaglione D, Lanteri S, Acquadro A, Lai Z, Knapp SJ, Rieseberg L, Portis E (2012) Large-scale transcriptome characterization and mass discovery of SNPs in globe artichoke and its related taxa. Plant Biotechnol J 10:956–969. https://doi.org/10.1111/j.1467-7652.2012.00725.x

Scaglione D, Reyes-Chin-Wo S, Acquadro A, Froenicke L, Portis E, Beitel C, Tirone M, Mauro R, Lo Monaco A, Mauromicale G, Faccioli P, Cattivelli L, Rieseberg L, Michelmore R, Lanteri S (2016) The genome sequence of the outbreeding globe artichoke constructed de novo incorporating a phase-aware low-pass sequencing strategy of F_1 progeny. Sci Rep 6:19427

Sonnante G, D'Amore R, Blanco E, Pierri CL, De Palma M, Luo J, Tucci M, Martin C (2010) Novel hydroxycinnamoyl-coenzyme A quinate transferase genes from artichoke are involved in the synthesis of chlorogenic acid. Plant Physiol 153:1224–1238. https://doi.org/10.1104/pp.109.150144

Tang J, Chu C (2017) MicroRNAs in crop improvement: fine-tuners for complex traits. Nat Plants 3:17077. https://doi.org/10.1038/nplants.2017.77

Voinnet O (2009) Origin, biogenesis, and activity of plant microRNAs. Cell 136:669–687. https://doi.org/10.1016/j.cell.2009.01.046

Yang L, Wu G, Poethig RS (2012) Mutations in the GW-repeat protein SUO reveal a developmental function for microRNA-mediated translational repression in Arabidopsis. Proc Natl Acad Sci USA 109:315–320. https://doi.org/10.1073/pnas.1114673109

Yu Y, Jia T, Chen X (2017) The 'how' and 'where' of plant microRNAs. New Phytol 216:1002–1017. https://doi.org/10.1111/nph.14834

Zhang Y, Xia R, Kuang H, Meyers BC (2016) The diversification of plant *NBS-lrr* defense genes directs the evolution of microRNAs that target them. Mol Biol Evol 33:2692–2705. https://doi.org/10.1093/molbev/msw154

Genome Resequencing

12

Alberto Acquadro, Lorenzo Barchi, Ezio Portis and Sergio Lanteri

Abstract

The first reference genome sequence of the highly heterozygous globe artichoke has been recently released (January 2016, www.artichokegenome.unito.it), organized in 17 pseudomolecules and including about 27,000 predicted genes. Since the release of the first genome draft, it was clear that, due to the its homozygous profile, the plant selected for the draft assembly should be flaked by other assembled genomes cultivated and highly heterozygous. Acquadro and colleagues (Sci Rep 7:5617, 2017) resequenced and assembled six *Cynara cardunculus* genotypes (40×), which are representative of the four main varietal types, as well as one genotype of cultivated cardoon (*C. cardunculus* var. *altilis*). The genomes were annotated revealing an analogous number of genes in all the genotypes, ranging from 27,121 in the 'Catanese' to 28,029 in the 'Romanesco.' The five proteomes were also analyzed/categorized for RGA features. A total of ~23.5 M SNPs/indels were discovered, and a functional SNPs analysis was established. Recently, to improve the quality metrics of the first reference genome assembly (v1.0), a resequencing of the genotype was attempted through the implementation of a Hi-C library and the production of a v2.0 of the assembly with improved assembly metrics. This new high-quality reference genome, together with the genomic data from the other resequenced genotypes, representing a one-stop resource for *C. cardunculus* genomics will enable map-based cloning of useful genes from the species and will enhance the taxonomic breadth of the data available for comparative plant genomics in Compositae.

12.1 The Genome Resequencing Way

Cynara cardunculus (Asteraceae family) genomics has long remained poorly examined in respect of other species of the family, for which sequencing efforts were already attempted. In January 2016, the first genome sequencing has been publicly released; since the release of the first draft, it was clear that, due to the its artificial homozygous profile, the plant selected for the draft assembly should be flaked by other assembled genomes cultivated and highly heterozygous. Moreover, the survey through resequencing of *C. cardunculus* genetic resources is critical for understanding the species evolution, to carry out genetic studies and to implement new breeding strategies.

A. Acquadro · L. Barchi · E. Portis (✉) · S. Lanteri
DISAFA Plant Genetics and Breeding, University of Torino, Largo P. Braccini 2, 10095 Grugliasco, Torino, Italy
e-mail: ezio.portis@unito.it

E. Portis et al. (eds.), *The Globe Artichoke Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-030-20012-1_12

Table 12.1 Details of the *Cynara cardunculus* genotypes studied (from Acquadro et al. (2017), used under CC BY 4.0)

Accession name	Group	Code	Propagation
'2C' (reference)	Artichoke breeding line	2C	Seed
'Altilis 41'	Cultivated cardoon	A41	Seed
'Violetto di Sicilia'	Catanesi	VS	Shoot
'Violetto di Toscana' ('Tema')	Violetti	VT	Seed
'Romanesco C3'	Romaneschi	C3	Shoot
'Spinoso di Palermo'	Spinosi	SP	Shoot

Italy is the globe artichoke top producing country and harbors the richest primary cultivated gene pool, classified based on capitulum traits (Dellacecca et al. 1976; Porceddu et al. 1976; Vanella et al. 1981) in (i) '*Spinosi*': with long sharp spines on bracts and leaves; (ii) '*Violetti*': with medium-sized, violet-colored capitula and fleshy thorns; (iii) '*Romaneschi*': with spherical or sub-spherical green capitula; (iv) '*Catanesi*': with relatively small, elongated capitula with more or less marked violet streaks. Acquadro and colleagues (2017) considered six *C. cardunculus* genotypes (Table 12.1) for the resequencing analysis, comprising one cultivated cardoon ('A41'—Altilis 41) and five globe artichoke: '2C' (reference genome), VS ('Violetto di Sicilia'), VT ('Violetto di Toscana'), C3 ('Romanesco C3'), and SP ('Spinoso di Palermo'), representative of the core varietal types. All the genotypes were derived from propagated clones except for A41 and VT, which were seed propagated.

Sequencing was conducted using the Illumina platform with the 2PE × 150 bp chemistry to reach an average coverage of 35×, ranging from 24.4× for A41 to 45.3× for SP. Using a genome reconstruction approach based on iterative mapping and reference-guided assembly (Gan et al. 2011), the resequenced genomes were assembled and reconstructed at a chromosome scale. In detail, de novo assembly was performed with ABySS 1.9.0 assembler (Simpson et al. 2009) after a step of k-mers optimization, to reach the optimal assembly performances. Then, the IMR/DENOM pipeline (Gan et al. 2011) was adopted to reconstruct the genome of each varietal type, starting from the initial assembled set of contigs using the globe artichoke genome reference sequence as a guide. This approach avoided the requirement of multiple libraries construction (e.g., mate pair libraries) and resulted to be less demanding in terms of bioinformatics resources in respect of a de novo assembly step. The average percentage of reconstruction was close to 98.9%, analogous to the one of the reference genomes, but smaller compared to the estimated genome size of *C. cardunculus* (1,084 Mb). This fact is probable ascribable to the absence of portions of highly repetitive loci in the reference genome assembly; nevertheless, globe artichoke incorporated about 95% of the existing genes (Scaglione et al. 2016).

12.2 Gene Prediction and Gene Families

The MAKER-P suite (Campbell et al. 2014) was used to conduct the structural annotation of the genomes of the five resequenced genotypes (globe artichoke: 2C, SP, C3, VS, and VT) and cultivated cardoon (A41). Indeed, by repeating the gene prediction pipeline of the reference genome sequence, but applying a more severe AED threshold, 28,310 features were spotted, which correspond to a 5% of increase in respect of the 26,889 previously predicted (Scaglione et al. 2016). The six genomic sequences, together with the corpus of information related to structural and functional annotations, are accessible from the Web site www.artichokegenome.unito.it, where six independent JBrowse (Skinner et al. 2009) interfaces have been ad hoc built to facilitate the comparative search of genes/features within the species.

The gene comparative analysis showed, for all the analyzed plants, a similar number of genic features (Fig. 12.1, Table 12.2), where 2C showed the upmost (28,310) number of genes (AED $\leq$ 0.5), while VS the lowest one (27,121). The OrthoMCL pipeline (Li et al. 2003) was able to group together 161,855 sequences into 24,417 gene families, considering gene from the six genotypes (excepting singletons, Fig. 12.1). A total of 18,826 gene families (i.e., 138,098 genes) appeared shared between the six considered genotypes, whereas only 426 (i.e. 2,145 genes) were common if only the globe artichoke genotypes were included in the analysis. Overall, a limited number of gene families (i.e. 11, for a total of 24 genes) were exclusive of the reference genome, while one (i.e. 2 genes) appeared unique for both VS/SP genotypes. Analyzing the shared gene families, 12 (i.e. 36 genes) were in common among the non-spiny genotypes (VS/VT/C3), while 159 gene families (i.e. 318 genes) were shared by the 'pigmented' types (VS/VT). The oblong capitula types (SP, VS, and VT) showed to share a whole of 149 gene families (i.e. 449 genes). Finally, sequences not belonging to any cluster (singletons) varied from 149 (in the reference 2C) to 999 (in A41).

Some significant GO enrichments for genes in common among the five globe artichoke genotypes were observed. Notably, the top-level enrichments in GO processes were translation (GO:0006412) and gene expression (GO:0010467), although the latter was absent in C3 genotype. Considering the GO functions (F), enrichments were highlighted in structural molecule activity (GO:0005198) and structural constituent of ribosome (GO:0003735). Some GO components (C) enrichments were present for ribosome (GO:0005840) and ribonucleoprotein (GO:0030529) complexes. The 80% of the predicted genes showed a minimum of one IPR domain (Table 12.2), when analyzed with InterProScan. The most abundant domain of the top 20 SUPERFAMILY (de Lima Morais et al. 2011) domains was 'P-loop containing nucleoside triphosphate hydrolase' (SSF52540) which belongs to several UniPathway, like chlorophyll or coenzyme A biosynthesis. The second most retrieved SUPERFAMILY IDs were (i) 'protein Kinase-like domain' (SSF56112), which contains genes involved in signaling and regulatory processes in the eukaryotic cell, (ii) Armadillo-type fold (SSF48371), and (iii) 'Leucine-rich repeat domain, L domain-like' (SSF52058). The latter two are involved in translation factor activity and defense response, respectively.

12.3 Focus on Resistance Genes

A long time of cultivation and selection of globe artichoke germplasm, mainly concentrated on desirable commercial traits, give rise to an increase of losses due to the emergence of several diseases in both globe artichoke and cultivated cardoon. In this regard, the proteins encoded by resistance genes analogs (RGAs) have a crucial role in boosting (or not) the defense mechanism used by plants to recognize and react to pathogens. RGAs are generally divided into three main groups. The first one includes the so-called NLRs, which consist of a non conserved N-terminus, together with (NB)-ARC and LRR domains (Hulbert et al. 2001), as well as other extra motifs which usually are not kinase domains. They can be further divided into two major subgroups (TNLs and non-TNLs) based on the presence/absence of a TOLL/interleukin-1 receptor (TIR) domain at the N-terminus (Takken and Goverse 2012). The second group includes the receptor-like kinases (RLKs), while the third one groups the receptor-like proteins (RLPs); both contain a transmembrane domain (TM), but only RLKs exhibit a kinase cytosolic domain (Monaghan and Zipfel 2012).

The RGAs were mined and characterized in the five reconstructed genomes, together with the reference one. Candidates RGAs were identified using a Blastp approach starting from a set of 2,860 Arabidopsis RGAs downloaded from PRGDB (Plant Resistance Genes database (Sanseverino et al. 2010), http://prgdb.crg.eu/) against all the six proteomes, and classified following HMMER (http://hmmer.janelia.org/) analyses in six groups. The results are depicted in Table 12.3.

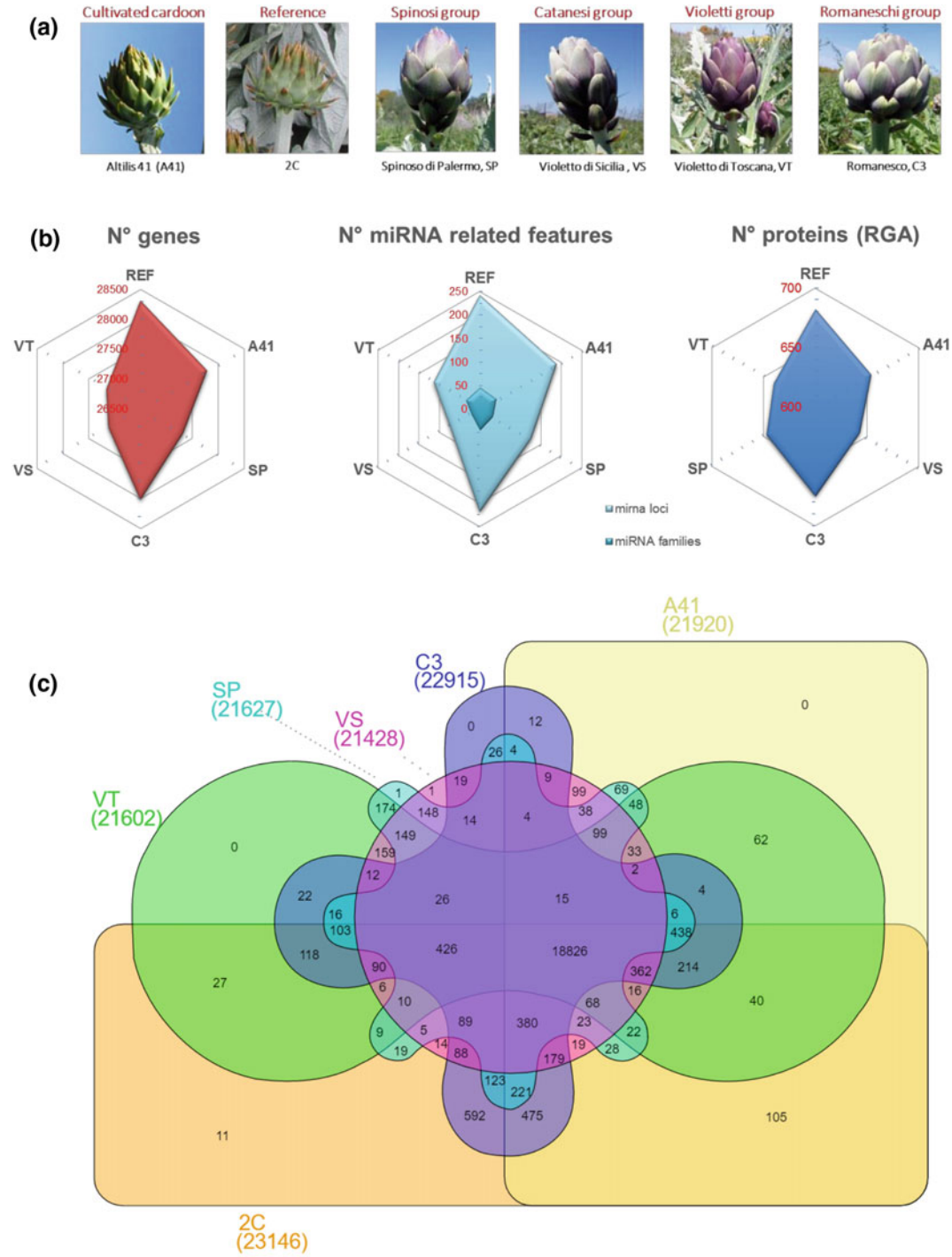


Fig. 12.1 **a** Capitula of the six genotypes. **b** N° of genes, miRNAs, and RGAs loci. **c** Orthologous gene families (from Acquadro et al. (2017), used under CC BY 4.0)

Table 12.2 Sequencing, assembly and structural annotation statistics

Sequencing data	2C	A41	VS	VT	C3	SP
SRA codes	SRR1914377; SRR1914378	SRR1826176; SRR1826114; SRR1914331	SRP055806	SRP055806	SRR1826175; SRR1825940; SRR1914330	SRP055806
Number of raw reads	–	90,410,254 (×2)	148,872,150 (×2)	129,452,237 (×2)	126,585,508 (×2)	174,120,908 (×2)
Number of reads	–	88,593,112 (×2)	138,616,098 (×2)	121,283,190 (×2)	123,535,166 (×2)	163,030,615 (×2)
Total amount sequence (Gb)	–	110	82.1	64	46	93.6
Estimated fold coverage	–	24.6×	38.5×	33.7×	34.3×	45.3×
ABySS assembly						
Number of contigs	–	5,741,441	6,988,492	6,242,434	8,456,162	7,566,149
Total length (contigs, Mb)	–	1,106.4	1,001.4	922.1	1,409.1	1,116.3
IMR/DENOM reconstruction						
Number of sequences	79,681	95,970	74,740	74,498	77,535	74,317
Sequences/Mb	121.6	147.2	115.7	115.5	118.7	115.3
Total length (contigs, Mb)	654.6	651.6 (99.5%[a])	645.9 (98.6%[a])	644.7 (98.4%[a])	652.8 (99.7%[a])	644.3 (98.4%[a])
Total length (scaffold, Mb)	724.7	721.9 (99.6%[a])	714.6 (98.6%[a])	713.1 (98.4%[a])	722.9 (99.7%[a])	712.3 (98.3%[a])
L_{50} (kb)	17.5	13.5	8.9	8.9	9.5	8.9
N_{50}	10,596	13,964	20,425	20,491	19,621	20,504
L_{90} (kb)	3.4	1.3	1.3	1.3	1.4	1.4
N_{90}	41,711	46,781	46,036	45,970	45,776	45,799
G + C %	32.00	35.18	35.04	35.08	35.28	35.01
N° of sequences > 10 kb	20,561	20,897	19,975	19,915	20,454	19,922
Number of genes	28,310	27,785	27,121	27,160	28,029	27,326
Number of proteins with IPR	22,571 (79.7%)	22,199 (79.9%)	21,898 (80.7%)	21,888 (80.6%)	22,406 (80%)	21,997 (80.5%)

[a]New genomes were compared with metrics deriving from the reference (2C) (from Acquadro et al. (2017), used under CC BY 4.0)

Overall, the majority of resistance genes belonged to the RLK family (up to 180 in C3), followed by those enclosing Pkinase and TM domains (other-KTM, up to 157 in 2C), and RLP (up to 86 in 2C). On the other side, a few RGAs including at least one NB-ARC domain (NBS such as TNL, CNL, and TN) were poorly represented. It has been described that NBS-LRR genes, after speciation events, experienced a gene expansion in several species such as *Arabidopsis* (Meyers et al. 2003; Ding et al. 2007) and rice (Zhou et al. 2004). On the contrary, these results seem to confirm a species-specific evolution of TNLs in Asterids, as hypothesized by Kim et al. (2012), and are in accord with what Christopoulou et al. (2015) recently reported in their study, where RLK class is the most represented in lettuce. Interestingly, several identified

Table 12.3 Classification of RGAs identified in the six genotypes

Identified domains	Acronym	2C	A41	VS	VT	C3	SP
LRR:NB-ARC	NL	9	7	7	6	8	7
LRR:NB-ARC:CC	CNL	4	4	4	6	4	5
LRR:NB-ARC:TIR	TNL	–	–	1	–	–	–
LRR:NB-ARC:TM	Other-TMNL	1	–	1	–	1	–
LRR:NB-ARC:TM:CC	Other-LNTMC	–	–	–	–	–	1
NB-ARC	N	9	8	9	8	9	11
NB-ARC:CC	CN	7	7	8	8	8	10
NB-ARC:TIR	TN	1	1	1	1	1	1
NB-ARC:TM	Other-NT	4	3	3	2	4	2
NB-ARC:TM:CC	Other-NTMC	3	2	1	2	3	1
LRR:Pkinase:TM	RLK	173	175	175	170	180	173
LRR:TM	RLP	86	72	65	71	77	65
Pkinase	Other-K	59	58	57	54	58	60
Pkinase:CC	Other-KC	9	9	8	9	9	8
Pkinase:TM	Other-KTM	157	146	149	156	155	148
Pkinase:TM:CC	Other-KTMC	4	4	3	3	4	5
LRR:Pkinase	KL	1	3	2		1	1
LRR:Pkinase:TM:CC	Other-LRTMC	5	5	4	6	6	6
LRR:TM:CC	Other-CLTM	1	–	–	–	–	–
TIR	T	2	2	–	2	2	2
LRR	L	22	27	24	17	19	23
ABC_tran	Other-A	1	1	1	2	1	2
ABC_tran:Pkinase:TM	Other-AKTM	–	1	–	–	–	–
ABC_tran:TM	Other-ATM	49	46	46	47	50	46
ABC_tran:TM:CC	Other-ACTM	4	4	5	4	3	4
Total	–	611	585	574	574	603	581
% of Blastp genes identified with HMM domains	–	0.9	0.9	0.9	0.9	0.9	0.9
Blastp genes identified	–	685	654	643	640	675	647

RGA proteins classified based on domain identification together with their frequency (from Acquadro et al. (2017), used under CC BY 4.0)

putative RGAs showed missing domains compared to 'canonical' NBS, RLPs and RLKs, in line with already reported by several authors (Büschges et al. 1997; Brandwagt et al. 2007; Wretblad et al. 2007; Cao et al. 2011; Gururani et al. 2012; Wang et al. 2013). The identified resistance genes represent, on average, about 2% of the total predicted genes for each of the genotypes in study. Although RGAs appeared quite uniformly distributed along the 17 chromosomes, they were preferentially located on six chromosomes, namely 1, 2, 3, 5, 10, and 13, suggesting a sort of specialization in resistance pathways. Furthermore, some distinct cluster of RGAs were identified, as on chromosome 10, where a cluster of genes with NB domain in a region of 2 Mb (16–18 Mb) and one of other-KTM genes at 10–11 M were spotted. On the other side, no clusters for RLK genes were highlighted.

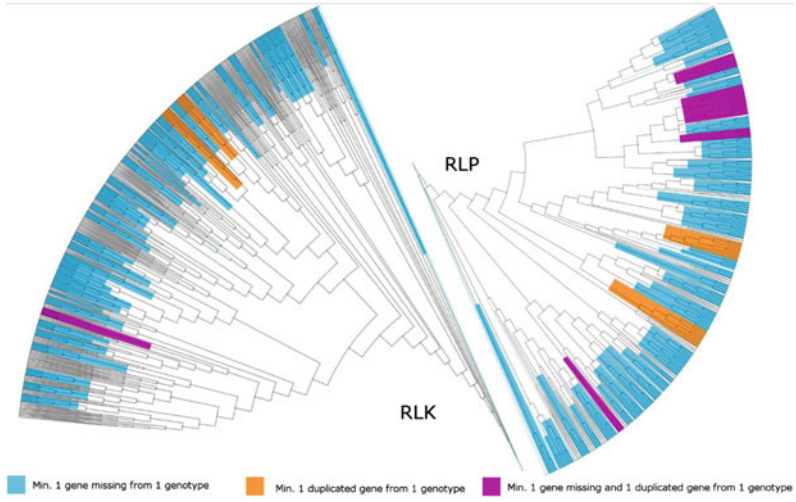


Fig. 12.2 Genetic tree of the resistance genes (RLP and RLK) of the six *C. cardunculus* genotypes. Clusters with one or more duplicated genes in one genotype are depicted in orange; clusters with at least one gene missing from one genotype and furthermore having at least one duplicated gene from one genotype are in purple; taxa lacking one or more RGA from at least one genotype are decorated in light blue (from Acquadro et al. (2017), used under CC BY 4.0)

RAxML (Stamatakis 2014) analyses, carried out on RGA amino acid sequence alignments, allowed to generate phylogenetic trees for each of the RGA classes in study (Fig. 12.2). Each resistance gene predicted in one genome clustered together with its orthologs from the other five genomes, although some taxa were orphan of one or more orthologs. As an example, 46 (24%) and 45 (50%) were the taxa lacking at least a RGA from no less than one genotype for the families RLK and RLP, respectively. Furthermore, some clusters for both RLKs and RLPs were found, containing at least one duplicated gene from one genotype as well as one missing gene from a genotype and a duplicated gene from another genotype.

It is known that the maintenance of many NBS/resistance genes has a potential impact on fitness costs (Tian et al. 2003; Orgil et al. 2007), and it has been suggested that microRNAs are used by plants to modulate the expression of NBS gene (Zhai et al. 2011; Eckardt 2012; Shivaprasad et al. 2012; Kallman et al. 2013; Fei et al. 2013). Indeed, it was found that in the five reconstructed genomes together with the reference one, from nine (in VS) to 41 (in 2C) identified RGAs (by Blastp approach) were putatively targeted by a miRNA, suggesting that also this post-transcriptional mechanism could exist in *C. cardunculus*.

The positioning of RGAs on the 6 genomes is a valuable tool for the development of a high-density genome-wide RGA genetic map for the species, which is critical for projecting diagnostic markers, identifying quantitative trait loci (QTL) or markers linked to plant disease resistance loci.

12.4 Diversity Evaluation at a Genomic Scale

Resequencing data lead to highlight a large set of variants across different accessions. At first, all the reads were aligned against the globe artichoke reference genome (2C), with a high average mapping rate (about 96.5%); in a second step, the SNP calling pipeline highlighted, all in all, the presence of 23.5 M SNP/indel. Variants ranged from 815 k in 2C (reference) up to almost 14.5 M in VS (Table 12.4).

Despite the low variant number rescued in the reference, all in heterozygous state, a comparable amount of SNP/indel was identified in all the

Table 12.4 Statistics about the highlighted variants (from Acquadro et al. (2017), used under CC BY 4.0)

Genotypes	2C	A41	VS	VT	C3	SP
SNP	781,530	5,900,934	13,440,135	11,860,358	11,937,400	13,241,315
SNP rate (%)	0.11	0.81	1.85	1.64	1.65	1.83
SNP/1000 bp	1.07	8.13	18.53	16.35	16.46	18.26
1 SNP every (bp)	927.6	122.8	53.9	61.1	60.7	54.7
Indel	34,232	443,603	1,055,545	987,949	910,23	1,150,341
Indel rate (%)	0.0047	0.06	0.15	0.14	0.13	0.16
Indel/1000 bp	0.05	0.61	1.46	1.36	1.26	1.59
1 indel every (bp)	21,179	1,634	686.8	733.8	796.5	630.2
SNP/indel	815,853	6,344,545	14,495,680	12,848,307	12,847,630	14,391,656
SNP/indel rate (%)	0.11	0.87	1.99	1.77	1.77	1.98
SNP/indel/1000 bp	1.12	8.75	19.99	17.72	17.72	19.85
1 SNP/indel every (bp)	892.85	114.27	50.01	56.42	56.43	50.37

resequenced globe artichoke genotypes: SP and VS showed over 14 M of SNP/indel, while C3 and VT showed over 12.8 M of SNP/indel. The number of variants in the cultivated cardoon (A41) genotype cut up to about 6 M. Among the globe artichoke genotypes, the observed SNP frequency ranged from about 1/54 (VS) to 1/122 bp (A41), while indel frequency was comprised between about 1/630 (SP) and 1/1,634 bp (A41). In a prior study (Scaglione et al. 2012), which relied merely on a reduced complexity library (RAD tag), the SNP frequency of the species was assessed as 1/179 bp, while indel frequency as 1/5,000 bp. These observations slightly differ to the ones measured using whole-genome derived data and could be ascribed to the use of one methylation sensitive enzyme adopted in the RAD-tag library protocol used, likely introducing a bias toward the un-transcribed regions of the genome. Overall, the globe artichoke SNP/indel frequency seems very high also in respect of the one found in other crops, such as grapevine (Cardone et al. 2016) or tomato (Causse et al. 2013).

The globe artichoke reference genome (2C) showed, as expected, the lowest heterozygosity value (0.11%, frequency of 1/893 bp), which is similar to the one showed by other horticultural dicots inbred species (e.g., eggplant; Delledonne et al. 2015). Indeed, the 2C genotype was ad hoc self-fertilized three times in order to reduce residual heterozygosity, thereby permitting a reliable genome assembly (Fig. 12.3, track B). The four globe artichoke varietal types, reflecting a different breeding history (Portis et al. 2005, 2012), showed a high level of heterozygosity (tenfold higher), exhibiting heterozygous SNPs ranging from 7.4 M in VT to 9.3 M in VS. Indeed, the lesser heterozygosity found in the VT ecotype was likely due to specific breeding practices conducted during time to stabilize its commercial production, which certainly contributed to fix many loci in a homozygous state. In a similar way, in the cultivated cardoon (A41), which is seed propagated, an even smaller number (4.2 M) of heterozygous SNP was found. Some SNP-dense genomic regions were highlighted, as well as genomic regions with a low frequency of SNPs in heterozygous state, while exhibiting SNPs fixed in homozygous; many of them occurred in gene-dense regions in a genotype-specific fashion (Fig. 12.3). Some other regions were in common between genotypes (e.g., ch.1 in VS, C3 and VT): As example, an extensive region in chromosome 7, with many homozygous SNPs, appeared shared between two genotypes (SP and VT). Most likely, selection conducted during time by farmers could have

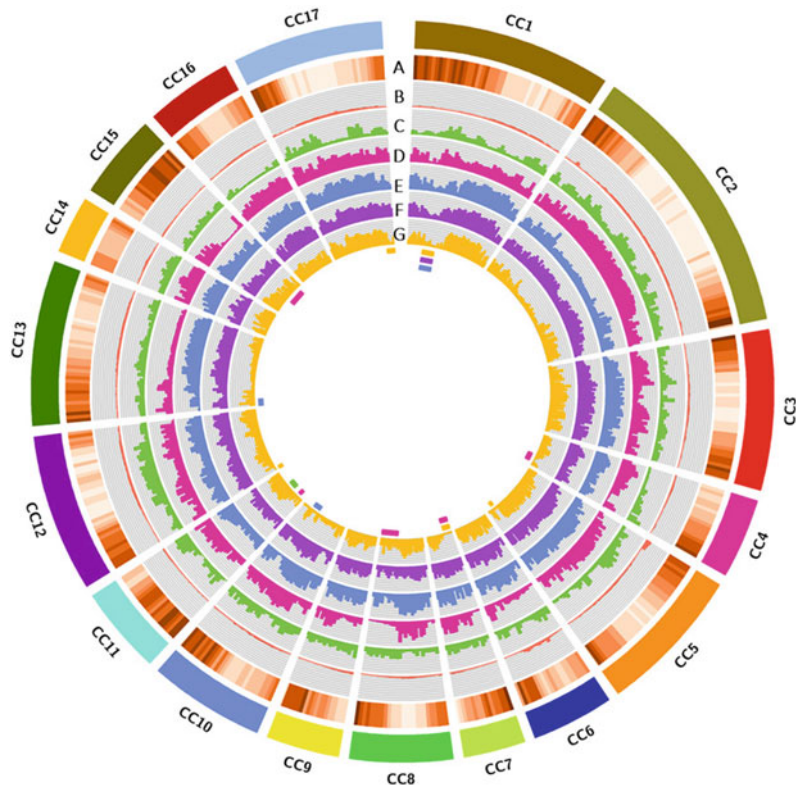


Fig. 12.3 Circos representation of gene/SNP density: (A) gene density heat map (reference); density (1M window) of SNPs (in heterozygous state) for 2C (B), A41 (C), SP (D), VS (E), C3 (F), VT (G) genotypes (from Acquadro et al. (2017), used under CC BY 4.0)

acted differentially to fix specific genomic regions, guiding genotype varietal diversification.

SNPs were functionally annotated using gene coordinates and showed 500 k exon-based polymorphisms, which represent about the ~2% of the overall detected SNPs (23 M). In detail (Fig. 12.4), only a little fraction of total mined homozygous SNPs were located inside coding sequences (1.41%), while the majority (2/3) of the variants were located outside the gene space (intergenic region: 57.6%; intronic region: 9.3%) Among the homozygous SNPs found in exons, 53.20 and 46.19% coded for synonymous and non-synonymous amino acid changes, respectively, while 0.62% appeared as non-sense mutations. Considering the heterozygous SNP/indel dataset, the upmost number of variants was placed in intergenic and intronic region (65.0 and 8.1%), while only 1.35% was in CDS. Considering coding SNPs/indels, only the 1.05% were categorized as non-sense mutation, while non-synonymous synonymous mutations were the 49.60 and the 49.35% of total variants, respectively.

A survey on specific missense SNPs located in genes involved in the biosynthesis of specific secondary metabolites (caffeoylquinic acids and sesquiterpene lactones) was ad hoc investigated (Fig. 12.5), and their impact on the gene biological function was predicted. In particular, deleterious amino acid substitutions on genes related to dicaffeoylquinic acids (CQAs) and sesquiterpene lactones (SLs) biosyntheses were searched. The CQAs pathway did not show relevant deleterious coding SNPs, while a prevalence of upstream/downstream located variations. The SLs pathway, as opposite, showed ~54 deleterious SNPs/indels, even if in heterozygosis, in 12 over 17 SLs-related enzymatic gene functions. As example, the germacrene A synthase (GAS), a key gene leading to cynaropicrin, as one of the major representatives in *C. cardunculus* (Eljounaidi et al. 2014) acting in leaves as antifeedants (Chadwick et al. 2013), was the

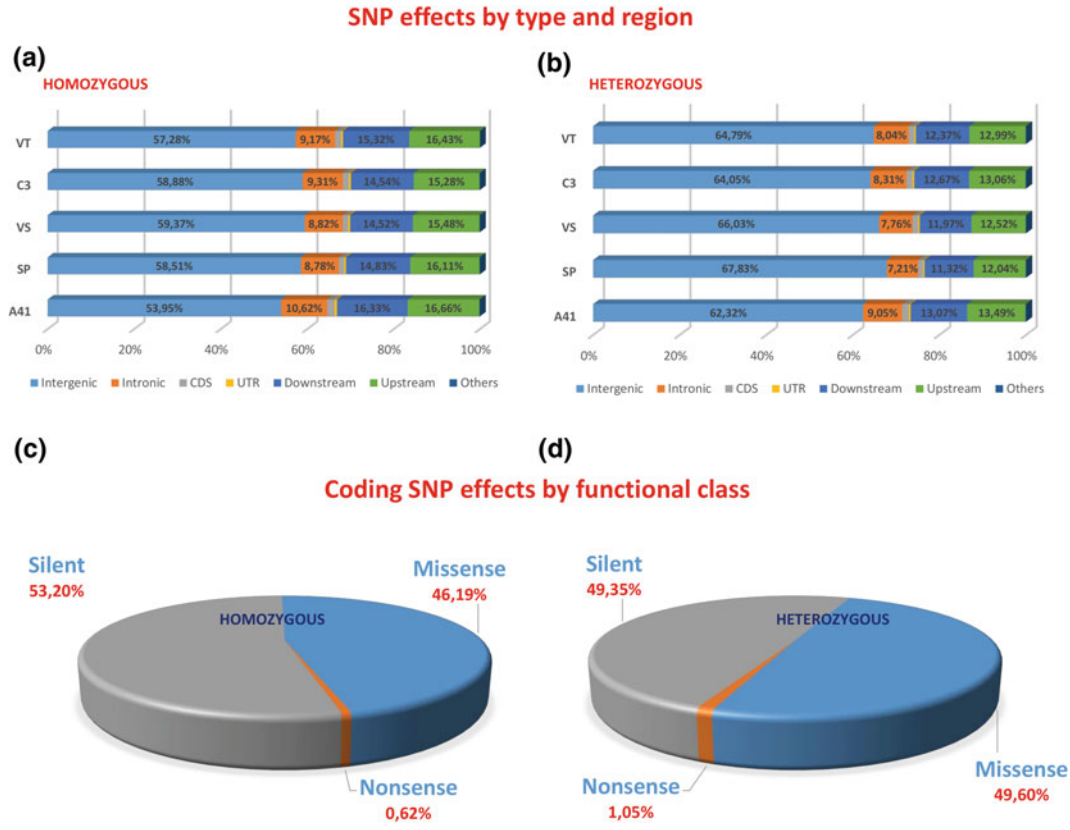


Fig. 12.4 SNP categorization of coding and non-coding variants. SNP effect by type and region for homozygous (**a**) and heterozygous (**b**) regions. SNP effect by functional class for homozygous (**c**) and heterozygous (**d**) regions (from Acquadro et al. (2017), used under CC BY 4.0)

most affected gene by deleterious mutations. Cynaropicrin was found to accumulate in leaves and to a lesser extent in the edible head bracts in both globe artichoke and cultivated cardoon (Schneider and Thiele 1974; Menin et al. 2012; Ramos et al. 2013), being responsible of the peculiar bitter taste. Curiously, the deleterious mutation was selectively detected in the globe artichoke genotypes but not in cultivated cardoon, which indeed has not been bred for head quality fitting human consumption tastes (e.g., sweetness vs. bitterness), but merely to obtain flashy stalks.

12.5 Improving the Current Reference Assembly: The Globe Artichoke Genome v2.0

A new genome assembly (v2.0) obtained from a Hi-C (Dovetail™) genomic library and based on the previously generated sequencing data has been recently reported (Acquadro et al. 2018). Hi-C is a proximity ligation-based method and, like other chromatin capture-based methods, relies on the fact that, after fixation, fragments of DNA in local proximity inside the cell nucleus

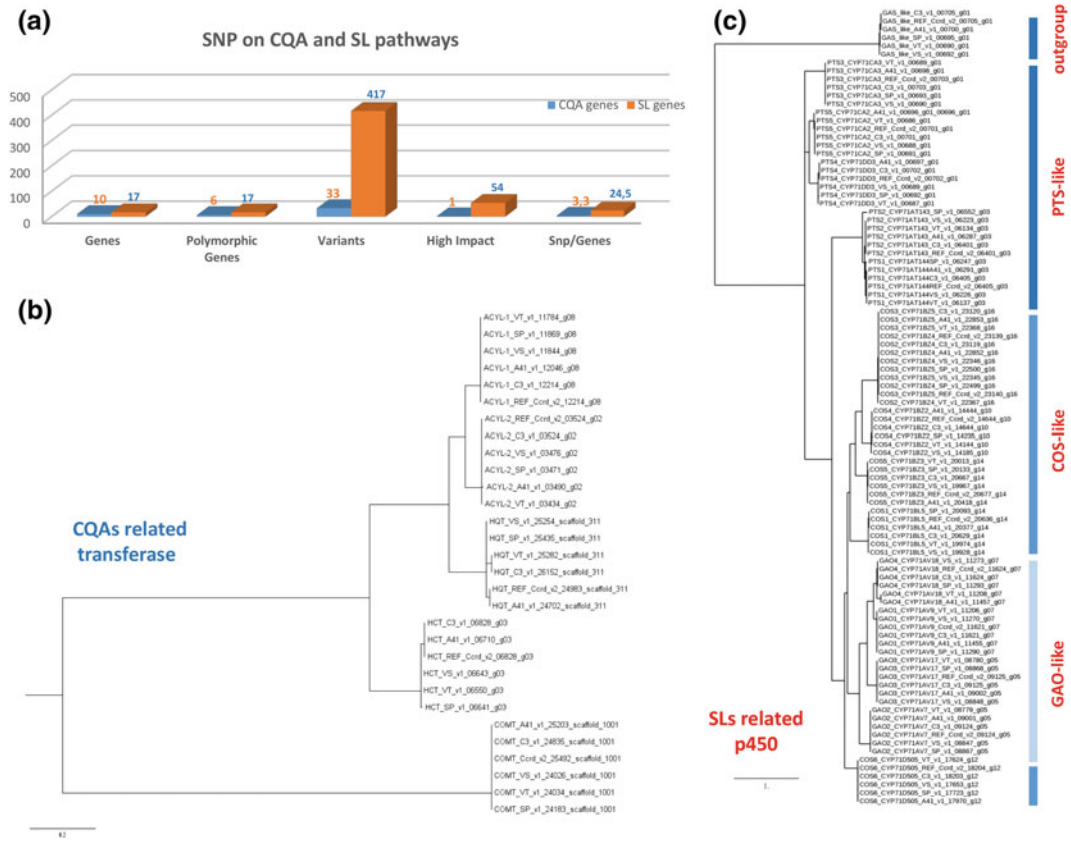


Fig. 12.5 **a** Coding SNPs data in CQAs and SLs pathways. Phylogenetic trees of proteins belonging to CQA-related transferase genes from the six genotypes analyzed (**b**) and to SL-related p450 genes from the same accessions (**c**) (from Acquadro et al. (2017), used under CC BY 4.0)

have more chances to be ligated together, and so sequenced as pairs, than distant regions. Consequently, the number of read pairs between intra-chromosomal regions can be described by a slowly decreasing function of the genomic distance between them (Putnam et al. 2016). Hi-C data were analyzed with the HiRise assembly pipeline, which enabled an accurate assembly of the globe artichoke genome up to the chromosome level.

In respect of first reference genome assembly (v1.0), the scaffold N_{50} was improved from 126 kb to 44.8 Mb (~356-fold increase) and N_{90} from 29 kb to 17.8 Mb (~685-fold increase). The L_{90} of the v1.0 sequence included 6123 scaffolds, while the new v2.0 just 16 super-scaffolds, a number close to the haploid chromosome number of the species. The newly generated super-scaffolds were assigned to pseudomolecules using reciprocal blast procedures. The cumulative size of unplaced scaffolds, in the version 1.0, was about 199 Mb, of which 165 Mb have been attributed to super-scaffolds in the 2.0 version, thus reducing to 33 Mb the unplaced scaffolds and increasing to 94% the anchored genome sequence. In the version 2.0, the annotation pipeline based on MAKER-P (Campbell et al. 2014) predicted 28,632 genes, a number which largely exceeded the one found in version 1.0 (i.e., 26,889; Scaglione et al. 2016), and which is very close to the one we recently obtained following the genome reconstruction of globe artichoke genotypes (i.e., 28,310; Acquadro et al. 2017).

Notable, chromosome conformation capture (Hi-C) is a high-throughput technology for physical mapping, which has emerged over labor-intensive and time-consuming traditional genetic/physical mapping. It has been applied to many species such as Arabidopsis (Jiao et al. 2017), medicago (Moll et al. 2017), lettuce (Reyes-Chin-Wo et al. 2017), barley (Mascher et al. 2017), amaranth (Lightfoot et al. 2017), quinoa (Jarvis et al. 2017), black raspberry (Jibran et al. 2018), and now globe artichoke (Acquadro et al. 2018). This new high-quality reference genome not only enhances the taxonomic breadth of the data available for comparative plant genomics in Compositae, but also is going to facilitate map-based cloning of economically important genes from the *C. cardunculus* species.

References

Acquadro A, Barchi L, Portis E, Mangino G, Valentino D, Mauromicale G, Lanteri S (2017) Genome reconstruction in *Cynara cardunculus* taxa gains access to chromosome-scale DNA variation. Sci Rep 7:5617. https://doi.org/10.1038/s41598-017-05085-7

Acquadro A, Barchi L, Portis E, Valentino D, Lanteri S (2018) Globe artichoke genome 2.0: improvement of its assembly quality & contiguity with proximity ligation library. In: Proceeding 62° SIGA annual congress, Verona, 25–28 Sep 2018

Brandwagt BF, Kneppers TJA, Nijkamp HJJ, Hille J (2007) Overexpression of the tomato Asc-1 gene mediates high insensitivity to AAL toxins and fumonisin B1 in tomato hairy roots and confers resistance to *Alternaria alternata* f. sp. *lycopersici* in *Nicotiana umbratica* plants. http://dx.doi.org/101094/MPMI200215135

Büschges R, Hollricher K, Panstruga R, Simons G, Wolter M, Frijters A, van Daelen R, van der Lee T, Diergaarde P, Groenendijk J, Töpsch S, Vos P, Salamini F, Schulze-Lefert P (1997) The barley *Mlo* gene: a novel control element of plant pathogen resistance. Cell 88:695–705. https://doi.org/10.1016/S0092-8674(00)81912-1

Campbell MS, Law M, Holt C, Stein JC, Moghe GD, Hufnagel DE, Lei J, Achawanantakun R, Jiao D, Lawrence CJ, Ware D, Shiu S-H, Childs KL, Sun Y, Jiang N, Yandell M (2014) MAKER-P: a tool kit for the rapid creation, management, and quality control of plant genome annotations. Plant Physiol 164:513–524

Cao A, Xing L, Wang X, Yang X, Wang W, Sun Y, Qian C, Ni J, Chen Y, Liu D, Wang X, Chen P (2011) Serine/threonine kinase gene Stpk-V, a key member of powdery mildew resistance gene *Pm21*, confers powdery mildew resistance in wheat. Proc Natl Acad Sci 108:7727–7732. https://doi.org/10.1073/pnas.1016981108

Cardone MF, D'Addabbo P, Alkan C, Bergamini C, Catacchio CR, Anaclerio F, Chiatante G, Marra A, Giannuzzi G, Perniola R, Ventura M, Antonacci D (2016) Inter-varietal structural variation in grapevine genomes. Plant J 88:648–661. https://doi.org/10.1111/tpj.13274

Causse M, Desplat N, Pascual L, Le Paslier M-C, Sauvage C, Bauchet G, Bérard A, Bounon R, Tchoumakov M, Brunel D, Bouchet J-P (2013) Whole genome resequencing in tomato reveals variation associated with introgression and breeding events. BMC Genom 14:791. https://doi.org/10.1186/1471-2164-14-791

Chadwick M, Trewin H, Gawthrop F, Wagstaff C (2013) Sesquiterpenoids lactones: benefits to plants and people. Int J Mol Sci 14:12780–12805. https://doi.org/10.3390/ijms140612780

Christopoulou M, Wo SR-C, Kozik A, McHale LK, Truco M-J, Wroblewski T, Michelmore RW (2015) Genome-wide architecture of disease resistance genes in lettuce. G3: Genes Genomes Genet 5:2655–2669. https://doi.org/10.1534/g3.115.020818

de Lima Morais DA, Fang H, Rackham OJL, Wilson D, Pethica R, Chothia C, Gough J (2011) SUPERFAMILY 1.75 including a domain-centric gene ontology method. Nucleic Acids Res 39:D427–D434

Dellacecca V, Magnifico V, Marzi V, Porceddu E, Mugnozza GT (1976) Contributo alla conoscenza delle varietà di carciofo coltivate nel mondo. In: Atti 2° Congresso Internazionale di Studi sul carciofo. Edizioni Minerva Medica, Turin, Italy, pp 199–315

Delledonne M, Dal Molin A, Minio A, Ferrarini A, Venturini L, Avanzato C, Toppino L, Sala T, Bassolino L, Barchi L, Comino C, Acquadro A, Portis E, Rinaldi R, Scaglione D, Francese G, D'alessandro A, Mennella G, Perrone D, Acciarri N, Pietrella M, Aprea G, Sulli M, Lanteri S, Rotino GL, Giuliano G (2015) A high quality eggplant (*Solanum melongena* L.) genome draft allows the mapping of phenotypic and metabolic QTLs. In: The 12th Solanaceae conference, Bordeaux, 25–29 Oct

Ding J, Zhang W, Jing Z, Chen J-Q, Tian D (2007) Unique pattern of R-gene variation within populations in Arabidopsis. Mol Genet Genomics 277:619–629. https://doi.org/10.1007/s00438-007-0213-5

Eckardt NA (2012) A microRNA cascade in plant defense. Plant Cell 24:840. https://doi.org/10.1105/tpc.112.240311

Eljounaidi K, Cankar K, Comino C, Moglia A, Hehn A, Bourgaud F, Bouwmeester H, Menin B, Lanteri S, Beekwilder J (2014) Cytochrome P450s from *Cynara cardunculus* L. CYP71AV9 and CYP71BL5, catalyze distinct hydroxylations in the sesquiterpene lactone biosynthetic pathway. Plant Sci 223:59–68. https://doi.org/10.1016/j.plantsci.2014.03.007

Fei Q, Xia R, Meyers BC (2013) Phased, secondary, small interfering RNAs in posttranscriptional regulatory networks. Plant Cell 25:2400–2415. https://doi.org/10.1105/tpc.113.114652

Gan X, Stegle O, Behr J, Steffen JG, Drewe P, Hildebrand KL, Lyngsoe R, Schultheiss SJ, Osborne EJ, Sreedharan VT, Kahles A, Bohnert R, Jean G, Derwent P, Kersey P, Belfield EJ, Harberd NP, Kemen E, Toomajian C, Kover PX, Clark RM, Rätsch G, Mott R (2011) Multiple reference genomes and transcriptomes for *Arabidopsis thaliana*. Nature 477:419–423. https://doi.org/10.1038/nature10414

Gururani MA, Venkatesh J, Upadhyaya CP, Nookaraju A, Pandey SK, Park SW (2012) Plant disease resistance genes: current status and future directions. Physiol Mol Plant Pathol 78:51–65. https://doi.org/10.1016/j.pmpp.2012.01.002

HMMER. http://hmmer.janelia.org/

Hulbert SH, Webb CA, Smith SM, Sun Q (2001) Resistance gene complexes: evolution and utilization. Annu Rev Phytopathol 39:285–312. https://doi.org/10.1146/annurev.phyto.39.1.285

Jarvis DE, Ho YS, Lightfoot DJ, Schmöckel SM, Li B, Borm TJA, Ohyanagi H, Mineta K, Michell CT, Saber N, Kharbatia NM, Rupper RR, Sharp AR, Dally N, Boughton BA, Woo YH, Gao G, Schijlen EGWM, Guo X, Momin AA, Negrão S, Al-Babili S, Gehring C, Roessner U, Jung C, Murphy K, Arold ST, Gojobori T, van der Linden CG, van Loo EN, Jellen EN, Maughan PJ, Tester M (2017) The genome of *Chenopodium quinoa*. Nature 542:307–312. https://doi.org/10.1038/nature21370

Jiao W-B, Accinelli GG, Hartwig B, Kiefer C, Baker D, Severing E, Willing E-M, Piednoel M, Woetzel S, Madrid-Herrero E, Huettel B, Hümann U, Reinhard R, Koch MA, Swan D, Clavijo B, Coupland G, Schneeberger K (2017) Improving and correcting the contiguity of long-read genome assemblies of three plant species using optical mapping and chromosome conformation capture data. Genome Res 27:778–786. https://doi.org/10.1101/gr.213652.116

Jibran R, Dzierzon H, Bassil N, Bushakra JM, Edger PP, Sullivan S, Finn CE, Dossett M, Vining KJ, VanBuren R, Mockler TC, Liachko I, Davies KM, Foster TM, Chagné D (2018) Chromosome-scale scaffolding of the black raspberry (*Rubus occidentalis* L.) genome based on chromatin interaction data. Hortic Res 5:8. https://doi.org/10.1038/s41438-017-0013-y

Kallman T, Chen J, Gyllenstrand N, Lagercrantz U (2013) A significant fraction of 21-nucleotide small rna originates from phased degradation of resistance genes in several perennial species. Plant Physiol 162:741–754. https://doi.org/10.1104/pp.113.214643

Kim J, Lim CJ, Lee B-W, Choi J-P, Oh S-K, Ahmad R, Kwon S-Y, Ahn J, Hur C-G (2012) A genome-wide comparison of NB-LRR type of resistance gene analogs (RGA) in the plant kingdom. Mol Cells 33:385–392. https://doi.org/10.1007/s10059-012-0003-8

Li L, Stoeckert CJ, Roos DS (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res 13:2178–2189

Lightfoot DJ, Jarvis DE, Ramaraj T, Lee R, Jellen EN, Maughan PJ (2017) Single-molecule sequencing and Hi-C-based proximity-guided assembly of amaranth (*Amaranthus hypochondriacus*) chromosomes provide insights into genome evolution. BMC Biol 15:74. https://doi.org/10.1186/s12915-017-0412-4

Mascher M, Gundlach H, Himmelbach A, Beier S, Twardziok SO, Wicker T, Radchuk V, Dockter C, Hedley PE, Russell J, Bayer M, Ramsay L, Liu H, Haberer G, Zhang X-Q, Zhang Q, Barrero RA, Li L, Taudien S, Groth M, Felder M, Hastie A, Šimková H, Staňková H, Vrána J, Chan S, Muñoz-Amatriaín M, Ounit R, Wanamaker S, Bolser D, Colmsee C, Schmutzer T, Aliyeva-Schnorr L, Grasso S, Tanskanen J, Chailyan A, Sampath D, Heavens D, Clissold L, Cao S, Chapman B, Dai F, Han Y, Li H, Li X, Lin C, McCooke JK, Tan C, Wang P, Wang S, Yin S, Zhou G, Poland JA, Bellgard MI, Borisjuk L, Houben A, Doležel J, Ayling S, Lonardi S, Kersey P, Langridge P, Muehlbauer GJ, Clark MD, Caccamo M, Schulman AH, Mayer KFX, Platzer M, Close TJ, Scholz U, Hansson M, Zhang G, Braumann I, Spannagl M, Li C, Waugh R, Stein N (2017) A chromosome conformation capture ordered sequence of the barley genome. Nature 544:427–433. https://doi.org/10.1038/nature22043

Menin B, Comino C, Portis E, Moglia A, Cankar K, Bouwmeester HJ, Lanteri S, Beekwilder J (2012) Genetic mapping and characterization of the globe artichoke (+)-germacrene A synthase gene, encoding the first dedicated enzyme for biosynthesis of the bitter sesquiterpene lactone cynaropicrin. Plant Sci 190:1–8. https://doi.org/10.1016/j.plantsci.2012.03.006

Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW (2003) Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis. Plant Cell Online 15:809–834. https://doi.org/10.1105/tpc.009308

Moll KM, Zhou P, Ramaraj T, Fajardo D, Devitt NP, Sadowsky MJ, Stupar RM, Tiffin P, Miller JR, Young ND, Silverstein KAT, Mudge J (2017) Strategies for optimizing BioNano and Dovetail explored through a second reference quality assembly for the legume model, *Medicago truncatula*. BMC Genomics 18:578. https://doi.org/10.1186/s12864-017-3971-4

Monaghan J, Zipfel C (2012) Plant pattern recognition receptor complexes at the plasma membrane. Curr Opin Plant Biol 15:349–357. https://doi.org/10.1016/j.pbi.2012.05.006

Orgil U, Araki H, Tangchaiburana S, Berkey R, Xiao S (2007) Intraspecific genetic variations, fitness cost and benefit of RPW8, a disease resistance locus in *Arabidopsis thaliana*. Genetics 176:2317–2333. https://doi.org/10.1534/genetics.107.070565

Porceddu E, Dellacecca V, Bianco VV (1976) Classificazione numerica di cultivar di carciofo. In: Atti 2°

Congresso Internazionale di Studi sul carciofo. Edizioni Minerva Medica, Turin, Italy, pp 1105–1119

Portis E, Mauromicale G, Barchi L, Mauro R, Lanteri S (2005) Population structure and genetic variation in autochthonous globe artichoke germplasm from Sicily Island. Plant Sci 168:1591–1598. https://doi.org/10.1016/j.plantsci.2005.02.009

Portis E, Scaglione D, Acquadro A, Mauromicale G, Mauro R, Knapp SJ, Lanteri S (2012) Genetic mapping and identification of QTL for earliness in the globe artichoke/cultivated cardoon complex. BMC Res Notes 5:252. https://doi.org/10.1186/1756-0500-5-252

Putnam NH, O'Connell BL, Stites JC, Rice BJ, Blanchette M, Calef R, Troll CJ, Fields A, Hartley PD, Sugnet CW, Haussler D, Rokhsar DS, Green RE (2016) Chromosome-scale shotgun assembly using an in vitro method for long-range linkage. Genome Res 26(3):342–350

Ramos PAB, Guerra ÂR, Guerreiro O, Freire CSR, Silva AMS, Duarte MF, Silvestre AJD (2013) Lipophilic extracts of *Cynara cardunculus* L. var. *altilis* (DC): a source of valuable bioactive terpenic compounds. J Agric Food Chem 61:8420–8429. https://doi.org/10.1021/jf402253a

Reyes-Chin-Wo S, Wang Z, Yang X, Kozik A, Arikit S, Song C, Xia L, Froenicke L, Lavelle DO, Truco M-J, Xia R, Zhu S, Xu C, Xu H, Xu X, Cox K, Korf I, Meyers BC, Michelmore RW (2017) Genome assembly with in vitro proximity ligation data and whole-genome triplication in lettuce. Nat Commun 8:14953. https://doi.org/10.1038/ncomms14953

Sanseverino W, Roma G, De Simone M, Faino L, Melito S, Stupka E, Frusciante L, Ercolano MR (2010) PRGdb: a bioinformatics platform for plant resistance gene analysis. Nucleic Acids Res 38:D814–D821. https://doi.org/10.1093/nar/gkp978

Scaglione D, Lanteri S, Acquadro A, Lai Z, Knapp SJ, Rieseberg L, Portis E (2012) Large-scale transcriptome characterization and mass discovery of SNPs in globe artichoke and its related taxa. Plant Biotechnol J 10:956–969. https://doi.org/10.1111/j.1467-7652.2012.00725.x

Scaglione D, Reyes-Chin-Wo S, Acquadro A, Froenicke L, Portis E, Beitel C, Tirone M, Mauro R, Lo Monaco A, Mauromicale G, Faccioli P, Cattivelli L, Rieseberg L, Michelmore R, Lanteri S (2016) The genome sequence of the outbreeding globe artichoke constructed de novo incorporating a phase-aware low-pass sequencing strategy of F_1 progeny. Sci Rep 6:19427

Schneider G, Thiele K (1974) Die verteilung des bitterstoffes cynaropicrin in der artischocke. Planta Med 26:174–183. https://doi.org/10.1055/s-0028-1097986

Shivaprasad PV, Chen H-M, Patel K, Bond DM, Santos BACM, Baulcombe DC (2012) A microRNA superfamily regulates nucleotide binding site-leucine-rich repeats and other mRNAs. Plant Cell 24:859–874. https://doi.org/10.1105/tpc.111.095380

Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJM, Birol I (2009) ABySS: a parallel assembler for short read sequence data. Genome Res 19:1117–1123

Skinner ME, Uzilov AV, Stein LD, Mungall CJ, Holmes IH (2009) JBrowse: a next-generation genome browser. Genome Res 19:1630–1638. https://doi.org/10.1101/gr.094607.109

Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Takken FL, Goverse A (2012) How to build a pathogen detector: structural basis of NB-LRR function. Curr Opin Plant Biol 15:375–384. https://doi.org/10.1016/j.pbi.2012.05.001

Tian D, Traw MB, Chen JQ, Kreitman M, Bergelson J (2003) Fitness costs of R-gene-mediated resistance in *Arabidopsis thaliana*. Nature 423:74–77. https://doi.org/10.1038/nature01588

Vanella B, Porceddu E, De Pace C (1981) Applicazioni di metodi di analisi numerica per il miglioramento genetico del carciofo. In: Atti 3rd Congresso Internazionale di Studi sul Carciofo, Laterza, Bari, pp 797–807

Wang W, Zhang Y, Wen Y, Berkey R, Ma X, Pan Z, Bendigeri D, King H, Zhang Q, Xiao S (2013) A comprehensive mutational analysis of the Arabidopsis resistance protein *rpw8.2* reveals key amino acids for defense activation and protein targeting. Plant Cell 25:4242–4261. https://doi.org/10.1105/tpc.113.117226

Wretblad S, Bohman S, Dixelius C (2007) Overexpression of a *Brassica nigra* cDNA gives enhanced resistance to *Leptosphaeria maculans* in *B. napus*. http://dx.doi.org/101094/MPMI2003166477

Zhai J, Jeong D-H, De Paoli E, Park S, Rosen BD, Li Y, Gonzalez AJ, Yan Z, Kitto SL, Grusak MA, Jackson SA, Stacey G, Cook DR, Green PJ, Sherrier DJ, Meyers BC (2011) MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs. Genes Dev 25:2540–2553. https://doi.org/10.1101/gad.177527.111

Zhou T, Wang Y, Chen JQ, Araki H, Jing Z, Jiang K, Shen J, Tian D (2004) Genome-wide identification of NBS genes in japonica rice reveals significant expansion of divergent non-TIR NBS-LRR genes. Mol Genet Genomics 271:402–415. https://doi.org/10.1007/s00438-004-0990-z

Genome Database (www.artichokegenome.unito.it)

13

Ezio Portis, Flavio Portis, Luisa Valente, Lorenzo Barchi, Sergio Lanteri and Alberto Acquadro

Abstract

The first high-quality genome assembly of the globe artichoke has been produced within the Compositae Genome Project and the resequencing analyses of four globe artichoke genotypes, representative of the core varietal types, as well as a genotype of the related taxa cultivated cardoon was, later on, carried out. The Web site "www.artichokegenome.unito.it" hosts all the available genomic sequences, together with their structural/functional annotations and project information are presented to users via the open-source tool JBrowse, allowing the analysis of collinearity and the discovery of genomic variants, thus representing a one-stop resource for *Cynara cardunculus* genomics. Pseudomolecules as well as unmapped scaffolds were used for the bulk mining of SSR markers and for the construction of the first globe artichoke microsatellite marker database. A database, called "*Cynara cardunculus* MicroSatellite DataBase" (*CyMSatDB*) was developed to provide a searchable interface to the SSR data. *CyMSatDB* facilitates the retrieval of SSR markers, as well as suggested forward and reverse primers, on the basis of genomic location, genomic versus genic context, perfect versus imperfect repeat, motif type, motif sequence, and repeat number.

13.1 Introduction

The globe artichoke genome draft was assembled from ~133-fold next-generation sequencing data into 13 K scaffolds (N50 = 125 Kbp, L50 = 1411), which represent 725 Mb of genomic sequence. Furthermore, a de novo gene prediction identified a total of 26,906 gene models (see Chap. 9). Recently, the resequencing analyses (~35X) of four globe artichoke genotypes, representative of the core varietal types, as well as a genotype of the related taxa cultivated cardoon was carried out. The genomes were reconstructed at a chromosomal scale and structurally/functionally annotated. Gene prediction indicated a similar number of genes, while distinctive variations in miRNAs and resistance gene analogues (RGAs) were detected. Overall, 23,5 M SNPs/indels were discovered (see Chap. 12).

The Web site "www.artichokegenome.unito.it" hosts all the available genomic sequences, together with their structural/functional annotations and project information are presented to users via the

E. Portis · L. Barchi · S. Lanteri · A. Acquadro (✉)
DISAFA Plant Genetics and Breeding, University of Torino, Largo P. Braccini 2, 10095 Grugliasco, Torino, Italy
e-mail: alberto.acquadro@unito.it

F. Portis · L. Valente
Yebokey User Experience Design and Web Development, Via Michele Buniva 2b, 10134 Torino, Italy

E. Portis et al. (eds.), *The Globe Artichoke Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-030-20012-1_13

open-source tool JBrowse, allowing the analysis of collinearity and the discovery of genomic variants, thus representing a one-stop resource for *Cynara cardunculus* genomics. The genome sequence of globe artichoke has been also used to catalogue the genome's content of simple sequence repeat (SSR) markers for the construction of the first globe artichoke microsatellite marker database, called *Cynara cardunculus* MicroSatellite DataBase (*CyMSatDB*). Simple sequence repeats (SSRs), which are ubiquitous and highly polymorphic, are particularly favored in a plant breeding context largely because they are usually codominantly inherited and can be readily assayed by conventional PCR. Perfect, imperfect, and compound SSRs were in silico mined using the SciRoKo SSR-search module (http://kofler.or.at/bioinformatics/SciRoKo). A minimum of four repetitions together with a minimum length of 15nt was requested; accordingly, any sequence was considered as a perfect SSR whenever a motif was repeated at least 15 times (1nt motif), eight times (2nt), five times, (3nt) or four times (4–6nt), allowing for only one mismatch. For compound repeats, the maximum default interruption (spacer) length was set at 100 bp. *CyMSatDB*, here described is a user-friendly and freely accessible tool, which offers chromosome wise as well as location wise search of primers by implementing Primer3.

13.2 JBrowse Resources

JBrowse (Buels et al. 2016), in contrast to its predecessor GBrowse (Stein 2013) is a fast, scalable genome browser built completely with JavaScript and HTML5. One of its key advantages relies on the ability to run on a desktop machine, as a stand-alone app, on all the operating system available (Windows, Mac OS, and Linux). Furthermore, it can be easily integrated into an already available public Web site with a highly intuitive interface. The JBrowse has been developed keeping in mind to be highly cross-platform. Indeed, it can be displayed on WebKit browsers (e.g., Safari and Chrome), Mozilla-based browsers (e.g., Firefox), and Microsoft Internet Explorer.

The JBrowse supports a wide range on input files, including GFF3, BED, FASTA, BAM, and VCF. Moreover, it can easily manage high-dimension genomes as well as deep-coverage sequencing. It allows to store several genomic sequences (Fig. 13.1a): user can switch from a genome to another from the main interface. Pre-computed genomic data sets (e.g., gene structures, orthologs from related plant species, expressed sequence tag (EST), repeats, etc.) can be easily displayed. Finally, it allows to pan and zoom as well as to select and organize tracks to be visualized via drag-and-drop functionality (Fig. 13.1b).

The setup of JBrowse is facilitated by the availability of scripts provided with the program. Genomic sequences are imported using the *prepare-refseqs.pl* script, which indexes a FASTA file containing the reference sequences. Track data are made available using JSON formatted files that user can import from feature data like flat files, or databases that have Bio::DB::* Perl interfaces. In case of flat files like GFF3 or BED, the provided script *flatfile-to-json.pl* represents an easy-to-use solution. In addition, the high customization of the JBrowse allows to taylor the appearance of the added tracks based on user preferences. After loading feature data, a special index of feature names has to be generated using *generate-names.pl* script; this allows the user to find features by typing feature names or IDs in the autocompleting search box. Finally, the JBrowse can display alignments directly from BAM files, with no preprocessing necessary.

The Web site "www.artichokegenome.unito.it" hosts all the available genomic sequences, together with their structural/functional annotations and project information are presented to users via the open-source tool JBrowse allowing the analysis of collinearity and the discovery of genomic variants, thus representing a one-stop resource for

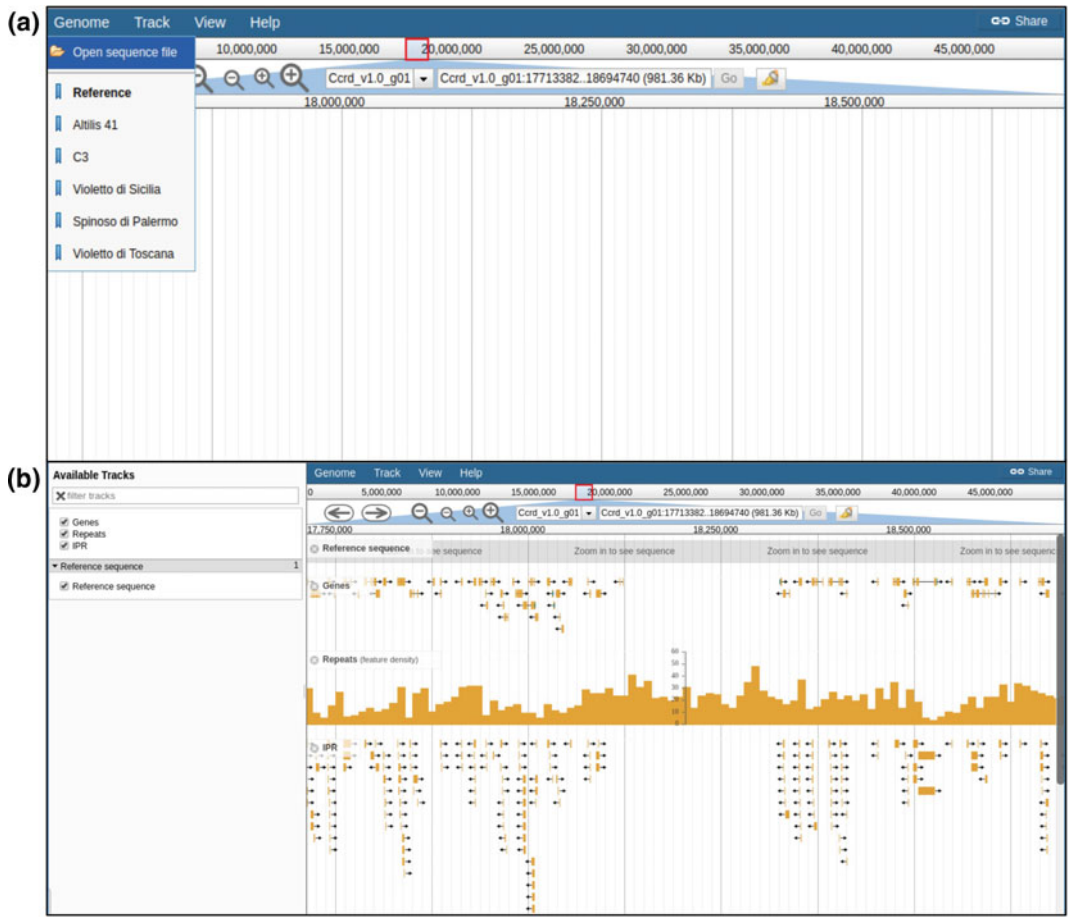


Fig. 13.1 Artichoke JBrowse. **a** Select one of the 6 genome sequences available on the "Genome" selector tool on the left. **b** Select tracks (Genes, Repeats or IPR domain) to view from the 'Available Tracks' selector on the left. Drag the mouse or use the convenient pan/zoom controls at the top to navigate around the browser to visualize a certain locus of interest

C. cardunculus genomics. Apart from representing the genome annotation of the 6 sequenced genomes, the artichoke JBrowse displays tracks of transposable elements and protein IPR domains (Fig. 13.1b) obtained using Interproscan5 for each of the genotypes available. Track data can be also be downloaded by the user in widely used biological formats, such as GFF3, BED (for genomic intervals), and FASTA (for sequences), referring either to a specific sequence feature or to the underlying genome sequence region (making use of the "Reference Sequence" track). The platform will be progressively implemented as the new features become publicly available.

13.3 CyMSatDB Construction, Content, and System Architecture

The public domain (*Cynara cardunculus* MicroSatellite DataBase (CyMSatDB), available at www.artichokegenome.unito.it/cymsatdb/, provides browsable access to all the SSRs identified in the globe artichoke genome. Figure 13.2 provides a flow of user operation, interaction, and various features and utilities. SSRs can be retrieved on the basis of simple characteristics, such as "genomic location" (chromosome number), "SSR feature" (whole

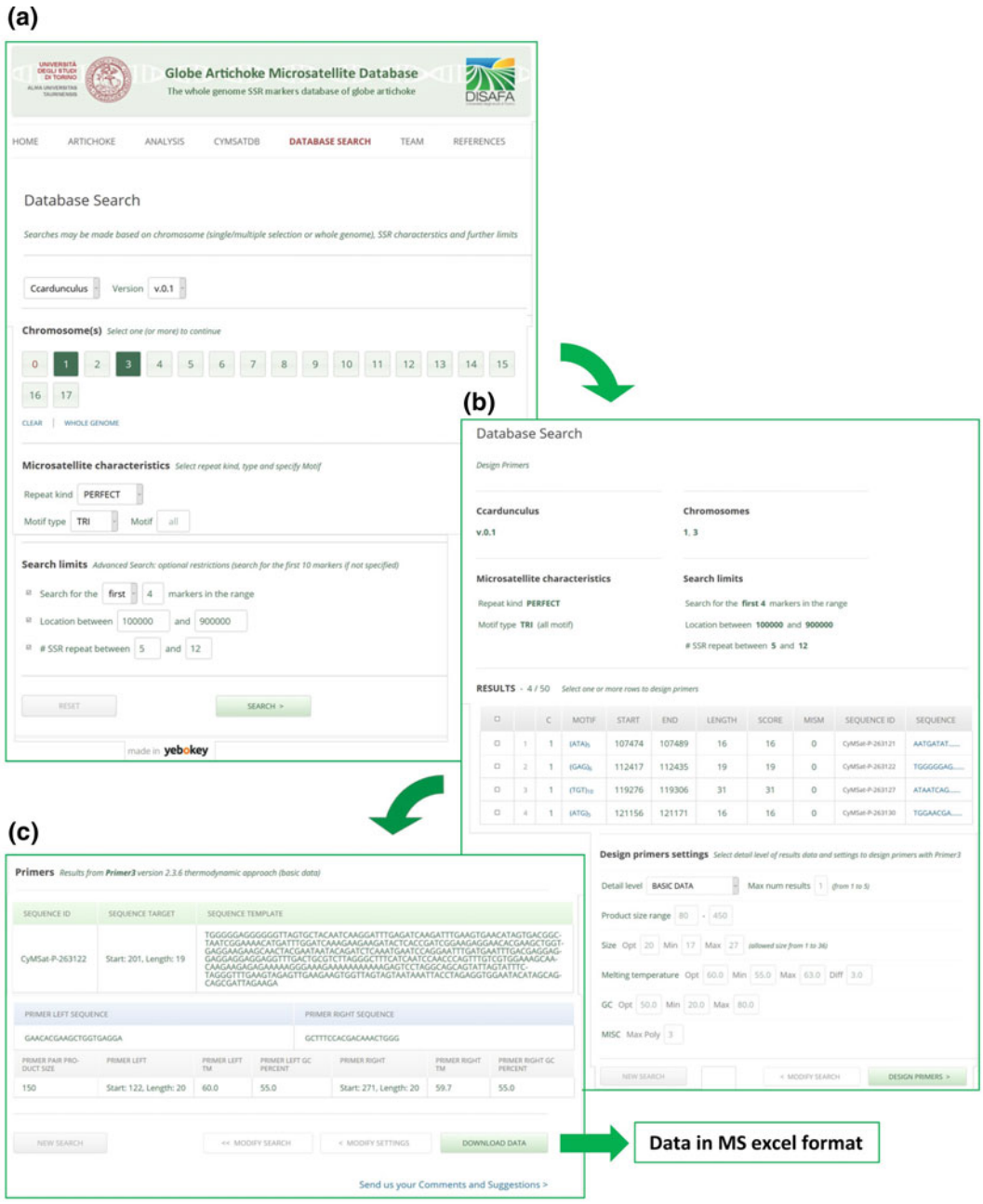


Fig. 13.2 A worked example of an SSR search and primer design using CyMSatDB. **a** Settings given for chromosome selection and SSR search. **b** The SSR output and settings given for the design of primers. **c** Suggested primers and the downloading of the result. From Portis et al. (2016), used under CC BY 4.0

genomic or only genic SSR), "repeat kind" (perfect vs imperfect), or advanced characteristics, such as "motif type" (mono- to hexanucleotide), "specific motif sequence", "repeat number", and "specific genomic location." Multiple parameters can be also combined to search for a specific set of SSRs as per user requirement. Microsatellites can be mined based on the choice of chromosome, where more than one

chromosome may also be selected (Fig. 13.2a). The flexibilities provided will enable the researcher to select markers of choice at the desired genomic interval, by limiting the search based on chromosomal location as well as the number of markers in that range.

The output lists all the SSRs meeting the user-selected parameters(s) in tabular form, along with SSR identifiers, LG number, motif type and length, genomic location (start and end position), SSR length, and an option for downloading the flanking sequences (Fig. 13.2b). Primers can be readily designed to amplify selected SSR primers, using the Primer3 stand-alone tool. The user may go for primer designing as a default setting or modify the standard Primer3 parameters. The "Design Primers" button directs the use to a list of up to five possible primer pairs, with their melting temperatures (Tm), their GC content, and the expected length of the amplicon. The primer information can be readily exported into Microsoft Excel (Fig. 13.2c).

CyMSatDB was specifically designed to work efficiently with any mobile device and smartphone, so to rapidly access from iPhone, iPad, Android smartphones, and tablets, easily searched and for custom primer designing. When displayed on a smaller device, CyMSatDB will automatically resize and optimize to look optimal also for small devices. Designed for domain experts, this interface implements error prevention and autocorrection functions. In particular, it: (i) reminds the default values/ranges for each field; (ii) autocorrects data on key-press instead of on mouse-out; (iii) explains errors using detailed descriptive tooltips; (iv) signals success when the field is correctly completed. The system offers two kinds of data export of the designed primers: a synthetic version showing only the most relevant data and a complete one. The data download provides an excel file specifically formatted for higher readability. A management interface is also available for administrators to ease the data import in the system. The conversion of the SciRoKo output into CyMSatDB compatible data is made simple through immediate interface operations instead of using command line scripts.

The marker information housed in CyMSatDB will be of high value for linkage mapping and for facilitating the marker saturation of genomic regions. Several plug-ins have been implemented to generate primers at a user-defined chromosomal location. The database could be used to design assays sampling more than one SSR locus in a single PCR, which require both compatibilities with respect to primer annealing temperature and a non-overlap in amplicon size. Such multiplex assays are of particular relevance to varietal identification, especially in the context of plant variety protection and varietal release.

13.4 The SSR Content of the Globe Artichoke Genome

The ~725 Mbp of globe artichoke genome sequence yielded 177,207 perfect SSRs, equivalent to an overall density across the genome of 244.5 SSRs/Mbp; of these, about 21% (37,748) were compound loci. The most common motifs (~73% of all SSRs) were dinucleotides, followed by the tri- (~11%), the tetra- (~6%), and the mononucleotides (~5%). Both penta- and hexanucleotide repeats were uncommon (together <5%). More than 224,000 imperfect SSRs were detected (Table 13.1). Among this group, the occurrence of mono-, di-, tri-, and tetranucleotide motifs was quite similar to that seen among the group of perfect SSRs, but that of the larger motifs was rather higher: together the penta- and hexanucleotide SSRs represented ~18% of the set of imperfect SSRs.

The distribution of repeat unit number among the set of perfect SSRs is summarized in Fig. 13.3a. For all SSR classes, the frequency was inversely proportional to the number of repeat units; for example, SSRs composed of ten or fewer repeat accounted for 49.2% of the set, while those harboring >20 repeats represented <5% of the set. The decline in frequency with greater length was least well marked for the mono- and dinucleotide SSRs, and most marked for the pentanucleotide SSRs. As a result, the mean number of repeat units found among the set of dinucleotide SSRs (12.7) was nearly double

Table 13.1 The chromosome-by-chromosome distribution of perfect, compound and imperfect SSRs

Linkage groups	Total Mbp	Perfect								Compound		Imperfect	
		Mono-	Di-	Tri-	Tetra-	Penta-	Hexa-	Total	SSRs/Mbp	Total	SSR/Mbp	Total	SSRs/Mbp
LG01	49.71	772	9637	1345	872	366	368	13,360	268.8	2714	54.6	17,309	348.2
LG02	70.34	747	10,834	1993	893	359	386	15,212	216.3	3150	44.8	19,282	274.1
LG03	40.26	566	8660	1206	774	302	335	11,843	294.2	2581	64.1	14,919	370.6
LG04	20.15	304	4553	507	349	171	155	6039	299.7	1304	64.7	7593	376.9
LG05	37.16	523	8476	1081	688	288	275	11,331	304.9	2552	68.7	14,238	383.1
LG06	20.61	284	4683	517	414	161	133	6192	300.4	1472	71.4	7627	370.0
LG07	15.56	267	4059	376	367	122	140	5331	342.6	1209	77.7	6693	430.1
LG08	25.92	369	5894	752	537	244	189	7985	308.0	1808	69.7	9888	381.4
LG09	18.33	270	4502	456	349	145	167	5889	321.3	1277	69.7	7380	402.7
LG10	29.10	434	5797	781	447	184	227	7870	270.4	1700	58.4	9974	342.7
LG11	22.01	447	5410	629	508	208	191	7393	336.0	1522	69.2	9298	422.5
LG12	39.65	474	7436	1087	651	217	238	10,103	254.8	2199	55.5	12,575	317.2
LG13	41.51	568	8227	1066	718	279	334	11,192	269.6	2353	56.7	14,165	341.3
LG14	14.48	241	4002	342	347	136	113	5181	357.9	1219	84.2	6518	450.3
LG15	21.26	414	5456	574	519	178	153	7294	343.0	1572	73.9	9256	435.3
LG16	21.91	295	4921	594	416	165	141	6532	298.1	1485	67.8	8216	374.9
LG17	37.69	366	6066	1077	502	201	211	8423	223.5	1798	47.7	10,808	286.8
LG00	199.02	1052	20,777	5526	1391	576	715	30,037	150.9	5833	29.3	38,323	192.6
Total	724.67	8393	129,390	19,909	10,742	4302	4471	177,207	244.5	37,748	52.1	224,062	309.2

From Portis et al. (2016), used under CC BY 4.0

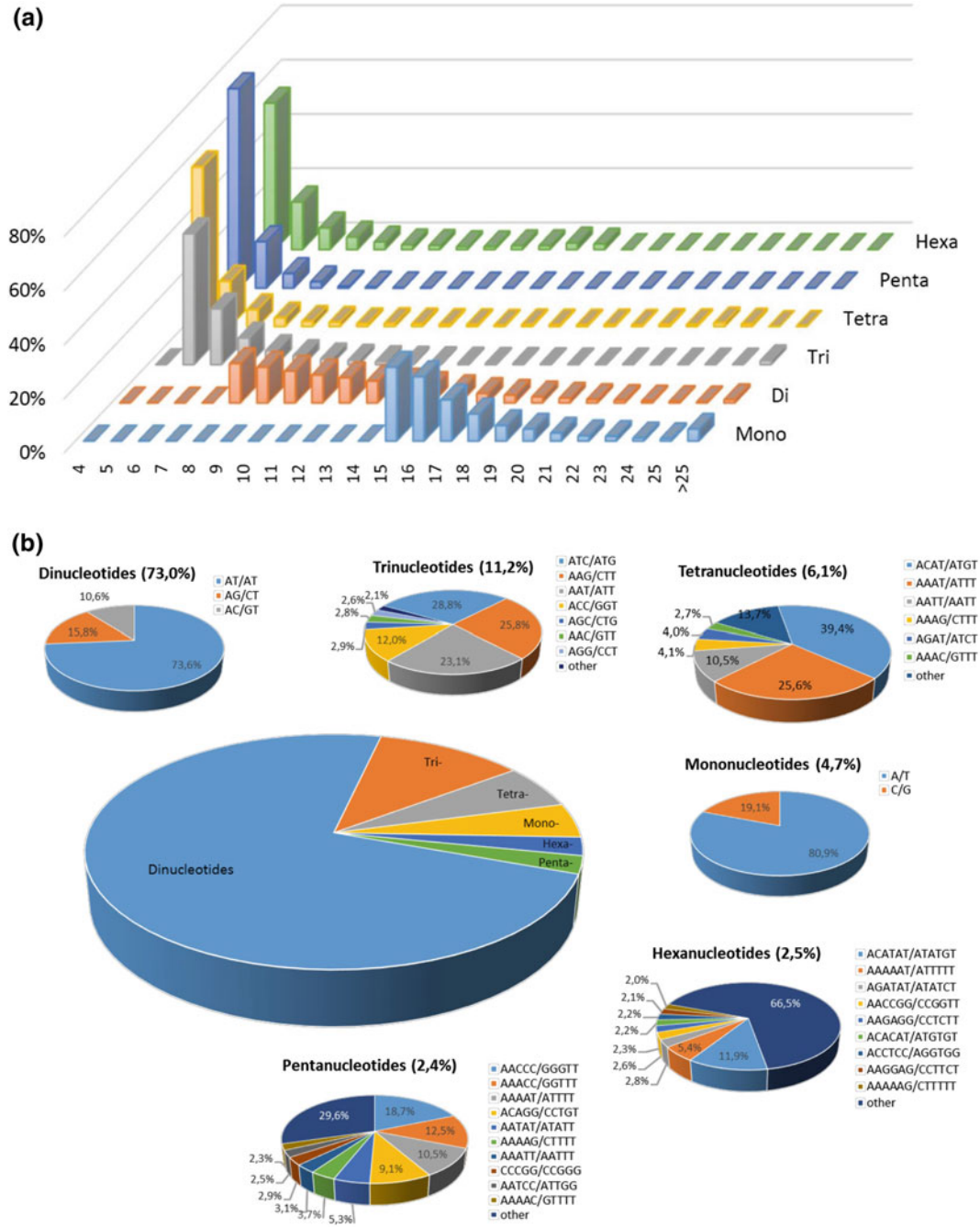


Fig. 13.3 **a** Relative frequency of perfect SSRs motifs in the globe artichoke genome. **b** The distribution of the major repeat types in the globe artichoke genome. From Portis et al. (2016), used under CC BY 4.0

that of in either the tri- or the tetranucleotide SSRs (7.4 and 6.3, respectively) and more than double that in the penta- and hexanucleotide ones (4.8 and 5.8, respectively).

A detailed analysis of individual repeat motifs for each type of SSR found in the globe artichoke genome sequence was conducted (Fig. 13.3b). Among the mononucleotide types, A/T predominated

heavily (80.9%), as did AT/TA among the dinucleotide types (73.6%); the most frequently occurring motif types among the trinucleotide SSRs were ATC/ATG (28.8%), among the tetranucleotide types ACAT/ATGT (39.4%), among the pentanucleotide types AACCC/GGGTT (18.7%), and among the hexanucleotide types ACATAT/ATATGT (11.9%). AT/TA was not just the predominant dinucleotide motif, but was also the most frequent motif globally, represented in 53.7% of the SSRs; in contrast, CG/CG was very rare (0.002%). Among the trinucleotides, ATC, AAG, and AAT were the most abundant motifs (present in ~77% of SSRs), whereas the GC-rich motifs ACG and CCG were the least common. Similarly, the AT-rich motifs ACAT, AAAT, and AATT dominated the set of tetranucleotide SSRs (Fig. 13.3b).

The SSR loci identified were further classified on the basis of their repeat motif and distribution over each pseudomolecule. Although the genome sequence has been organized into 17 pseudomolecules, representing the haploid globe artichoke set, ~17% of the SSRs were detected in not anchored scaffolds (LG00). Of the remaining ~83%, the mean number per LG of perfect and imperfect SSRs was, respectively, 8657 and 10,926. The greatest number of SSRs (15,212 perfect, 19,282 imperfect) was assigned to the longest LG (LG02, 70.34 Mbp), and the lowest (5181 perfect, 6518 imperfect) to the shortest (LG14, 14.48 Mbp), resulting in a range of perfect SSR density from 216.3 to 357.9 per Mbp (Table 13.1).

13.5 Genic SSRs

The genomic distribution of SSRs and the relationship with annotated genomic components (gene space) were also analyzed, on the basis of the available and assembled globe artichoke pseudomolecules. In all, 4761 perfect and 6583 imperfect SSRs were identified as present in the genic sequence, distributed within 3781 genes (about 14% of the full gene complement). The overall microsatellite density in globe artichoke genes was significantly lower than in genomic sequence and their distribution was coherent with that of genes and opposite to that of the repetitive elements in all the pseudomolecules analyzed, in agreement with the reports of Morgante et al. (2002). As has been shown for other genomes (Toth et al. 2000; Morgante et al. 2002; Mun et al. 2006), genic SSRs tend to be composed of either tri- or hexanucleotides. A likely reason for this prevalence is that, in coding regions, frameshift mutations generated by replication slippage are likely to experience negative selection: alterations in the repeat number of a tri- or hexanucleotide repeat, depending on the SSR's position within the sequence, do not induce frameshift. With respect to the surviving dinucleotide motifs in globe artichoke, while AT/TA repeats tend to prevail in non-transcribed regions, AG/CT is commonest in genic sequence, which is the case in other species as well (Morgante et al. 2002; Cavagnaro et al. 2010; Shi et al. 2013). There was generally a preference for GC-rich motif in the trinucleotide repeats present in genic DNA: the GC content was ~48%, whereas it was only ~32% in the global set of trinucleotide repeats. SSRs embedded in genic DNA are generally more likely to be conserved across (related) species boundaries than those in non-genic DNA, a feature which makes them attractive for comparative mapping (Varshney et al. 2005). On the other hand, since genic DNA is under positive selection, these SSRs are typically less polymorphic. Given that a high repeat number is correlated with allelic variation, it was therefore unsurprising to find that nearly 5% of the global globe artichoke SSRs harbored >20 repeats, while this proportion among the genic SSRs was under 2%.

The genes harboring one or more SSRs were grouped into the three GO main categories ["biological processes" (BP), "molecular functions" (MF), and "cellular components" (CC)], and thereafter in the 24 sub-GO categories. Over-representation was noted for a number of genes. Those allocated to BP belonged to the sub-categories "gene expression," "regulation of

gene expression," "regulation of transcription, DNA-dependent," and "cellular macromolecule metabolic process." Similarly, on the basis of MF, enrichment was observed for "nucleic acid binding," "sequence-specific binding level," "transcription regulator activity," and "protein dimerization activity." Finally, with respect to CC, the only identified sub-category was "nucleus." A prominent class of genes were those encoding transcription factors; a similar concentration of SSRs within specific classes of genic sequence has been observed in other species (Martin et al. 2002; Zhang et al. 2006; Yu et al. 2010; Kujur et al. 2013; Liu et al. 2015). The implications of transcription factor sequences harboring an SSR have been alluded to elsewhere (Liu et al. 2015) and should be addressed in future analyses of diversification within globe artichoke.

13.6 A Cross-Species Comparison

The content and distribution of SSRs in the globe artichoke genome was compared with that present in 14 other plant genomes. The globe artichoke genome included almost twice as many perfect SSRs as in tomato (88,781), four times as many as in sweet orange (43,147), and ten times as many as in Arabidopsis (17,289), but fewer than in sunflower (269,635), a fellow Compositae species. The total length of the sequence represented by repeats in globe artichoke (~4.4 Mbp) was equivalent to 0.61% of the assembled genome, a proportion comparable to that in both grapevine (0.67%) and black cottonwood (0.57%), but was significantly higher than in the other 12 species. A negative relationship between genome size and SSR density has been reported by Morgante et al. (2002). Even though the largest genomes analyzed (cotton and sunflower) indeed displayed the lowest SSR density, the SSR density in globe artichoke (244.5 per Mbp) was more than double that in tomato (113.6 per Mbp), in spite of their comparable genome sizes, but analogous to one of the grape wine and black cottonwood (263.7 and 227.1 SSRs/Mbp, respectively), which own smaller genomes. Arabidopsis also deviated from the trend, as previously noted by Cavagnaro et al. (2010) its genome size is small, but its SSR density is lower (144.4 per Mbp) than in both horseweed (double its genome size) and globe artichoke (six times its size). In respect to the other genomes in the study, another signature of the globe artichoke genome was the high frequency (73%) of dinucleotide perfect motifs. Together, mono-, di-, and trinucleotide repeats formed the bulk of the SSR content in all 15 genomes analyzed (ranging from 74% in cotton to 92% in Arabidopsis). In globe artichoke, tomato, common bean, date palm, and banana, dinucleotide motifs predominated, followed by the tri-, tetra-, and mononucleotide ones (Fig. 13.4). However, in both of the other two Compositae species (sunflower and horseweed), the frequency of mononucleotide repeats was higher than that of tetranucleotide repeats. The distribution of SSR types in globe artichoke was most unlike that in Arabidopsis and grapevine (where mononucleotide motifs predominated), and in sweet orange and rice (where trinucleotide motifs were by far the commonest type).

As in most plant species, the base composition of the globe artichoke SSR motifs was AT-rich. With respect to the mononucleotide motifs, A/T predominated in 14 of the 15 species (the exception was rice, where the proportion was only ~38%). Among the dinucleotide motif types, AT/TA was the most frequent, except in date palm, where the commonest motif was AG/CT. The frequency of the CG/GC motif rose above 1% only in rice. It has been suggested by Cavagnaro et al. (2010) that AT-rich repeats predominate in dicotyledonous but not in monocotyledonous species, and that this difference reflects their genomes' nucleotide composition: the mean GC% content of dicotyledonous species is 34.6%, while that of monocotyledonous species is 43.7%. However, this difference could not satisfactorily account for the disparate frequency of certain motifs between these two phylogenetic classes. Rather uniquely, the commonest tri- and pentanucleotide motifs in globe artichoke were ATC and AACCC, (and not AAT and AAAAT)

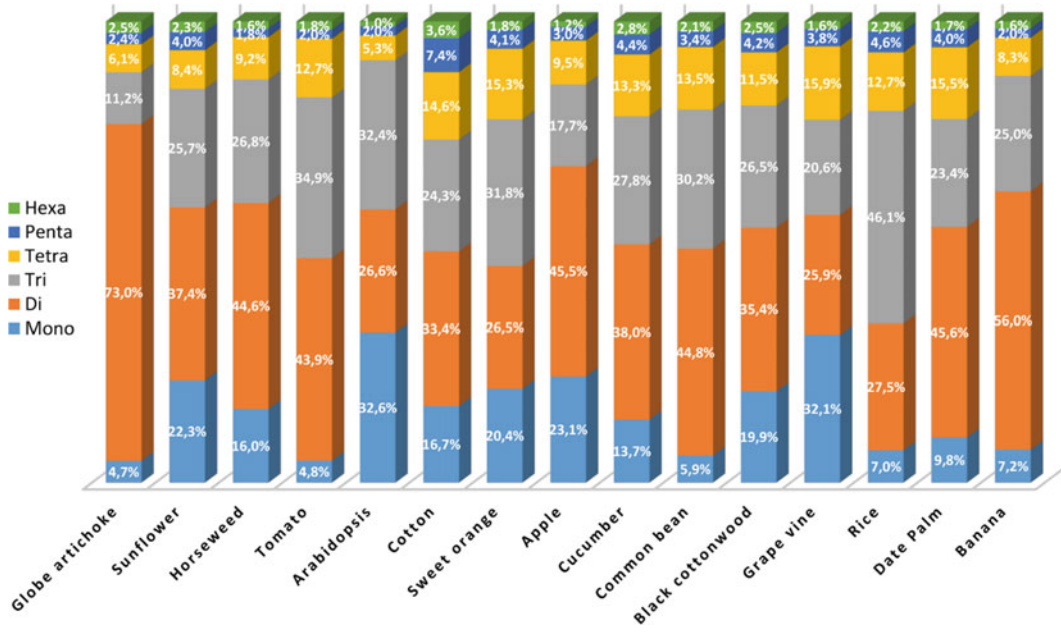


Fig. 13.4 The representation of different SSR motifs across 15 plant genomes. From Portis et al. (2016), used under CC BY 4.0

while an abundance of the tetra- ACAT repeats was found only in globe artichoke and sunflower. The hexanucleotides ACGTAG/ACGTCT and AGGGCC/CCCTGG were unique to globe artichoke, while AGCGAT/ATCGCT was restricted to the three Compositae species globe artichoke, sunflower, and horseweed.

References

Buels R, Yao E, Diesh CM, Hayes RD, Munoz-Torres M, Helt G, Goodstein DM, Elsik CG, Lewis SE, Stein L, Holmes IH (2016) JBrowse: a dynamic web platform for genome visualization and analysis. Genome Biol 17:66

Cavagnaro PF, Senalik DA, Yang L, Simon PW, Harkins TT, Kodira CD, Huang S, Weng Y (2010) Genome-wide characterization of simple sequence repeats in cucumber (*Cucumis sativus* L.). BMC Genomics 11:569

Kujur A, Bajaj D, Saxena MS, Tripathi S, Upadhyaya HD, Gowda CL, Singh S, Jain M, Tyagi AK, Parida SK (2013) Functionally relevant microsatellite markers from chickpea transcription factor genes for efficient genotyping applications and trait association mapping. DNA Res Int J Rapid Publ Rep Genes Genomes 20:355–374

Liu W, Jia X, Liu Z, Zhang Z, Wang Y, Liu Z, Xie W (2015) Development and characterization of transcription factor gene-derived microsatellite (TFGM) markers in *Medicago truncatula* and their transferability in leguminous and non-leguminous species. Molecules 20:8759–8771

Martin P, Makepeace K, Hill SA, Hood DW, Moxon ER (2002) Microsatellite instability regulates transcription factor binding and gene expression. PNAS 102:3800–3804

Morgante M, Hanafey M, Powell W (2002) Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. Nat Genet 30:194–200

Mun JH, Kim DJ, Choi HK, Gish J, Debellé F, Mudge J, Denny R, Endré G, Saurat O, Dudez AM, Kiss GB, Roe B, Young ND, Cook DR (2006) Distribution of microsatellites in the genome of *Medicago truncatula*: a resource of genetic markers that integrate genetic and physical maps. Genetics 172:2541–2555

Portis E, Portis F, Valente L, Moglia A, Barchi L, Lanteri S, Acquadro A (2016) A genome-wide survey of the microsatellite content of the globe artichoke genome and the development of a web-based database. PLoS ONE 11(9):e0162841

Shi J, Huang S, Fu D, Yu J, Wang X, Hua W et al (2013) Evolutionary Dynamics of microsatellite distribution in plants: insight from the comparison of sequenced Brassica, Arabidopsis and other angiosperm species. PLoS ONE 8:e59988

Stein LD (2013) Using GBrowse 2.0 to visualize and share next-generation sequence data. Brief Bioinform 14:162–171

Toth G, Gaspari Z, Jurka J (2000) Microsatellites in different eukaryotic genomes: survey and analysis. Genome Res 10:967–981

Varshney RK, Graner A, Sorrells ME (2005) Genic microsatellite markers in plants: features and applications. Trend Biotech 23:48–55

Yu JK, Paik H, Choi JP, Han JH, Choe JK, Hur CG (2010) Functional domain marker (FDM): an in silico demonstration in Solanaceae using simple sequence repeats (SSRs). Plant Mol. Biol. Rep. 28: 352–356

Zhang L, Zuo K, Zhang F, Cao Y, Wang J, Zhang Y, Sun X, Tang K (2006) Conservation of noncoding microsatellites in plants: implication for gene regulation. BMC Genom 7:323

Future Prospects

14

Alberto Acquadro, Ezio Portis and Sergio Lanteri

Abstract

The availability of the genome sequence of globe artichoke and cultivated cardoon, and the development of physical maps in both taxa, have provided new opportunities to study genetic variation between varietal types to an unprecedented scale, thus contributing to dissect the path from sequence variation to phenotype. Furthermore, they have also provided new tools for performing phylogenetic studies within the *Asteraceae* family. However, genetic and genomic studies in *Cynara cardunculus* are just at the beginning. This chapter provides food for thought on which are the main problems to be addressed and the main objectives that should be pursued in the near future.

Recently, the genome sequence of globe artichoke as well as the one of the cultivated cardoon have been made available (Scaglione et al. 2016, Chap. 9), and transcriptome sequencing information (Scaglione et al. 2009, 2012) has guided all the gene prediction exercises during genome annotation. Thanks to the previous construction of molecular maps of both taxa, the genome sequences have been anchored to chromosomes allowing the development of physical maps (Acquadro et al. 2017, Chap. 12). In globe artichoke, studies have been already performed to locate genes and QTL for yield and morphological traits on genetic maps (Portis et al. 2014, 2015, Chap. 6), however, with the goal to integrate QTLs analyses into breeding programs, the power, precision and accuracy of the QTL effects should be confirmed in larger numbers of progenies and multiple genetic backgrounds also through genome-wide association studies (GWAS). The availability of the genome sequences has already allowed to develop thousands of microsatellite markers uniformly distributed in the genome (Portis et al. 2016, Chap. 13), and by aligning to the reference genome data obtained from low coverage resequencing of globe artichoke genotypes, millions of SNPs have been spotted (Acquadro et al. 2017), enabling to identify the genetic variation between and within globe artichoke varietal types to an unprecedented scale. The future resequencing of genomic regions or genes in individuals of different landraces/varietal types will contribute to dissect the path from sequence variation to phenotype, providing "training sets" for the establishment of marker assisted and genomic selection models. Likewise, sequence data from RNA samples will contribute to discover new RNA species, to measure levels of gene expression and to study the transcriptional state of different cells or tissues also in

A. Acquadro · E. Portis · S. Lanteri (✉)
DISAFA Plant Genetics and Breeding, University of Torino, Largo P. Braccini 2, 10095 Grugliasco, Torino, Italy
e-mail: sergio.lanteri@unito.it

E. Portis et al. (eds.), *The Globe Artichoke Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-030-20012-1_14

response to biotic and abiotic stresses. Eventually, the DNA sequences of globe artichoke and cultivated cardoon will allow future remarkable progresses in phylogeny studies within the genus *Cynara* as well as within the Compositae (Asteraceae) family, which accounts for ~ 10% of all angiosperm species. Comparative studies based on genomic or transcriptome data will enable the identification of gene orthologies as well as perform synteny studies, thus to infer on polyploidization events and chromosomal rearrangements occurred during evolution and speciation. Recently, on the basis of gene sequences obtained from publicly available transcriptome datasets, including *Cynara cardunculus*, a robust Asteraceae phylogeny covering 73 species has been reconstructed (Huang et al. 2016) and the species divergence times and dating of whole genome duplications deduced.

At present, open-pollinated seed-propagated varieties and F_1 hybrids, which due to the marked inbreeding depression characterizing the species are not quite as uniform as the ones obtained from crosses between highly homozygous inbreds, are becoming popular in cultivation. This is mainly due to the reduced cost of planting and diffusion of virus and fungal infections, as well as the possibility to turn globe artichoke from perennial to annual crop. Nevertheless, mostly in the Mediterranean region, a wide number of varietal types are still vegetatively propagated and subjected to risk of extinction. It is thus necessary to increase efforts in developing strategies to convert current vegetatively propagated landraces into seed-propagated varieties to reduce the risk of losing a huge amount of germplasm. At present, studies conducted in this regard are rather limited and based on the application of selection schemes supported by a limited number of microsatellites and AFLP markers, however they have highlighted the possibility to identify genotypes which when inter-crossed guarantee a high and a stable production (Mauromicale et al. 2018).

Genome manipulation technology is an emerging field which has already caused a revolution in plant breeding. However, some uncertainties and concerns are associated with GM technology since copies of a gene may be integrated, additional fragments inserted, and gene sequences rearranged or deleted, which may result in lack of operation or instability of the genes introduced or their interference with other gene functions. Recently, new approaches defined "targeted genome engineering" allow the introduction of modifications in a specific region of the genome and may offer a valid alternative to previous approaches of genome manipulation. These techniques pave the path to address a wide range of goals also in globe artichoke and related taxa, not only to improve the productivity and quality of the production, but also resistance to biotic and abiotic stresses as well as to investigate the fundamental roots of its biology. In particular, the emergent Clustered Regularly Interspaced Short Palindromic Repeats associated to Cas9 protein (CRISPR/Cas9) technology is expected to play a key role in future efforts to improve crop traits. By CRISPR, it is possible to induce point mutations in a target sequence (also many simultaneously), or it can be possible to induce homology-directed recombination (HDR) to introduce new genetic variants in a specific region of the genome (Schaeffer and Nakata 2015). Unfortunately, at present, the possibility to apply this technique in globe artichoke is limited by the lack of an efficient technique for the regeneration of transformed plants from *calli*, and future efforts are needed for developing appropriate protocols. To overcome obstacles of in vitro regeneration would even make feasible the use of genome editing technique for the manipulation of protoplasts (Woo et al. 2015; Subburaj et al. 2016). The establishment of a cell suspension culture in globe artichoke by Ordas et al. (1991) paved the way for isolation and culture of protoplasts; however, the bottleneck of protoplast-to-plant regeneration still persists. Efforts should also be put in place to evaluate the potential of novel transformation platform technologies such as pollen magnetofection (Xiang et al. 2018). This technique aims to directly produce transgenic seeds without regeneration and represents a culture-free, simple, and fast method. In this system, in the

presence of a magnetic field, exogenous DNA loaded with magnetic nanoparticles can be delivered into pollen with the potential to facilitate the development of genetically modified varieties in recalcitrant species such as *Cynara cardunculus*.

With CRISPR/Cas9 technology, it is today possible to guide and reprogram plant development. The last frontier in globe artichoke could be the production of "cloned seeds" from original high-performance genotypes. This could be addressed by mimicking the approach adopted in rice (Khanday et al. 2019), namely by disabling some genes crucial for meiosis and triggering the development of a seed embryo using the "Baby Boom 1" gene without the needing of fertilization.

References

Acquadro A, Barchi L, Portis E, Mangino G, Valentino D, Mauromicale G, Lanteri S (2017) Genome reconstruction in *Cynara cardunculus* taxa gains access to chromosome-scale DNA variation. Sci Rep 7:5617. https://doi.org/10.1038/s41598-017-05085-7

Huang CH, Zhang C, Liu M, Hu Y, Gao T, Qi J, Ma H (2016 Nov) Multiple polyploidization events across Asteraceae with two nested events in the early history revealed by nuclear phylogenomics. Mol Biol Evol 33 (11):2820–2835

Khanday I, Skinner D, Yang B, Mercier R, Sundaresan V (2019) A male-expressed rice embryogenic trigger redirected for asexual propagation through seeds. Nature 565:91–95

Mauromicale G, Portis E, Acquadro A, Lo Monaco A, Pesce GR, Lanteri S (2018) An integrated model to accelerate the development of seed-propagated varieties of globe artichoke. Crop Breed Appl Biotechnol 18(1):72–80

Ordas RJ, Tavazza R, Ancora G (1991) Callus formation from isolated globe artichoke (*Cynara scolymus* L.) suspension protoplast. Plant Sci 77:253–259

Portis E, Mauro RP, Acquadro A, Moglia A, Mauromicale G, Lanteri S (2015) The inheritance of bract pigmentation and fleshy thorns on the globe artichoke capitulum. Euphytica 206(2):523–531

Portis E, Mauro RP, Barchi L, Acquadro A, Mauromicale G, Lanteri S (2014) Mapping yield-associated QTL in globe artichoke. Mol Breed 34(2):615–630

Portis E, Portis F, Valente L, Moglia A, Barchi L, Lanteri S, Acquadro A (2016) A genome-wide survey of the microsatellite content of the globe artichoke genome and the development of a web-based database. PLoS ONE 11(9):e0162841

Scaglione D, Acquadro A, Portis E, Taylor CA, Lanteri S, Knapp SJ (2009) Ontology and diversity of transcript-associated microsatellites mined from globe artichoke EST database. BMC Genom 10:454

Scaglione D, Lanteri S, Acquadro A, Lai Z, Knapp SJ, Rieseberg L, Portis E (2012) Large-scale transcriptome characterization and mass discovery of SNPs in globe artichoke and its related taxa. Plant Biotechnol J 10:956–969. https://doi.org/10.1111/j.1467-7652.2012.00725.x

Scaglione D, Reyes-Chin-Wo S, Acquadro A, Froenicke L, Portis E, Beitel C, Tirone M, Mauro R, Lo Monaco A, Mauromicale G, Faccioli P, Cattivelli L, Rieseberg L, Michelmore R, Lanteri S (2016) The genome sequence of the outbreeding globe artichoke constructed de novo incorporating a phase-aware low-pass sequencing strategy of F_1 progeny. Scientific Reports 6:19427

Schaeffer SM, Nakata PA (2015) CRISPR/Cas9-mediated genome editing and gene replacement in plants: transitioning from lab to field. Plant Sci 240:130–142

Subburaj S, Chung SJ, Lee C, Ryu SM, Kim DH, Kim JS (2016) Site-directed mutagenesis in Petunia x hybrida protoplast system using direct delivery of purified recombinant Cas9 ribonucleoproteins. Plant Cell Rep 35:1535–1544

Woo JW, Kim J, Kwon SI, Corvalan C, Cho SW, Kim H et al (2015) DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. Nat Biotechnol 33:1162–1164

Xiang Z, Zhigang M, Yan W, Wenjie C, Changjiao S, Bo C, Jinhui C, Manli Y, Zhanghua Z, Sandui G, Dan L, Jerry QC, Rui Z, Haixin C (2018) Pollen magnetofection for genetic modification with magnetic nanoparticles as gene carriers. Nature Plants 3:956–964 (2017)