Compendium of Plant Genomes
Series Editor: Chittaranjan Kole

Nirala Ramchiary · Chittaranjan Kole *Editors*

The Capsicum Genome

Springer

Compendium of Plant Genomes

Series editor

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

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

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

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

Nirala Ramchiary · Chittaranjan Kole
Editors

The Capsicum Genome

Springer

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

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

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

Library of Congress Control Number: 2018967743

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

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

Chittaranjan Kole

Preface to the Series

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

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

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

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

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

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

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

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

Plant genome science has emerged as an important subject in academia, and the present compendium of plant genomes will be highly useful both to students and teaching faculties. Most importantly, research scientists involved in genomics research will have access to systematic deliberations on the plant genomes of their interest. Elucidation of plant genomes is of interest not only for the geneticists and breeders, but also for practitioners of an array of plant science disciplines, such as taxonomy, evolution, cytology, physiology, pathology, entomology, nematology, crop production, biochemistry, and obviously bioinformatics. It must be mentioned that information regarding each plant genome is ever-growing. The contents of the volumes of this compendium are, therefore, focusing on the basic aspects of the genomes and their utility. They include information on the academic and/or economic importance of the plants, description of their genomes from a molecular genetic and cytogenetic point of view, and the genomic resources developed. Detailed deliberations focus on the background history of the national and international genome initiatives, public and private partners involved, strategies and genomic resources and tools utilized, enumeration on the sequences and their assembly, repetitive sequences, gene annotation, and genome duplication. In addition, synteny with other sequences, comparison of gene families, and, most importantly, the potential of the genome sequence information for gene pool characterization through genotyping by sequencing (GBS) and genetic improvement of crop plants have been described. As expected, there is a lot of variation of these topics in the volumes based on the information available on the crop, model, or reference plants.

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

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

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

Kalyani, India Chittaranjan Kole

Preface to the Volume

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

Despite so much of economic importance, the study toward the identification of genes or quantitative trait loci governing fruit traits (size, shape, and texture), beneficial metabolites, nutrient elements uptake, physiological traits, biotic and abiotic stress tolerance is still limited in capsicum. However, many of the genes/QTLs for those traits have been identified in tomato crop, considered as the model for fleshy fruit plants, which belongs to the same family Solanaceae as capsicum. This relatively less progress might be due to the complexity of capsicum genome, which is approximately about 3.5 Gigabase (Gb) in size, compared to the tomato genome of 900 Megabase (Mb). Nevertheless, classical breeding efforts could improve yield, fruit morphology and metabolites content, and resistance to biotic and abiotic stress tolerance in capsicum. Furthermore, advances in molecular biology, and advent of high-throughput genome and transcriptome sequencing technologies enhanced our understating of the capsicum genome structure and function through the development of genetic maps with different molecular markers, dissection of quantitative trait loci underlying economically important traits and subsequently to identify a few genes causing trait variations. And recently, with the help of advanced high-throughput genome sequencing technologies, the whole-genome sequences (both nuclear and organellar) of *Capsicum* species have been reported. The complete sequence of chloroplast genome of *Capsicum annuum*, *C. chinense*, *C. frutescens*, *C. tovarii*, *C. chacoense*, *C. baccatum*, *C. galapagoense, C. eximium* and *C. lycianthoids*, and mitochondrial genome of *C. annuum* has been reported since 2012, while the complete nuclear genome of *C. annuum*, *C. baccatum*, and *C chinense* has been reported since 2014, respectively. Furthermore, the identification of noncoding RNAs and whole genome cytosine methylation which directly or indirectly regulates the gene expression governing the traits of interest in the capsicum genome are being reported.

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

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

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

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

New Delhi, India

Dr. Nirala Ramchiary
Prof. Chittaranjan Kole

Contents

Abbreviations

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

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

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

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

The Capsicum Crop: An Introduction

1

Pasquale Tripodi and Sanjeet Kumar

Abstract

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

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

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

1.1 Origin and Diffusion

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

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

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

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

1.2 Economic and Culinary Importance

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

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

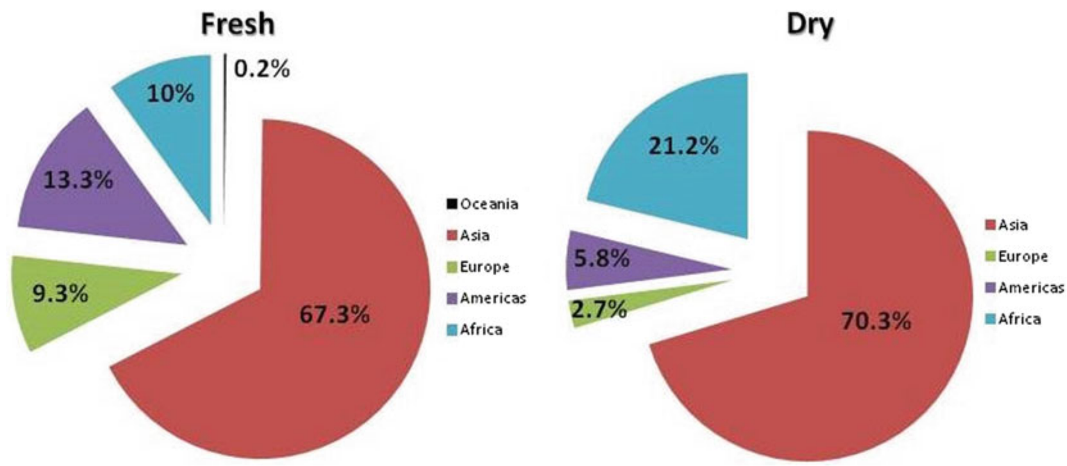

Fig. 1.1 Production share of dry and fresh pepper by region (FAOSTAT 2017)

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

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

1.3 The Properties of Pepper

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

Other than food uses are recognized as active ingredient in cosmetics, pharmaceuticals, and pest management (Bosland and Votava 1999). The extractable colors from fruits due to the presence of compounds unique in pepper such as capsanthin, capsorubin, and cryptocapsin are extensively used in the food processing industry as natural colorant for a wide range of products such meats, cheeses, and other foods. Non-food uses include (a) coloring and flavoring agents, from oleoresins (carotenoids) extracts or powder, as example, paprika powder can be used to inhibit lipid oxidation of pork meat while oleoresin is used to enhance physical and sensory properties of food products (Baenas et al. 2019); (b) ethno-botanical/traditional medicine, from fruit extracts and powders (pungent fruits); (c) modern medicine/pharmaceuticals, from extracts of capsaicinoids and carotenoids which can exert analgesic, antimicrobial, antioxidant, and anti-inflammatory effects; (d) insecticides/repellents and antibacterial effect from capsaicinoids extracts and organic acids (i.e., cinnamic, coumaric, ferulic, and caffeic); (e) spiritual, using whole fruits, e.g., "ristras"; (f) ornamental, using whole plants or fruits; (g) defense/punishment, using capsaicin extracts/or powder (Kumar et al. 2006). The use in cosmesis is favored by the presence of natural compounds which allow to avoid allergies and other side effects and are addressed to protect skin oxidative and UVA-mediated damage having thanks to the anti-wrinkle action and fighting against free radicals (Baenas et al. 2019). The industrial preparations are based on oleoresins rich of the above-mentioned bioactive compounds. Finally, among the most curious aspects of the *Capsicum* genus, there is certainly the rampant interest of many, searching and collecting, even in urban contexts, different species, characterized by a wide variety traits, as well as ornamental, aesthetically appreciated or rare varieties. This is evident in the rise of associations and websites dedicated to the subject.

1.4 Genetic Resources and Breeding

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

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

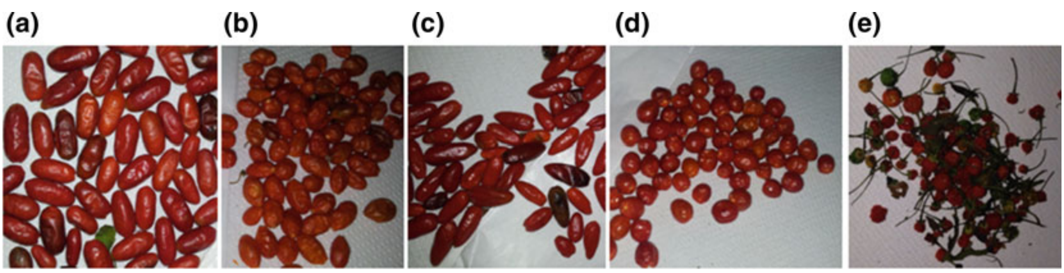

Fig. 1.3 Mature fruits of wild *Capsicum* species: **a** *C. chacoense*, **b** *C. praetermissum*, **c** *C. eximium*, **d** *C. annuum* var. *glabriusculum*, and **e** *C. flexuosum*

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

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

(continued)

Table 1.1 (continued)

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

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

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

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

References

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

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

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

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