Compendium of Plant Genomes
Series Editor: Chittaranjan Kole

Takashi Onozaki
Masafumi Yagi *Editors*

The Carnation 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 Prof. C. Kole, Series Editor, at ckoleorg@gmail.com

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

Takashi Onozaki • Masafumi Yagi
Editors

The Carnation Genome

Springer

Editors
Takashi Onozaki
Institute of Vegetable
and Floriculture Science
NARO
Tsukuba, Ibaraki, Japan

Masafumi Yagi
Institute of Vegetable
and Floriculture Science
NARO
Tsukuba, Ibaraki, Japan

ISSN 2199-4781 ISSN 2199-479X (electronic)
Compendium of Plant Genomes
ISBN 978-981-15-8260-8 ISBN 978-981-15-8261-5 (eBook)
https://doi.org/10.1007/978-981-15-8261-5

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

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, pheno- typic 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

Carnation (*Dianthus caryophyllus* L.) is one of the most important ornamental flowers in the world along with chrysanthemum and rose. A very complex hybridization process lies behind the modern carnation cultivars owing to the long history of breeding. The genome of carnation was sequenced by a Japanese research team during the 2013 end. Carnation has been genetically improved over the years and there are various types of flower colors, shapes, patterns, and sizes. The molecular mechanism of anthocyanin synthesis and transposable elements causing the diversity have been well studied. In addition, the breeding and physiological research for improving flower vase life, which is one of the most important traits in ornamentals, have been aggressively carried out in carnation as a model of ethylene susceptible flowers. To improve the selection efficiency, genomic analysis tools including DNA markers and genetic linkage maps have been developed. In carnation, mutant cultivars contributed to the expansion of flower colors and shapes, thus mutation breeding technology such as ion beams irradiation has been developed. Moreover, carnation is a scarce ornamental in which genetic engineering technology has been put into practical use, and the blue-violet genetically modified carnation has been widely distributed in the world. In this book, we summarize the recent progress in carnation genomic research for large-scale transcriptome analysis, draft genome sequence, DNA markers, and genome mapping. We also report the flower color, mutations, flower vase life, interspecific hybridization, fragrance for carnation, and discuss the future prospects in carnation genome researches.

Prof. Chittaranjan Kole, Series Editor of the book series entitled, Compendium of Plant Genomes recommended us to edit a book on *The Carnation Genome* and helped us all along. For publishing this book, we have been supported by Nareshkumar Mani, Praveen Anand, Fumiko Yamaguchi and Yuko Matsumoto of Springer, We are much grateful for the helpful suggestions and encouragements.

Tsukuba, Japan

Takashi Onozaki
Masafumi Yagi

Contents

Contributors

Yutaka Abe Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, Koganei, Tokyo, Japan;
Division of Food Additives, Division of Pharmacognosy, Phytochemistry and Narcotics, National Institute of Health Sciences, Kawasaki, Kanagawa, Japan

Steve Chandler School of Applied Sciences, Royal Melbourne Institute of Technology, RMIT University, Bundoora, VIC, Australia

Yoshihiro Hase Takasaki Advanced Radiation Research Institute, National Institutes for Quantum and Radiological Science and Technology (QST), Takasaki, Gunma, Japan

Kaoru Higuchi Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, Koganei, Tokyo, Japan

Hideki Hirakawa Facility for Genome Informatics, Kazusa DNA Research Institute, Kisarazu, Chiba, Japan

Makiko Hotta Aichi Agricultural Research Center, Nagakute, Aichi, Japan

Luna Iijima Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, Koganei, Tokyo, Japan

Yoshio Itoh Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, Koganei, Tokyo, Japan

Kyutaro Kishimoto Institute of Vegetable and Floriculture Science, NARO, Tsukuba, Ibaraki, Japan

Sanae Kishimoto Institute of Vegetable and Floriculture Science, NARO, Tsukuba, Ibaraki, Japan

Yuki Matsuba Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, Koganei, Tokyo, Japan;
North American Crop Protection Division, BASF, Durham, NC, USA

Taira Miyahara Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, Koganei, Tokyo, Japan;
Graduate School of Horticulture, Chiba University, Chiba, Chiba, Japan

Noriko Nakamura Suntory Global Innovation Center Ltd, Soraku-gun, Kyoto, Japan

Masayoshi Nakayama Institute of Vegetable and Floriculture Science, NARO, Tsukuba, Ibaraki, Japan

Yuzo Nishizaki Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, Koganei, Tokyo, Japan;
Division of Food Additives, Division of Pharmacognosy, Phytochemistry and Narcotics, National Institute of Health Sciences, Kawasaki, Kanagawa, Japan

Mikio Nimura Aichi Agricultural Research Center, Nagakute, Aichi, Japan

Jun Ogata Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, Koganei, Tokyo, Japan;
Division of Food Additives, Division of Pharmacognosy, Phytochemistry and Narcotics, National Institute of Health Sciences, Kawasaki, Kanagawa, Japan

Akemi Ohmiya Institute of Vegetable and Floriculture Science, NARO, Tsukuba, Ibaraki, Japan

Masachika Okamura Central Laboratories for Key Technologies, Kirin Company, Yokohama, Kanagawa, Japan

Yoshihiro Ozeki Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, Koganei, Tokyo, Japan

Nobuhiro Sasaki Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, Koganei, Tokyo, Japan;
Department of Food and Life Sciences, Toyo University, Ora-gun, Gunma, Japan

Shigeru Satoh Faculty of Agriculture, Ryukoku University, Otsu, Shiga, Japan

Yukiko Shinbo Suntory Global Innovation Center Ltd, Soraku-gun, Kyoto, Japan

Akane Suzuki-Wagner Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, Koganei, Tokyo, Japan

Takamasa Suzuki Department of Biological Chemistry, College of Bioscience and Biotechnology, Chubu University, Kasugai, Aichi, Japan

Yoshikazu Tanaka Suntory Global Innovation Center Ltd, Soraku-gun, Kyoto, Japan

Koji Tanase Institute of Vegetable and Floriculture Science, NARO, Tsukuba, Ibaraki, Japan

Takashi Tsujimoto Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, Koganei, Tokyo, Japan

Masafumi Yagi Institute of Vegetable and Floriculture Science, NARO, Tsukuba, Ibaraki, Japan

Abbreviations

2,4-PDCA	2,4-pyridinedicarboxylic acid
3-PCA	3-pyridinecarboxylic acid
4CL	4-coumarate:CoA ligase
AA5GT	Acyl-glucose dependent anthocyanin 3-*O*-glucoside 5-glucosyltransferase
AARC	Aichi Agricultural Research Center
ACC	1-aminocyclopropane-1-carboxylate
ACO	ACC oxidase
ACS	ACC synthase
AFLP	Amplified fragment length polymorphism
AHCT	Anthocyanin *O*-hydroxycinnamoyltransferase
ALA	5-aminolevulinic acid
AMalT	Malyl-glucose dependent anthocyanin acyltransferase
ANS	Anthocyanidin synthase
APM	Amiprophos-methyl
AVI	Anthocyanic vacuolar inclusion
BAMT	Benzoic acid carboxyl methyltransferase
BEAT	Benzylalcohol *O*-acetyltransferase
bHLH	Basic helix-loop-helix
BSMT	Benzoic acid/salicylic acid methyltransferase
C4H	Cinnamate 4-hydroxylase
CAO	Chlorophyllide *a* oxygenase
CARACC3	Carnation ACS 3
CARAO1	Carnation ACO 1
CARAS1	Carnation AC*S* 1
CAS	Caffeine synthase
CBW	Carnation bacterial wilt
CCDs	Carotenoid cleavage dioxygenases
cDNA	Complimentary DNA
CHGT	UDP-glucose dependent chalcone glucosyltransferase
CHI	Chalcone isomerase
CHS	Chalcone synthase
CHYB	Carotenoid β-ring hydroxylase
CML37	Calmodulin-like protein 37
COL16	Constans-like 16
CP	Cysteine proteinase

CRTISO	Carotenoid isomerase
CWP	Cerise Westpearl
DAT	Deacetylvindoline 4-*O*-acetultransferase
DB	Database
DD	Differential display
ddRAD-Seq	Double digest restriction site–associated DNA sequencing
DFR	Dihydroflavonol 4-redulctase
DMAPP	Dimethylallyldiphosphate
DREB3	Dehydration-responsive element binding protein 3
EBGs	Early anthocyanin biosynthetic genes
ER	Ethylene receptor
EST	Expressed sequence tag
F3’5’H	Flavonoid 3’,5’-hydroxylase
F3H	Flavone 3-hydroxylase
FONS	Fully-open and non-senescent
GA	Gibberellic acid
GGPP	Geranylgeranyl diphosphate
GH1	Glucoside hydrolase family 1
GMO	Genetically modified organism
GO	Gene ontology
GST	Glutathione *S*-transferases
GT1	Family 1 glucosyltransferases
GWAS	Genome-wide association study
HCAR	7-hydroxymethyl-chlorophyll *a* reductase
HCBT	Anthranilate *N*-hydroxycinnamoyl/benzoyltransferase
HPLC	High-performance liquid chromatography
IAMT	Indole acetic acid methyltransferase
InDels	Insertions and deletions
INSDC	International Nucleotide Sequence Database Collaboration
IPI	IPP isomerase
IPP	Isopentenyl diphosphate
JMTs	Jasmonic acid carboxyl methyltransferases
KOG	Eukaryotic Orthologous Groups
LBG	Late biosynthetic gene
LCYB	Lycopene β-cyclase
LCYE	Lycopene ε-cyclase
LG	Linkage group
LOD	Logarithm of odds
LPB	Light Pink Barbara
LR	Left border
LTR	Long terminal repeat
MAS	Marker-assisted selection
MATE	Multidrug and toxic compound extrusion
MEP	Methylerythritol 4-phosphate
MgPMT	Mg-protoporphyrin IX methyltransferase
MP	Mate-pair
MVA	Mevalonate
NARO	National Agriculture and Food Research Organization

NGS	Next-generation sequencer/ next-generation sequencing
NIVFS	Institute of Vegetable and Floriculture Science, NARO
NOL	NYC1-like
NSY	Neoxanthin synthase
NYC1	Non-yellow coloring 1
ODS	Octadecylsilane
OMT	*O*-methyl transferase
ORF	Open reading frame
PAL1	Phenylalanine ammonia-lyase 1
PaO	Pheophorbide *a* oxygenase
PDCA	Pyridinedicarboxylic acid
PDS	Phytoene desaturase
PE	Paired-end
PGDBj	Plant Genome Database of Japan
PPH	Pheophytinase
PSY	Phytoene synthase
PYP1	Pale yellow petal 1
QTL	Quantitative trait locus
RAD	Restriction site–associated DNA
RAD-seq	RAD-sequencing
RAPD	Random amplified polymorphic DNA
RB	Right border
RCC	Red chlorophyll catabolite
RCCR	RCC reductase
RFLP	Restriction fragment length polymorphism
RNA-i	RNA-interference
RNA-seq	RNA-sequence
ROS	Reactive oxygen species
SAGE	Serial analysis of gene expression
SAM	*S*-adenosyl-l-methionine
SAMT	Salicylic acid carboxyl methyltransferase
SAT	1-*O*-sinapoyl-β-glucose:anthocyanin sinapoyltransferase
SCPL	Serine carboxypeptidase-like
SE	Single-end
SGR	*Stay-green*
SMT	1-*O*-sinapoyl-β-glucose:L-malate sinapoyltransferase
SNP	Single nucleotide polymorphism
SR	Senescence-related
SS1	Sucrose synthase 1
SSH	Suppression subtractive hybridization
SSR	Simple-sequence repeat
SST	1-*O*-sinapoyl-β-glucose:1-*O*-sinapoyl-β-glucose sinapoyltransferase
STP4	Sugar transport protein 4
STS	Sequence-tagged site
STS	Silver thiosulfate
SVs	Structural variants
T-DNA	Transfer-DNA

TE	Transposable element
THC	Tetrahydroxychalcone
TLC	Thin-layer chromatography
TPM	Transcripts per million
UA3GT	UDP-glucose dependent anthocyanin 3-O-glucosyltransferase
UGT	UDP-glucose dependent glycosyltransferase
VDE	Violaxanthin de-epoxidase
XES	Xanthophyll esterase
XET	Xyloglucan endotransglucosylase
XG	Xyloglucan
XGO	Xyloglucan oligosaccharide
XTH	Xyloglucan endotransglucosylase/hydrolase
ZDS	ζ-carotene desaturase
ZEP	Zeaxanthin epoxidase
Z-ISO	ζ-carotene isomerase

1 Draft Genome Sequence

Hideki Hirakawa

Abstract

The genome sequences of various kinds of plant species, such as cereals, vegetables, fruits, and flowers, have been determined. The genome sequence of carnation (*Dianthus caryophyllus*) cultivar 'Francesco' was first determined as a floricultural crop using next-generation sequencing data obtained by Illumina HiSeq and 454 GS FLX+ platforms. The genome size of the carnation was estimated as 638.7 Mb. The draft genome sequence DCA_r1.0 was determined by a combination of three kinds of *de novo* assemblers. The total and N50 lengths of the draft genome sequence were 568,887,315 bp and 60,737 bp, respectively. Of 56,382 transcripts identified by the gene prediction, 43,491 were determined to be genes without transposable elements. The genes were annotated by similarity searches against the UniProtKB/TrEMBL, TAIR10, and RAP-DB databases and by domain searches against the InterPro database. The genes were mapped onto the KEGG pathways and classified into gene ontology (GO) groups and EuKaryotic Orthologous Groups (KOG) of proteins. The genome data are available from the Carnation DB (https://plant1.kazusa.or.jp/carnation/). The 496 DNA markers of carnation, including 491 SSR markers, have been released from the Plant Genome DataBase Japan (PGDBj) (http://pgdbj.jp). As genome data and DNA markers accumulate, studies of the molecular genetics of carnation would be promoted, and molecular breeding would be made more effective.

1.1 Introduction

The genome sequence of *Arabidopsis thaliana* was first determined in 2000 as a model plant by an international collaboration (Arabidopsis Genome Initiative 2000). After that, the genome sequences of plant species were determined for rice (*Oryza sativa*) in 2002 (International Rice Genome Sequencing Project 2005), poplar (*Populus trichocarpa*) in 2006 (Tuskan et al. 2006), *Lotus japonicus* in 2006 (Sato et al. 2008), and so on. These genome sequences were determined by using a Sanger sequencer, such as ABI3730xl (Applied Biosystems, Foster City, CA, USA). In 2009, the genome sequence of cucumber (*Cucumis sativus*) (Huang et al. 2009) was determined by a combination of Sanger and Illumina platforms (Illumina) known as Next-Generation Sequencers (NGSs). The genome sequence of wild strawberry (*Fragaria vesca*) was determined by a combination of Sanger, 454,

H. Hirakawa (✉)
Facility for Genome Informatics, Kazusa DNA Research Institute, 2-6-7 Kazusakamatari, Kisarazu 292-0818, Chiba, Japan
e-mail: hh@kazusa.or.jp

T. Onozaki and M. Yagi (eds.), *The Carnation Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-981-15-8261-5_1