Compendium of Plant Genomes
Series Editor: Chittaranjan Kole

Chittaranjan Kole *Editor*

The *Catharanthus* Genome

Springer

Compendium of Plant Genomes

Series Editor

Chittaranjan Kole, Raja Ramanna Fellow, Department of Atomic Energy, Government of India, ICAR-National Institute for 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 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.

More information about this series at https://link.springer.com/bookseries/11805

Chittaranjan Kole
Editor

The *Catharanthus* Genome

Springer

Editor
Chittaranjan Kole
Raja Ramanna Fellow
Department of Atomic Energy
Government of India
ICAR-National Institute for Plant
Biotechnology
New Delhi, India

ISSN 2199-4781 ISSN 2199-479X (electronic)
Compendium of Plant Genomes
ISBN 978-3-030-89268-5 ISBN 978-3-030-89269-2 (eBook)
https://doi.org/10.1007/978-3-030-89269-2

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 F2 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 three 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 to both 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, particularly Dr. Christina Eckey and Dr. Jutta Lindenborn for the earlier set of volumes and presently Ing. Zuzana Bernhart for all their timely help and support.

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.

Chittaranjan Kole

Preface

Catharanthus roseus (L.) G. Don, a tropical perennial medicinal subshrub of the family Apocynaceae, is a source of pharmaceutically important indole alkaloids. *C. roseus*, commonly known as Madagascar periwinkle, Cape periwinkle, rosy periwinkle, sadabahar (in Hindi) and Chang Chung Hua (in Chinese), is a common ornamental in tropical and subtropical regions found growing in humid dry, hot and cool locations. It is now becoming increasingly popular system for the molecular and physiological studies specially for its secondary metabolism. This plant also enjoys the distinction of a taxon possessing the largest number of alkaloids in the plant kingdom. However, it is widely known as a miraculous medicinal plant containing an array of potentially bioactive phytochemicals including flavonoids, saponins and alkaloids of which the alkaloids are the main therapeutic chemical constituents. A large number, claimed to be over 100 to 400 in different reports, of alkaloids are present in Catharanthus, which have pharmaceutical, food additive and agrochemical properties. Of these, vinblastine and vincristine are well-known source of commercial anticancer drugs. The alkaloid ajmalicine present in the root is used in the treatment of high blood pressure, and other leaf alkaloids have antineoplastic activities. This subshrub has been used for hundreds of years to treat a variety of ailments and was a favorite ingredient of medical component within the middle ages because of the multifarious activities of its components including antitumor, antibacterial, antihyperglycemic, antihypertensive, diabetic wound healing activity, asthma, constipation and menstrual problems. Large demand, low yields of alkaloids and rapid life cycle of the species are prompting significant research efforts on *C. roseus*.

Catharanthus has been a leading plant in researches in several disciplines including basic botany, genetics, genomics, biochemistry, physiology, medicine and horticulture. Some of the useful features of *Arabidopsis* are shared by *C. roseus,* and since latter is most studied for its chemistry and pharmacology, it is an apparent choice for studies on the functional genetics of alkaloids. However, an array of novel concepts and strategies of the post-genomics era has facilitated identification of the therapeutically important genes and their functions and more importantly their monitoring and manipulations aiming at developing genomically designed plant types containing quantitatively and qualitatively improved anticancer alkaloids. This book on The Catharanthus genome will hopefully present the basic and applied information on this “plant for health.”

This book contains ten chapters covering various aspects. Chapter 1 of this book enumerates the taxonomy, distribution, nomenclature, habitat, morphology, anatomy, cytology, chemical constituents and their medicinal properties of Catharanthus. Chapter 2 discusses the plant-based cancer treatment with special reference to Catharanthus and different alkaloids present in the plant mainly vincristine and vinblastine and regulatory aspects of these anticancer herbal drugs. Chapter 3 deliberates on the systematic genetic improvement efforts in Catharanthus since its first identification in the Madagascar island forests, to its first cultivation at France and London, and its wide spread to tropical world. Driven by high commercial boost in development of ornamental cultivars for the landscape industry by the public sector plant breeders to its ongoing but limited continuous efforts in medicinal breeding for high concentration of monoterpene alkaloids supplemented with it targeted trait genetics and accomplishments. Somatic embryogenesis is a popular strategy of mass propagation particularly in horticultural plants. However, its success depends on the type of explants, growth promoters used and other regulating factors. These are explained in Chapter 4 along with details on somatic embryo development, structure and progress, i.e., maturation, germination of embryo. Chapter 5 presents an account of the effects of various media components and *in vitro* culture conditions governing alkaloid synthesis and a vivid depiction of alkaloid biosynthetic pathway in Catharanthus that could be made commercially viable.

The two major anticancer terpenoid indole alkaloids, vinblastine and vincristine, are naturally produced by Catharanthus. These are synthesized from vindoline via the catalysis of deacetylvindoline 4-*O*-acetyltransferase (DAT) in the last reaction of terpenoid indole alkaloid pathway. Chapter 6 presents the results of *Agrobacterium*-mediated genetic transformation for developing transgenic lines carrying the C. roseus DAT (CrDAT) gene to increase vincristine and alkaloid biosynthesis in the transgenic plants. Molecular breeding has emerged as a routine strategy to compliment traditional breeding. For this purpose, genetic linkage maps are constructed using different molecular markers and mapping populations, and genes and quantitative trait loci of interest are mapped to detect tightly linked molecular markers. Chapter 7 deliberates on these efforts made in Catharanthus along with prospects of marker-assisted breeding for improvement of the content of anticancer alkaloids. The chapter also deals with different omics approaches including transcriptomics, metabolomics and phenomics that could facilitate effective genomics-aided breeding.

Catharanthus is one of the few medicinal plants for which whole-genome sequencing has been accomplished. Information from genome sequencing not only provides the number and structure of genes but also depicts their phylogenetic relationships and evolutionary pathways. Besides, useful information on the function of the genes pave the way for formulation of strategies for their quantitative and qualitative improvement. The Chapter 8 enunciates the achievements of sequencing of nuclear and plastid genomes of Catharanthus and its various fundamental and applied implications. Chapter 9 deliberates on the biological pathways and regulatory genes and transcription factors

involved in biosynthesis of various terpenoid indole alkaloids in Catharanthus, functional genes and transcription factors involved in transport of terpenoid indole alkaloids and on the functional genomic tools for engineering the biosynthetic pathways in Catharanthus. Chapter 10 enumerates the achievements of genetic and genomics researches made so far in Catharanthus and highlights on the potential future research areas.

These chapters have been contributed by 40 eminent scientists from three countries including India, Mexico and Vietnam. I express my thanks and gratitude to them for their contributions.

Catharanthus has a special place in my heart. Firstly, I grew up watching this plant just beside a plant of holy basil, *Ocimum sanctum*, my mother used to worship! Secondly, it is the plant on which I had started my formal "genetic" work in a center of the Central Institute of Medicinal and Aromatic Plants of the Council of Scientific & Industrial Research of India at Bengaluru, India as an amateurish "researcher"! This is the source of my confidence to dare to contemplate editing this book and ultimately complete it. Hence, I will feel blessed if the readers find this book useful.

Chittaranjan Kole

Contents

Contributors

Rohit Sam Ajee Independent Researcher, Sagar, India

Apoorva Department of Botany, Center of Advance Study in Botany, Institute of Science, Banaras Hindu University (BHU), Varanasi, India

Yashika Bansal Cellular Differentiation and Molecular Genetics Section, Department of Botany, New Delhi, India

Ha Thi Bui Department of Biology, Thai Nguyen University of Medicine and Pharmacy, Thai Nguyen, Viet Nam

Mau Hoang Chu School of Biology, Thai Nguyen University of Education, Thai Nguyen, Viet Nam

Shweta Dwivedi Division of Plant Breeding and Genetic Resource Conservation, CSIR- Central Institute of Medicinal Plants (CIMAP), Lucknow, India

Bushra Ejaz Cellular Differentiation and Molecular Genetics Section, Department of Botany, New Delhi, India

Tripta Jhang Division of Plant Breeding and Genetic Resource Conservation, CSIR- Central Institute of Medicinal Plants (CIMAP), Lucknow, India

Shuchi Kaushik State Forensic Science Laboratory, Sagar, Madhya Pradesh, India

Sushil Kumar SKA Institute for Research, Education and Development (SKAIRED), New Delhi, India

Moien Qadir Malik Cellular Differentiation and Molecular Genetics Section, Department of Botany, New Delhi, India

Maneesha Mall CSIR-Central Institute of Medicinal and Aromatic Plants, Lucknow, India

Jyoti Mamgain Cellular Differentiation and Molecular Genetics Section, Department of Botany, New Delhi, India

Daya N. Mani CSIR-Central Institute of Medicinal and Aromatic Plants, Lucknow, India

A. Mujib Cellular Differentiation and Molecular Genetics Section, Department of Botany, New Delhi, India

Lan Thi Ngoc Nguyen School of Biology, Thai Nguyen University of Education, Thai Nguyen, Viet Nam

Quan Huu Nguyen School of Biology, Thai Nguyen University of Education, Thai Nguyen, Viet Nam

Krishna Kumar Rai Department of Botany, Center of Advance Study in Botany, Institute of Science, Banaras Hindu University (BHU), Varanasi, India;
Dr. Rajendra Prasad Central Agricultural University, Samastipur, Bihar, India

Sanjay Kumar Rai Dr. Rajendra Prasad Central Agricultural University, Samastipur, Bihar, India

Shashi Pandey Rai Department of Botany, Center of Advance Study in Botany, Institute of Science, Banaras Hindu University (BHU), Varanasi, India

Neetu Rani Department of Botany, Nirmala College, Ranchi University, Ranchi, Jharkhand, India

Saumya Shah Biotechnology Division, CSIR-Central Institute of Medicinal and Aromatic Plants, Lucknow, Uttar Pradesh, India

Neha Sharma Amity Institute of Biotechnology, Amity University, Noida, India

Sarika Sharma Division of Plant Breeding and Genetic Resource Conservation, CSIR- Central Institute of Medicinal Plants (CIMAP), Lucknow, India

Ajit K. Shasany CSIR-Central Institute of Medicinal and Aromatic Plants, Lucknow, India

Ajit Kumar Shasany ICAR-National Institute for Plant Biotechnology (NIPB), New Delhi, India

Ashutosh K. Shukla CSIR-Central Institute of Medicinal and Aromatic Plants, Lucknow, India

Indranil Singh Independent Researcher, Sagar, India

Pooja Singh CSIR-Central Institute of Medicinal and Aromatic Plants, Lucknow, India

Velusamy Sundaresan CSIR-Central Institute of Medicinal and Aromatic Plants, Research Centre, Bengaluru, India

Rukaya Syeed Cellular Differentiation and Molecular Genetics Section, Department of Botany, New Delhi, India

Divya Vashisth Biotechnology Division, CSIR-Central Institute of Medicinal and Aromatic Plants, Lucknow, Uttar Pradesh, India

Felipe A. Vázquez-Flota Unidad de Bioquímica y Biología Molecular de Plantas, Centro de Investigación Científica de Yucatán, Mérida, Yucatán, México

Lien Thi Kim Vu Institute of Theoretical and Applied Research, Duy Tan University, Hanoi, Viet Nam;
Faculty of Natural Sciences, Duy Tan University, Da Nang, Viet Nam

Seema Yadav CSIR-Central Institute of Medicinal and Aromatic Plants, Lucknow, India

Abbreviations

10HGO	10-Hydroxygeraniol oxidoreductase
16OMT	16-Hydroxytabersonine-16-O-methyltransferase
2,4-D	2,4-Dichlorophenoxyacetic acid
2D-DIGE	Two-dimensional difference gel electrophoresis
7DLS	7-Deoxyloganetic acid synthase
A	Ajmalicine
AACT	Acetoacetyl-CoA thiolase
ABA	Abscisic acid
ABC	ATP-binding cassette
ADME	Absorption, dispersion, metabolism and excretion
AFLP	Amplified fragment length polymorphism
AJM	Ajmalicine
Akt	Ak strain transforming
AP	Activator protein
AP2/ERF	APETALA2/ethylene-responsive factor
APM	Alkaloid producing media
AS	Anthranilate synthase
ATP	Adenosine triphosphate
AVLB	α-3',4'-Anhydrovinblastine
BAP	6-Benzylaminopurine
BCE	Before Common Era
bHLH	Basic helix-loop-helix protein
BIS	bHLH iridoid synthesis (transcription factor)
BSA	Bulked segregant analysis
BSLMM	Bayesian sparse linear mixed model
C	Catharanthine
cAMP	Cyclic adenosine monophosphate
CC BY	Creative Commons Attribution
cDNA-AFLP	Complementary DNA-AFLP
CHMIS-C	China's Comprehensive Herbal Medicine Information System for Cancer
CMK	4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase
COX-2	Cyclooxygenase 2 gene
CPR	Cytochrome P450 reductase
CR	Cathenamine reductase

CrDAT	*Catharanthus roseus* deacetylvindoline 4-O-acetyltransferase
CrGBF	*C. roseus* G box binding factor
CRISPR	Clustered regularly interspaced short palindromic repeats
croFGD	*C. roseus* Functional Genomics Database
CYP450	Cytochrome P 450
D17H	Desacetoxyvindoline-17-hydroxylase
D4H	Desacetoxyvindoline-4-hydroxylase
DAT	Deacetylvindoline-4-*O*-acetyltransferase
DL7H	7-Deoxyloganic acid-7-hydroxylase
DLGT	7-Deoxyloganetic acid glucosyltransferase
dsRNA	Double-stranded RNA
DVB	Double vindoline-bonded
DXPS	1-Deoxy-D-xylulose-5-phosphate synthase
DXR	1-Deoxy-D-xylulose-5-phosphate reductoisomerase
E. coli DH5α	*Escherichia coli* DH5 Alpha
EGFR	Epidermal growth factor receptor
EHRS	European Health Record System
EMMA	Efficient mixed-model association
EMMAX	Efficient mixed-model association eXpedite
EMS	Ethylmethane sulfonate
ERK	Extracellular-signal-regulated kinase
EST	Expressed sequence tag
F1	First filial (generation)
FaST-LMM	Factored spectrally transformed linear mixed model
FC	Flow cytometry
FDA	Food and Drug Administration (USA)
FM	Functional marker
G10H	Geraniol-10-hydroxylase
GA_3	Gibberellic acid
GACP	Guidelines of Good Agricultural and Collection Practices
GATA	GATA-binding protein
GC	Gas chromatography
GEMMA	Genome-wide efficient mixed model
GES	Geraniol-by-geraniol synthase
GES	Geraniol synthase
GI	Gradient index
GO	Gene ontology
GPPS	Geranyl pyrophosphate synthase
GTP	Guanosine triphosphate
GUS	β-Glucuronidase
GWAS	Genome-wide association studies
HDR	Hydroxymethylbutenyl-4-diphosphate reductase
HDS	Hydroxymethylbutenyl-4-diphosphate synthase
HIV	Human immunodeficiency virus
HMGR	Hydroxymethylglutaryl-CoA reductase
HMGS	Hydroxymethylglutaryl-CoA synthase

HPLC	High-pressure liquid chromatography
HRM	High-resolution melt analysis
IBA	Indole-3-butyric acid
IC_{50}	Half-maximal inhibitory concentration
IDI	Isopentenyl diphosphate isomerase
IDS	Iridoid synthase
IGR	Intergenic region
InDel	Insertion/deletion
IPAP	Internal phloem associated parenchyma
IPP	Isopentenyl diphosphate
IR	Inverted repeat
ISSR	Inter-simple sequence repeat
JAK-STAT	Janus kinase-STAT
Kn	Kinetin
LAMP	Loop-mediated isothermal amplification
LAMT	Loganic acid *O*-methyltransferase
LC	Liquid chromatography
LMM-Lasso	Linear mixed-model-Lasso
lncRNA	Long non-coding RNA
LSC	Large single copy
MAE	Microwave-assisted extraction
MAPK-ERK	Mitogen-activated protein kinase-ERK
MAS	Marker-assisted selection
MAT	Minovincine-O-acetyltransferase
MAT	Minovincinine-19-*O*-acetyltransferase
MATE	Multidrug and toxic compound extrusion
MCT	MEP cytidylyltransferase
MDR	Multidrug resistance
MECS	2-C-methyl-D-erythritol-2,4-cyclodiphosphate synthase
MeJA	Methyl jasmonate
MEP	Mevalonic acid pathway or methyl-D-erythritol-4 phosphate
MEP	2-C-Methyl-D-erythritol 4-phosphate
MFA	Metabolic flux analysis
MIA	Monoterpene indole alkaloid
miRNA	microRNA
MLMA	Model-based association analysis
MLMM	Multi-locus mixed model
MS	Murashige and Skoog (medium)
MS	Mass spectroscopy
MTIA	Monoterpene indole alkaloid
MTMM	Multi-trait mixed model
mTOR	Mammalian target of rapamycin
MVA	Mevalonate
NAA	Naphthalene acetic acid
NAC	No apical meristem

NCBI	National Center for Biotechnology Information
NCIM	National Collection of Industrial Microorganisms
NEU	N-Nitroso-N-ethyl urea
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
nm	Nearly median
NMR	Nuclear magnetic resonance
NMT	N-methyltransferase
NMT	16-methoxy-2,3-dihydro-3-hydroxy-tabersonine-*N*-methyltransferase
NPF2.9	Nitrate/peptide family strictosidine transporter 2.9
nsm	Nearly submedian
nst	Nearly subtelocentric
nt	Nucleotide
OMT	O-Methyl transferase
ORCA	Octadecanoid-responsive Catharanthus AP2-domain protein
P3D	Population parameters previously determined
PCR	Polymerase chain reaction
PDR	Pleiotropic drug resistance
PEG	Polyethylene glycol
PEP	Phosphoenolpyruvate
PF	Pink flower variety
P-gp	P-glycoprotein
PGR	Plant growth regulator
PI3K	Phosphoinositide-3-kinase
PIF	Phytochrome interacting factor
PnWBP	Peanut witches' broom
PRX	Peroxidase
PTP-1B	Protein tyrosine phosphatase-1B
QTL	Quantitative trait locus
QTLs	Quantitative trait loci
RAPD	Random amplified of polymorphic DNA
rCrDAT	Recombinant CrDAT
RFLP	Restriction fragment length polymorphism
RILs	Recombinant inbred lines
RNAi	RNA interference
rRNA	Ribosomal ribonucleic acid
SAC	Spindle assembly checkpoint
SDG	Strictosidine degluocosidase
SE	Somatic embryogenesis
SGD	Strictosidine β-D-glucosidase
SI	Symmetry index
SLS	Secologanin synthase
SNP	Single-nucleotide polymorphism
SSC	Small single copy
SSR	Simple sequence repeat