

Chittaranjan Kole *Editor*

Genomic Designing of Climate-Smart Oilseed Crops

 Springer

Genomic Designing of Climate-Smart Oilseed Crops

Chittaranjan Kole
Editor

Genomic Designing of Climate-Smart Oilseed Crops

 Springer

Editor
Chittaranjan Kole
ICAR-National Research Center on Plant
Biotechnology
New Delhi, Delhi, India

ISBN 978-3-319-93535-5 ISBN 978-3-319-93536-2 (eBook)
<https://doi.org/10.1007/978-3-319-93536-2>

Library of Congress Control Number: 2018965428

© Springer Nature Switzerland AG 2019

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

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



Prof. Thomas C. Osborn

*Vegetable Analytics and Pipeline Design Lead
at Bayer Crop Science in St. Louis, MO, USA
Previously Professor and Bascom Chair in
Agronomy in the University of Wisconsin at
Madison,*

*Where I worked with him on the Brassicaceae
system and learnt how to develop novel
concepts of plant molecular genetics and
side-by-side generate genetic resources for
crop improvement.*

Preface

The last 120 years have witnessed a remarkable evolution in the science and art of plant breeding culminating in quite a revolution in the second decade of the twenty-first century! A number of novel concepts, strategies, techniques, and tools have emerged from time to time over this period, and some of them deserve to be termed as milestones. Traditional plant breeding, immediately following the rediscovery of the laws of inheritance, has been playing a spectacular role in the development of innumerable varieties in almost all crops during this entire period. Mention must be made on the corn hybrids, rust-resistant wheat, and obviously the high-yielding varieties in wheat and rice that ushered the so-called green revolution. However, the methods of selection, hybridization, mutation, and polyploidy employed in traditional breeding during this period relied solely on the perceivable phenotypic characters. But most, if not all, of the economic characters in crops are governed by polygenes which are highly influenced by environment fluctuations, and hence, phenotype-based breeding for these traits has hardly been effective.

Historical discovery of DNA structure and replication in 1953 was followed by a series of discoveries in the 1960s and 1970s that paved the way for recombinant DNA technology in 1973 facilitating the detection of a number of DNA markers in 1980 onwards and their utilization in construction of genetic linkage maps and mapping of genes governing the simply inherited traits and quantitative trait loci controlling the polygenic characters in a series of crop plants starting with tomato, maize, and rice. Thus, new crop improvement technique called as molecular breeding started in later part of the twentieth century. On the other hand, genetic engineering made modification of crops for target traits by transferring alien genes, for example, the *Bt* gene from the bacteria *Bacillus thuringiensis*. A large number of genetically modified crop varieties have thus been developed starting with the commercialization of “flavr Savr” tomato in 1994.

Meantime, the manual DNA sequencing methodology of 1977 was being improved with regard to speed, cost-effectiveness, and automation. The first-generation sequencing technology led to the whole genome sequencing of *Arabidopsis* in 2000 and followed by rice in 2002. The next-generation sequencing technologies were available over time and used for sequencing of genomes of many

other model and crop plants. Genomes, both nuclear and organellar, of more than 100 plants have already been sequenced by now, and the information thus generated are available in public database for most of them. It must be mentioned here that bioinformatics played a remarkable role in handling the enormous data being produced in each and every minute. It can be safely told that the “genomics” era started in the beginning of the twenty-first century itself accompanying also proteomics, metabolomics, transcriptomics, and several other “omics” technologies.

Structural genomics has thus facilitated annotation of genes, enumeration of gene families and repetitive elements, and comparative genomics studies across taxa. On the other hand, functional genomics paved the way for deciphering the precise biochemistry of gene function through transcription and translation pathways. Today, genotyping-by-sequencing of primary, secondary, and even tertiary gene pools; genome-wide association studies; and genomics-aided breeding are almost routine techniques for crop improvement. Genomic selection in crops is another reality today. Elucidation of the chemical nature of crop chromosomes has now opened up a new frontier for genome editing that is expected to lead the crop improvement approaches in near future.

At the same time, we will look forward to replacement of genetically modified crops by cisgenic crops through transfer of useful plant genes and atomically modified crops by employing nanotechnology that will hopefully be universally accepted for commercialization owing to their human-friendly and environment-friendly nature.

I wish to emphatically mention here that none of the technologies and tools of plant breeding are too obsolete or too independent. They will always remain pertinent individually or as complementary to each other, and will be employed depending on the evolutionary status of the crop genomes, the genetic resources and genomics resources available, and above all the cost-benefit ratios for adopting one or more technologies or tools. In brief, utilization of these crop improvement techniques would vary over time, space, and economy scales! However, as we stand today, we have all the concepts, strategies, techniques, and tools in our arsenal to practice ‘genome designing’, as I would prefer to term it, of crop plants not just genetic improvement to address simultaneously ‘food, nutrition, energy, and environment security, briefly the FNEE security’, as I introduced the concept in 2013 and have been talking about for the last 5 years at different platforms.

Addressing FNEE security has become more relevant today in the changing scenario of climate change and global warming. Climate change will lead to greenhouse gas emissions and extreme temperatures leading to different abiotic stresses including drought or waterlogging, on the one hand, and severe winter and freezing, on the other hand. It will also severely affect uptake and bioavailability of water and plant nutrients and will adversely cause damage to physical, chemical, and biological properties of soil and water in cropping fields and around. It is also highly likely that there will be emergence of new insects and their biotypes and of new plant pathogens and their pathotypes. The most serious concerns are, however, the unpredictable crop growth conditions and the unexpected complex interactions among all the above stress factors leading to drastic reduction in crop yield and

quality in an adverse ecosystem and environment. Climate change is predicted to significantly reduce productivity in almost all crops. For example, in cereal crops, the decline of yield is projected at 12–15%. On the other hand, crop production has to be increased at least by 70% to feed the alarmingly growing world population, projected at over 9.0 billion by 2050 by even a moderate estimate.

Hence, the unpredictability of crop growing conditions and thereby the complexity of biotic and abiotic stresses warrant completely different strategies of crop production from those practiced over a century aiming mostly at one or the few breeding objectives at a time such as yield, quality, resistance to biotic stresses due to disease–pests, tolerance to abiotic stresses due to drought, heat, cold, flood, salinity, acidity, improved water and nutrient use efficiency, etc. In the changing scenario of climate change, for sustainable crop production, precise prediction of the above limiting factors by long-term survey and timely sensing through biotic agents and engineering devices and regular soil and water remediation will play a big role in agriculture. We have been discussing on “mitigation” and “adaptation” strategies for the last few years to reduce the chances of reduction of crop productivity and improve the genome plasticity of crop plants that could thrive and perform considerably well in a wide range of growing conditions over time and space. This is the precise reason for adopting genomic designing of crop plants to improve their adaptability by developing climate-smart or climate-resilient genotypes.

Keeping all these in mind, I planned to present deliberations on the problems, priorities, potentials, and prospects of genome designing for development of climate-smart crops in about 50 chapters, each devoted to a major crop or a crop group, allocated under five volumes on cereal, oilseed, pulse, fruit, and vegetable crops. These chapters have been authored by more than 250 of eminent scientists from over 30 countries including Argentina, Australia, Bangladesh, Belgium, Brazil, Canada, China, Egypt, Ethiopia, France, Germany, Greece, India, Ireland, Japan, Malaysia, Mexico, New Zealand, Kenya, Pakistan, Philippines, Portugal, Puerto Rico, Serbia, Spain, Sri Lanka, Sweden, Taiwan, Tanzania, Tunisia, Uganda, UK, USA, and Zimbabwe.

There are a huge number of books and reviews on traditional breeding, molecular breeding, genetic engineering, nanotechnology, genomics-aided breeding, and gene editing with crop-wise and trait-wise deliberations on crop genetic improvement including over 100 books edited by me since 2006. However, I believe the present five book volumes will hopefully provide a comprehensive enumeration on the requirement, achievements, and future prospects of genome designing for climate-smart crops and will be useful to students, teaching faculties, and scientists in the academia and also to the related industries. Besides, public and private funding agencies, policy-making bodies, and the social activists will also get a clear idea on the road traveled so far and the future roadmap of crop improvement.

I must confess that it has been quite a difficult task for me to study critically the different concepts, strategies, techniques, and tools of plant breeding practiced over the last 12 decades that also on a diverse crop plants to gain confidence to edit the chapters authored by the scientists with expertise on the particular crops or crop groups and present them in a lucid manner with more or less uniform outline of contents and formats. However, my experience gained over the last 7 years in the capacity of the Founding Principal Coordinator of the International Climate Resilient Crop Genomics Consortium (ICRCGC) was highly useful while editing these books. I have the opportunity to interact with a number of leading scientists from all over the world almost on regular basis. Organizing and chairing the annual workshops of ICRCGC since 2012 and representing ICRCGC in many other scientific meetings on climate change agriculture offered me a scope to learn from a large number of people from different backgrounds including academia, industries, policy-making bodies, funding agencies, and social workers. I must acknowledge here the assistance I received from all of them to keep me as a sincere student of agriculture specifically plant breeding.

This volume entitled *Genomic Designing of Climate-Smart Oilseed Crops* includes eight major crops including Soybean, Oilseed Rape, Groundnut, Sunflower, Flax, Rape and Mustard, Sesame, and Castor Bean. These chapters have been authored by 54 scientists from six countries including Australia, Canada, China, India, Serbia, and USA. I place on record my thanks for these scientists for their contributions and cooperation.

My own working experience on oilseed crops dates back to early 90s in the laboratory of Prof. Thomas C. Osborn in the Department of Agronomy of the University of Wisconsin-Madison. I must confess that this period of about 4 years through working on the Brassicaceae system in his lab and other two labs of his collaborating faculties including Prof. Paul H. Williams in the Department of Plant Pathology and Prof. Jiwan P. Palta in the Department of Horticulture had tailored my mind-set and enriched my expertise and helped me to grow as a science worker. Hence, I have dedicated this book to Prof. Osborn as a token of my respect, thanks, and gratitude.

New Delhi, India

Chittaranjan Kole

Contents

1 Approaches, Applicability, and Challenges for Development of Climate-Smart Soybean	1
Juhi Chaudhary, SM Shivaraj, Praveen Khatri, Heng Ye, Lijuan Zhou, Mariola Klepadlo, Priyanka Dhakate, Giriraj Kumawat, Guntant Patil, Humira Sonah, Milind Ratnaparkhe, Rupesh Deshmukh and Henry T. Nguyen	
2 Genetic Solutions to Improve Resilience of Canola to Climate Change	75
Harsh Raman, Rajneet Kaur Uppal and Rosy Raman	
3 Climate-Smart Groundnuts for Achieving High Productivity and Improved Quality: Current Status, Challenges, and Opportunities	133
Sunil S. Gangurde, Rakesh Kumar, Arun K. Pandey, Mark Burrow, Haydee E. Laza, Spurthi N. Nayak, Baozhu Guo, Boshou Liao, Ramesh S. Bhat, Naga Madhuri, S. Hemalatha, Hari K. Sudini, Pasupuleti Janila, Putta Latha, Hasan Khan, Babu N. Motagi, T. Radhakrishnan, Naveen Puppala, Rajeev K. Varshney and Manish K. Pandey	
4 Sunflower and Climate Change: Possibilities of Adaptation Through Breeding and Genomic Selection	173
Dragana Miladinović, Nada Hladni, Aleksandra Radanović, Siniša Jocić and Sandra Cvejić	
5 Progress Toward Development of Climate-Smart Flax: A Perspective on Omics-Assisted Breeding	239
S. M. Shivaraj, Priyanka Dhakate, Humira Sonah, Tri Vuong, Henry T. Nguyen and Rupesh Deshmukh	

6 Breeding *Brassica juncea* and *B. rapa* for Sustainable Oilseed Production in the Changing Climate: Progress and Prospects 275
Priya Panjabi, Satish Kumar Yadava, Nitin Kumar, Rajkumar Bangkim and Nirala Ramchiary

7 Potential for Adaptation to Climate Change Through Genomic Breeding in Sesame..... 371
Haiyang Zhang, Hongmei Miao and Ming Ju

8 Breeding for Climate Resilience in Castor: Current Status, Challenges, and Opportunities 441
Xuegui Yin, Jiannong Lu, Kwadwo Gyapong Agyenim-Boateng and Shuai Liu

Contributors

Kwadwo Gyapong Agyenim-Boateng Agricultural College, Guangdong Ocean University, Zhanjiang, Guangdong, China

Rajkumar Bangkim Department of Botany, University of Delhi North Campus, New Delhi, India

Ramesh S. Bhat University of Agricultural Sciences (UAS), Dharwad, India

Mark Burow Texas A & M AgriLife Research, Lubbock, USA;
Department of Plant and Soil Science, Texas Tech University, Lubbock, USA

Juhi Chaudhary Division of Plant Sciences, University of Missouri, Columbia, MO, USA

Sandra Cvejić Institute of Field and Vegetable Crops, Novi Sad, Serbia

Rupesh Deshmukh National Agri-Food Biotechnology Institute, Mohali, India

Priyanka Dhakate National Institute of Plant Genome Research, New Delhi, India

Sunil S. Gangurde International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, India

Baozhu Guo Crop Protection and Management Research Unit, USDA-ARS, Tifton, USA

S. Hemalatha BS Abdur Rahman University, Chennai, India

Nada Hladni Institute of Field and Vegetable Crops, Novi Sad, Serbia

Pasupuleti Janila International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, India

Siniša Jocić Institute of Field and Vegetable Crops, Novi Sad, Serbia

Ming Ju Henan Sesame Research Center, Henan Academy of Agricultural Sciences, Zhengzhou City, Henan, China

Hasan Khan University of Agricultural Sciences (UAS), Raichur, India

Praveen Khatri National Agri-Food Biotechnology Institute, Mohali, India

Mariola Klepadlo Division of Plant Sciences, University of Missouri, Columbia, MO, USA

Nitin Kumar Department of Bioengineering and Technology, Institute of Science and Technology, Gauhati University, Guwahati, Assam, India;
School of Life Sciences, Jawaharlal Nehru University, New Delhi, India

Rakesh Kumar International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, India

Giriraj Kumawat Indian Institute of Soybean Research, Indore, India

Putta Latha Regional Agricultural Research Station (RARS), Acharya NG Ranga Agricultural University (ANGRAU), Tirupati, India

Haydee E. Laza Department of Plant and Soil Science, Texas Tech University, Lubbock, USA;
USDA-ARS-CSRL, Lubbock, USA

Boshou Liao Oil Crops Research Institute (OCRI) of Chinese Academy of Agricultural Sciences (CAAS), Wuhan, China

Shuai Liu Agricultural College, Guangdong Ocean University, Zhanjiang, Guangdong, China

Jiannong Lu Agricultural College, Guangdong Ocean University, Zhanjiang, Guangdong, China

Naga Madhuri Regional Agricultural Research Station (RARS), Acharya NG Ranga Agricultural University (ANGRAU), Tirupati, India

Hongmei Miao Henan Sesame Research Center, Henan Academy of Agricultural Sciences, Zhengzhou City, Henan, China;
Department of Agronomy, Zhengzhou University, Zhengzhou City, Henan, China

Dragana Miladinović Institute of Field and Vegetable Crops, Novi Sad, Serbia

Babu N. Motagi University of Agricultural Sciences (UAS), Dharwad, India

Spurthi N. Nayak University of Agricultural Sciences (UAS), Dharwad, India

Henry T. Nguyen Division of Plant Sciences, University of Missouri, Columbia, MO, USA

Arun K. Pandey International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, India

Manish K. Pandey International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, India;
BS Abdur Rahman University, Chennai, India

Priya Panjabi Department of Botany, University of Delhi North Campus, New Delhi, India

Gunvant Patil Division of Plant Sciences, University of Missouri, Columbia, MO, USA

Naveen Puppala New Mexico State University-Agricultural Science Center at Clovis, New Mexico, USA

Aleksandra Radanović Institute of Field and Vegetable Crops, Novi Sad, Serbia

T. Radhakrishnan ICAR-Directorate of Groundnut Research (ICAR-DGR), Junagadh, India

Harsh Raman NSW Department of Primary Industries, Wagga Wagga Agricultural Institute, Wagga Wagga, NSW, Australia

Rosy Raman NSW Department of Primary Industries, Wagga Wagga Agricultural Institute, Wagga Wagga, NSW, Australia

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

Milind Ratnaparkhe Indian Institute of Soybean Research, Indore, India

S. M. Shivaraj University Laval, Quebec City, Canada

Humira Sonah University Laval, Quebec City, Canada;
National Agri-Food Biotechnology Institute, Mohali, India

Hari K. Sudini International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, India

Rajneet Kaur Uppal NSW Department of Primary Industries, Wagga Wagga Agricultural Institute, Wagga Wagga, NSW, Australia

Rajeev K. Varshney International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, India

Tri Vuong Division of Plant Sciences, University of Missouri, Columbia, MO, USA

Satish Kumar Yadava Center for Genetic Manipulation of Crop Plants, University of Delhi South Campus, New Delhi, India

Heng Ye Division of Plant Sciences, University of Missouri, Columbia, MO, USA

Xuegui Yin Agricultural College, Guangdong Ocean University, Zhanjiang, Guangdong, China

Haiyang Zhang Henan Sesame Research Center, Henan Academy of Agricultural Sciences, Zhengzhou City, Henan, China

Lijuan Zhou Division of Plant Sciences, University of Missouri, Columbia, MO, USA

Abbreviations

^{13}C	Isotope carbon 13
$\Delta^{13}\text{C}$	Carbon isotope discrimination
ABA	Abscisic acid
AC	Ash content
ADH	Alcohol dehydrogenase
AFLP	Amplified fragment length polymorphism
ALA	α -linolenic acid
ALS	Acetolactate synthase
AMOVA	Analysis of molecular variance
ANOVA	Analysis of variance
APX	Ascorbate peroxidase
AQP	Aquaporin
ARF	Auxin-response factor
ATP	Adenosine triphosphate
ATR	Atrazine
ATT	Acquired thermotolerance
BABA	β -Aminobutyric acid
BAC	Bacterial artificial chromosome
<i>Bar</i>	Bialaphos resistance gene
BC	Back cross
BH	Branching habit
BIM	Bayesian interval mapping
BL	Bayesian LASSO
BLAST	Basic local alignment search tool
BLUP	Best linear unbiased prediction
BPMV	Bean pod mottle virus
BRR	Bayesian ridge regression
BSA	Bulked segregant analysis
<i>Bt</i>	<i>Bacillus thuringiensis</i>
C10	C+1 alleles with capsule open at tip

CaMV	Cauliflower mosaic virus
Cas9	CRISPR-associated 9 protein
CAT	Catalase
CC	Climate change
CDS	Coding DNA sequence
CEG	Core eukaryotic genes
CGMCP	Centre for Genetic Manipulation of Crop Plants
CGRIS	Chinese Genetic Resources Information System
CID	Carbon isotope discrimination
CIM	Common Information Model
CIM	Composite interval mapping
CL	Capsule length
cM	CentiMorgan
CNV	Copy number variant
CO ₂	Carbon dioxide
CMS	Cytoplasmic male sterility
CN	Capsule number per plant
CNN	Capsule node number
CNPA	Centro Nacional de Pesquisa de Algodao
CNS	Capsule number per stem
CNV	Copy number variant
CR	Clubroot resistance gene
CRISPR	Clustered regularly interspaced short palindromic repeats
CRP	Coordinated research project
CRR	Charcoal rot resistance
<i>cryIAcF</i>	Delta-endotoxin of <i>Bacillus thuringiensis</i> gene (1AcF)
<i>cryIEC</i>	Delta-endotoxin of <i>Bacillus thuringiensis</i> gene (1EC)
CS	Climate smart
CTD	Canopy temperature depression
CWR	Crop wild relative
CZL	Capsule zone length
DAG	Diacylglycerols
DALP	Direct amplification of length polymorphism
DAP	Days after planting
DArT	Diversity arrays technology
DAS	Days after sowing
DDBJ	DNA Databank of Japan
DEG	Differentially expressed gene
DGAT	Diacylglycerol acetyltransferase
dgatA	Acyl-CoA:diacylglycerol acyltransferase A
DH	Doubled haploid
DI	Disease index
DMO	Dicamba monooxygenase
DREB	Dehydration responsive element binding (protein)
DREB2A	Drought responsive element binding protein 2A

DS	Determinate sesame
DSB	Double-stranded break
DSF	Days from sowing to flowering
dsRNA	Double-stranded RNA
Dt	Determinate
DTF	Days to flowering
Dw	Dwarf
<i>Dwf</i>	Dwarfing gene
ECP/GR	European Cooperative Programme for Crop Genetic Resources Network
EDB	European Brassica Database
ELS	Early leaf spot
EMF	Embryonic flower
EMS	Ethyl methanesulphonate
EPA	Eicosapentaenoic acid
EPA	Environmental Protection Agency (USA)
EPSP	5-Enolpyruvylshikimate-3-phosphate
ESCORENA	European Co-operative Research Network on Flax and other Bast Plants
ESPS	5-Enolpyruvylshikimate-3-phosphate synthase
EST	Expressed sequence tag
ETI	Effector-triggered immunity
F ₂	Second filial generation
FA	Fatty acid
FA	Flowers per leaf axil
FAO	Food and Agriculture Organization
FC	Fiber content
FCL	Length of the lateral capsule
FCT	Thickness of the lateral capsule
FCW	Width of the lateral capsule
FDA	Food and Drug Administration (USA)
<i>FLC</i>	Flowering Locus C
FNEE	Food, nutrition, energy and environment
FNI	Fast neutron irradiation
FOS	<i>Fusarium oxysporum</i> f.sp. <i>sesami</i>
FSD	Fresh seed dormancy
FT	Flowering locus T
G × E	Genotype × environment
GAB	Genomics-assisted breeding
GBS	Genotyping-by-sequencing
GCA	General combining ability
GE	Genetically engineered
GEAC	Genetic Engineering Appraisal Committee (India)
GEBV	Genome-estimated breeding value
GFF	General feature format

GFP	Green fluorescent protein
GM	Genetically modified
GMHRA	Glyphosate acetyltransferase and modified soybean acetolactate synthase
GN	Grain number per capsule
GO	Gene ontology
GP	Gene pool
GR	Glyphosate resistant
GRD	Groundnut rosette disease
GRDC	Grains Research and Development Corporation
GRIN	Germplasm Resources Information Network (USA)
GRU	Germplasm Resources Unit
GS	Genomic selection
GSO	Seamless capsule open at tip
GSS	Genome survey sequence
GUS	β -glucuronidase gene
GWAS	Genome-wide association study
H	Index of genetic diversity
HAAS	Henan Academy of Agricultural Sciences
HDR	Homology-directed repair
He	Average expected heterozygosity per locus
HFC	Height to the first capsule
Hi-C	Chromosome conformation capture
HOA	High OA
HPPD	4-Hydroxyphenylpyruvate dioxygenase
HR	Highly resistant
HS	Highly susceptible
Hsfs	Heat shock transcription factors
HSP	Heat shock protein
HSRC	Henan Sesame Research Center
I	Shannon's information index
IAEA	International Atomic Energy Agency
IBC	Institute of Biodiversity Conservation
IBPGR	International Bureau of Plant Genetic Resources
ICGR-CAAS	Institute of Crop Germplasm Resources, Chinese Academy of Agricultural Sciences
ICRISAT	International Crops Research Institute for the Semi-Arid Tropics
ID	Indehiscent
IFDB	International Flax Database
IFVCNS	Institute of Field and Vegetable Crops
IL	Internode length
ILs	Interspecific lines
IND	Improved nondehiscent
InDel	Insertion/Deletion
INTA	Instituto Nacional de Tecnología Agropecuaria

IOD	Iodine value
IP	Intellectual property
IPCC	Intergovernmental Panel on Climate Change
ISSR	Inter-simple sequence repeat
ITPGRFA	The International Treaty for Plant Genetic Resources for Food and Agriculture
KASP	Kompetitive allele-specific polymerase chain reaction
KEGG	Kyoto Encyclopedia of Genes and Genomes
LD	Linkage disequilibrium
LD	Long day
LEA	Late embryogenesis abundant
LG	Linkage group
LIN	Linolenic acid
LIO	Linoleic acid
LIS-1	Linum insertion sequence 1
LLS	Late leaf spot
LN	Leaf number per plant
LncRNA	Long noncoding RNA
LOA	Low OA
LOD	Logarithm of odds
LRR	Leucine-rich repeat
MABC	Marker-assisted backcrossing
MAGIC	Multiparent advanced generation intercross
MAPKK	Mitogen-activated protein kinase kinase
MARS	Marker-assisted recurrent selection
MAS	Marker-assisted selection
MBB	Mexican bean beetle
MBPG	Multinational Brassica Genome Sequencing Project
MCL	Length of the central capsule
MCT	Thickness of the central capsule
MCW	Width of the central capsule
MDA	Malondialdehyde
MG	Maturity group
MIM	Multiple interval mapping
MLS	Multilateral System
MoEF&CC	Ministry of Environment, Forest and Climate Change (India)
MR	Moderately resistant
MSD	Main stem diameter
MSIL	Length of main stem internode
MSNN	Node number of main stem
MTA	Marker-trait association
MTA	Material Transfer Agreement
Na	Number of alleles
NAM	Nested association mapping
NARS	National Agricultural Research System (India)

NBPGR	National Bureau of Plant Genetic Resources (India)
NBS	Nucleotide-binding site
NCBI	National Center for Biotechnology Information
NDVI	Normalized difference vegetation index
Ne	Effective number of alleles
NGS	Next-generation sequencing
NHEJ	Non-homologous end joining
NIL	Near-isogenic lines
NN	Node number
<i>nptII</i>	Neomycin phosphotransferase II gene
NTG	N-methyl-N'-nitro-N-nitrosoguanidine
NUE	Nutrient use efficiency
OA	Osmotic adjustment
OA	Oxalic acid
OC	Oil content
OLE	Oleic acid
<i>OLP</i>	Osmotin-like protein gene
PAGE	Parametric analysis of gene set enrichment
PAL	Palmitic acid
PAT	Phosphinothricin acetyltransferase
PAV	Presence/absence variants
PCR	Polymerase chain reaction
PE	Paired end
PEG	Polyethylene glycol
PGR	Pod growth rate
PH	Plant height
PHB	Polyhydroxybutyrate
PIABS	Photosynthetic efficacy index
PIC	Polymorphic information content
PiHS	Population-based integrated haplotype score
<i>Pl</i>	Downy mildew resistance gene
PLCP	Papain-like cysteine protease
PLH	Potato leafhopper
PO	Protein content
POX	Peroxidase
PPO	Polyphenol oxidase
PR	Pathogenesis-related
PRH	Bearing height of primary raceme
PUFA	Polyunsaturated fatty acids
PVE	Phenotypic variation explained
QTL	Quantitative trait locus
QTLs	Quantitative trait loci
R	Resistance gene
RAD	Restriction site-associated DNA
RAPD	Random amplified polymorphic DNA

RCA	<i>R. communis</i> agglutinin
<i>RcPAL</i>	<i>Ricinus communis</i> phenylalanine ammonialyase gene
<i>RcPEPC</i>	<i>Ricinus communis</i> phosphoenolpyruvate carboxylase gene
<i>Rf</i>	Fertility restoration gene
RFLP	Restriction fragment length polymorphism
RGB	Red, green and blue
RGC	Resistance gene candidate
RHL	Residual heterozygous line
RIL	Recombinant inbred line
RNAi	RNA interference
ROD	Reduction of density
RR-BLUP	Ridge regression best linear unbiased prediction
RRGS	Reduced-representation genome sequencing
RRS	Reduced-representation sequencing
RSA	Root system architecture
RSAMPL	Random selective amplification of microsatellite polymorphic locus
RSLs	Recombinant substitution lines
RT-PCR	Real-time PCR
RT-PCR	Reverse transcription PCR
RXBS	Rongxian black sesame
SAM	Sequence alignment map
SAT	Semi-arid tropics
SBA	Soybean aphid
SbDV	Soybean dwarf virus
SBL	Soybean looper
<i>SbNHX1</i>	<i>Salicornia brachiata</i> reverse transporter protein gene
SC	Sesamin content
SCAR	Sequence-characterized amplified region
scFv	Single-chain variable fragment
SCMR	SPAD chlorophyll meter reading
SCoT	Start codon targeted polymorphism
SD	Short day
SDS	Sudden death syndrome
SEA	Singular enrichment analysis
SFW	Sesame Fusarium wilt
SG	Selective genotyping
SGMD	Soybean Genomics and Microarray Database
SGP	The Sesame Genome Project
SGWG	The Sesame Genome Working Group
SHA	Shattering
SIM	Simple interval mapping
SiNPs	Silicon nanoparticles
SLA	Specific leaf area
SLAF	Specific length amplified fragment

SLAF-seq	Specific length amplified fragment sequencing
SMA	Single marker analysis
SMG	Suppressor with morphogenetic effects on genitalia
SMV	Soybean mosaic virus
SN	Seed number per plant
SNC	Seed number per capsule
SNP	Single nucleotide polymorphism
SOD	Superoxide dismutase
SoyGD	Soybean Genome Database
SR	Shatter resistant
SRAP	Sequence-related amplified polymorphism
SSCP	Single-strand conformational polymorphism
SSH	Semi-shattering
SSR	Simple sequence repeat
STE	Stearic acid
STF	Days from sowing to flowering
STS	Sequence tagged site
SUS	Super-shattering
TAG	Triacylglycerols
TALEN	Transcription activator-like effector nuclease
TE	Transpiration efficiency
TFL-like	Terminal flower-like
TIGR	The Institute for Genomic Research
TILLING	Targeting induced local lesions in genomes
TIR	Temperature induction response
TL	Tip length without the capsule
TP	Training population
TRAP	Target region amplification polymorphism
TRAP	Tartrate-resistant acid phosphatase
TSS	Total soluble sugars
TSW	Thousand seed weight
TSWV	Tomato spotted wilt virus
TT	Triazine tolerant
TUFGEN	Total Utilization Flax Genomics
UGM	Ungrouped matches
UPM	The Universidad Politécnica de Madrid
UPOV	International Union for the Protection of New Varieties of Plants
USDA	United States Department of Agriculture
UTRs	Untranslated regions
VBC	Velvet bean caterpillar
VIR	Vavilov Institute of Plant Industry
VNIIMK	All-Russia Research Institute of Oil Crops
VPD	Vapor pressure deficit
WGR	Whole genome re-sequencing
WGS	Whole genome shotgun

WSC	Water soluble carbohydrates
WUE	Water use efficiency
ZFN	Zinc-finger nucleases

Chapter 1

Approaches, Applicability, and Challenges for Development of Climate-Smart Soybean



**Juhi Chaudhary, SM Shivaraj, Praveen Khatri, Heng Ye,
Lijuan Zhou, Mariola Klepadlo, Priyanka Dhakate,
Giriraj Kumawat, Gunvant Patil, Humira Sonah,
Milind Ratnaparkhe, Rupesh Deshmukh and Henry T. Nguyen**

Abstract Soybean (*Glycine max* L.) is an economically important crop providing a great source for vegetable oil and protein. Yield losses of soybean under current climate change keep increasing, despite the progressive increase in yield through breeding and management practices since the 1960s. Conventional breeding facilitated the development of high-quality soybeans with enhanced tolerance to severe environmental fluctuations such as drought, flooding, heat, and salinity. However, conventional approaches are laborious, time consuming, and looks inefficient to fulfill the increasing demands of the growing world population. The advances in marker-assisted and genomics-assisted breeding, sequencing technologies, and bioinformatics tools have enabled the soybean improvement at a faster pace. The rapidly accumulating genomic resources have enabled the development of molecular markers associated with many important quantitative trait loci, provided a clear picture of genomic variations in soybean germplasm, and identified key genes for genetic engineering. This knowledge is being utilized to facilitate the development of climate-smart soybeans. In this chapter, we discuss and summarize the advances

Juhi Chaudhary and S. M. Shivaraj—These authors contributed equally to this work.

J. Chaudhary · H. Ye · L. Zhou · M. Klepadlo · G. Patil · H. T. Nguyen (✉)
Division of Plant Sciences, University of Missouri, Columbia, MO, USA
e-mail: nguyenhenry@missouri.edu

S. Shivaraj · H. Sonah
University Laval, Quebec City, Canada

P. Khatri · H. Sonah · R. Deshmukh (✉)
National Agri-Food Biotechnology Institute, Mohali, India
e-mail: rupesh0deshmukh@gmail.com

P. Dhakate
National Institute of Plant Genome Research, New Delhi, India

G. Kumawat · M. Ratnaparkhe
Indian Institute of Soybean Research, Indore, India

in soybean improvement through conventional and genomics-assisted breeding, genetic engineering approaches, and available bioinformatics tools for soybean. This chapter also highlights soybean genetic resources, diversity analysis, association mapping, as well as recent strategies such as gene editing and nanotechnology application in soybean breeding programs. This information could facilitate the incorporation of climatic-smart traits in breeding for more stable soybean production with the changing climate.

Keywords Soybean · Climatic change · Genetic diversity · Marker-assisted breeding · Genetics · Genomics · Stress resistance · Gene editing

1.1 Introduction

Soybean is a self-pollinated plant that belongs to the family Fabaceae and *Glycine* genus. The *Glycine* genus is subsequently divided into subgenera, *Glycine* and *Soja*. The subgenus *Soja* has two highly recognized species including the cultivated soybean *Glycine max* and the wild soybean *Glycine soja*. Soybean is an economically important legume crop that is rich in seed protein (40%) and oil (20%), which provides sources of starch, dietary fiber, protein, lipids, and essential minerals for human as well as for livestock (Chaudhary et al. 2015). It is widely grown as a grain legume and oilseed crop in the world including the countries United States, Brazil, Argentina, China, and India. The US is the leading producer with 35% (119.5 Million Metric Tons) of the global production (340.9 Million Metric Tons) of soybean (SoyStats 2018 www.soystats.com).

Soybean is in high demand not only for food and feed consumption, but also it can potentially serve as a future fuel feedstock and biodegradable plastics (Candeia et al. 2009; Song et al. 2011). Furthermore, soybean is also used in industrial and pharmaceutical applications as well as in the production of biodiesel (Goldberg 2008). Due to diverse uses, soybean has become a highly desirable crop and its demand is rapidly increasing. However, the increasing global population will need doubled food production by the year 2050 and we can only achieve ~55% of the required production at the current rate of yield improvement. It is expected to be more difficult to produce sufficient yield with the changing climate (Ray et al. 2013; Deshmukh et al. 2014). Climate change and extreme weather conditions have a negative impact on crop yield, because solar radiation, temperature, and precipitation are the main drivers of crop growth. Additionally, climate change influences the plant diseases and pest infestations, as well as the supply of and demand for irrigation water (Rosenzweig et al. 2001). Therefore, the emphasis must be given toward the production of high-yielding soybeans with high nutritional value, which are environmentally stable and resistant to extreme weather conditions.

Plant breeding has undoubtedly improved soybean yield and resistance to biotic and abiotic stresses to achieve the current level of demand, but the main challenge is to continue to increase the production under the current scenario of climate

change. In general, breeding for a complex trait is challenging due to their control by multiple genes and they are also greatly influenced by the environment. The conventional breeding procedures such as backcrossing, single pod descent, pedigree breeding, and bulk population breeding are used in order to develop improved varieties of soybean (Poehlman et al. 1995). To facilitate breeding advances, it is necessary to employ modern breeding techniques such as marker-assisted breeding, recombinant DNA technology, genome editing and “omics” (genomics, transcriptomics, proteomics, metabolomics, ionomics) to improve the soybean quality and yield. In addition, the concerns about environmental stress due to climate change and demand of ample supply have instilled a new urgency into accelerating the rates of genetic gain in breeding programs. Therefore, regardless of the conventional breeding efforts, it is essential to integrate the next-generation molecular and omics approaches for the production of high-yielding soybeans with enhanced adaptation to various environmental stresses.

1.2 Prioritizing Climate-Smart Traits

1.2.1 Flowering Time and Maturity

Plants can perceive various environmental signals, such as photoperiod, temperature, and stresses, to flower, and thus control seed production. In soybeans, flowering time and maturity are important agronomic traits, which are useful for developing soybean cultivars with a wider geographical adaptation. Soybean is a short-day plant. Short days induce flowering while long-day conditions delay flowering. Photoperiod and in-season temperature are the primary factors that dictate the region where a soybean variety is adapted. Soybean can grow in a wide range of latitudes, from 50°N to 35°S (Watanabe et al. 2012). The adaptability of soybean in a wide latitude is caused by natural variations of many genes controlling flowering time and maturity. The study on the flowering and maturity controlling mechanism of soybean can provide a theoretical basis and genetic materials for soybean breeding, especially under the climate change scenario. Understanding the regulatory mechanisms of flowering time and maturity allows us to modify the growth cycles of soybean to overcome or avoid different stresses by manipulation of the two traits.

1.2.1.1 Overview of Flowering and Maturity Regulating Genes in Soybean

Flowering time (days to R1) and maturity (days to R8) in soybean have been reported to be highly correlated traits (Mansur et al. 1996). Photoperiod insensitivity, flowering time, and maturity were found to be controlled by the same genes or by tightly clustered genes in the same chromosomal region (Tasma et al. 2001).

In 1927, a major gene locus was detected to control maturity (Owen 1927). Subsequent research work found that the *E1* locus is largely responsible for the variation in flowering time among cultivars (Bernard 1971; Abe et al. 2003). To date, ten genes related to flowering and maturity have been reported including nine *E* genes (*E1–E9*) and one *J* gene (Bernard 1971; Buzzell and Voldeng 1980; Bonato and Vello 1999; Cober et al. 2010; Kong et al. 2014). Six *E* genes, *E1*, *E3*, *E4*, *E7*, *E8*, and *E9*, can specifically participate in photoperiod response (Cober et al. 1996; Cober and Voldeng 2001; Cober et al. 2010), with *E1*, *E3*, *E4*, *E7*, and *E8* as recessive loci (Watanabe et al. 2012; Kong et al. 2014). Introgression of these early flowering alleles results in earlier flowering under long day and improved adaptation to short summers at high latitudes. The *J* locus was identified in the progeny of crosses between standard and late flowering cultivars with a long-juvenile habit, whose recessive allele causes late flowering under short days (Ray et al. 1995). In general, the trait of “delayed juvenile” is useful for adaptation to low latitudes and spring sowings at the lower latitudes (Tomkins and Shipe 1997).

1.2.1.2 Cloning Genes Underlying the Flowering and Maturity Traits

Efforts were made to clone the underlying genes of the loci to understand the mechanisms of flowering and maturity in soybean. The *E1* gene was map-based cloned to encode a B3-like protein, which is belonging to a family of plant-specific transcription factors. *E1* from soybean shows high similarity to other legumes, such as *Medicago truncatula* and *Lotus corniculatus*. However, the *E1* gene does not exist in the model plants *Arabidopsis* and rice. The *E2* gene was identified as the homolog of *GIGANTEA (GI)*, the unique plant-specific nuclear clock-associated protein, which contributes to the maintenance of circadian period length and amplitude, and regulates flowering time and hypocotyl growth in response to day length (Watanabe et al. 2011). *E2* can enhance the photoperiod response of soybean, and it is closely related to the early flowering phenotype of soybean and the light adaptability. *E3* and *E4* encode the phytochrome (phy) family of photoreceptors PHYA3 and PHYA2, respectively (Liu et al. 2008; Tsubokura et al. 2013). Soybean contains four *PHYA* genes that consist of two pairs of homologs. *E3* and *E4* represent in different homolog pairs. The homolog of *E4*, *PHYA1*, is apparently functional, whereas the homolog of *E3* carries a deletion and is probably a pseudogene (Watanabe et al. 2009). *E9* was identified as *FT2a*, an ortholog of *Arabidopsis FLOWERING LOCUS T*, through fine-mapping, sequencing, and expression analysis. Recessive allele of *E9* delays flowering because of lower transcript abundance that is caused by allele-specific transcriptional repression.

1.2.1.3 Application of Classification of Maturity Group (MG) in Soybean

Understanding the mechanism of soybean flowering time and maturity diversity and adaptation is very important for breeding for high productivity in diverse latitudes. Many soybean cultivars were bred with different maturity to adapt various ecological environments. For the convenience of breeding layout, 13 MGs from 000, 00, 0, I, II, to X were classified based on photoperiod and yield trial in North America (Zhang et al. 2007). Maturity group zones represent defined areas, where a cultivar is best adapted. But the classification of maturity group is still not internationally unified. Based on the knowledge mentioned above, flowering and maturity were highly controlled by major genes in soybean. Therefore, flowering time and maturity can be adjusted by soybean genetic change through breeding efforts and genetic engineering. Although photoperiod remains constant, climatic conditions, management practices, and soybean genetics have changed during the past decades. Maturity group adaptation zones need to be understood, applied and adjusted for the breeding benefit (Mourtzinis and Conley 2017).

1.2.2 Seed Composition

Soybean is a major crop for oil and protein resources, accounts for 56% of total oilseed production in the world (Wilson 2008). The seed quality is determined by the seeds' composition, including protein, oil, sugars, and minerals. Soybean seeds contain 40% protein, 20% oil, 15% soluble carbohydrate, and 15% fiber on a dry weight basis. Protein and oil are the most abundant and valuable compositions in soybean.

1.2.2.1 Oil

Soybean seed contains about up to 230 g kg⁻¹ of oil on a dry weight basis and the oil contents are constituted by 16% saturated, 23% monounsaturated, and 58% polyunsaturated fatty acids (Bellaloui et al. 2015). The major unsaturated fatty acids in soybean are the polyunsaturated alpha-linolenic acid (7–10%), linoleic acid (51%), and the monounsaturated oleic acid (23%) (Poth 2000; Ivanov et al. 2010). This makes soybean oil valuable in terms of human healthy diets. However, soybean oil has approximately 24% monounsaturated fatty acids (C18:1), which are significantly less than competing oils such as canola (61%) and olive (40%) (Terés et al. 2008). Oleic Acid (C18:1), a monounsaturated omega-9 fatty acid typically makes up 55–83% of total oil content in olive. Monounsaturated fats are resistant to high heat, making extra virgin olive oil a healthy choice for cooking.

Genetic Regulation of Seed Oil Production

The oil concentration in soybean seeds is a quantitative trait governed by a number of genes mostly with small effects and under influence of the environment. A negative relationship between seed oil and protein was well documented, which makes it difficult for breeders to develop high-oil soybean genotypes while retaining a high level of protein (Wilcox and Shibles 2001; Hyten et al. 2004b). There are >130 quantitative trait loci (QTLs) reported to be associated with oil content in soybean (Qi et al. 2011), since the first documented report to detect oil QTL (Diers et al. 1992). Among these oil QTLs, only a few have been detected in multiple genetic backgrounds or environments, and none have been widely used in marker-assisted selection (MAS) for high oil in soybean breeding programs. This could be due to several factors affecting the usefulness of QTL, including large confidence intervals, QTL \times environment, and QTL \times genetic background interactions, which all impede the use of QTL in breeding programs (Qi et al. 2011).

Except QTLs, some transcription factors have been reported to modify the seed oil content in soybean, such as transcription factors, *LEC1*, *LEC2*, *ABI3*, and *FUS3*, which are master regulators of seed development, and thus regulate oil content (Mendoza et al. 2005). Besides, overexpression of *GmDOF4* and *GmDOF11* increased lipid content in seeds of transgenic Arabidopsis plants via direct activation of lipid biosynthesis genes and the repression of storage protein genes (Wang et al. 2007). Transcription factor *GmbZIP123*, also elevated lipid contents in seeds of transgenic Arabidopsis plants by activating Suc-transporter genes and cell-wall-invertase genes for sugar translocation and sugar breakdown, respectively (Song et al. 2013).

Metabolic Engineering of Fatty Acid Composition

Most domesticated oilseed crops have been successfully modified through either breeding or genetic engineering approaches to optimize the ratio of endogenous fatty acids in the storage oil for specific end uses (Drexler et al. 2003). For example, suppression of the oleate D12-desaturase gene in soybean, sunflower, cotton, and canola has resulted in the production of oils with a high C18:1 fatty acids, which have a greater oxidative stability and improved performance in high-temperature cooking applications. Oils with a high C18:1 ratio are also desired by the chemical industry, as C18:1 can be used in a variety of applications including detergents, soaps, lubricants, cosmetics, and emulsifying agents, and as a source of C9 monomers for plastics (Metzger and Bornscheuer 2006). Buhr et al. (2002) described the development of transgenic soybean events in which the expression of *FAD2-1* and *FatB* was simultaneously downregulated in a seed-specific fashion, thereby generating soybean oil with a reduced content of C16:0 (<5%) and significantly increased C18:1 content (>85%) (Buhr et al. 2002). Recently, naturally occurred mutant alleles of *FAD2-1A* and *FAD2-1B* in soybean plant introduction (PI) collections were identified (Pham et al. 2011). The traditionally bred soybean