

Chittaranjan Kole *Editor*

# Genomic Designing of Climate-Smart Cereal Crops

 Springer

# Genomic Designing of Climate-Smart Cereal Crops

Chittaranjan Kole  
Editor

# Genomic Designing of Climate-Smart Cereal Crops

 Springer

*Editor*

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

ISBN 978-3-319-93380-1                      ISBN 978-3-319-93381-8 (eBook)

<https://doi.org/10.1007/978-3-319-93381-8>

© Springer Nature Switzerland AG 2020

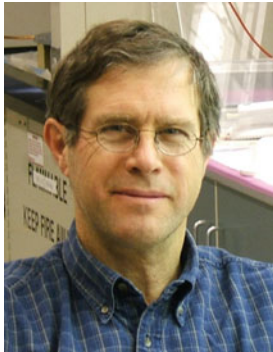
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission 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, expressed 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

*Dedicated to*



*Late Dr. Evgueni V. Ananiev*

*Formerly Head of the Laboratory of Plant Molecular Genetics and Genetic Engineering, N. I. Vavilov Institute of General Genetics, USSR Academy of Sciences, Moscow for mentoring me for my first research works on plant biotechnology through cloning, sequencing, and utilizing plant genes for phylogenetic analysis and evolution. He remains as a model to us and the world as an outstanding scientist and a wonderful human being mingled in a single persona!!*

# 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 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 onward 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 the 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 models 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 have 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 the near future.

At the same time, we will look forward to the replacement of genetically modified crops by cisgenic crops through the 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 is 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, I 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 one hand and severe winter and freezing on the other. 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 an 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 a 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 about 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, or improved water and nutrient use efficiency. 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 a 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, and 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 Cereal Crops* includes seven major crops including Rice, Wheat, Maize, Oat, Sorghum, Pearl Millet, and Finger Millet. These chapters have been authored by 52 scientists from 13 countries including China, Ethiopia, India, Ireland, Kenya, Mexico, Philippines, Spain, Sweden, UK, USA, Zambia, and Zimbabwe. I place on record my thanks for these scientists for their contributions and cooperation.

I feel myself proud that I could start my research works under the supervision of Late Dr. Evgueni V. Ananiev at his Laboratory of Plant Molecular Genetics and Genetic Engineering, N. I. Vavilov Institute of General Genetics, USSR Academy of Sciences, Moscow. As a member of his team, I could be a party for making some original contributions in the fields of cloning, sequencing, and utilizing some plant genes for phylogenetic analysis and evolution in cereals crops under the supervision of Late Dr. Ananiev, an outstanding scientist and wonderful human being. Hence, I have dedicated this book to (Late) Dr. E. V. Ananiev as a token of my respect, thanks, and gratitude. I wish to put on record my deep regards and acknowledgement to his wife, Dr. Olga N. Danilevskaya, for providing me the picture of Late Dr. Ananiev I have used in the dedication page of this book.

New Delhi, India

Chittaranjan Kole

# Contents

<b>1 Genomics-Assisted Breeding of Climate-Smart Inbred and Hybrid Rice Varieties</b> . . . . .	1
Jauhar Ali, Anumalla Mahender, G. D. Prahalada, Ma. Anna Lynn Sevilla, Angelito Galang, Erik Jon De Asis, Madonna Dela Paz, Corinne Mira Marfori-Nazarea, Katrina Leslie Nicolas and Ricky Vinarao	
<b>2 Advanced Genomics and Breeding Tools to Accelerate the Development of Climate Resilient Wheat</b> . . . . .	45
Yuanfeng Hao, Awais Rasheed, Robert Jackson, Yonggui Xiao, Yong Zhang, Xianchun Xia and Zhonghu He	
<b>3 Increasing Genetic Gains in Maize in Stress-Prone Environments of the Tropics</b> . . . . .	97
B. M. Prasanna, Sudha K. Nair, Raman Babu, Manje Gowda, Xuecai Zhang, Yunbi Xu, Mike Olsen, Vijay Chaikam, Jill E. Cairns, Mainassara Zaman-Allah, Yoseph Beyene, Amsal Tarekegne and Cosmos Magorokosho	
<b>4 Genomic Approaches for Climate Resilience Breeding in Oats</b> . . . . .	133
Julio Isidro-Sánchez, Elena Prats, Catherine Howarth, Tim Langdon and Gracia Montilla-Bascón	
<b>5 Genomic Designing for Climate Smart Sorghum</b> . . . . .	171
Ganapathy Kuyyamudi Nanaiah and Sujay Rakshit	
<b>6 Genomic Designing of Pearl Millet: A Resilient Crop for Arid and Semi-arid Environments</b> . . . . .	221
Desalegn D. Serba, Rattan S. Yadav, Rajeev K. Varshney, S. K. Gupta, Govindaraj Mahalingam, Rakesh K. Srivastava, Rajeev Gupta, Ramasamy Perumal and Tesfaye T. Tesso	

**7 Genomic Designing for Climate Smart Finger Millet . . . . . 287**  
Damaris Achieng Odeny, Adnan Niazi, Kassahun Tesfaye,  
Dagnachew Lule, Solomon Wanyonyi and Josephine Sarah Kunguni

# Abbreviations

AATF	African Agricultural Technology Foundation
ABA	Abscisic Acid
ABCILs	Advanced backcross introgression lines
ABRE	ABA-responsive cis-elements
AFLP	Amplified fragment length polymorphism
AGCR	Agronomically important cloned rice genes
AGT	Advanced genomic technologies
AI	Artificial intelligence
AICPMIP	All India Coordinated Pearl Millet Improvement Program
AICRIP	All India Coordinated Rice Improvement Project
AM	Association mapping
ANN1	<i>Annexin1</i> protein
ARCH	Agricultural Research Center-Hays
ARS	Agricultural Research Service
ASI	Anthesis-silking interval
B&MGF	Bill & Melinda Gates Foundation
BAC	Bacterial artificial chromosome
<i>BanII</i>	<i>Bacillus aneurinolyticus</i> II
BC	Backcross
BGI	Beijing Genomics Institute
<i>Bgt</i>	<i>Blumeria graminis</i> f. sp. <i>tritici</i>
BLB	Bacterial leaf blight
BLUP	Best linear unbiased prediction
BMR	Brown midrib
BPH	Brown planthopper
BSST	Bioinformatics and statistical software tools
BYDV	Barley yellow dwarf virus
CA	Carbonic anhydrase
CAAS	Chinese Academy of Agricultural Sciences
CAH	Carbonic anhydrase

CBC	Customized breeding chip
CFT	Confined field trial
CGIAR	Consultative Group on International Agricultural Research
CIM2GTAILs	Second-generation tropically adapted inducer lines
CIMMYT	International Maize and Wheat Improvement Center
cM	CentiMorgan
CMIP	Climate Research Program's Coupled Model Intercomparison Project
CML	CIMMYT Maize Line
CMS	Cytoplasmic male sterile/ sterility
CNV	Copy number variations
COG	Cluster of orthologous genes
CRISPR	Clustered regularly interspaced short palindromic repeats
CRU	Climatic Research Unit
CS	Climate-smart
CSCWR	Single-copy genes between wheat and rice
CSH	Climate-smart hybrid
CSRH	Climate-smart rice hybrids
CSRVs	Climate-smart rice (inbred and hybrid) varieties
DArT	Diversity arrays technology
DH	Doubled haploid
DHN	Dehydrin
DM	Downy mildew
DP	Donor parent
DQP	Designed quantitative trait loci pyramiding
DREB	Dehydration responsive element binding proteins
DRESTs	Drought-regulated expressed sequence tags
DT	Drought-tolerant
DTMA	Drought-Tolerant Maize for Africa Project
<i>Dwf</i>	Dwarfing gene
<i>EcNAC1</i>	<i>EleusinecoracanaNo Apical Meristem1</i>
EDI	Ear digital imaging
EPA	Environmental Protection Agency
<i>Eps</i>	Earliness <i>per se</i>
ERF	Ethylene responsive factors
EST	Expressed sequence tag
F <sub>2</sub>	Second filial generation
FAO	Food and Agriculture Organization
FAOSTAT	Food and Agriculture Organization Statistics
FAW	Fall armyworm
FHB	Fusarium head blight
GAB	Genomics-assisted breeding
GA	Gibberellic acid
GARSS	Genomics-assisted recurrent selection scheme
GB	Genetic background

GBLUP	Genomic best linear unbiased prediction
GBS	Genotyping-by-sequencing
GBTM	Genomics-based trait mapping
GBTS	Genotyping-by-target sequencing
GC	Guanine and cytosine content
GCA	General combining ability
GCA	Global climatic alterations
GCM	Global Climate Model
GEBV	Genome-estimated breeding value
GET	Genome editing tools
GIS	Geographic Information System
GISH	Genomic in situ hybridization
GLAD	Green leaf area duration
GLH	Green leafhopper
GM	Genetically modified
GO	Gene ontology
GP	Gene pool
GPT	Genomic prediction tools
GRIN	Germplasm Resources Information Network (USA)
GS	Genomic selection
GSR-BT	Green super rice-breeding strategy
GSR	Green super rice
GST	Glutathione-S-transferase
GT	Genotyping
GWAS	Genome-wide association study
GXE	Genotype-by-environment
GY	Grain yield
HapMap	Haplotype map
HG	Heterotic group
HHZ	Huanghuazhan
HIR	Haploid induction rate
<i>His</i>	Histidine-rich calcium-binding gene
HRDC	Hybrid Rice Development Consortium
HSP	Heat shock protein
<i>HTL</i>	<i>Hypersensitive to Light</i>
HTMA	Heat-Tolerant Maize for Asia Project
HTPG	High-throughput platform for genotyping
HTP	High-throughput phenotyping
IBPGR	International Bureau of Plant Genetic Resources
IBP	Interconnected breeding populations
ICMP	ICRISAT pearl millet
ICRISAT	International Crops Research Institute for the Semi-arid Tropics
IITA	International Institute of Tropical Agriculture
ILs	Introgression lines
IMAS	Improved Maize for African Soils Project

InDels	Insertions–Deletions
INV	Invertase gene
IPCC	Intergovernmental Panel on Climate Change
IPM	Integrated pest management
IRD	Institute of Research for Development
IRGSP	International Rice Genome Sequencing Project
IRMA	Insect Resistant Maize for Africa Project
IRRI	International Rice Research Institute
ISSR	Inter-simple sequence repeat
IVDMD	In vitro dry matter digestibility
KAI2	KARRAKIN-INSENSITIVE2
KALRO	Kenya Agricultural and Livestock Research Organization
KASP	Kompetitive allele specific PCR
KBLUP	Kinship-enhanced best linear unbiased prediction
KSU	Kansas State University
LAMP	Leucine aminopeptidase
LD	Linkage disequilibrium
LEA	Late embryogenesis abundant
LG	Linkage group
LGS	Low germination stimulant
LIC	Low-input check
LiDAR	Light detection and ranging
LOX	Lipoxygenase
<i>Lr</i>	Leaf rust
LRR	Leucine-rich repeat
<i>Ltn</i>	Leaf tip necrosis
MABC	Marker-assisted backcrossing
MAGIC	Multi-parent advanced generation intercross
MARS	Marker-assisted recurrent selection
MAS	Marker-assisted selection
MCMV	Maize chlorotic mottle virus
MCR	Multi-copy rice genes
MDH	Malate dehydrogenase
ME	Malic enzyme
ME	Mega-environment
MET	Multi-environment trials
MLN	Maize lethal necrosis
<i>mlo</i>	Modulator of defense and cell death gene
MRI	Magnetic resonance imaging
<i>MspI</i>	Moraxella species I
MSV	Maize streak virus
MTA	Material transfer agreement
NAC	No apical meristem
NAM	Nested association mapping
NARS	National Agricultural Research System (India)

NBS	Nucleotide-binding site
NCBI	National Center for Biotechnology Information
NDVI	Normalized difference vegetation index
NGS	Next-generation sequencing
NIL	Near-isogenic lines
NLR	Nucleotide-binding and leucine-rich repeat
NUE	Nitrogen use efficiency
OA	Osmotic adjustment
<i>Opm</i>	<i>Opaque2</i> modifiers
OPVs	Open-pollinated varieties
PAGE	Parametric analysis of gene set enrichment
PAV	Presence-absence variants
PBCM	Process-based crop simulation model
PBMA	Prediction-based model approach
PCR	Polymerase chain reaction
PDLs	Pyramided lines
PEG	Polyethylene glycol
PEPC	Phosphoenolpyruvate carboxylase
PET	Positron emission tomography
<i>PFT</i>	Pore-forming toxin-like
PGD	Phosphogluconate dehydrogenase
PgDwarf8	<i>Pennisetum glaucum</i> Dwarf8 gene
PgHd3a	<i>Pennisetum glaucum</i> heading date 3a gene
PGM	Phosphoglucomutase
<i>PgPHYC</i>	<i>Pennisetum glaucum</i> phytochrome C gene
PGR	Plant genetic resources
<i>Pgt</i>	<i>Puccinia graminis</i> f. sp. <i>tritici</i> n
PHG	Practical haplotype graph
PL	Pyramiding line
PMiGAP	Pearl millet inbred germplasm association panel
<i>Pm</i>	Powdery mildew
<i>Ppd</i>	Photoperiod
<i>PstI</i>	Providencia stuartii I
<i>Pst</i>	<i>Puccinia striiformis</i> f. sp. <i>tritici</i> n
<i>Pt</i>	<i>Puccinia triticina</i>
PYT	Preliminary yield trial
QA/QC	Quality assurance/Quality control
QPM	Quality protein maize
QTL	Quantitative trait locus
QTLs	Quantitative trait loci
R4D	Research for development
RAD	Restriction site-associated DNA sequencing
RAPD	Random amplified polymorphic DNA



RCGS	Rapid cycle genomic selection
rDNA	Recombinant DNA
<i>Rf</i>	Fertility restoration gene
RFLP	Restriction fragment length polymorphism
RGB	Red, green, and blue
RGT	Rapid generation turn over
<i>Rht</i>	Reduced height
RILs	Recombinant inbred lines
ROS	Reactive oxygen species
RP	Recipient Parent
RP	Recurrent parent
Rr1	Resistance gene 1
<i>R</i>	Resistance
RRL	Reduced representation libraries
RRS	Reduced representation sequencing
RS	Recurrent selection
RTK	Real-time kinematic
RUBISCO	Ribulose 1, 5-bisphosphate carboxylase
SAM	Sequence alignment map
SAT	Semiarid tropics
SCA	Specific combining ability
SCAR	Sequence-characterized amplified region
SCA	Sugarcane aphid
SCMR	SPAD chlorophyll meter reading
SCMV	Sugarcane mosaic virus
SCR	Single-copy rice genes
SC	Salinity check
sgRNA	Single-guide RNA
SHW	Synthetic hexaploid wheat
SIM	Simple interval mapping
SKDH	Shikimate dehydrogenase
SMA	Single marker analysis
SME	Small-and medium-enterprise
SNP	Single nucleotide polymorphism
SOC	Soil organic carbon
SRAP	Sequence-related amplified polymorphism
<i>Sr</i>	Stem rust
SSA	Sub-Saharan Africa
SSCP	Single-strand conformational polymorphism
SSD	Single seed descent
SSR	Simple sequence repeat
STARP	Semi-thermal asymmetric reverse PCR
STMA	Stress-Tolerant Maize for Africa Project
STS	Sequence tagged site

TAILs	Tropically adapted inducer lines
TALENs	Transcription activator-like effector nucleases
TB	Transgenic breeding
tGBS	Tunable genotyping-by-sequencing
TKW	Thousand-kernel weight
TLB	Turicum leaf blight
TOI	Target trait of interest
TPEs	Target populations of environments
TRAP	Target region amplification polymorphism
TSILs	Trait-specific introgression lines
TSW	Thousand seed weight
TTE	Target trait environment
UAV	Unmanned aerial vehicle
USAID	US Agency for International Development
USDA	United States Department of Agriculture
VIP	Vegetative insecticidal proteins
<i>Vrn</i>	Vernalization
WA	West Africa
WB	Wheat blast
WCA	West and Central Africa
WEMA	Water Efficient Maize for Africa Project
WGRS	Whole-genome re-sequencing
WGS	Whole-genome sequencing
WTR1	Weed-Tolerant Rice 1
WUE	Water use efficiency
XRF	X-ray fluorescent
<i>Yr</i>	Yellow rust/leaf rust
ZFNs	Zinc-finger nucleases
ZFP	Zinc-finger protein

# Chapter 1

## Genomics-Assisted Breeding of Climate-Smart Inbred and Hybrid Rice Varieties



**Jauhar Ali, Anumalla Mahender, G. D. Prahalada,  
Ma. Anna Lynn Sevilla, Angelito Galang, Erik Jon De Asis,  
Madonna Dela Paz, Corinne Mira Marfori-Nazarea,  
Katrina Leslie Nicolas and Ricky Vinarao**

**Abstract** Global climatic alterations (GCA) such as drought, flood, salinity, submergence, and high/low temperature pose serious threats to crop productivity in several Asian and African countries. As a mitigation option to GCA, we need to develop and deploy climate-smart rice (inbred and hybrid) varieties (CSRVs) resilient to biotic and abiotic stresses. Genomics-assisted breeding (GAB) by integrating advanced genomics tools and an improved green super rice (GSR) breeding strategy (GSR-BT) is one of the leading reliable strategies to develop CSRVs. This provides a high-quality genome sequence and SNPs as allelic variants from the advanced genomics tools that help in understanding the molecular and physiological mechanisms underlying trait expression. It also plays a key role in the quick introgression of the desired genetic variants with the highest precision and less or no genetic drag through an innovative GSR-BT. Under the GSR program at the International Rice Research Institute (IRRI), first, BC<sub>1</sub>F<sub>2</sub> populations derived from Huanghuazhan (HHZ), Weed Tolerant Rice 1 (WTR-1), and TME80518 (recipient parents) and 16 donors were developed. Later, these introgression lines (ILs) underwent simultaneous screenings over three rounds for different abiotic and biotic stresses as well as evaluation under normal irrigated conditions, resulting in the identification of trait-specific ILs. A total of 1333 (HHZ-ILs) + 2232 (WTR-1-ILs) + 1408 (TME80518-ILs) ILs were developed, which were further used either for varietal improvement or as parental lines for the development of pyramided lines (PDLs). Furthermore, a designed QTL pyramiding (DQP) approach began for stacking traits/genes derived from the generated trait-specific ILs. These more productive efforts resulted in the development of 1280 (HHZ-PDLs) + 850 (WTR-1-PDLs) PDLs that showed a significantly superior performance over the tolerant checks. Notably, we developed and distributed

---

J. Ali (✉) · A. Mahender · G. D. Prahalada · Ma. Anna L. Sevilla · A. Galang · E. J. De Asis ·  
M. D. Paz · C. M. Marfori-Nazarea · K. L. Nicolas · R. Vinarao  
Rice Breeding Platform, International Rice Research Institute, 4031, Los Baños,  
Laguna, Philippines  
e-mail: [J.Ali@irri.org](mailto:J.Ali@irri.org)

more than 240 CSRVs with multiple abiotic and biotic stress tolerance to Asia and Africa without compromising on grain yield and quality. A total of 26 IRRI-bred CSRVs were released, and 91 such cultivars were nominated into national cooperative yield trials from three recipient parents within a short span of 7 years. The released CSRVs are being cultivated on more than 2.7 million ha on a cumulative basis in Asia and Africa.

**Keywords** Breeding strategy · Genomics · Climate change · QTL and GWAS · MAS · GSR varieties

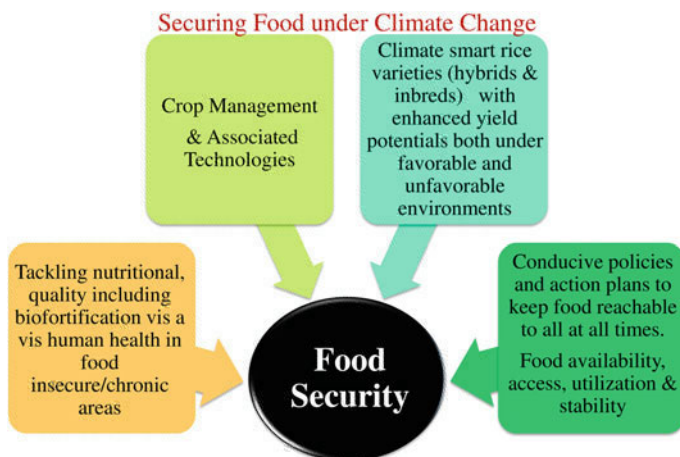
## 1.1 Introduction

Providing sufficient nutritious food to feed the increasing world population is a significant challenge for agriculture (Abberton et al. 2016). The global population is predicted to reach 9.7 billion by 2050, thus requiring agricultural food production to increase by 60.0–110.0% to meet the additional food demand (Tilman et al. 2011). Moreover, global agricultural productivity is severely threatened by biotic and abiotic stresses, reduced arable land, and, most alarmingly, global climate change (Kole et al. 2015; Hampton et al. 2016). The fluctuation in rainfall regimes, day/night temperature on land and sea, and weather patterns; elevated carbon dioxide concentrations; and greenhouse gas emissions in particular regions globally are the impacts of global climatic alterations (GCA) (Batley and Edwards 2016; Beckage et al. 2018). The Intergovernmental Panel on Climate Change (IPCC) estimated that the global surface temperature of Earth is expected to increase by 1.4–5.8 °C by 2100 (IPCC 2007), thereby resulting in a decrease in precipitation in the subtropics and an increase in the possibility of frequent occurrence of extreme climatic events such as drought, high/low temperature, floods, and cyclones. Furthermore, Nelson et al. (2014), Rosegrant et al. (2014), and Iizumi et al. (2017) conducted a combining analysis of biophysical and economic models by using multiple climatic conditions. The analysis concluded that about 11.0% of the grain yield decline in the four major crops (maize, soybeans, rice, and wheat) by 2050 would eventually affect global food and nutritional security, especially in developing countries.

Rice (*Oryza sativa* L.) is one of the major staple food crops of the world that provide the daily calorie intake for more than 50.0% of the world population (Muthayya et al. 2014). This crop is grown on 159 million hectares annually, and more than 90.0% of this rice is produced by 200 million smallholding farmers in Asia and Africa (Mottaleb et al. 2012). It is a semiaquatic annual plant that includes 22 *Oryza* species, of which two species, Asian rice (*O. sativa*) and African rice (*O. glaberrima*), are well domesticated for human consumption, and the other 20 are wild species (Londo et al. 2006). Both the cultivated species have unique domestication histories and include subspecies/subpopulations with distinctive morphological and physiological traits (Choi et al. 2017; Singh et al. 2018). The majority of the rice from *O. sativa* is well described and further classified into five

subpopulations: *aus*, *indica*, temperate *japonica*, tropical *japonica*, and aromatic rice (Garris et al. 2005). The origin of *O. sativa* was first observed in Southeast Asia (India, Myanmar, Thailand, North Vietnam, and China) between 8,000 and 15,000 years ago (Normile 1997; Wei et al. 2012; Civián and Brown 2017), whereas *O. glaberrima* was domesticated from its wild ancestor *O. barthii* in the floodplains of the Niger River in Africa about 3,000 years ago (Portères 1976).

Global rice production is severely affected by global climate change, which in turn challenges the food security of at least 70.0% of the Asian population (Chang et al. 2013; Ali et al. 2017). It has resulted in drought, salinity, flood, submergence, elevated CO<sub>2</sub>, and high/low temperature. Drought is one of the severe constraints that affect rice production. Nearly 23 million hectares of rainfed rice is facing drought stress by showing negative effects on plant growth, physiology, and grain development (Serraj et al. 2011; Fahad et al. 2017). Salinity is the second most abiotic stress after a drought that causes severe hindrance to rice production and productivity depending upon the level of lethality (Gregorio 1997). Similarly, about 20 million hectares of rice land is affected by flooding and submergence leading to complete crop loss (<https://irri.org/our-impact/increase-food-security/flood-tolerant-rice-saves-farmers-livelihoods>). Matthews et al. (1997) revealed that a 1 °C increase in mean daily temperature could lead to a decrease in rice yield by 5–7%, whereas an increase of 1 °C in night temperature could lead to a 10.0% yield reduction (Peng et al. 2004; Yang et al. 2017). Importantly, an increase in temperature during flowering and pollination can drastically reduce pollen viability and ultimately reduce seed set and grain yield (Hatfield and Prueger 2015; Kumar et al. 2015a). In addition to this, elevated CO<sub>2</sub> concentration in the atmosphere resulting from GCA leads to increased damaged rice grains and decreased grain quality (Madan et al. 2012; Liu et al. 2017). Moreover, it has been predicted that worldwide rice consumption would increase from 465 million tons to 487 million tons from 2012 to 2020 (Mohanty 2009) and 25.0% more rice might be required for ensuring food security in 2030 vis-à-vis rice consumption in 2010 (Seck et al. 2012). Hence, there is an urgent need for the introgression of climate-smart (CS) traits such as tolerance to drought, salinity, flood and submergence, elevated CO<sub>2</sub> and temperature, and other stresses like biotic stresses. It is essential to develop climate-smart rice varieties (CSRVs) with CS traits, which is crucial to make them withstand increase or decrease in day and night temperature, elevated CO<sub>2</sub>, and resistant to multiple broad-spectrum stresses including biotic stresses with low input use efficiency to ensure high quality and quantity of rice production. Securing food under changing climatic conditions requires addressing four key elements. First is the development of CSRVs (inbreds and hybrids) that provide stable yields under both favorable and marginal environments. Second is all crop management technologies associated with this. The third is by tackling nutritional quality, including biofortification, vis-à-vis human health. Last is ensuring food availability, accessibility, and use with stability (Fig. 1.1). However, all these require conducive policies and an action plan to keep food accessible to all at all times. Research efforts to cope with climate change are currently integrated into several national agricultural research programs globally.



**Fig. 1.1** Key strategies to secure food under climate change

The increase in rice yield along with resistance to biotic stress and tolerance abiotic stresses has been driven by conventional rice breeding approaches. Till date, most of the ruling mega-varieties were developed through conventional breeding approaches. These varieties were high yielding and possessed key traits to withstand major insect pests and diseases like IR62 and IR64 (Khush 1987). These varieties showed their resistance to major biotic stresses such as blast, bacterial leaf blight, tungro, brown planthopper (BPH), green leafhopper (GLH), and stem borer. Notably, these varieties are still being used as the biotic stress-resistant varieties and checks for BPH and GLH bioassay experiments. On the other hand, for the abiotic stress tolerance, several conventionally bred rice varieties were also released. One among them is Sahbhagi Dhan (IR74371-70-1-1) drought-tolerant rice variety released for cultivation in India, in 2010. It was introduced under direct-seeded upland as well as for lowland conditions. The same variety was released in Bangladesh (as “BRRI Dhan 56”) and in Nepal (as “Sookha Dhan 3”) (Mandal et al. 2009; Dar et al. 2012; Anantha et al. 2016).

Recent advances in genomics technologies, whole-genome sequencing (WGS) using next-generation sequencing (NGS) platforms, genotyping-by-sequencing (GBS), fixed Affymetrix genotyping platforms, and genomic selection (GS) have led to rapid success in discovery and deployment of novel genomic regions that confer useful traits. Understanding the molecular and physiological basis, especially the traits that influence climate resilience, is essential for developing an appropriate breeding strategy (Kole et al. 2015; Rasheed et al. 2017; Ali et al. 2018).

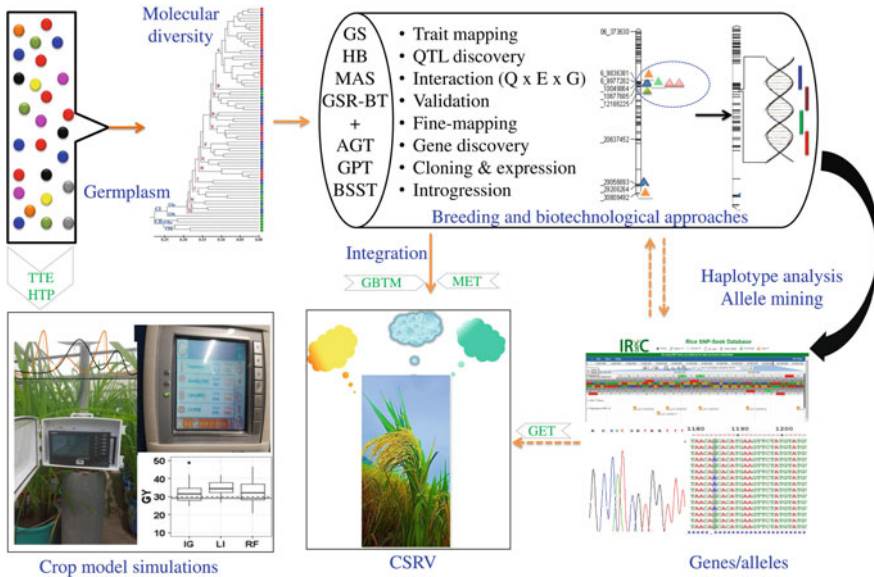
The development of CSRVs needs to explore the genetic variability of the target traits and introduce variation into genotypes of interest through rapid breeding cycles. Hence, there is a need to develop a novel breeding strategy that helps in generating breeding products within a short period by precise introgression of the

expected genetic segments. In this review, we showcase the potential of the integration of novel efficient breeding strategies developed by green super rice (GSR) at IRRI and advanced genomics tools such as genomics-assisted breeding (GAB) to develop and deploy CSRVs in Asian and African countries.

## 1.2 Rice Genomics

Rapid advances in structural and functional genomics helped in cloning over 2000 genes controlling key agronomic traits and partially characterizing their molecular biological mechanisms (Li et al. 2018). Rice has an immense treasure of genomics and bioinformatics resources that can accelerate the development of useful products. It is the first monocot species whose whole-genome sequence (WGS) was known. Among the major world food crops, rice is a model plant for genomics studies (Tyagi and Khurana 2003). It has the smallest genome size (330–430 Mb) compared to other cereals such as sorghum (750 Mb), maize (3000 Mb), barley (5000 Mb), and wheat (16,000 Mb) (Tyagi et al. 2004). Improved de novo sequencing technologies helped in generating a draft rice genome sequence of two rice subspecies, *O. sativa* subsp. *japonica* and *indica*, using Nipponbare and 93–11 varieties, respectively (Goff et al. 2002; Yu et al. 2002). The WGS of Nipponbare published by the International Rice Genome Sequencing Project (IRGSP) with a 6X coverage is being used as a reference genome sequence for the sequencing of several varieties because of its high sequence quality (Yu et al. 2005). The reference genome sequence served as a powerful tool in the mining of useful traits by capturing exact allelic variants, presence–absence variants (PAV), and copy number variations (CNV), along with the dissection of complex traits (Huang and Han 2014). The availability of enormous information on the genomics of rice helped not only in understanding the predicted and annotated rice genes (30,000–50,000) but also in unraveling the genetic interaction controlling the essential agro-morphological traits under varied environmental conditions.

Advanced and cost-effective NGS methods facilitated the rapid dissection of the complex genetics of agronomically essential traits through quantitative trait locus (QTL) mapping and genome-wide association study (GWAS) by developing a large number of genetic markers, including insertion–deletion (InDel) and single nucleotide polymorphism (SNP) markers (Fig. 1.2). NGS platforms increase the resolution of the mapping, which in turn increases the success of precise localization of genetic loci/QTLs (Guo et al. 2014; Yang et al. 2017; Descalsota et al. 2018; Varshney et al. 2018). One of the most important applications of rice genomics is carrying out precise marker-assisted selection (MAS), and GAB approaches. This resulted in the successful introgression of several beneficial traits with less or no genetic drag. These remarkable capabilities of rice genomics studies helped in the rapid understanding of molecular genetics, enhancement of breeding selection strategies, functional genomics through characterizing the specific genes, and also in the use as a reference approach for different high-throughput genome



**Fig. 1.2** Generalized schematic diagram showing different GAB strategies for developing CSRVs

sequencing technologies in other staple food crops (He et al. 2014; Das et al. 2017). In summary, rice genomics studies paved the way for increased genomics studies, particularly in exploring the hidden genetic diversity existing in a large number of cultivars, landraces, and wild accessions that are the richest sources of hidden genetic diversity and shaping it into farmers' chosen varieties. The available bioinformatics tools and databases which are being used to explore structural and functional genomics technologies are presented in Table 1.1.

### 1.3 Different Genomics Approaches Used for Rice Crop Improvement

Developing CSRVs is a major concern for ensuring global food security under erratic changes in climatic conditions (Tong et al. 2016; Varshney et al. 2018). To overcome these challenges, efficient and advanced genomics approaches have been deployed in recent decades to understand the molecular genetics and physiological responses of complex traits and in designing novel breeding strategies for improving genetic gains in rice (Takagi et al. 2013; Bhat et al. 2016; Sun et al. 2018). Although conventional breeding strategies put forward popular rice varieties, development of these varieties through traditional breeding strategies is time-consuming, and it is difficult to dissect the hidden genetic variation existing for essential agronomic traits, which mostly have low heritability and are highly



**Table 1.1** Distribution of SNPs in different customized breeding chip (CBCs) used for molecular genetic studies and breeding programs in rice

Sl. no	SNP array	No. of SNPs	Function	Avg. SNP call rate (%)	No. of genotypes	References
1	50 K	51,478	Molecular genetic diversity and breeding	99.9	801 <sup>b</sup>	Chen et al. (2014a)
2	50 K	50,051	Molecular genetics and phylogenetics	99.9	192 <sup>b</sup>	Singh et al. (2015)
3	44 K	44,100	Molecular genetic diversity and breeding	80.0	500 <sup>b</sup>	Tung et al. (2010)
4	44 K	44,100	Pyramiding for useful genes	92.8	31 <sup>b</sup>	Kurokawa et al. (2016)
5	6 K	5,274	Molecular genetic diversity and breeding	80.0	258 <sup>b</sup>	Thomson et al. (2017)
6	5 K	5,246	Heat stress tolerance	89.0	272 <sup>a</sup>	Ps et al. (2017)
7	6 K	5,102	Genetics for yield-attributed traits	80.0	197 <sup>a</sup>	Tan et al. (2013)
8	6 K	5,102	Grain shape	80.0	197 <sup>a</sup>	Hu et al. (2013)
9	6 K	6,000	Salinity stress tolerance	89.0	220 <sup>b</sup>	Kumar et al. (2015b)
10	6 K	5,291	Molecular genetic diversity	90.0	471 <sup>b</sup>	Xu et al. (2016)
11	6 K	5,274	Heat tolerance	90.0	167 <sup>a</sup>	Ye et al. (2015)
12	6 K	6,000	Salinity stress tolerance	89.0	94 <sup>a</sup>	Gimhani et al. (2016)
13	6 K	5,000	Low-temperature stress tolerance	80.0	230 <sup>a</sup>	Najeeb et al. (2019)
14	6 K	5,000	Weed-competitive ability	80.0	167 <sup>a</sup>	Dimaano et al. (unpublished)

<sup>a</sup>Biparental mapping populations, <sup>b</sup>Diverse rice accessions

influenced by the environment. Hence, the application of advanced genomics strategies is necessary to achieve the colossal task of feeding the alarming growing global population more quickly. The following section describes different genomics approaches used for rice crop improvement.

### 1.3.1 Whole-Genome Sequencing (WGS) of Diverse Rice Genotypes

The rapid development of omics technologies provides novel opportunities to understand the complexity of molecular, physiological, and biochemical network pathways in crops. In recent decades, the WGS of rice accessions has held great potential to identify the genetic background of rice accessions, map the genetic

locus associated with the target trait of interest (TOI), and understand the molecular function of the mapped locus (Li et al. 2014; Duitama et al. 2015; Rahman et al. 2017). The omics technologies including transcriptomics, metabolomics, and proteomics surely benefited from the advancement in NGS technologies as they give high-quality sequence information that helps in better understanding gene design and its mechanism (Rasheed et al. 2017). Increasing the genetic gain of yield and yield-associated traits and tolerance of the major stresses under GCA is crucial for the development of CSRVs. This can be achieved through the identification and incorporation of novel sources of genetic variation as genes/alleles from various genetic resources of rice germplasm (McCouch et al. 2013; Xie et al. 2015; Ali et al. 2017). The current trend in NGS platforms of WGS revolutionized the discovery of novel genomic regions derived from even wild *Oryza* species, which possess quite unique genome characteristics compared with *O. sativa* cultivated varieties, through linkage (QTL mapping) and association mapping (GWAS) approaches, which help in unlocking the hidden novel alleles in them (Varshney et al. 2014; Kole et al. 2015). Under the GSR breeding program, we used a set of 500 varieties (mini-core), which were whole-genome sequenced as a part of the 3K Rice Genomes project, to explore the genetic essence of the breeding lines developed. Among them, only 50 varieties were used for the identification of genomic regions conferring multiple stress tolerance and, eventually, for the development of CSRVs (Ali et al. unpublished). Simultaneous screening of BC<sub>1</sub>F<sub>2</sub> bulk populations in Huanghuazhan (HHZ) recipient parent and eight donor parents over three rounds of multiple abiotic and biotic stresses in comparison with checks and parental lines led to the development of 495 backcross introgression lines (BC<sub>1</sub>F<sub>6</sub> BILs) and these were sequenced using an NGS platform of WGS by the BGI (Ali et al. unpublished). Similarly, WGS platforms were also used for understanding the sequential variations among the genotypes, the diversity pattern of rice accessions, population structure studies, and trait association analysis for QTLs/genes for TOI (Guo et al. 2014; Tang et al. 2016). Genome-wide SNP discovery was initially undertaken from *japonica* (Nipponbare), and then *indica* (93–11) varieties and found 408,898 SNPs and PAV between the two subspecies (Feltus et al. 2004; Shen et al. 2004). Further, a total of 20 diverse varieties and landraces were sequenced in the OryzaSNP project, and 160,000 high-quality polymorphic SNPs were discovered (McNally et al. 2009). In addition, 23 million SNPs, from 94 *O. sativa* varieties and 10 wild relatives (Duitama et al. 2015); 22,682 polymorphic SNPs from 12 parental varieties (Tang et al. 2016); 3.6 million SNPs from 517 rice landraces (Huang et al. 2010); 8 million SNPs from 446 diverse *O. rufipogon* accessions; 794,297 SNPs from 576 introgression lines (ILs) from 11 donors (Ali et al. 2018); and 29 million SNPs and 2.4 million PAV from 3,010 diverse Asian cultivated rice varieties were discovered from the WGS (Wang et al. 2018). Finally, these large numbers of SNPs were used in analyzing the diversity pattern and further validation of these SNPs. The same information can be helpful to understand the functional genomics and transcriptomics for identifying superior alleles and their mechanism in future breeding programs. In summary, WGS provides the highest density of widely distributed SNPs, which enhances the resolution to achieve precise localization of

the putative QTLs/genes conferring tolerance of several biotic and abiotic stresses, which resulted from climate change, and the introgression of these genomic regions would facilitate the development of climate-resilient rice varieties.

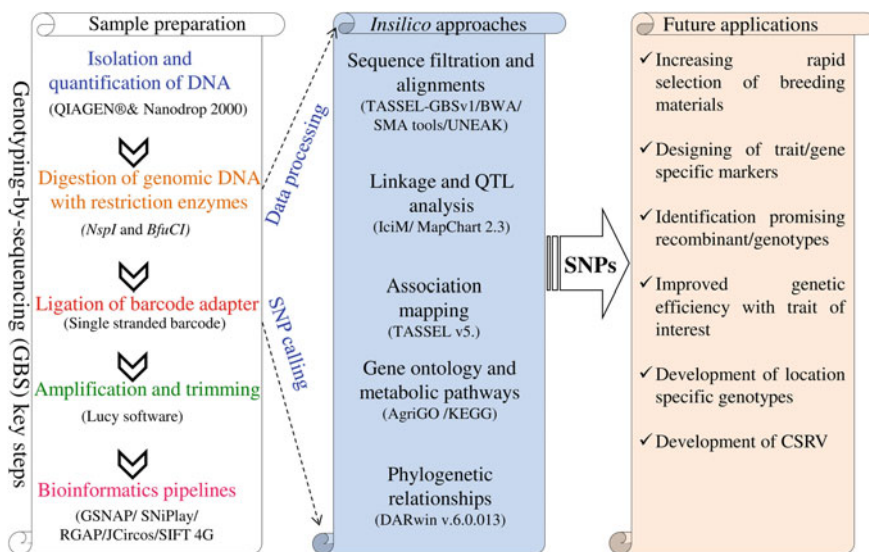
### ***1.3.2 3K Genomes and Their Use in Developing Climate-Smart Breeding Products***

Improving genetic gain by developing CSRVs is essential for increasing global rice production. Harnessing the hidden genetic diversity existing in cultivars through efficient NGS platforms is the prime objective of GAB to accelerate breeding cycles and enhance genetic potential (Seck et al. 2012; Fahad et al. 2014; Ali et al. 2018). Globally, about 780,000 accessions of *Oryza* are maintained in gene banks. Germplasm of Asian origin occupies a larger portion of the total accessions maintained in gene banks across the world (Jockson 2010; Jacob et al. 2015). To explore genetic variation and use more efficiently, the Chinese Academy of Agricultural Sciences (CAAS), Beijing Genomics Institute (BGI) Shenzhen, and IRRI launched a 3,000 (3K) rice genomes project for sequencing 3024 gene bank rice accessions (The 3000 Rice Genomes Project 2014). Among the total 3K rice accessions, 92.0–94.0% of the genome sequence aligned to the reference genome of Nipponbare, which indicated the conserved genomic regions among the total 3K rice accessions. A total of 17 terabytes of raw genomic sequence data generated through various bioinformatics pipeline databases, 40 million SNP variants, and 2.4 million short InDels ( $\leq 50$  bases long) were retrieved from the whole 3K genomes analysis (Alexandrov et al. 2014). The 3K genome sequences provide high-quality, reliable information on the large set of valuable genes for the development of novel CSRVs for higher grain yield, higher tolerance of biotic and abiotic stresses, and higher nutritional grain quality.

Hence, to develop CSRVs by exploring genetic variation, a total of 500 elite varieties belonging to the mini-core collection and part of the 3K set were systematically used under the GSR-BT. The donor and recurrent parents of the GSR-BT underwent WGS using efficient NGS platforms to isolate a large set of reliable SNPs. Interestingly, among the several donor and recipient parents sequenced, only three recipient parents (HHZ, WTR-1, and TME80518) and 16 donor parents were fully used for the early backcross introgression-breeding program of GSR-BT. In this program, the BC<sub>1</sub>F<sub>2</sub> populations derived from HHZ, WTR-1, and TME80518 (recipient parents) and 16 donors at IRRI were screened simultaneously over three rounds for different abiotic and biotic stresses and normal irrigated conditions. This resulted in the identification of 1333 (HHZ-ILs) + 2232 (WTR-1-ILs) + 1408 (TME80518-ILs) trait-specific introgression lines (ILs) and, further, by designed QTL pyramiding efforts, 2023 HHZ-PDLs + 661 (WTR-1-PDLs) pyramiding lines (PDLs) were developed and found superior to the checks for all of the traits studied (Sasaki 2017). The enormous genome sequence

information on these lines was also used for the discovery of genomic regions conferring resistance to several biotic stresses, tolerance of multiple abiotic stresses, nutrient-use efficiency (NUE), and high CO<sub>2</sub> concentrations. Notably, several QTLs governing complex traits were identified by genome sequencing of 495 ILs in an HHZ background and tGBS for 575 ILs in a WTR-1 background (Garcia-Oliveira et al. 2009; Yorobe et al. 2016; Ali et al. 2017, 2018; Pang et al. 2017a; Taghavi et al. 2017; Feng et al. 2018). The same WGS information from the GSR breeding lines will be used again for characterizing the detected genetic loci for molecular and physiological characterization (Fig. 1.3). Interestingly, out of 50 breeding lines sequenced, only three breeding lines were used for gene discovery and varietal improvement. This information opens up the way forward for tapping the insights of the genetic value of the remaining breeding lines for identifying novel traits/genomic regions and developing more efficient CSRVs.

In addition to the several improved breeding lines, we developed and distributed more than 240 rice materials with multiple abiotic and biotic stress tolerance to Asia and Africa without compromising on grain yield and quality. Most convincingly, 26 IRRI-bred climate-resilient materials were released, and 91 such cultivars were nominated into national cooperative yield trials within a short span of 7 years. These varieties are now being cultivated on more than 2.7 million ha on a seed distribution basis alone to farmers in Asia and Africa (Ali et al. 2017; Feng et al. 2018). Newly developed materials breaking three yield barriers and having multiple stress tolerance (GSR IR2-8-Y14-SU3-R2, GSR IR2-5-L10-Y1-Y2, and GSR IR2-1-R5-N1-Y3) are now being shared, which give an average of 5.9 and 18.3%



**Fig. 1.3** Schematic representation of GBS technology, approaches, and applications in developing CSRVs

higher grain yield than the best hybrid check (Mestiso 6) and NSIC Rc38 under MET, respectively (Sasaki 2017). Such high-yielding materials would further augment the ongoing hybrid rice breeding program at IRRI by using them as parental lines. Success might be attributed to innovative introgression breeding and designed QTL pyramiding efforts combined with appropriate cross-tolerance screening and selection techniques that tapped the hidden genetic diversity from a large number of donors into an adaptable varietal background such as HHZ, WTR-1, and TME80518.

### ***1.3.3 GBS and Its Use for QTL Discovery of Climate-Smart Traits***

High-throughput genome sequencing technologies have paved the way to novel approaches to reveal genetic information on an unparalleled scale compared to Sanger sequencing (Pareek et al. 2011; Reuter et al. 2015). Numerous genotyping technologies are widely used by several researchers based on the cost of sequencing, time, genome coverage, reduction in sequence depth, and missing rate (Sonah et al. 2013). GBS, restriction site-associated DNA (RAD) sequencing, and reduced representation libraries (RRL) are a few of the high-throughput genotyping methods that reduce cost and retrieve high-quality SNPs (Davey et al. 2011; Sonah et al. 2013). As compared with RAD and RRL, GBS or next-generation genotyping (NGG) has revolutionized crop genotyping methodology to discover thousands to millions of SNPs across the wide range of species for the identification of genetic variations underlying promising agronomic traits (Huang et al. 2009; Torkamaneh et al. 2016; Scheben et al. 2017; Wickland et al. 2017). It was developed as a rapid, robust, and cost-effective genotyping approach that is capable of extracting numerous SNPs, which are markers of choice and abundantly distributed across the genome (Scheben et al. 2017). Tunable Genotyping-By-Sequencing (tGBS<sup>®</sup>) is a novel technique that involves the ligation of single-strand oligonucleotides and two restriction enzymes to produce overhangs in opposite orientations, which can be amplified and sequenced. As compared with numerous GBS methods, tGBS libraries produce higher average read depth, a low percentage of missing rate, and a higher rate of SNP calling accuracy (>97.0–99.0%) across populations (He et al. 2014; Ott et al. 2017; Pang et al. 2017a; Ali et al. 2018). Hence, the tGBS approach allows the generation of high-quality SNP markers, which are excellent sources in GAB and genetic applications for dissecting many complex traits (Torkamaneh et al. 2016; Bhatia et al. 2018; Feng et al. 2018). Figure 1.3 represents the methodology, different approaches, and applications in GBS.

GBS has the greatest application in gene discovery and its characterization. It has revolutionized association mapping, which relies on historical recombination events, as it increases the statistical power of detection of the rarest allele (Kole et al. 2015). Because of the enhanced resolution of the genetic map, GBS enables

the identification of rare alleles conferring TOI through GWAS. GBS helps in identifying the rarest recombination, which happened during the several recombination events in history. The following are some of the case studies that proved the successful application of GBS in GWAS analysis. Huang et al. (2010) and Huang et al. (2012) discovered several QTLs using GBS as a genotyping platform in GWAS analysis (49 QTLs for 14 agronomic traits, 32 QTLs for flowering time, and 10-grain yield-related QTLs). Through combined gene-based association and haplotype analysis, Wang et al. (2017b) identified 72 QTLs using GWAS. The functional annotation of the detected loci indicated 19 candidate genes associated with seven important QTL regions that were significantly affecting grain quality traits.

On the other hand, GBS has also been used in discovering the genomic regions influencing TOI in biparental mapping populations. As compared to traditional QTL mapping using a low resolving marker system, GBS-based QTL mapping ensures the high-resolution mapping of the genomic regions conferring TOI (Guo et al. 2014), thereby enhancing the dependability of the detected putative QTLs. Considering these advantages, we have also employed GBS as a genotyping platform for the identification of genomic regions conferring resistance to multiple stresses, including biotic and abiotic stresses, NUE, Fe toxicity, and deficiency traits in GSR materials. The developed biparental population that was generated through unique GSR-BT underwent GBS genotyping. Numerous high-quality SNPs extracted from GBS were used for the marker–trait association study using QTL analysis. Ali et al. (unpublished) identified several promising novel loci for the biotic stresses BPH, GLH, blast, BLB, and RGSV. Furthermore, Pang et al. (2017a), Feng et al. (2018), and Ali et al. (unpublished) also proved the effectiveness of the detected QTLs through QTL pyramiding for broadening the genetic base by using the designed QTL pyramiding approach. Research is ongoing to dissect the detected QTLs to identify candidate genes through functional validation by a map-based cloning approach.

In summary, recent trends in molecular breeding and sequencing-based association studies (QTL and GWAS) derived from WGS and/or tGBS have become more popular for dissecting complex agronomic traits and identifying haplotypes in various crops because of the higher success rate and statistical power. Hence, the high-density GBS-SNP marker assay provides crucial insights into the novel genetic architecture that further helps us to identify potential candidate genes/alleles for the development of CSRVs in future breeding programs.

### ***1.3.4 Allele Mining***

Advanced genomics tools, WGS, and GBS platforms provide information about genome sequence variations as favorable alleles, which allows isolation of these alleles that contribute to trait expression. The favorable alleles within the detected