Compendium of Plant Genomes
Series Editor: Chittaranjan Kole

Tapan K. Mondal
Robert J. Henry *Editors*

The Wild Oryza Genomes

Springer

Compendium of Plant Genomes

Series editor

Chittaranjan Kole, Raja Ramanna Fellow, Department of Atomic Energy, Government of India, Kalyani, India

Whole-genome sequencing is at the cutting edge of life sciences in the new millennium. Since the first genome sequencing of the model plant Arabidopsis thaliana in 2000, whole genomes of about 70 plant species have been sequenced and genome sequences of several other plants are in the pipeline. Research publications on these genome initiatives are scattered on dedicated web sites and in journals with all too brief descriptions. The individual volumes elucidate the background history of the national and international genome initiatives; public and private partners involved; strategies and genomic resources and tools utilized; enumeration on the sequences and their assembly; repetitive sequences; gene annotation and genome duplication. In addition, synteny with other sequences, comparison of gene families and most importantly potential of the genome sequence information for gene pool characterization and genetic improvement of crop plants are described.

Interested in editing a volume on a crop or model plant? Please contact Dr. Kole, Series Editor, at ckole2012@gmail.com

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

Tapan K. Mondal · Robert J. Henry
Editors

The Wild Oryza Genomes

Springer

Editors
Tapan K. Mondal
IARI Campus
ICAR-National Research Center on Plant Biotechnology
New Delhi, Delhi
India

Robert J. Henry
Queensland Alliance for Agriculture and Food Innovation
University of Queensland
St Lucia, QLD
Australia

ISSN 2199-4781 ISSN 2199-479X (electronic)
Compendium of Plant Genomes
ISBN 978-3-319-71996-2 ISBN 978-3-319-71997-9 (eBook)
https://doi.org/10.1007/978-3-319-71997-9

Library of Congress Control Number: 2017962553

Printed on acid-free paper

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

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

Chittaranjan Kole

Foreword

Rice is a staple crop for mankind, and all concerted efforts are to improve rice with yields of more than 10 tons/ha possible under favorable environments. The increasing world population coupled with climate change and reduced availability of cultivable land demand continually improved rice cultivars. This calls for extensive research efforts to decode rice genome helping to exploit the maximum genetic potential of the domesticated rice gene pool to improve rice. However to move past the current genetic limits of rice, we must look for some alternative genomic resources.

Wild species are a potential source of many useful genes that may not be present in the gene pool of the domesticated species. They are great sources of many alternative useful alleles. Although there are 24 species available in the genus *Oryza*, only two are currently domesticated species. The other species are largely unexplored, though every species has agronomically useful traits that could be introduced in rice. Several of the species have high drought, salinity or lodging tolerance, and disease and pest resistance, the degree of which is much higher than that in the most tolerant or resistant genotype of rice. Some of the traits such as acid soil tolerance, shade or low light intensity tolerance, and high micronutrient content are unique to wild species. However to exploit these traits, the prerequisite is to generate large-scale genetic and genomic resources. Some international initiatives have been taken up such as OMAP, IOMAP to decode the genomes of these species, to accelerate breeding efforts. To date, 11 species have been sequenced and released into the public domain. Further, several breeding techniques to transfer genes from wild species have been developed and some improved cultivars of rice with the DNA of these wild species have been commercialized. Several genes from wild species have been cloned for various traits and are in process to make transgenic rice which will come to the market in the future. These efforts are collectively generating huge amounts of information on breeding, genetics, genomics, and in other OMICS areas.

I am also pleased to note the effort of the series editor, Prof. Chittaranjan Kole, an internationally acclaimed scientist, in identifying Dr. Tapan K. Mondal, Principal Scientist, ICAR-National Research Centre on Plant Biotechnology, India, and Prof. Robert J. Henry, Director of Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, Australia, who have contributed significantly to wild rice genome

sequencing, for this synthesis volume. I am sure that this book will be very useful to researchers not only working on wild species of rice and rice itself, but also in the field of wild crop genomics, in addition to having utility among science managers and policy makers. I feel humble to write the foreword for this important book.

New Delhi, India
August 2017

Trilochan Mohapatra, Ph.D., FNA, FNASc. FNAAS
Secretary and Director General
Department of Agricultural Research and Education
and Indian Council of Agricultural Research
Ministry of Agriculture and Farmers Welfare
Government of India

Preface to the Series

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

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

1970s was extended to facilitate sequencing of these genomic resources in the last decade of the twentieth century.

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

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

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

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

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

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

Plant genome science has emerged as an important subject in academia, and the present compendium of plant genomes will be highly useful both to students and teaching faculties. Most importantly, research scientists involved in genomics research will have access to systematic deliberations on the plant genomes of their interest. Elucidation of plant genomes is not only

of interest for the geneticists and breeders, but also for practitioners of an array of plant science disciplines, such as taxonomy, evolution, cytology, physiology, pathology, entomology, nematology, crop production, biochemistry, and obviously bioinformatics. It must be mentioned that information regarding each plant genome is ever-growing. The contents of the volumes of this compendium are therefore focusing on the basic aspects of the genomes and their utility. They include information on the academic and/ or economic importance of the plants, description of their genomes from a molecular genetic and cytogenetic point of view, and the genomic resources developed. Detailed deliberations focus on the background history of the national and international genome initiatives, public and private partners involved, strategies and genomic resources and tools utilized, enumeration on the sequences and their assembly, repetitive sequences, gene annotation, and genome duplication. In addition, synteny with other sequences, comparison of gene families, and, most importantly, potential of the genome sequence information for gene pool characterization through genotyping by sequencing (GBS) and genetic improvement of crop plants have been described. As expected, there is a lot of variation of these topics in the volumes based on the information available on the crop, model, or reference plants.

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

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

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

Kalyani, India Chittaranjan Kole

Preface

Economically, rice is the most important staple cereal feeding the world from the ancient times. Academically, it is a model crop plant, which is used to address several basic questions in plant biology, yet its wild relatives are an untapped reservoir of agronomically important alleles that are absent in the rice gene pool. Although the genus *Oryza* is well known due to the importance of rice, the genus includes 25 species, among which only the Asiatic rice (*Oryza sativa*) and the African rice (*Oryza glaberrima*) are used to feed the world. The genus is diverse as indicated by a wide range of chromosome numbers ($2n$ = 24 to 48), different ploidy levels, and genome sizes. Although research efforts from conventional breeding to functional genomics are advanced in rice, comparatively little is known about the wild species. However, many of the wild species of rice are already known for their tolerance to biotic and abiotic stress. Additionally, some of them harbor specific growth and developmental attributes such as profuse tillering, low shattering of mature seeds, presence of a salt gland to pump excessive salt, which upon transfer to cultivated rice could lead to increased productivity and profitability of rice cultivation. However, a prerequisite to understand the genetic factors, which are responsible for such desirable characters present in wild species, is sequencing of the genomes to identify the genes involved. Following the first sequencing of rice in 2002, the genome of 11 species has been sequenced so far and made available for public use recently. The genomes of the remaining species are expected to be sequenced soon. These vast genomic resources will be extremely useful for addressing some of the basic questions about the origin of the genus, evolutionary relationship among the species, domestication, and environmental adaptation and also will be useful to substantiate molecular breeding as well as pre-breeding work to introgress useful characters horizontally from those wild species to cultivated rice.

This book collates the latest state-of-the-art information from a wide range of research domains such as cytology, breeding, physiology, genomics, and proteomics. The volume places emphasis on the latest genomic related works of the 23 species of the genus *Oryza*, public as well as private genomic resources and their impact on genetic improvement research which will be useful to the international research community at large helping to feed 7 billion people in a sustainable manner. The current volume entitled, "The Wild Oryza Genomes" covers genomics and its application for the

improvement of rice breeding. It includes 25 chapters comprising three general topics with the rest on specific species. Chapter 1 deals with an overview of the different species of *Oryza*, their genomic and genetic resources and conservation, and is written by two eminent rice breeders. Due to availability of genome sequence of several wild species, the requirements to manage the vast sequence-related data have also increased and to make them more user-friendly, development of databases has become essential. Chapter 2 describes the various Web resources that are available for wild species of *Oryza*. Speciation remains always attractive to biologist as it is a source of the variation that is the basis for genetic improvement. Chapter 3 dealt with various aspects of the evolutionary relationships among the different species of *Oryza* based on the conserved nuclear-encoded genes and organelle genomes.

Chapters 4–20 have been written on individual species including a brief account of academic interest, the trait they have, and discussion on their chromosomes, breeding potential, availability of genomic and genetic resources, progress on the genome sequence, organelle genome sequences. Each of the chapters also provides a photograph of the adult plant of a particular species and its geographical distribution. Finally, each chapter ends by discussing the research gaps that need to be filled in future. These chapters are presented in the order of the alphabetic name of the species as follows: *O. alta* (Chap. 4), *O. australiensis* (Chap. 5), *O. barthii* (Chap. 6), *O. brachyantha* (Chap. 7), *O. coarctata* (Chap. 8), *O. glaberrima* (Chap. 9), *O. glumaepatula* (Chap. 10), *O. grandiglumis* (Chap. 11), *O. granulate* (Chap. 12), *O. latifolia* (Chap. 13), *O. longiglumis* (Chap. 14), *O. longistaminata* (Chap. 15), *O. meridionalis* (Chap. 16), *O. meyeriana* (Chap. 17, *O. minuta* (Chap. 18), *O. neocaledonica* (Chap. 19), *O. nivara* (Chap. 20).

The officinalis complex consists of ten different species but Chap. 21 describes seven perennial species (*O. officinalis, O. rhizomatis, O. eichengeri, O. minuta, O. malampuzhaensis, O. punctate*, and *O. schweinfurthiana*) in details as the other three species have been described in individual chapters primarily due to the availability of more information about these three species. Considerable success has been achieved in terms of developing genetic stocks through wide hybridization and mapping agronomically important genes from the species of *O. officinalis* complex which have been depicted in the chapter.

Further, four chapters have been deliberated species-wise. They are *O. perenni* (Chap. 22), *O. rhizomatis* (Chap. 23), *O. ridleyi* (Chap. 24) and *O. rufipogon* (Chap. 25). Among these, *O. rufipogon*—the progenitor of present-day cultivated rice, *O. sativa* is one of the most studied wild species of rice. These four chapters provide more details with a special emphasis on genomics and breeding. Every effort has been taken to include all the published reports on these species and to discuss their future potential.

The last chapter, i.e., Chap. 26 with the name, "An account of unclassified species (*Oryza schlechteri*), subspecies (*Oryza indandamanica* Ellis, *Oryza sativa* f. *spontanea* Baker) and ortho-group species (*Leersia perrieri*) of *Oryza*," describes one unclassified species, i.e., *O. schlechteri*, ortho-group species, i.e., *Leersia perrieri*, and two subspecies, i.e., *O. indandamanica* and

O. sativa f. spontanea. The purpose of this chapter is to sensitize the readers with the fact that the *Oryza* genome is very dynamic, still evolving, and ultimately will give birth to new species with useful trait in future.

We also thank our family members for bearing with us throughout the process of editing and finalization of this book. We are also thankful to all 67 authors from 14 different countries who despite their busy schedules have shared their research experience on their respective species in the form of a chapter. We also express our gratitude to Springer-Verlag and its entire staffs particularly Ms. Abirami Purushothaman and Mr. Naresh Kumar for their kind help and understanding in publication and promotion of this book. We made every effort to include all the published research papers in the area of genomics but apologies for those works, if any, which did not appear in this volume despite a detailed search worldwide. Finally, we also thank Prof. C. R. Kole to give an opportunity to edit this important volume.

We are confident that this book will be useful to researchers, both in academia and industry, and policy makers working on not only wild species of rice but also domesticated rice and cereals as a whole.

New Delhi, India — Tapan K. Mondal
St Lucia, Australia — Robert J. Henry

Contents

Abbreviations

AAC	Apparent amylose content
ABC	Advanced backcross
ADB	Africa Development Bank
Adh	Alcohol dehydrogenase
AFLP	Amplified fragment length polymorphism
AILs	Alien introgression lines
AMPs	Advanced mapping populations
ASH	Asymmetric somatic hybridization
BAC	Bacterial artificial chromosome
BADH	Betaine aldehyde dehydrogenase
BB	Bacterial blight
BES	BAC end sequences
BGI	Beijing Genomics Institute
BILs	Backcross inbred lines
BLB	Bacterial leaf blight
BPH	Brown planthopper
BSA	Bulked segregant analysis
CGH	Comparative genomic hybridization
CGIAR	Consultative Group for International Agricultural Research
C_i	Intercellular CO_2 concentration
CIAT	International Center for Tropical Agriculture
CMS	Cytoplasm male sterility
CSSLs	Chromosome segment substitution lines
csSSR	Candidate gene-based SSR
CTAB	Cetyltrimethyl ammonium bromide
CWR	Common wild rice
DArT	Diversity Arrays Technology
DDC	Degeneration divergence complementation
DGE	Digital gene expression technology
DH	Double haploid
EBI	The European Bioinformatics Institute
eIF	Eukaryotic translation initiation factor
ERF	Ethylene responsive factor
ESTs	Expressed sequence tags
FAO	Food and Agriculture Organization
FISH	Fluorescent in situ hybridization
FLcDNA	Full-length cDNA
FPC	Fingerprint contig

GISH	Genome in situ hybridization
g_m	Mesophyll conductance
GMS	Genealogy Management System
GRH/GLH	Green rice/leafhopper
g_s	Stomatal conductance
GSV	Grassy stunt virus
GWAS	Genome-wide association studies
H_e	Genetic diversity
HL-CMS	Honglian cytoplasmic male sterility
H_o	Heterozygosity
IAfRIC	International African Rice Improvement Consortium
ICIS	International Crop Information System
IGS	Intergenic spacer
ILs	Introgression lines
IOMAP	International *Oryza* Map Alignment Project
IRs	Inverted repeats
IRGSP	International Rice Genome Sequencing Project
ISSR	Inter-simple sequence repeat
ITS	Internal transcribed spacer
KASPar	Competitive allele-specific PCR
LD	Linkage disequilibrium
LINEs	Long interspersed nuclear elements
LRR	Leucine-rich repeat
LSC	Large single copy
LTRs	Long terminal repeats
MAALs	Monosomic alien addition lines
MAB	Marker-assisted breeding
MALDI-TOF	Matrix-assisted laser desorption ionization time of flight
MAPK	Mitogen-activated protein kinase
MARs	Matrix attachment regions
MAS	Marker-assisted selection
MENERGEP	Methodologies and new resources for genotyping and phenotyping
miRNA	microRNA
MITEs	Miniature inverted-repeat transposable elements
MLS	Multilateral system
NBS-LRR	Nucleotide binding sites and leucine-rich repeats
NCBI	National Center for Biotechnology Information
NCGR	National Center for Genome Resource
NCGRP	National Center for Genetic Resources Preservation
NERICA	New Rice for Africa
NGS	Next-generation sequencing
NIG	National Institute of Genetics
NILs	Near isogenic lines
OEC	Oxygen-evolving complex
OGEP	*Oryza* genome evolution
OMAP	*Oryza* Map Alignment Project
OSCA	Hyperosmolality-gated Ca-permeable channels
PAGE	Polyacrylamide gel electrophoresis

PAL	Phenylalanine ammonia lyase
PGRC	Plant Genetic Resources Centre
Pn	Net photosynthetic rate
P_n/c_i	Carboxylation efficiency
P_n/g_s*(WUEi)*	Intrinsic water use efficiency
P_n/T	Transpiration efficiency
PReDA	Plant Repeat Database
PS	Photosystem
PV	Phenotypic variance
QTL	Quantitative trait loci
RAM	Rapid Alleles Mobilization
RAP	Rice Annotation Project
RAPD	Random amplified polymorphic DNA
RAP-DB	Rice Annotation Project Database
RcbL	Rubisco large subunit
Rf	Fertility restoration
RGKbase	Rice Genome Knowledgebase
RILs	Recombinant inbred lines
RiTE-db	Rice TE database
RT-PCR	Real-Time polymerase chain reaction
RYMV	Rice yellow mottle virus
SGSV	Svalbard Global Seed Vault
SINEs	Short interspersed nuclear elements
SNP	Single nucleotide polymorphism
SOAP	Short Oligonucleotide Analysis Package
SSC	Small single copy
STMS	Sequence-tagged microsatellite
TDM	Total dry matter
TEs	Transposable elements
TFs	Transcription factors
TRF	Tandem Repeat Finder
TSS	Transcription start site
VCF	Variant Call Format
WA-CMS	Wild abortive cytoplasmic male sterility
WBPH	White-backed Planthopper
WGS	Whole-genome shotgun
WPM	Woody Plant Medium

Wild Relatives of Rice: A Valuable Genetic Resource for Genomics and Breeding Research

1

Darshan Singh Brar and Gurdev S. Khush

Abstract

Worldwide, more than 3.5 billion people depend on rice for more than 20% of their daily calories. Global rice demand is estimated to rise from 723 million tons in 2015 to 763 million tons in 2020 and to further increase to 852 million tons in 2035, an overall increase of 18% or 129 million tons in the next 20 years. World rice production has more than doubled from 257 million tones in 1966 to 680 million tons in 2010. This was mainly achieved through the application of principles of classical Mendelian genetics and conventional plant breeding. Further, rice productivity is continually threatened by several diseases (bacterial blight, blast, tungro virus, rice yellow mottle virus, sheath blight, etc.) and insects (plant hoppers, stemborer, and gall midge) including many abiotic stresses (drought, salinity, submergence, cold, heat, soil toxicities, etc.). To overcome these constraints particularly in the context of global climatic changes, there is urgent need to broaden the gene pool of rice; one of the options is to exploit wild species of *Oryza* which are reservoirs of useful genes/QTLs for rice improvement. Interspecific hybrids, alien introgression lines (AILs), chromosomal segmental substitution lines (CSSLs) have been produced. Several genes/QTLs governing agronomic traits have been transferred from wild species into rice, and a few of these tagged with molecular markers and used in marker-assisted selection (MAS). Some breeding lines of rice derived from wide crosses have been released as varieties.

1.1 Wild Relatives of Rice

The genus *Oryza* has two cultivated and 24 wild species ($2n$ = 24, 48 chromosomes) representing 11 genomes: AA, BB, CC, BBCC, CCDD, EE, FF, GG, HHJJ, KKLL (Vaughan 1989, 1994; Aggarwal et al. 1997; Ge et al. 1999; Table 1.1). Of the two cultivated species, *O. sativa* ($2n$ = 24, AA) commonly referred as 'Asian rice' is high yielding and cultivated worldwide, whereas *O. glaberrima* ($2n$ = 24, AA) known as 'African rice' is low yielding and grown in a limited area in West Africa. The wild species are grass-like plants that are weedy and inferior in morphological traits, having poor plant type, poor grain characteristics, low grain yield and are shattering in nature. These wild species differ markedly in morphological characteristics such as growth

D. S. Brar (✉)
School of Agricultural Biotechnology, Punjab Agricultural University, Ludhiana, Punjab, India
e-mail: darshanbrar@pau.edu

G. S. Khush
University of California, Davis, CA, USA

T. K. Mondal and R. J. Henry (eds.), *The Wild Oryza Genomes*,
Compendium of Plant Genomes, https://doi.org/10.1007/978-3-319-71997-9_1

habit, height, flowering, leaf size, morphology, and panicle size, panicle branching, awning, and seed size and adaptation to different habitats (Fig. 1.1; Table 1.1). In spite of their weedy nature, these wild species are important genetic resource for breeding and genomics research and are reservoir of useful genes/QTLs for tolerance to major biotic (diseases, insects) and abiotic (drought, salinity, heat) stresses, yield-related traits including weed-competitive ability, new source of cytoplasmic male sterility (CMS), and other traits related to rice improvement.

Oryza probably originated at least 130 million years ago and spread as a wild grass in Gondwanaland, the super continent that eventually broke up and drifted apart to become Asia, Africa, the Americas, Australia, and Antarctica. The genus *Oryza* has been classified into four species complexes: (i) *sativa* complex, (ii) *officinalis* complex, (iii) *meyeriana* complex, and (iv) *ridleyi* complex (Table 1.1). *O. schlechteri* and *O. coarctata* are placed in an unclassified group.

1.1.1 *O. sativa* Complex

This complex consists of 8 diploid species ($2n = 24$) belonging to AA genome, of which two are cultivated (*O. sativa* and *O. glaberrima*) and six are wild species (*O. nivara*, *O. rufipogon*, *O. breviligulata*, *O. longistaminata*, *O. meridionalis*, and *O. glumaepatula*) (Table 1.1). These species form the primary gene pool and are easily crossable with rice commonly used in transfer of genes into rice cultivars. *O. rufipogon* is distributed in South and Southeast Asia including Northern Territory and Queensland of Australia, commonly found in swamps, marshes, swampy grasslands, and in deepwater rice fields. *O. nivara* is an annual form distributed in India, Nepal, Cambodia, Laos, Thailand, whereas *O. rufipogon* exists as perennial populations. The three taxa–*O. sativa, O. nivara,* and *O. rufipogon*–together with the weedy race (*O. sativa* f. *spontanea*) form a large species complex.

O. longistaminata, a native of Africa, is closely related to its annual relative *O. barthii* which grows in swampy areas, river, or stream sides. These species, closely related to *O. glaberrima,* are somewhat easier to distinguish from each other. *O. glaberrima* is cultivated in a limited area in West Africa in upland and rain-fed ecology and also in deepwater fields. Perennial and annual relatives of *O. glaberrima* are *O. longistaminata* and *O. breviligulata,* respectively. However, now many scientists consider that *O. barthii* was domesticated to produce *O. glaberrima*. *O. glaberrima* is distinguished from *O. sativa* by its short, rounded ligule, panicle-lacking secondary branches, and almost glabrous lemma and palea. *O. glaberrima* is not as variable as *O. sativa.*

1.1.2 *O. officinalis* Complex

The *O. officinalis* complex consists of 12 species: 6 diploid and 6 allotetraploid species (*O. schweinfurthiana*, *O. minuta*, *O. malampuzhaensis*, *O. latifolia*, *O. alta,* and *O. grandiglumis*) (Table 1.1). Some of the species grow in partial shade or moist soil; others are adapted to swamps and seasonal pools of water and open habitat. This complex has related species groups in Asia, Africa, and Latin America. In Asia, the most common species is *O. officinalis,* widely distributed in South and Southeast Asia and South and Southwest China. *O. officinalis* thrives in partial shade or full sun. In the Philippines, it is called bird rice.

O. minuta is distributed in the Philippines and is sympatric with *O. officinalis.* It grows in shade or partial shade along stream edges. Only a few populations of tetraploid *O. officinalis*, *now classified as O. malampuzhaensis,* have been found localized in neighboring parts near the town of Malampuzha of Kerala and Tamil Nadu in South India. A new species from Sri Lanka, *O. eichingeri,* was distributed in Sri Lanka (Vaughan 1989).

In Africa, the two species of the *O. officinalis* complex are *O. punctata* and *O. eichingeri.* The American species of this complex (*O. latifolia, O. alta,* and *O. grandiglumis)* are allotetraploid ($2n = 48$) with a CCDD genome. *O. latifolia* is widely distributed, growing in Central and South America as well as in the Caribbean islands;

Table 1.1 Chromosome number, genomic composition and distribution of *Oryza* species, and their useful traits

Species	2*n*	Genome	No. of accessions[a]	Distribution	Useful traits
O. sativa complex					
O. sativa L.	24	AA	116,751	Worldwide	Cultigen, highly productive, grown worldwide
O. glaberrima Steud.	24	A^gA^g	1655	West Africa	Cultigen; low yielding but has tolerance to drought, acidity, iron toxicity; resistance to blast, RYMV, African gall midge, nematodes; weed competitiveness
O. nivara Sharma et Shastry	24	AA	1503	Tropical and subtropical Asia	Resistance to grassy stunt virus, BB, blast, BPH, yield-enhancing loci(QTLs)
O. rufipogon Griff.	24	AA	1034	Tropical and subtropical Asia, tropical Australia	Resistance to BB, BPH, tungro virus; moderately tolerant to Shb, tolerance to aluminum and soil acidity; source of CMS, yield-enhancing loci (QTLs)
O. breviligulata A. Chev.et Roehr (*O. barthii*)	24	A^gA^g	207	Africa	Resistance to GLH, BB; drought avoidance
O. longistaminata A. Chev. et Roehr .	24	A^IA^I	216	Africa	Resistance to BB, nematodes, stemborer, drought avoidance
O. meridionalis Ng	24	A^mA^m	53	Tropical Australia	Elongation ability; drought avoidance
O. glumaepatula Steud.	24	$A^{gp}A^{gp}$	54	South and Central America	Resistance to blast, elongation ability; source of CMS
O. officinalis complex					
O. punctata Kotschy ex Steud	24	BB	46	Africa	Resistance to BPH, zigzag leafhopper
O. schweinfurthiana Prodoehl	48	BBCC	35	Africa	Resistance to BPH, zigzag leafhopper
O. minuta J.S. Presl. ex C.B. Presl.	48	BBCC	62	the Philippines and Papua New Guinea	Resistance to BB, blast, BPH, GLH
O. malampuzhaensis Krishnasw. & Chandrasekh	48	BBCC	13	India	Resistance to thrips, BPH, GLH, WBPH, BB, stem rot
O. officinalis Wall ex Watt	24	CC	275	Tropical and subtropical Asia, tropical Australia	Resistance to thrips, BPH, GLH, WBPH, BB, stem rot
O. rhizomatis Vaughan	24	CC	20	Sri Lanka	Drought avoidance
O. eichingeri A. Peter	24	CC	22	South Asia and East Africa	Resistance to BPH, WBPH, GLH
O. latifolia Desv.	48	CCDD	58	South and Central America	Resistance to BPH, high biomass production
O. alta Swallen	48	CCDD	12	South and Central America	Resistance to striped stemborer; high biomass production

(continued)

Table 1.1 (continued)

Species	2*n*	Genome	No. of accessions[a]	Distribution	Useful traits
O. grandiglumis (Doell) Prod.	48	CCDD	10	South and Central America	High biomass production
O. australiensis Domin.	24	EE	36	Tropical Australia	Resistance to BPH, BB, blast; drought avoidance
O. brachyantha A. Chev. et Roehr.	24	FF	17	Africa	Resistance to BB, yellow stemborer, leaf folder, whorl maggot; tolerance to laterite soil
O. meyeriana complex					
O. granulata Nees et Arn. ex Watt	24	GG	23	South and Southeast Asia	Shade tolerance; adaptation to aerobic soil
O. meyeriana (Zoll. et (Mor. ex Steud.) Baill.	24	GG	9	Southeast Asia	Shade tolerance; adaptation to aerobic soil
O. neocaledonica Morat	24	GG	1	New Caledonia	–
O. ridleyi complex					
O. longiglumis Jansen	48	HHJJ	6	Irian Jaya, Indonesia, and Papua New Guinea	Resistance to blast, BB
O. ridleyi Hook. F.	48	HHJJ	14	South Asia	Resistance to blast, BB, blast, stemborer, whorl maggot
Unclassified					
O. schlechteri Pilger	48	HHKK	1	Papua New Guinea	Stoloniferous
O. coarctata Roxb.	48	KKLL	1	South Asia	Salt tolerance

BPH brown planthopper; *GLH* green leafhopper; *WBPH* white-backed planthopper; *BB* bacterial blight; *Shb* sheath blight; *CMS* cytoplasmic male sterility; *RYMV* rice yellow mottle virus

[a]Accessions maintained in rice genebank at IRRI, the Philippines

Modified from Khush (1997), Brar and Singh (2011). Taxonomy changes harmonized with recent literature and GRIN (Courtesy of Dr. S. Hamilton, IRRI, the Philippines)

- *O. breviligulata* is now *O. barthii*
- Tetraploid *O. punctata* is now *O. schweinfurthiana*
- Tetraploid *O. officinalis* is now *O. malampuzhaensis*
- *Porteresia coarctata* is now *O. coarctata*

O. alta and *O. grandiglumis* grow only in South America, primarily in the Amazon basin. A diploid species of this complex, *O. australiensis*, (EE) occurs in northern Australia in isolated populations and grows under wet places, seasonally dry pools, river levees, and also in open habitats.

O. brachyantha (2*n* = 24, FF) is distributed in the African continent. It grows in the Sahel zone and in East Africa in ponds, in shallow water, and in open habitats in granite/laterite soils. It is often sympatric with *O. longistaminata*. Of all the species, it is most closely related to the genus *Leersia*. This species has a small, narrow spikelet with long awns (6–17 cm).

1.1.3 *O. ridleyi* Complex

This complex comprises two tetraploid species, (*O. ridleyi* and *O. longiglumis* 2*n* = 48, HHJJ),

Fig. 1.1 Plants of some wild species of *Oryza*

which usually grow in shaded habitats beside rivers, streams, or pools. The *O. ridleyi* complex is primarily found in Southeast Asia (Cambodia, Malaysia, Myanmar) and New Guinea, whereas *O. longiglumis* is distributed in Indonesia (Iran Jaya) and Papua New Guinea. *O. ridleyi* and *O. longiglumis* are very similar in morphology and ecology.

1.1.4 *O. meyeriana* Complex

Species of *O. meyeriana* complex differ from other three complexes (i.e., *O. sativa*, *O. officinalis, O. ridleyi*) in that they are small-sized plants and have unbranched panicles with small spikelets. This complex has two species (*O. meyeriana* and *O. granulata,* 2*n* = 24, GG). *O. granulata* grows in South Asia, Southeast Asia, and Southwest China, whereas *O. meyeriana* is found in Southeast Asia (Ellis 1985). Unlike other complexes, *O. meyerina* and *O. granulata* species do not grow in permanently or seasonally flooded water. This species complex grows in the shade or partial shade in forests. *O. granulata* is called forest rice by tribal people of Kerala, South India; peacock rice in parts of Vietnam; and bamboo rice in the Philippines. Recently, a new diploid species, *O. neocaledonica,* has been included in this complex (Table 1.1)

1.1.5 Unclassified Complex

O. schlechteri: This is a tetraploid species (2*n* = 48, HHKK) distributed in Indonesia (Irian Jaya) and Papua New Guinea. Richard Schlechter first collected it in 1907 from Northeast New Guinea. Vaughan and Sitch (1991) recollected it as living material from the same location. Naredo et al. (1993) reported that the presence of a sterile lemma and a striated spikelet epidermal (abaxial) surface lacking siliceous triads in *O. schlechteri* allies this species with other *Oryza* species. It is a tufted perennial, 30–40 cm tall, with an erect, 4- to 5-cm panicle and small, unawned spikelets. It is a stoloniferous species and can grow either under full or partial shade.

Porteresia coarctata is now classified as *O. coarctata*; it is a tetraploid species (2*n* = 48 KKLL) commonly found in coastal areas of South Asia. It has unusual anatomy, including glands to secrete salts. It has rough, erect leaves and occurs in brackish water. The species is characterized by large caryopses with a somewhat bent apex, a large embryo relative to the

endosperm, and a short petiole attachment at the base. The leaf blade is coriaceous with prickly tuberculate margins and has a peculiar arrangement of vascular bundles; each rib contains one smaller vascular bundle and below it a larger one.

1.1.6 Related Genera

Besides *Oryza,* the tribe *Oryzeae* contains 10 other genera. Vaughan (1989, 1994) has given a brief description of these different genera. These genera include *Chikusichloa, Hygroryza, Leersia, Luziola, Prosphytochloa, Rhynchoryza, Zizania, Zizaniopsis, Potamophila,* and *Porteresia.*

1.2 Exploration and Conservation of Wild Species of Oryza

Rice genetic resources comprise different landraces, modern and obsolete varieties, genetic stocks, and the wild *Oryza* species. The International Rice Gene Bank at the International Rice Research Institute (IRRI), the Philippines, conserves the largest and most diverse collection of rice germplasm comprising 116,751 accessions of rice (*O. sativa*), 3728 accessions of wild species, and 1655 accessions of *O. glaberrima* (Table 1.1). Seeds of wild species including other rice accessions can be obtained from IRRI. The facilities of the gene bank ensure the long-term conservation of this valuable gene pool. The seeds are stored at −20 °C for long-term storage and 2–4 °C as an active collection for distribution and use in research. The IRRI gene bank has 3728 accessions of 24 wild species of *Oryza*, 1655 accessions of *O. glaberrima* (cultivated African rice), and 116,751 accessions of *O. sativa* (cultivated Asian rice). The largest number of wild species (3067) accessions is of the *O. sativa* complex followed by 606 (*O. officinalis* complex), 33 (*O. meyerina* complex), 20 accessions of the *O. ridleyi* complex and one accession each of *O. schlechteri* and *O. coarctata* (Table 1.1).

Since the establishment of IRRI, over the last 50 years, these genetic resources have been maintained and shared with the global scientific community. These wild species have also been characterized based on morphological, cytological, biochemical, and molecular markers.

Besides IRRI, other centers also maintain and conserve seeds of some selected wild *Oryza* species in their respective gene banks including Africa Rice Center, Cotonou, Benin (West Africa); International Institute of Tropical Agriculture (IITA), Nigeria; National Bureau of Plant Genetic Resources (NBPGR), New Delhi, India; National Institute of Agrobiological Resources (NIAR), Japan; China National Rice Research Institute, Huangzou, China; Cuu Long Delta Rice Research Institute (CLRRI) Omon, Cantho, Vietnam; Kaesart University, Bangkok, Thailand; Malaysian Agricultural Research and Development Institute (MARDI), Malaysia; University of Queensland, Australia. Some of these institutes maintain wild species germplasm for active use or only for short-term storage.

1.2.1 Exploration and Collection of Germplasm

Records of early occasional exploration and collection for wild species of rice in different countries can be found in various publications (Oka 1988). Specimens of wild *Oryza* species can be found in many herbaria, particularly in India, China, Indonesia, Thailand, and Malaysia. Efforts for collection and conservation of wild *Oryza* species were initiated in the late 1950s by many National Agricultural Research Systems (NARS) on small scales, along with conservation programs for cultivated rice varieties and landraces. India and China are the world's two largest rice-producing and rice-consuming countries, and wild *Oryza* species are found abundantly in both countries. Exploration and collection of wild species in China can be traced back to as early as 1917 when Dr E. D. Merrill and colleagues first found *O. rufipogon* at Lofu Mountain and the Shilong Plain in Guangdong

Province (Wu 1990). Professor Ding Ying initiated the systematic exploration and collection of wild *Oryza* species in China. In 1926, he found *O. rufipogon* and collected its samples in many more sites, such as those in Guangzhou, Heiyang, Zhengcheng, Qingyuan, and Sashui in Guangdong. In the 1960s, many Chinese scientists from various agricultural research stations and universities organized several expedition trips, which covered larger areas of the wild *Oryza* species 'natural range.' Several exploration and collection missions at different scales were undertaken in the late 1980s and early 1990. Lu and Sharma (2003) have reviewed the exploration and collection of wild *Oryza* species in different countries/regions.

IRRI has been actively involved in collaboration with NARS and International Board for Plant Genetic Resources in organizing workshops on collection, conservation, and utilization of rice genetic resources. Two workshops were organized in 1977 and 1983. The third international workshop established new initiatives for international cooperation and conservation activities of wild *Oryza* species. Since then collection of wild *Oryza* species has gradually received more attention by the NARS, particularly in Asian countries, and more intensive and systematic collecting activities have been conducted in different countries. Based on reports by several NARS programs, a nominal number of seed samples of wild *Oryza* specks were collected in the early 1980s. Lu (1998, 1999) emphasized conservation of wild species of rice in Nepal and other countries.

Much germplasm exploration for rice was completed by the early 1990s. By the end of 1962, the IRRI varietal collection contained 6867 accessions from 73 countries. By 1972, the collection had grown to 14,600 accessions (Chang 1972, 1976). By the early 1980s, the number of accessions in the IRRI gene bank reached 49,027. More than 200 accessions of rice were collected during the second half of 1995 from Lao People's Democratic Republic (PDR). The IRRI gene bank now contains 3728 accessions of 24 wild species of *Oryza,* besides 1655 accessions of African cultivated rice (*O. glaberrima*).

1.2.2 Genetic Erosion

Wild species of *Oryza* are important reservoirs of useful genes for rice improvement. Through millions of years of evolution and genetic adaptation to variable environments, wild species have accumulated abundant biodiversity. Genetic erosion or loss of biodiversity of rice varieties has been recognized as a problem since the 1960s. Factors such as the adoption of high-yielding rice varieties, farmer's increased integration into the markets, change of farming systems, industrialization, human population increases, and cultural change have significantly accelerated the continual erosion of the rice gene pool (Bellon et al. 1998). A similar situation has also been observed for wild *Oryza* species. In many places of Asia, populations of wild *Oryza* species are becoming extinct or are threatened because their natural habitats are seriously damaged by extension of cultivation areas, expansion of communication systems such as road construction, and urban pressures. According to unpublished data collected by the Chinese Academy of Sciences in 1994, nearly 80% of the common wild rice (*O. rufipogon*) sites recorded during the 1970s have already disappeared (cited by Lu and Sharma 2003). The size of some surviving *O. rufipogon* populations was also found to be significantly reduced. A similar situation has been observed in other countries such as the Philippines, Vietnam, Thailand, Nepal, Indonesia, Malaysia, India, and Bangladesh.

The problems of genetic erosion are severe, but international efforts to conserve rice genetic resources, in which IRRI has taken a leading role, have led to the establishment of several gene banks in Asia. These joint efforts between national, regional, and international organizations ensure the long-term conservation of the biodiversity of the rice gene pool.

1.2.3 Conservation Strategy

For many plant species, ex situ conservation of seeds is safe and cost-efficient, provided proper attention is paid to seed drying and storage

conditions. Fortunately, rice seeds exhibit orthodox storage behavior and can be dried to a low moisture content of ca. 6% and stored at −20 °C, retaining their viability for decades, if not longer. Vaughan (1994) has elaborated on Herbarium specimens of various wild *Oryza* species preserved in many herbaria of different countries. Jackson (1997) has suggested strategies for conservation of genetic resources. There are two basic approaches to germplasm conservation (ex situ and in situ conservation).

1.2.3.1 Ex Situ Conservation

In this approach, genetic resources are actually removed from their original habitat or natural environment. Ex situ conservation provides efficient means for germplasm preservation, utilization, exchange, and information generation through effective management and value-added research of the conserved wild rice species. However, seed samples placed under ex situ conservation in a gene bank become isolated from the *Oryza* ecosystem where they originated and grew. The expected microevolution of these *Oryza* species in their original environment is stopped, particularly the adaptive variations that could occur during change in environmental conditions. Therefore, in evolutionary terms, ex situ conservation is static (Bellon et al. 1998). Concerns have been raised following the observation that static conservation may reduce the adaptive potential of wild *Oryza* species and their populations in the future. Thus, ex situ conservation cannot be considered the only approach for conserving biodiversity of wild *Oryza* species. Complementary dynamic approaches such as in situ conservation are also necessary.

The long-term conservation of rice genetic resources is the principal aim of the International Rice Gene Bank (IRG). The gene bank has operated since 1977, although genetic conservation activities started in the early 1960s. For several countries, including Sri Lanka, Cambodia, Lao PDR, and the Philippines, the germplasm conserved in the IRG represents a more or less complete duplicate of their national collections. For other countries, such as India and the People's Republic of China, only part of their national collections are duplicated at IRRI. Nevertheless, the IRG has provided an important safety net for national conservation efforts. On several occasions, it has been possible to restore rice germplasm that had been lost in national gene banks with accessions already conserved at IRRI. IRRI maintains an active collection for medium-term storage and distribution of rice germplasm, at +2 °C in sealed, laminated aluminum foil packets, and long-term (50–100 years) conservation, at −20 °C, each with two vacuum-sealed aluminum cans.

The germplasm collection is held in trust by IRRI under the auspices of FAO in an International Network of ex situ collections. Duplicate storage of the IRG collection is maintained at the National Seed Storage Laboratory (NSSL), Fort Collins, USA, and about 75% of the collection is currently stored under black-box conditions. Duplicate storage of African rices is shared between IRRI, the International Institute of Tropical Agriculture (IITA) in Nigeria, and the Africa Rice Center (earlier named as WARDA) in Benin, Africa.

1.2.3.2 In Situ Conservation

This method attempts to preserve the integrity of genetic resources by conserving them within the evolutionary dynamic ecosystems of their original habitat or natural environment. Under in situ conservation, local control of traditional rice varieties will ensure that benefits accrue to farmers and communities that have developed them. For long-term and dynamic conservation, the in situ approach has great value. However, for some reasons, in situ conservation has, in general, received the least attention and even been rejected. Limited scientific and financial inputs are constraints in in situ conservation and its design and management for wild *Oryza* species.

1.3 Gene Transfer from Wild Species into Rice

Recent advances in tissue culture, genetic engineering, molecular cytogenetics, comparative genetics, and genomics, particularly in rice

genome sequencing, have opened new opportunities to develop improved rice germplasm with novel genetic properties, and in understanding the function of rice genes. Breeders have successfully used conventional breeding methods and exploited the rice (*O. sativa*) gene pool to develop high-yielding improved rice varieties resistant to pests and abiotic stresses with improved quality characteristics. The major emphasis has been on utilizing indica, japonica, and javanica germplasm through intraspecific hybridization (indica × indica, japonica × japonica, indica × japonica). In several cases, genetic variability for target agronomic traits is limited in the cultivated rice gene pool. Under such situations, interspecific hybridization is an important plant-breeding approach to introduce novel genes for different agronomic traits from wild species into rice.

The genus *Oryza* has 24 wild species ($2n$ = 24, 48). The number has been 22 for many years; later 2 were reclassified (tetraploid *O. punctata* and teterploid *O. officinalis*), added as *O. schweinfurthiana* and *O. malampuzhaensis,* respectively, and became 24. These wild species are reservoirs of many useful genes, particularly for resistance to major biotic and abiotic stresses (Table 1.1). However, these wild species are associated with several weedy traits, such as grain shattering, poor plant type, poor grain characteristics, and low seed yield. Besides, several incompatibility barriers such as pre- and post-fertilization barriers, hybrid sterility, limited recombination between the chromosomes of cultivated and wild species, hybrid breakdown and linkage drag limit the transfer of useful genes from wild species into cultivated species (Brar and Khush 1986, 1997). The major consideration in alien gene transfer is to selectively transfer agronomically important genes from wild species, avoiding linkage drag. To achieve precise transfer of genes from wild to cultivated species, strategies involving a combination of conventional plant-breeding methods with tissue culture and molecular approaches have become important (Brar and Khush 2002, 2006). Advances in tissue culture, molecular marker technology, genomics, and genomic in situ hybridization (GISH) have opened new opportunities to tap and characterize alien genetic variability even from distant genomes of *Oryza* through interspecific hybridization. Useful genes for resistance to BPH, BB, blast, and tungro and tolerance to acid sulfate conditions and cytoplasmic male sterility have been transferred from wild species of rice. Some of the breeding lines derived from wide crosses have been released as varieties.

1.3.1 Strategies for Alien Gene Transfer

Strategy to transfer genes from wild species into rice depends on the nature of the target trait(s), relatedness of the wild species and incompatibility barriers. Several protocols are available to overcome such barriers (Brar and Khush 2002). Some of the steps involved are described below:

Growing of wild species: The main source of wild species seed is the International Rice Gene Bank located at IRRI, Manila, the Philippines. Seeds can also be obtained from other institutes involved in rice research who have their own gene banks in rice-growing countries. Wild species are usually maintained and multiplied by vegetative (tiller) propagation. However, growing wild species from seeds needs careful handling during germination and raising of seedlings to maturity. Also, special care is needed to bag the panicle and to collect seeds as spikelet shattering is the most common problem. Vaughan (1994) has given details on: (i) how to obtain seeds, (ii) how to store seeds, (iii) where, when, and how to grow wild species (iv) breaking seed dormancy, (v) care during germination and raising of seedlings to maturity and selfing and harvesting of panicles/seeds.

Search for useful genetic variability for target traits: Wild species are known to be natural reservoir of useful genes for rice improvement; however, it is not uncommon that some accessions of a particular species may have limited genetic variability for the target agronomic trait of interest to breeders. Thus, it is essential to thoroughly screen and phenotype several

accessions of different species to ensure sufficient genetic variability. As an example, out of 6000 accessions of cultivated rice and wild species screened, none was found to be resistant to grassy stunt virus except one accession (IRGC 101508) of *O. nivara* collected from Uttar Pradesh (UP), India; in fact, only few plants of these accessions were resistant. Priority should be placed on identifying variability in the closely related AA genome species (*O. sativa* complex $2n = 24$, AA), followed by *O. officinalis* complex CC genome species, and later on distantly related species such as *O. brachyantha* (FF), *O. granulata* (GG), and *O. ridleyi* (HHJJ), which show limited homoeologous pairing with the AA genome of rice.

Production of hybrids and alien introgression lines (AILs): Interspecific hybrids are produced between elite breeding lines with the wild species accessions carrying useful genes for target traits of immediate interest to the breeder. Such hybrids are produced through direct crosses between rice and AA genome wild species. However, embryo rescue is required to produce interspecific hybrids and AILs between rice and all the other 17 wild species of *Oryza* except the AA genome wild species. Backcrossing with the recurrent rice parent is used to develop introgression lines. MAS is practiced to accelerate breeding through transfer of wild species genes/QTLs for traits linked with molecular markers.

Evaluation of introgression lines for transfer of target traits: Advanced introgression lines generated through backcrossing are evaluated for the transfer of target traits. This involves extensive laboratory, screen house, and field testing using various screening and inoculation protocols and testing in hotspot nurseries for major biotic and abiotic stresses. Evaluation for target trait(s) is carried out at multilocations and also across regions/countries in collaboration with NARS partners.

Molecular characterization of alien introgression: Molecular markers are used to characterize introgression from wild species during backcross breeding. The availability of dense molecular maps of rice comprising simple sequence repeat (SSR) and more recently SNP markers has facilitated large-scale analysis to determine the extent and process of alien introgression. SNPs are proving to be great value in construction of CSSL.

Mapping of introgressed alien genes/QTL: Monosomic alien addition lines (MAALs $2n = 25$) can be used to locate the wild species gene on a specific chromosome. Different types of mapping populations are generated through wide crosses such as BILs, recombinant inbred lines (RIL), doubled haploid (DH), and near-isogenic alien introgression lines, including other segregating population F_2, F_3, etc. Introgressed alien genes/QTL are mapped and tagged with molecular markers for use in MAS. Genome in situ hybridization (GISH) has become popular in characterization of parental genomes in interspecific progenies and for locating introgressed segments on rice chromosomes.

1.3.2 Alien Introgression Lines (AILs) and Chromosomal Segmental Substitution Lines (CSSLs)

Interspecific hybrids have been produced between rice (AA genome) and wild species of *Oryza,* representing 10 of the 11 genomes (AA, BB, CC, BBCC, CCDD, EE, FF, GG, HHJJ, KKLL) either through direct crosses or through embryo rescue (Brar et al. 1991; Brar and Khush 2002, 2006). A large number of AILs have been produced from crosses of rice with different wild species of *Oryza* by IRRI and several other NARS institutes (Brar and Singh 2011). As an example, at the Punjab Agricultural University (PAU), India, large sets of AILs have been produced from crosses of rice varieties (PR 114, Pusa 44) with six (AA genome) wild species: *O. glaberrima*, *O. nivara, O. rufipogon, O. longistaminata, O. glumaepatula,* and *O. barthii.* A large number of wide cross derivatives have been used by different institutes to transfer useful genes into rice (Table 1.2). Some examples include transfer of resistance to BPH, BB, blast, and tungro and tolerance to acid sulfate conditions, including introgression of genes for cytoplasmic male

sterility from different wild species into rice. Besides these traits, AILs have also been important genetic resources for the identification of QTL's for yield-related traits. One of the QTL, qSP2.2 from *O. longistaminata,* governing increased number of spikelets per panicle is being introgressed into Basmati rice which otherwise has panicle with a few spikelets.

Besides AILs, monosomic alien addition lines (MAALs $2n = 25$) representing 24 chromosomes of rice and a single extra chromosome of wild species have been produced. MAALs have been established from 7 wild species (CC, BBCC, CCDD, EE, FF, GG, HHJJ genome) (Jena and Khush 1990; Brar and Khush 2002). Similarly, CSSL having 24 chromosomes of rice but with one or two substituted segments of wild species chromosomes has also been produced. These substitution lines are being characterized using molecular markers, and a number of a genes/QTLs introgressed from wild species have been mapped. Doi et al. (2003) developed a set of CSSL with segments of *O. glaberrima, O. glumaepatula,* and *O. meridionalis* in the background of japonica cultivar (Taichung 65) and identified genes governing several traits. Surapaneni et al. (2017) developed CSSLs having segments of *O. nivara* substituted in the rice variety Swarna to identify major effect QTLs for yield-related traits.

Yoshimura et al. (2010) reported construction of CSSL from different AA genome species in the background of Taichung 65. Similarly, many institutes in China, Japan, India, USA, and France have developed CSSL and mapped and cloned genes/QTLs Ramos et al. (2016), produced CSSL of *O. longistaminata* in the background of Taichung 65 and reviewed development of CSSL of wild species by various other institutes.

1.3.3 Introgression from AA Genome Species of *O. sativa* complex

Crosses between rice and six other wild species of the *O. sativa* complex having the AA genome can be easily made and the genes have been transferred through conventional crossing and backcrossing procedures. Among the classical examples are the introgression of a gene for grassy stunt virus resistance from *O. nivara* to cultivated rice varieties (Khush 1977) and the transfer of a cytoplasmic male sterile (CMS) source from wild rice, *O. sativa* f. *spontanea*, (Lin and Yuan 1980). Other useful genes, such as *Xa21* for bacterial blight resistance, were transferred into rice from *O. longistaminata*, and new CMS sources from *O. perennis* and *O. glumaepatula*. More recently, genes for BPH, green leafhopper (GLH), green rice leafhopper (GRH) blast, tungro tolerance and tolerance to salinity, and acid sulfate soil conditions have been transferred from *O. sativa* complex species into indica rice cultivars (Table 1.2). Some breeding lines with genes transferred from wild species have been released for commercial cultivation in rice-growing countries (Table 1.3). Some examples on the specific traits transferred from wild species into rice are discussed below.

Introgression of gene(s) for resistance to grassy stunt virus: The grassy stunt virus is a serious disease transmitted by the vector brown planthopper (*Nilaparvata lugens*) BPH. Severe yield losses or even total loss may occur under epidemic conditions. Of the 6000 accessions of cultivated rice and several wild species screened, only one accession of *O. nivara* (accession 101508) was found to be resistant (Ling et al. 1970). A single dominant gene *Gs* confers resistance in *O. nivara*. Following four backcrosses with improved rice varieties, the gene for grassy stunt resistance was transferred into cultivated germplasm (Khush 1977). The first set of grassy stunt resistant varieties, IR28, IR29, and IR30, was released for cultivation in 1974. Subsequently, many such varieties, e.g., IR34, IR36, IR38, IR40, IR48, IR50, IR56, and IR58, have been released. These varieties have been widely grown and grassy stunt-infected plants are rarely seen in farmer's field now.

Introgression of gene(s) for resistance to tungro disease: Rice tungro disease (RTD) is another most serious viral disease in South and Southeast Asia. It is transmitted by the vector