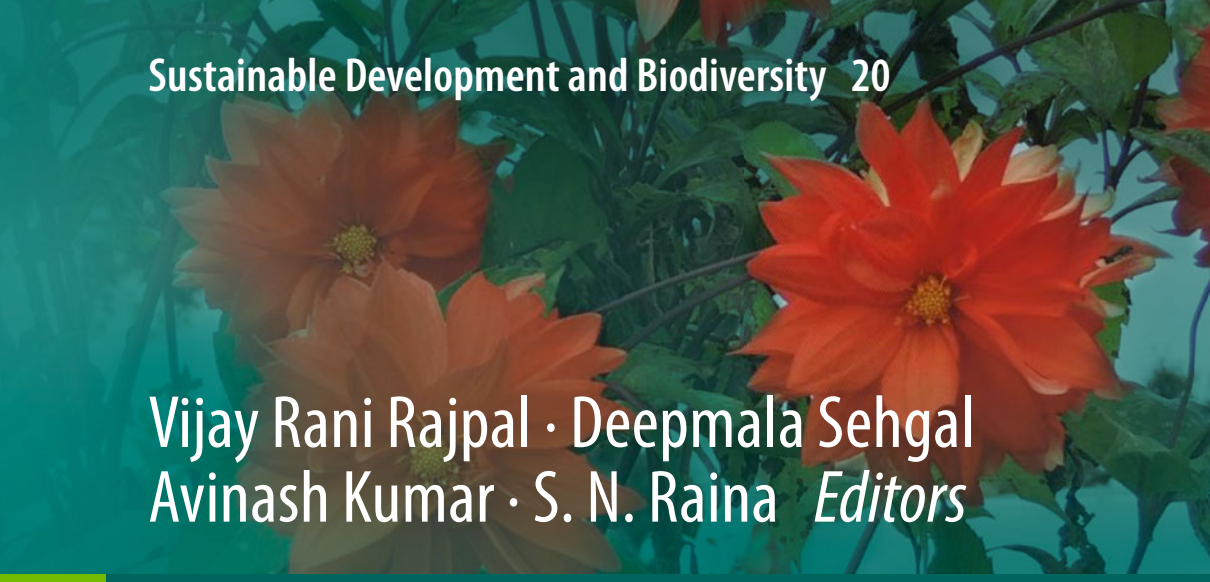
Sustainable Development and Biodiversity 20



Vijay Rani Rajpal · Deepmala Sehgal
Avinash Kumar · S. N. Raina *Editors*

Genetic Enhancement of Crops for Tolerance to Abiotic Stress: Mechanisms and Approaches, Vol. I

Springer

Sustainable Development and Biodiversity

Volume 20

Series editor

Kishan Gopal Ramawat
Botany Department, M.L. Sukhadia University, Udaipur, India

This book series provides complete, comprehensive and broad subject based reviews about existing biodiversity of different habitats and conservation strategies in the framework of different technologies, ecosystem diversity, and genetic diversity. The ways by which these resources are used with sustainable management and replenishment are also dealt with. The topics of interest include but are not restricted only to sustainable development of various ecosystems and conservation of hotspots, traditional methods and role of local people, threatened and endangered species, global climate change and effect on biodiversity, invasive species, impact of various activities on biodiversity, biodiversity conservation in sustaining livelihoods and reducing poverty, and technologies available and required. The books in this series will be useful to botanists, environmentalists, marine biologists, policy makers, conservationists, and NGOs working for environment protection.

More information about this series at http://www.springer.com/series/11920

Vijay Rani Rajpal · Deepmala Sehgal
Avinash Kumar · S. N. Raina
Editors

Genetic Enhancement of Crops for Tolerance to Abiotic Stress: Mechanisms and Approaches, Vol. I

Springer

Editors
Vijay Rani Rajpal
Department of Botany
Hansraj College
University of Delhi
Delhi, India

Deepmala Sehgal
CIMMYT Headquarters
El Batán, Veracruz, Mexico

Avinash Kumar
Vinoba Bhave University
Hazaribag, Jharkhand, India

S. N. Raina
Amity Institute of Biotechnology
Amity University
Noida, Uttar Pradesh, India

ISSN 2352-474X ISSN 2352-4758 (electronic)
Sustainable Development and Biodiversity
ISBN 978-3-319-91955-3 ISBN 978-3-319-91956-0 (eBook)
https://doi.org/10.1007/978-3-319-91956-0

Library of Congress Control Number: 2018951914

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

Preface: Volume I

The changing climate change scenarios have gripped humanity for a long time and are expected to worsen in the coming decades. Agriculture is already feeling the effects of climate change by reduced crop productivity, heavy yield losses, scarcity of water for farming, reduced rate of precipitation, and the list goes on. In staple crops, particularly wheat, rice, maize, soybean, barley, and sorghum, research has shown about 30% of the yearly variation in agricultural yields due to changes in rainfall and temperature.

Of all the threats that agriculture is exposed to due to climate change, abiotic stresses such as drought (water deficit), extreme temperatures (cold, frost, and heat), and/or salinity (sodicity) are the most devastating ones, causing more than 50% of crop yield losses. Mineral (metal and metalloid) toxicity is an additional abiotic factor, which is becoming a big threat for both major and minor crops. Thus, improving tolerance to these abiotic stresses is a global plant breeding target. A lot of research has been conducted to investigate plants' responses to these stresses at the structural, physiological, transcriptional, and molecular level and on the resistance mechanisms allowing them to adapt and survive these stressful events. A major research target has also been cross talk among various mechanisms, in case of multiple stresses faced by plants.

Precise analysis of proteome and metabolome is essential for understanding the fundamentals of stress physiology and biochemistry. Scientists have utilized 'omics' platforms to unravel the influence of abiotic stresses on levels of different protein groups and metabolite classes and to pinpoint candidate genes underneath. In addition, chromatin modifications, nucleosome positioning, and DNA methylation have been recognized as important components in plants' adaptations to stresses. The potential of improving stress tolerance in crops by enhancing the stress memory through the activation of priming responses or the targeted modification of the epigenome has been a burning research topic.

This book provides a consolidated and an updated account of the research being conducted in above-mentioned areas by plant scientists all over the world. It is an invaluable resource for researchers and educators in the areas of tools and technologies to unravel plant's responses to abiotic stresses. The outcomes presented on

staple crops will be useful to a broad community of scientists working in similar areas and can provide useful leads to build strategies to generate abiotic stress tolerant varieties. Students will find this book handy to clear their concepts and to get an update on the research conducted in various crops at one place.

New Delhi, India — Vijay Rani Rajpal
El Batán, Mexico — Deepmala Sehgal
Hazaribag, India — Avinash Kumar
Noida, India — S. N. Raina

The original version of the book was revised: Book editor's affiliation has been updated in Copyright page. The correction to the book is available at https://doi.org/10.1007/978-3-319-91956-0_12

Acknowledgements

We sincerely thank all authors not only for agreeing to contribute chapters despite their hectic schedules and other work commitments but also for putting in their sincere efforts in providing up-to-date information. We are thankful to Prof. K. G. Ramawat for inspiring us to take up this assignment. All the three editors, Vijay Rani Rajpal, Deepmala Sehgal, and Avinash Kumar are thankful to their Mentor and Guide Prof. S. N. Raina for guidance and inspiration always. Vijay Rani Rajpal is also thankful to the two other co-editors for their active involvement from inception to compilation and editing process through the course of this book. The book would not have been possible without their active involvement and whole-hearted support. The editors Deepmala Sehgal and Avinash Kumar are also thankful to senior editors Profs. S. N. Raina and Vijay Rani Rajpal who had been their teachers and Ph.D. supervisors too for providing them the great opportunity to be an editor.

Last but not least, editors gratefully acknowledge their families for their understanding, patience, and emotional support. Our sincere thanks to the whole Springer team who was tirelessly involved in the production process. We particularly appreciate Dr. Valeria and Dr. Ineke for their continued support.

We are very hopeful that this book will attract readers who are crop scientists and to even undergraduates and postgraduates of agricultural universities and institutes that are interested in the genetic improvement of crop plants using modern tools.

Contents

Contributors

Nasheeman Ashraf Plant Biotechnology Division, CSIR-Indian Institute of Integrative Medicine, Srinagar, Jammu and Kashmir, India

Shoib Ahmad Baba Plant Biotechnology Division, CSIR-Indian Institute of Integrative Medicine, Srinagar, Jammu and Kashmir, India

Renu Deswal Molecular Plant Physiology and Proteomics Laboratory, Department of Botany, University of Delhi, Delhi, India

Om Prakash Dhankher Stockbridge School of Agriculture, University of Massachusetts, Amherst, MA, USA

Manoj K. Dhar Genome Research Laboratory, School of Biotechnology, University of Jammu, Jammu, India

T. R. Ganapathi Nuclear Agriculture and Biotechnology Division, Bhabha Atomic Research Centre, Mumbai, India

Brijesh K. Gupta Plant Stress Biology, International Centre for Genetic Engineering and Biotechnology, New Delhi, India

Deepti Jain Department of Plant Molecular Biology, Interdisciplinary Centre for Plant Genomics, Delhi University, South Campus, New Delhi, India; Department of Botany, Bangalore University, Bangalore, India

Rohit Joshi Plant Stress Biology, International Centre for Genetic Engineering and Biotechnology, New Delhi, India

Prakriti Kashyap Molecular Plant Physiology and Proteomics Laboratory, Department of Botany, University of Delhi, Delhi, India

Sanjana Kaul Genome Research Laboratory, School of Biotechnology, University of Jammu, Jammu, India

J. P. Khurana Department of Plant Molecular Biology, Interdisciplinary Centre for Plant Genomics, Delhi University, New Delhi, India

Rakesh Kumar ICAR-Indian Institute of Wheat and Barley Research, Karnal, India

Smita Kumar CSIR-National Botanical Research Institute, Council of Scientific and Industrial Research (CSIR-NBRI), Lucknow, India

H. M. Mamrutha ICAR-Indian Institute of Wheat and Barley Research, Karnal, India

Manjari Mishra School of Life Sciences, Jawaharlal Nehru University, New Delhi, India

Shikha Mittal Division of Genetics, ICAR-Indian Agricultural Research Institute, New Delhi, India

Girish Chandra Pandey Department of Bioscience and Biotechnology, Banasthali University, Banasthali, India

Ashwani Pareek Stress Physiology and Molecular Biology Laboratory, School of Life Sciences, Jawaharlal Nehru University, New Delhi, India

Prashanti Patel Nuclear Agriculture and Biotechnology Division, Bhabha Atomic Research Centre, Mumbai, India

Suprasanna Penna Nuclear Agriculture and Biotechnology Division, Bhabha Atomic Research Centre, Mumbai, India

Asha Rani Department of Molecular Biology, Biotechnology and Bioinformatics, Chaudhary Charan Singh Haryana Agricultural University, Hisar, India

Suchismita Roy School of Life Sciences, Jawaharlal Nehru University, New Delhi, India

Davinder Sharma ICAR-Indian Institute of Wheat and Barley Research, Karnal, India

Indu Sharma ICAR-Indian Institute of Wheat and Barley Research, Karnal, India

Rahul Sharma Genome Research Laboratory, School of Biotechnology, University of Jammu, Jammu, India

M. N. Shiva Kameshwari Department of Botany, Bangalore University, Bangalore, India

Mohan B. Singh Plant Molecular Biology and Biotechnology Laboratory, University of Melbourne, Parkville, VIC, Australia

Nidhi Singh Division of Genetics, ICAR-Indian Agricultural Research Institute, New Delhi, India

Rajender Singh ICAR-Indian Institute of Wheat and Barley Research, Karnal, India

Richa Singh Department of Molecular Biology, Biotechnology and Bioinformatics, Chaudhary Charan Singh Haryana Agricultural University, Hisar, India

Sneh L. Singla-Pareek Plant Stress Biology, International Centre for Genetic Engineering and Biotechnology, New Delhi, India

Jyoti Taunk Department of Molecular Biology, Biotechnology and Bioinformatics, Chaudhary Charan Singh Haryana Agricultural University, Hisar, India

Nepolean Thirunavukkarasu Division of Genetics, ICAR-Indian Agricultural Research Institute, New Delhi, India

Ratan Tiwari ICAR-Indian Institute of Wheat and Barley Research, Karnal, India

Prabodh Kumar Trivedi CSIR-National Botanical Research Institute, Council of Scientific and Industrial Research (CSIR-NBRI), Lucknow, India

Karnam Venkatesh ICAR-Indian Institute of Wheat and Barley Research, Karnal, India

Parivartan Vishal Genome Research Laboratory, School of Biotechnology, University of Jammu, Jammu, India

Karuna Yadav Nuclear Agriculture and Biotechnology Division, Bhabha Atomic Research Centre, Mumbai, India

Neelam R. Yadav Department of Molecular Biology, Biotechnology and Bioinformatics, Chaudhary Charan Singh Haryana Agricultural University, Hisar, India

Ram C. Yadav Department of Molecular Biology, Biotechnology and Bioinformatics, Chaudhary Charan Singh Haryana Agricultural University, Hisar, India

Chapter 1
Functional Genomics Approach Towards Dissecting Out Abiotic Stress Tolerance Trait in Plants

Rohit Joshi, Brijesh K. Gupta, Ashwani Pareek, Mohan B. Singh and Sneh L. Singla-Pareek

Abstract Plant functional genomics has revolutionized not only the methodologies for identification and elucidation of key genes' function but also in designing strategies for improving tolerance towards abiotic stresses. Leveraging various approaches has demonstrated the robustness and versatility in their application to study gene/genome function and engineering abiotic stress tolerance in plants. With the emergence of novel high throughput technologies in this area, functional genomics can contribute immensely in understanding the gene regulatory networks operating under stress, thereby benefiting crop improvement programs. This chapter provides recent findings in the field of functional genomics, thus offering several efficacious methodologies such as next generation sequencing, genome-wide hybridization, gene-inactivation and genome-editing-based strategies in addition to metabolite analysis for discovery as well as validation of the candidate genes. Further, methodologies such as gene expression microarrays, insertional mutagenesis, map-based cloning and various genomic-assisted methods are evaluated critically and discussed in the light of integration of the information obtained through functional genomics with practical application in crop breeding.

Keywords Functional genomics · Mutants · Crops · Transcriptomics
Gene-inactivation · Genome-wide hybridization · Genome-editing

Rohit Joshi and Brijesh K. Gupta have equally contributed to this work.

R. Joshi · B. K. Gupta · S. L. Singla-Pareek (✉)
Plant Stress Biology, International Centre for Genetic Engineering and Biotechnology, Aruna Asaf Ali Marg, New Delhi 110067, India
e-mail: sneh@icgeb.res.in

A. Pareek
Stress Physiology and Molecular Biology Laboratory, School of Life Sciences, Jawaharlal Nehru University, New Delhi 110067, India

M. B. Singh
Plant Molecular Biology and Biotechnology Laboratory, University of Melbourne, Parkville, VIC 3010, Australia

V. R. Rajpal et al. (eds.), *Genetic Enhancement of Crops for Tolerance to Abiotic Stress: Mechanisms and Approaches, Vol. I*, Sustainable Development and Biodiversity 20, https://doi.org/10.1007/978-3-319-91956-0_1

1.1 Introduction

Abiotic stresses such as cold, heat, waterlogging, drought, metal toxicity, salinity and sodicity reduce plants' growth and yield by as much as 50% in both natural and agricultural systems (Nakabayashi and Saito 2015). Improving tolerance to abiotic stresses, therefore, has become a major objective in plant breeding programs globally (Pareek et al. 2010). It has been estimated that a global increase in food production of 44 metric tons will be required each year to fulfill the food demand of rapidly increasing population, which will reach close to 10 billion by 2050 (Bohra et al. 2015; Wang et al. 2016). Plant's responses towards simultaneous occurrence of abiotic stresses, such as drought, heat and salinity, have gathered attention in various genomics studies (Singh et al. 2015a; Joshi et al. 2015b; Kushwaha et al. 2016). However, multigenic nature of abiotic-stress-tolerance trait(s) along with the lack of proficient selection techniques primarily hampers effective breeding strategies for abiotic stress tolerance (Ford et al. 2015). Furthermore, several reports have indicated differences between quantitative trait loci (QTL), some being linked to tolerance at one stage of plant's development while other linked to tolerance during some other stages (Yang et al. 2013). Dissection of the genetic basis of intra-specific variation in traits conferring abiotic stress tolerance will be useful for selecting and creating positive variations within the species. However, limited success has been achieved through traditional approaches such as inter- or intra-generic hybridizations, induced mutations and/or somaclonal variations (Chinnusamy et al. 2004; Bhullar and Gruissem 2013).

Recent advances in genomics and molecular biology have contributed significantly to the breeding programs by rapid identification and characterization of genes and genomic regions conferring abiotic stress tolerance. In this direction, one of the powerful approaches for gene discovery could be the exploration of naturally occurring genetic diversity between landraces and their wild relatives (Dwivedi et al. 2016). Thus, understanding the molecular basis of genetic diversity may help in identifying the key differences, which regulate the differential expression of same set of genes in contrasting genotypes. This may aid in unraveling the novel mechanisms underpinning abiotic stress tolerance in crops (Mickelbart et al. 2015).

Recent genomics-based approaches combined with high throughput tools have led to a revolution in crop improvement approaches. These advanced technologies directly affect the applicability of crop improvement methods by translating the entire genomic regions deciphering molecular responses of plants (Bohra and Singh 2015; Edwards 2016; Gupta et al. 2015). Forward and reverse genetics approaches together elucidate the genes and their products involved in expression, signal transduction and stress tolerance (Urano et al. 2010). Since 1980s, functional genomics leapt from being hypothetical or innovative concept to a widely accepted part of science in the year 2000. In the post genomic era, extensive utilization of functional genomics tools has increased our knowledge of the complex networks

operating during stress tolerance and adaptation. These functional genomics strategies combined with phenomics will improve our understanding towards gene complementation, transcript regulation, protein complex formation and their evolutionary pathways regulating abiotic stress tolerance traits. After the initiation of whole genome sequencing programs in 1990s, astonishing advancements in DNA sequencing technologies have brought breakthroughs in this area (Wheeler and Wang 2013). Already completed genome sequences of various model organisms including protists (Armbrust et al. 2004), fungi (Wood et al. 2002; Galagan et al. 2003) and eukaryotic plants (Li et al. 2014a, b; Hirakawa et al. 2014; Varshney et al. 2017) have confirmed the feasibility and efficacy of sequencing large genomes. Further, functional genomics provides the next step towards the biological revolution assigning the function to previously identified genes at organizational level that can control the genetic pathways defining the physiology of an organism (Rahman et al. 2016).

Several interrelated strategies enable the survival of tolerant genotypes under abiotic stresses. However, these strategies are less evolved in agricultural crop species, perhaps due to crop domestication. Abiotic stress tolerance in these plants can be achieved at the molecular level by engineering genes regulating chaperone production, osmoprotectant accumulation, reactive oxygen species (ROS) scavenging mechanisms and/or efficient transporter systems for exclusion or compartmentation of ions (Jan et al. 2013; Gupta and Huang 2014). In addition, several genes and their products act simultaneously at transcriptional and translational levels (Joshi et al. 2015b; Gupta et al. 2015; Guo et al. 2016). Functional validation of these genes can help in untangling the stress tolerance network and also in designing various functional markers for marker-assisted breeding.

Genetic transformation approaches offer a rapid way to improve plant stress tolerance. With the advent of high throughput techniques, functional genomics strategies went through a paradigm shift from single gene discovery to many thousands. Development of expressed sequence tags (ESTs) from cDNA libraries of abiotic stress-treated seedlings of plants as well as their complete genome sequence information provides an additional resource for gene discovery. In addition, strategies including promoter trapping, mutagenesis and gene complementation have led to the identification of key gene pools and hence, have provided valuable inputs towards the functional characterization of stress responsive genes and their underlying mechanisms (Hasanuzzaman et al. 2015). In this chapter we discussed current strategies in the field of functional genomics for improving abiotic stress tolerance in plants. Further, we discuss the role of model species and mutant populations in molecular mapping of abiotic stress tolerance determinants for crop improvement.

1.2 Stress Networks and Signaling Pathways Operative Under Abiotic Stresses in Plants

Stress perception as well as its signaling are the two critical components determining the adaptive response of the plant under unfavorable environmental conditions (Muthurajan and Balasubramanian 2009; Gupta et al. 2015). Osmotic and oxidative stresses induced in plants are a common consequence of abiotic stresses sharing many intermediate components of their signaling cascades (Rejeb et al. 2014). Thus, signaling sensors are now becoming main targets for genetic engineering, as they are the principle transducing elements right from the perception of the signal. One of the important stress sensors in higher plants is the Two Component System (TCS) which consists of histidine kinase (HK) sensor and response regulator (RR) (Pareek et al. 2006; Singh et al. 2015b). The investigations on different plant species such as maize and rice confirmed the role of TCS members in response to abiotic stresses (Liu et al. 2014a; Sharan et al. 2017). During abiotic stress, few of the members of TCS family show up-regulation i.e., *AHK1*, *OsHK3*, *GmHK7*, *GmHP3*, *GmHP6*, *GmRR1*, while others are down regulated i.e., *AHK2*, *AHK3*, *AHK4*, *AHP1*, *AHP3*, *AHP5*, *ARR8*, *ARR9*, *OsHK4*, *GmHK10*, *GmHK12 and GmPHP2* (Le et al. 2011; Nishiyama et al. 2013; Gahlaut et al. 2014).

Through yeast-two-hybrid assay, it was revealed that under cold and salt stress, Mitogen Activated Protein Kinase (MAPK) pathway involves MAPK/ERK kinase kinase-1 (MEKK1) which acts upstream to MAP kinase kinase-1 (MKK1), MAP kinase kinase-2 (MKK2), MAP kinase-4 (MPK4) and Mitogen-activated protein kinase-6 (MPK6) (Sinha et al. 2011). In *Arabidopsis*, the signals received by 80 MAPKKKs are transduced downstream from 10 MAPKKs to 20 MAPKs providing an opportunity for crosstalk at different points (Sinha et al. 2011). Similarly, under drought stress, it was reported that AtMEKK1 and AtMPK3 in *Arabidopsis* and OsMSRMK2 and OsMAPK5 in rice show higher expression (Sinha et al. 2011; Ara and Sinha 2014). Pitzschke et al. (2014) revealed MYB44 transcription factor as the interacting partner of MKK4, which in turn interacts with another MPK3-regulated transcription factor VIP1. These results further confirm that MAPK cascade is playing a central point of crosstalk during stress signaling (Pitzschke 2015; Wen et al. 2015). In addition, another important component during osmotic stress signaling pathway is the calcium-dependent protein kinase (CDPK). Overexpression of rice *OsCDPK7* gene was found to confer tolerance against salt, drought and chilling stress (Boudsocq and Sheen 2013).

Various stresses can occur individually, or in combination with others, at any developmental stage of plant and these vary by location and time, which can negatively affect photosynthetic efficiency and alter the source-sink relationship. Further, it can affect the remobilization of solutes, which is a limiting factor for grain weight and yield. Combinations of various traits contribute towards overall plant tolerance against abiotic stresses (Roy et al. 2011). However, it is still unknown how certain plants maintain yield under abiotic stress conditions

(Tripathi et al. 2012). Identifying key regulatory elements playing roles during multiple stress interactions through gene expression profiling is an important aspect of functional genomics. A number of transcription factors (TFs) differentially regulated during environmental stresses have already been analyzed using genome-wide transcriptome analysis (Hoang et al. 2014; Joshi et al. 2016a). These TFs show a very complex expression pattern, which suggests that stress resistance and tolerance are regulated by an extremely intricate gene regulatory network at transcriptional level. Amongst all, bZIP (Basic Leucine Zipper), MBF1 (Multiprotein bridging factor 1), WRKY, MYB (myeloblastosis) and NAC (NAM, ATAF1,2 and CUC2) transcription factors are the largest transcriptional regulators controlling growth, development, physiological processes, and abiotic stress responses in plants (Sahoo et al. 2013; Baloglu et al. 2014).

Rasmussen et al. (2013) employed microarray analysis to detect plant responses to multiple stress exposures, in combination, or individually and found that 25% of transcripts showed similar responses during individual stresses, but act differentially under stress combinations. Twenty-three transcripts were found to be specifically upregulated in the transcriptome analysis of *Arabidopsis* plants using triple combination of heat stress, drought and virus infection (Prasch and Sonnewald 2013). Of these, DREB2A (Dehydration-responsive element-binding protein 2A) and GBF3 (G-box-binding factor 3) were upregulated, whereas Rap2-9 (Related to APETALA2-9) was strongly down regulated. Transcript profiling of *Arabidopsis* plants revealed 43 drought, cold and salinity stress-inducible transcription factor genes including DREB, ERF (Ethylene Responsive Factor), zinc finger containing factors, MYBs, bHLHs (basic helix-loop-helix), bZIPs, NAC and WRKY (Umezawa et al. 2006). Similarly, transcript expression of whole WRKY family of rice showed 17 WRKY genes to be highly induced in both leaf and root under drought stress (Tripathi et al. 2014). Yang et al. (2011) showed that ABI5-Like1 (ABL1) gene regulates ABA and auxin responses by altering ABRE-containing WRKY genes' response in rice. In addition, it was reported recently that among the 9 members of AREB/ABFs in *Arabidopsis*, AREB1/ABF2, AREB2/ABF4 and ABF3 (Abscisic acid responsive elements-binding factor 3) are highly upregulated by osmotic stress and ABA treatments in vegetative tissues (Yoshida et al. 2014). Similarly, through 24K Affymetrix Genechip array, a total of 514 CBF2 (Centromere-binding factor 2) genes were identified under cold stress in *Arabidopsis*, including co-regulated genes like zinc finger proteins (CZF1 and CZF2), MYB73, RAV1 (related to ABI3/VP1 1), ZAT10 and ZAT12 (Vogel et al. 2005; Park et al. 2015). A genome-wide analysis of paper mulberry in response to cold stress showed that 794 TFs, belonging to 47 families were involved in the cold stress response (Peng et al. 2015). Among the differentially expressed TFs, one bHLH, two ERFs and three CAMTAs were involved in signal transduction at early stages followed by 5 bHLH, 14 ERFs, one HSF, 4 MYBs, 3 NACs, and 11 WRKYs in providing cold resistance. The late responsive group consisted of 3 ARR-B, C3H, 6 CO-like, 2 G2-like, 2 HSFs, 2 NACs and TCP. These results indicated towards a much greater cross-talk among different stresses during signaling

processes. The key regulators among this complex network are bHLH, bZIP, MYB and AP2 transcription factor families (Peng et al. 2015).

1.3 Functional Genomics Approaches

In the present scenario, direct introduction of genes through genetic engineering is coming up as a more rapid and reliable technique for improving stress tolerance in plants, in comparison to traditional breeding and marker-assisted selection approaches (Bohra et al. 2015). Current engineering strategies aim to functionally characterize the critical genes participating in either signaling or biochemical pathway to understand their distinctive roles in plant development and physiology (Teotia et al. 2016). The products of these genes are either stress-induced proteins or enzymes for osmoprotectant or scavengers of ROS that directly or indirectly provide tolerance against different environmental stresses (Joshi and Chinnusamy 2014; Khan et al. 2015). In addition, various transcription factor genes controlling the expression of different stress regulatory proteins are also unveiled (Wang et al. 2016). It is now necessary to study the abiotic stress tolerance in a collective manner on a genome-wide scale, which can be further utilized for elucidation of abiotic stress networks. With the availability of various omics tools including genomics, transcriptomics, and proteomics, major progress has been made for understanding the interaction and complexity of the stress adaptive mechanisms and their respective signaling pathways (Liu et al. 2014b). By using transcript profiling and allocation of small responsive elements in promoter regions, the determination of regulatory regions in chromatin structure, and the distribution of *cis*-regulatory elements and transcription factors can be predicted computationally.

One of the major challenges in the post-genomic era is to understand the function of genes. Recent high throughput biotechnological advances have facilitated the discovery of new genes and their functions. Unraveling gene functions and their interactions with other regulatory networks have long been exploited for generation of improved varieties (Akpınar et al. 2013). While the functions of several genes are still unknown, their function can often be correlated in association with other known genes, which provide even better understanding for the whole signaling network. We are now able to obtain a complete overview at the cellular level through transcript, protein and metabolite profiling. These approaches allow a deeper understanding of the complex cellular functioning during different physiological processes (Cramer et al. 2011).

Reconstruction of complex networks at whole genome level is achieved by characterizing and quantifying from genotype to phenotype (Feist and Palsson 2008). Understanding only the basic function of the gene in an organism does not provide an insight to its specific role under stress conditions. Sequence analysis of *Arabidopsis* showed that 13 and 20% of the genes are implicated either in signal transduction or in stress/defense responses, respectively (Mahalingam et al. 2003). Another exhaustive screening of more than 1,500 TFs revealed that almost 40 TFs

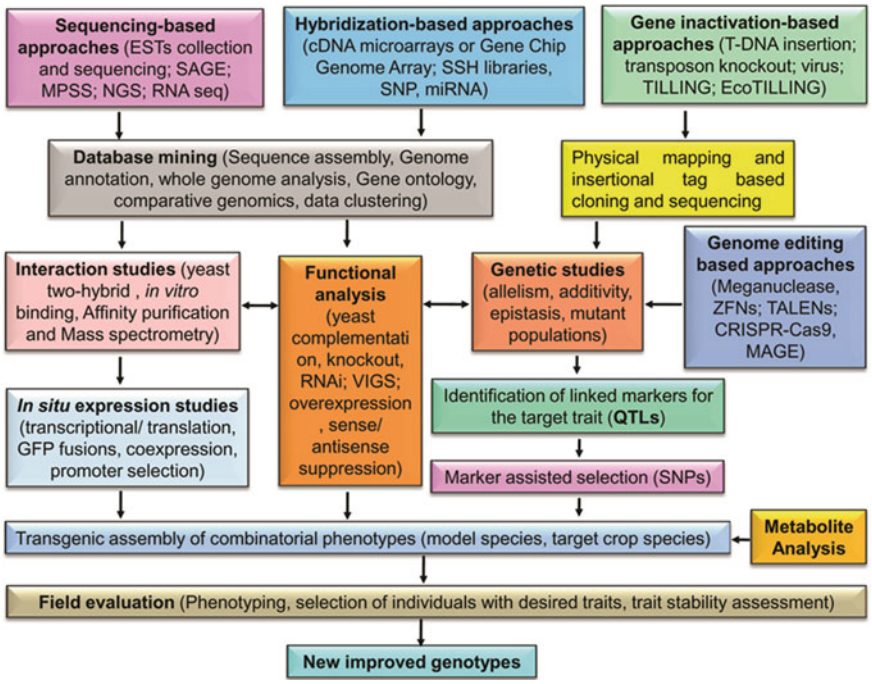


Fig. 1.1 Flow chart of the overview of functional genomics approach for plant improvement

were involved in improving stress tolerance in *Arabidopsis* (Nelson et al. 2007). The major outcome of the current plant genome research is the functional characterization of almost 54% of higher plant genes by comparing them with other known sequences (Sofi and Trag 2006). Parallel studies on the functional genomics in other organisms will also contribute significantly to understand their gene functions in coming years. Functional genomics strategies mainly utilize methodologies, which are sequence-based, hybridization-based, gene inactivation-based or genome-editing based (Fig. 1.1) as discussed below.

1.3.1 Sequencing-Based Approaches

One of the major approaches used to discover abiotic stress expressed gene catalogue is based on ESTs generated from various cDNA libraries expressing transcripts from several stresses in different tissues and developmental stages (Rahman et al. 2016). These libraries have been successfully developed to identify several specific and stress-responsive transcripts, but they under-represent rare transcripts or unexpressed transcripts under certain conditions. EST libraries are major focus of functional studies because they provide easier strategies for gene discovery and

genome annotation (Varshney et al. 2006). However, to get more information on polymorphism, EST sequences must be cautiously overlapped onto similar contigs to gather detailed information on the configuration of parental cDNA as in polyploid species like wheat (Rudd 2003). Despite all these factors, EST sequencing is a convincing strategy with already reported potential in gene discovery by aligning with collinear genotypes exposed to control and stress conditions (Ergen and Budak 2009).

Several attempts have been made using model plant species to characterize stress-specific transcripts using EST sequencing in higher glycophytes under salinity stress. National Center for Biotechnology Information (NCBI) dbEST database indexes rapidly growing libraries containing several ESTs generated from various crops and other plant species. However, large-scale cDNA sequencing programs from stress-treated plants of various species at different growth stages are still essential to enrich plant EST datasets. Additionally, the information on gene number as well as number of gene families playing significant roles in abiotic stress responses can be established by clustering the sequences of ESTs obtained through respective stress-treated cDNA libraries (Li et al. 2014a, b). Similar gene-indexing database Swissprot provides important information associated with stress-responsive genes among different plants and is frequently used to assign putative functions to stress-responsive genes (Sreenivasulu et al. 2007). In addition, data clustering produces consensus contigs, which is a more reliable approach than ESTs. Extensive attempts have been made in glycophytes, such as *Arabidopsis*, rice and halophytes to compare the abundance of expressed ESTs in their respective cDNA libraries (Wang et al. 2004; Baisakh et al. 2008; Li et al. 2014a, b). Extensive EST sequencing is still in progress for developmental stage-specific, tissue-specific and stress-specific cDNA libraries obtained from *Arabidopsis* and rice. Analyzing these EST databases will pave our way to specify stress regulated genes that can assist further to unravel the underlying regulatory metabolic pathways (Rahman et al. 2016).

Another approach which enables simultaneous quantitation of thousands of transcripts is SAGE (Serial Analysis of Gene Expression), in which mRNA is oligo (dT)-trapped and reverse transcribed to form cDNA, then small sequence tags are extracted and ligated to form long concatemeric chain and sequenced, leading to complete quantification of gene expression (Vega-Sánchez et al. 2007). Due to the recent advancements in next-generation DNA sequencing technologies, SAGE analysis has emerged as a high throughput, sensitive and cost-effective approach in comparison to Sanger sequencing approaches (Cheng et al. 2013). During the past several years, SAGE has been extensively used in plants with the availability of extensive EST databases of different species (Breyne and Zabeau 2001). Additionally, by combining 5′ RACE (Rapid amplification of cDNA ends) and SAGE (Serial analysis of gene expression) analysis, transcription start sites were also identified (Wei et al. 2004). Later on, several modifications such as SuperSAGE and DeepSAGE became available, in which the tag size is expanded providing greater efficiency to the annotation (Nielsen et al. 2006; Matsumura et al. 2012). Previous studies using SAGE in plants not only revealed new expressed

regions in the plant genome but also implied their novel functions including stress response in crops (Cheng et al. 2013).

Massively Parallel Signature Sequencing (MPSS) is also a powerful method enabling the parallel analysis of millions of transcripts on a genome-wide scale (Akpınar et al. 2013). In MPSS, transcription profiling is done using similar tag-based approach, where tagged PCR products obtained from cDNA are amplified so that each mRNA molecule produce ~100,000 of PCR products with a unique tag that are ligated to microbeads and sequenced (Kudapa et al. 2013). After several rounds of ligation-based sequencing, a 16–20 bp sequence signature is identified from each bead resulting into ~1 million sequence signatures. Because of high throughput analysis and longer tags, MPSS can detect novel transcripts particularly in species lacking whole genome sequence, in addition to identifying genes efficiently (Hamilton and Buell 2012). MPSS has also been utilized in small RNA expression studies (Nobuta et al. 2007) along with mRNA transcription studies in plants which are much correlated with abiotic stress responses (Sunkar et al. 2007). Publicly available plant MPSS database (http://mpss.udel.edu/) contains expression data for several genotypes, including economically important crops such as soybean, maize and rice (Nakano et al. 2006). In addition, NGS platforms have expanded genome-wide sequence expression analysis, in which sequencing of RNA populations and quantification of transcripts can be achieved through RNA-seq (Sánchez-León et al. 2012). The efficiency of Illumina-based digital gene expression system for high-throughput transcriptome sequencing has been demonstrated in crops under abiotic stress conditions in different tissues (Tao et al. 2012; Pandey et al. 2014).

1.3.2 Hybridization-Based Approaches

In response to abiotic stress, plants respond and adapt by altering physiological and biochemical processes resulting in altering responses of thousands of genes. Transcriptome analysis using gene chips and microarray technology provides an important experimental opportunity to unravel key biological processes and to provide information about unknown functional genes conferring abiotic stress tolerance (Gul et al. 2016). In principle, DNA sequences of complete genes of an organism are placed on microchips and used as substrates for hybridization for quantifying expression of different genes in a sample (Joshi et al. 2012). This gives the complete quantitative information about the relative expression of genes corresponding to their response towards various abiotic stresses along with the fold change in different developmental processes like germination, vegetative and flowering stages (Wu et al. 2015). In contrast to sequence-based approaches, array-based technique is a targeted approach where sequence is required to design probes (Rahman et al. 2016). Extensive microarray expression data already exists in public domain (www.genevestigator.com/gv/plant.jsp) with complete genome sequences of several model species including *Arabidopsis* and rice (Hruz et al.

2008; Urano et al. 2010). These gene expression databases provide deeper insight of the complex gene regulatory pathways under various stress responses. Furthermore, genes encoding several regulatory and functional proteins are now known, and the complex mechanisms of multi-gene regulation under abiotic stress response are partly deciphered. Several technical limitations including cross-hybridization and background noise etc. affect microarray analysis investigating stress responsive genes. Through oligo microarray, several model plants and economic crops have been analyzed, including *Arabidopsis* (Richards et al. 2012), rice (Jung et al. 2013), wheat (Quijano et al. 2015), corn (Allardyce et al. 2013), soybean (Le et al. 2012) and tomato (Martínez-Andújar et al. 2012).

Another strategy for RNA hybridization and comparative gene expression in tissues/genotypes is the GeneChip Genome Array. Several studies in model crops have employed these GeneChip Genome Arrays to detect expression of several genes at the same time in the whole genome (Verdier et al. 2013; Wu et al. 2015). In contrast to microarrays, gene chips are created by synthesizing several hundred thousand oligonucleotides on a miniature support using photolithography (Joshi et al. 2012). Further, by using this technique, it is feasible to visualize gene chips that represent an entire plant genome. For example, in soybean, gene chip array characterized genome-wide expression pattern, and identified drought-responsive candidate genes (Saxena et al. 2011). During recent years, a large amount of genome data has been obtained in rice using various chips with different specifications, including BGI/Yale 60K chip (Ma et al. 2005), Agilent 44K chip (Ghaffar et al. 2016), NSF45K chip (Jung et al. 2008), Affymetrix 57K chip (Russell et al. 2012) and NimbleGen (Fenart et al. 2013).

1.3.3 Gene Inactivation Based Approaches

Though the reports pertaining to genome-wide expression analysis in diverse plants are increasing on an exponential rate, only a few studies have focused on overexpression or suppression of these differentially expressed genes for their functional characterization. Currently, two main approaches are being utilized to knockout the desired genes, namely T-DNA insertion mutation and TILLING (Targeted Induced Local Lesions In Genomes). TILLING enables high-throughput genome-wide analysis of point mutations in target genomes to generate novel mutant alleles for crop improvement (Lee et al. 2014). It is applicable to the genomes of almost all species of plants including diploids and allohexaploids (Chen et al. 2014). The TILLING populations can be traditionally screened for phenotypic or genotypic variations under abiotic stresses (de Lorenzo et al. 2009).

Another modified method, called EcoTILLING, is also high-throughput, time-saving and cost-effective technique, developed to identify SNPs and small indels (Bajaj et al. 2016). EcoTILLING is applicable in polyploid species for differentiating among alleles of paralogous and homologous genes (Akpınar et al. 2013). It not only provides information on allelic variants for various genes but also

helps in unravelling the complexity of abiotic stress tolerance pathways. Recently, it has been used to detect SNPs involved in salt stress response in domestic rice genotypes (Negrão et al. 2011). Naredo et al. (2009) detected several SNPs in both lowland and upland rice cultivars involved in drought stress tolerance. Similarly, 46 INDELs (insertions/deletions) and 185 SNPs (single nucleotide polymorphisms) were identified using EcoTILLING while conducting allele mining for drought related genes in 96 barley genotypes (Cseri et al. 2011). Similarly, using EcoTILLING approach 1133 novel SNP allelic variants were discovered from diverse coding and regulatory sequence components of 1133 transcription factor genes by genotyping 192 diverse desi and kabuli chickpea genotypes (Bajaj et al. 2016).

T-DNA insertional mutagenesis can be utilized as a tool to study functional genomics in *Arabidopsis* and other higher plants (Jung and An 2013). Agrobacterium-mediated T-DNA transformation can also provide an efficient opportunity to target candidate genes into plant cells. Random insertion of T-DNA fragments in either exon or intron results in the target gene inactivation. During *Arabidopsis* functional genomics initiative, huge number of sequence-indexed T-DNA insertion lines was obtained, which are available currently in the public domain libraries (https://www.arabidopsis.org/portals/mutants/stockcenters.jsp) of *Arabidopsis* (Alonso et al. 2003). Similarly, during International Rice Genome Sequencing project 172,500 flanking sequence tags (FSTs) were submitted in Rice Functional Genomic Express database (RiceGE, http://signal.salk.edu/cgi-bin/RiceGE), which are also available from Rice *Tos17* Insertional Mutant Database (https://tos.nias.affrc.go.jp/). These T-DNA insertion mutants are a rich source for elucidating metabolic/signaling pathways and for functional analysis of genes in plants (Gao and Zhao 2012). In addition, gene inactivation can also be done by using RNAi technology. Using knockdown approach it was confirmed that SOS2 (a serine/threonine type protein kinase) and SOS3 (a calcium binding protein) loci are present in *Arabidopsis*, rice, wheat and *Brassica* (Kumar et al. 2009; Yang et al. 2009; Kushwaha et al. 2011; Feki et al. 2014). Now it is well documented that SOS3 interacts with SOS2 after receiving cytoplasmic calcium signals produced under high Na^+ concentrations. The SOS3-SOS2 complex further activates SOS1, a Na^+/H^+ antiporter gene to maintain homeostasis (Sharma et al. 2015).

1.3.4 Genome Editing Based Approaches

Currently available tools for genome editing provide intriguing possibilities for introducing targeted mutation, INDEL and sequence modifications to a predetermined location within the genome to functionally characterize plant genes and for improvement of abiotic stress tolerance in plants (Strange and Petolino 2012). Due to low homologous recombination frequency in plants, successful gene targeting is very difficult and inefficient (Xie and Yang 2013). Most commonly used genome editing tools are Zinc finger nucleases (ZFNs), transcriptional activator-like effector

nucleases (TALENs) and clustered regularly interspaced short palindromic repeat (CRISPR)-Cas9 (CRISPR-associated nuclease9) (Kumar and Jain 2015). TALENs have emerged as an alternative to ZFNs for genome editing and for introducing targeted double-strand breaks. TALENs have showed a very high success rate, but their large size may limit their delivery by recombinant adeno-associated viruses (AAV) (Gaj et al. 2013).

ZFNs are designed nucleases that induce targeted double strand breaks at specific genomic loci, thereby, allowing successful targeted mutagenesis and transgene integration in plants (Petolino et al. 2010). They are fusions of the nonspecific cleavage domain from the *Fok*I restriction endonuclease with custom-designed Cys_2-His_2 zinc-finger proteins. These chimeric nucleases produce sequence-specific DNA double-strand breaks that are repaired by error-prone non-homologous end joining to induce small alterations at targeted genomic loci (Gaj et al. 2012). They can be designed to cleave any DNA sequence and thus offer a wide range of sequences to be deleted. Using ZFNs, majority of targeted genome modifications have been performed including point mutations, deletions, insertions, inversions, duplications and translocations in several organisms and cell types (Joung and Sander 2013). The latest ground-breaking technology for genome editing is the type II clustered regularly interspaced short palindromic repeat (CRISPR)/Cas9 system from *Streptococcus pyogenes* (Bortesi and Fischer 2015). The CRISPR/Cas9 system is composed of Cas9 nuclease and customizable sgRNA which guides Cas9 to recognize target DNA and creates double strand breaks to initiate non-homologous end joining and homologous recombination repair pathways, resulting in genome modifications (Zhang et al. 2016). Since its discovery, CRISPR-Cas9 system has shown robustness and versatility in applications for genome editing in various biological contexts and has opened a new door to plant functional genomics research. This technology can be utilized for analysis of loss-of-function, gain-of-function and gene expression, along with modifications in spatio-temporal gene expression. It can also contribute in understanding gene function, gene regulatory networks and engineering abiotic stress tolerance in a variety of plants (Liu et al. 2016; Khatodia et al. 2016).

1.3.5 Metabolite Analysis

Metabolomics has now emerged as a relatively new area of functional genomics that contributes to our understanding of the complex molecular interactions in biological systems (Bino et al. 2004). Several reviews published earlier have described the role of metabolomics in functional genomics research (Hall et al. 2002; Sumner et al. 2003; Schauer and Fernie 2006; Saito and Matsuda 2010). Several reports are available on its applicability for abiotic stress tolerance in plants (Jorge et al. 2015; Nakabayashi and Saito 2015; Okazaki and Saito 2016; Sun et al. 2016). Integrated metabolomics and transcriptomics studies in model plants have significantly increased our knowledge on signal transduction pathways in different

crops under stress. Recently, a report on metabolite profiling in two contrasting rice genotypes i.e., FL478 (salt-tolerant) and IR64 (salt-sensitive) found 92 primary metabolites in the leaves and roots under control and salt stress conditions (Zhao et al. 2014). In general, 6 metabolites (phenylalanine, threonine, citric acid, raffinose, melicitose and galactinol) were induced in the leaves or roots, while 11 metabolites i.e., lysine, threonine, isoleucine, proline, valine, isocitric, sucrose, lactose, sorbitol, mannitol and galactopyranoside were increased specifically in leaves or roots under stress conditions. These compounds regulating sugar and amino acid metabolism pathways will increase our understanding of the physiological mechanisms underpinning salt tolerance. Similarly, comparative proteomic analysis in the shoots of IR64 and its mutant lines resulted in identification of 34 unique proteins expressed during salt stress exposure (Ghaffari et al. 2014). Similarly, Liu et al. (2014b) detected 83 proteins in roots and 61 proteins in leaves to be differentially expressed and reported of having their significant contributions against salinity stress in rice. Protein alterations upon external stimuli are vital, and thus proteomic analysis provides deep knowledge on key aspects of plant metabolic and regulatory pathways against abiotic stress (Kim et al. 2014). These differentially expressed proteins can act as an abiotic stress tolerance marker for plants (Zhang 2014). Our understanding of metabolite adaptation to abiotic stress in plants is still incomplete. Thus, it is necessary to deepen our knowledge further with targeted comprehensive metabolomics studies with more emphasis on primary and secondary metabolic pathways.

1.4 Role of Model Species and Mutant Populations

Although functional adaptation mechanisms are highly conserved among stress susceptible genotypes, the tolerant genotypes, however, evolved additional regulatory mechanisms that enhance their ability to cope with severe abiotic stresses (Joshi et al. 2016b). Whole genome sequencing of rice and *Arabidopsis* has increased our understanding of the genes playing a crucial role in providing multiple abiotic stress tolerance (Mustafiz et al. 2011; Kumar et al. 2012; Singh et al. 2012; Tripathy et al. 2012; Kaur et al. 2014). For example, in *Arabidopsis* early stages of heat stress triggers decay of 25% of the transcriptome and is catalyzed by the 5′-3′ exonuclease XRN4. cDNA libraries prepared from 21 days old heat stressed seedlings shows 19,804 distinct loci accounting for 76% of the total *Arabidopsis* genes. Out of these, only 801 (4%) were found to be upregulated, which represents proteins involved in heat and abiotic stress response, and 4,745 (25%) were found to be down-regulated (Merret et al. 2013). Similarly, RNA-Seq and digital gene expression (DGE) analysis in *Bryum argenteum*, a desiccation-tolerant moss found in largest cold desert (Gurbantunggut desert) of China, showed 4,081 and 6,709 differentially expressed genes after 2h and 24h rehydration, respectively. Further, upon rehydration, 142 TF transcripts were found to be up-regulated, including 23 members of ERF family (Gao et al. 2015).

By using modern genomics and genetic approaches, full-length cDNA populations and BAC sequences have been transferred from stress-tolerant genotypes to stress-sensitive ones to generate stress tolerant varieties with better growth and yield (Mir et al. 2012; Akpınar et al. 2013). As wheat, barley and rye are close relatives; their syntenic relationship can be utilized for positional cloning of important stress tolerant genes (Joshi et al. 2015a; Kole et al. 2015).

Mutant phenotype selection through mutational breeding is an old technique, which has successfully contributed in generating several important varieties of cereals. Using single base mismatches, several barley and wheat mutant populations have been developed for mutation studies and several projects are running throughout the globe for developing mutant populations of their diploid progenitors (Sikora et al. 2011; Dhakarey et al. 2016). Several sets of insertion mutants are already accessible for petunia, maize, snapdragon, rice and *Arabidopsis*. However, high degree of gene duplication and tight linkage between genes act as a major limiting factor to study gene function and genetic recombination in plants (Glover et al. 2015). One possible approach is to use either homologous recombination to eliminate tandem duplications by gene replacements or to introduce point mutations using RNA-DNA hybrids (Reams et al. 2012). This can also be achieved through inserting mutated sequences to generate stop codons within the conserved regions to produce null mutations in a multigene family. However, high throughput gene silencing on double-stranded RNA through bidirectional transcription of genes is broadly accepted, as it is easy to generate transgenic plants with drastic transcriptomic alterations (Zhang et al. 2015). Recently, CRISPR/Cas9 has emerged as a powerful tool to generate knock-in mutants or knock-out mutants with frameshift mutations in plants (Liu et al. 2016).

1.5 Mapping and Map-Based Cloning

Breeding programs of important crop species like rice and wheat functioning from several decades have broadened our knowledge in the mapping of several traits related to abiotic stress tolerance. Introduction of molecular marker techniques in conventional breeding gave further extension to mapping studies and in assessment of cultivated, land race and wild genotypes (Varshney et al. 2012). These studies led to identification of germplasm rich genotypes showing extensive variation at structural and expression levels under stress conditions. These variations are useful to confirm candidate genes for stress tolerance as well as for discovering alleles for further breeding programs (Ma et al. 2012). Majority of the known abiotic stress loci have been discovered as QTL, so a particular trait mapping in different genotypes using multiple populations can locate the common loci such as drought tolerance (Jaganathan et al. 2015). More than hundred abiotic stress related traits have already been mapped only in soybean in past years (Xia et al. 2013). Similarly, availability of whole genome sequence in rice, and its strong similarity with wheat and barley genomes makes rice a potential crop for marker generation from candidate loci.

Further strategy is positional cloning of functionally correlated genes for specific trait using forward genetics approach. Positional cloning may or may not identify target gene(s) associated with a particular phenotype directly. However, through complementation analysis, the target gene can be identified (Langridge and Fleury 2011). Another variant of positional cloning is map-based cloning, where chromosomal location of a gene is identified through genetic mapping using molecular markers (Kudapa et al. 2013). With faster and more accurate next-generation sequencing (NGS) technologies as well as advanced DNA polymorphism detection techniques, map based cloning and physical mapping using BAC libraries have now become more handy for different crops such as rice (Vij and Tyagi 2007), barley (Schulte et al. 2011), soybean (Fang et al. 2013; Song et al. 2016), *Brassica* (Mun et al. 2015) and wheat (Wang et al. 2015). High-density genetic linkage maps have been integrated with sequence-based physical map, thus resulting in improved resolution and accuracy of trait-specific genes/QTLs identification (Agarwal et al. 2016).

1.6 Conclusion

Functional genomics studies have played a central role in not only providing solutions to generate new varieties through genetic transformation but also have increased our understanding of cellular metabolism operating under abiotic stress. Several functionally characterized genes when inserted into crop plants have shown increased tolerance against various environmental stresses in comparison to wild type plants. These genetically engineered plants show higher osmolyte and protein accumulation and are generally more productive in terms of agricultural yield. Further, using genomic tools, stress- and organ-specific promoters have been identified and tested thoroughly for their specificity. Also, comparative genomics studies have identified genes that throw light upon conserved evolutionary mechanisms in plants

Huge wealth of data is now available for plant signaling in response to various abiotic stresses. Additionally, several transcription and signaling factors along with their interconnections and crosstalk mechanisms have increased our understanding of the intricate network that operates under stress. Despite this, full understanding of the genes controlling signaling pathways is lacking. The filtering of the huge data using bioinformatics tools and validation of the genes using advanced genomics tools like proteomics and metabolomics can alleviate this deficit. Mining of the data systematically for functional analysis by using mutants and overexpression analysis followed by microarray analyses can reveal interactions between signaling components and downstream targeted genes. Recent technological developments in functional genomics such as RNAi technology, gene editing and next generation genomics can help us uncover the variations integrated across diverse plant genomes. This can further be applied to manipulate crop species for enhanced defense strategies using conventional, marker assisted or transgenic approaches.