Plant Genomics

Edited by Ibrokhim Y. Abdurakhmonov

PLANT GENOMICS

Edited by **Ibrokhim Y. Abdurakhmonov**

Notice

Statements and opinions expressed in the chapters are these of the individual contributors and not necessarily those of the editors or publisher. No responsibility is accepted for the accuracy of information contained in the published chapters. The publisher assumes no responsibility for any damage or injury to persons or property arising out of the use of any materials, instructions, methods or ideas contained in the book.

We are IntechOpen,
the world's leading publisher of Open Access books
Built by scientists, for scientists

3,800+
Open access books available

116,000+
International authors and editors

120M+
Downloads

151
Countries delivered to

Our authors are among the
Top 1%
most cited scientists

12.2%
Contributors from top 500 universities

BOOK CITATION INDEX
CLARIVATE ANALYTICS
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Meet the editor

Ibrokhim Y. Abdurakhmonov received his B.S. Degree (1997) *in biotechnology* from the National University of Uzbekistan, M.S. degree i*n plant breeding* (2001) from Texas A&M University of USA, Ph.D. degree (2002) *in molecular genetics*, Doctor of Science degree (2009) *in genetics*, and full professorship (2011) i*n molecular genetics and molecular biotechnology* from the Institute of Genetics and Plant Experimental Biology, Academy of Sciences of Uzbekistan. He founded (2012) and is currently leading the Center of Genomics and Bioinformatics of Uzbekistan. He serves as an associate editor/editorial board member of several international and national journals on plant sciences. He received Government award, 2010 chest badge Sign of Uzbekistan," 2010 TWAS prize, and "ICAC Cotton Researcher of the Year 2013" for his outstanding contribution to cotton genomics and biotechnology. He was elected as the World Academy of Sciences (TWAS) Fellow (2014) on *Agricultural Science* and as a co-chair/chair of "Comparative Genomics and Bioinformatics" workgroup (2015) of the International Cotton Genome Initiative (ICGI).

Contents

Preface

Plants are the backbones of food chain for all living matters on Earth, which supply the humankind with food, feed products, and clothing and housing materials; balance agrobiosenosis and earth ecology; provide a medicine and cure for many diseases; and produce energy and biofuels. Plants are models to investigate and understand the life in our planet. All of these have been the main driving factors to care about plants, domesticate wild species with important properties, select and breed the best genotypes, and grow them for needs of humanity. This has consequently shaped plant sciences, genetics, and breeding. Furthermore, the advances made in understanding the genetics of phenotypic variations, equipped with optimized, targeted, and efficient selection, phenotyping, and statistical methods as well as agrotechnologies in past centuries, have revolutionized plant sciences and crop breeding.

The "Green Revolution" has significantly increased agricultural production for the past centuries because of success in conventional breeding and genetics and agro/chemical technologies. Even further, plant researchers have attempted to decipher the molecular basis of genetic diversities by cloning and sequencing the genes conditioning of the trait of interest and utilizing them in plant breeding as tools in vertical (i.e., hybridization) or even via revolutionizing horizontal (i.e., genetic engineering) transfers. There is no doubt that these efforts have significantly contributed to increase agricultural productions worldwide and enriched the diet of human well-beings. However, humanity still suffers with food deficiency that could be even more evident with a global human population increase, which is projected to reach ~9 billion by 2050. Moreover, current global climate changes with increased biosecurity threats could generate even worsened scenarios of more complex challenges for sustained agriculture and food security in the era of the societal globalization and advances of technologies in the twenty-first century and beyond.

There are needs to find better ways and solutions to mitigate the future challenges, which require enrichment and change of methodologies, technologies, and scientific views of the past. Here comes *Plant genomics*— a newly evolved discipline of plant sciences — targeting to decode, characterize, and study the genetic compositions, structures, organizations, functions, and interactions/networks of all plant genes in a genome-wide scale. Being evolved from plant molecular genetics, biology, and biotechnology, *Plant genomics* represent the key sub-divisions of structural, functional, comparative, evolutionary, physiological, and genetical genomics. Its development and advances, however, are tightly interconnected with plant science sub-disciplines such as proteomics, metabolomics, epigenomics, phenomics, metagenomics, transgenomics, breeding-assisted genomics, bioinformatics, and system biology as well as modern instrumentation and robotics sciences.

Aiming to develop genome-wide scaled high-throughput technologies and methodologies for plant science research and application, *Plant genomics* has significantly advanced over the past three decades in the land of inexpensive, high-throughput, and automated new-generation DNA sequencing methodologies and platforms. This resulted in decoding of more than 100 plant genomes of 64 megabases to 23 gigabases in lengths within past 10 years. These advances made have broad implications for every aspects of plant biology while generating many unexpected challenges and grand tasks ahead. The grand tasks ahead require more extended collaborations and integrated approaches as well as better computing and data storage/systematization capacity, ways to handle data with improved or novel analytical/bioinformatics tools, extended training and education of well-qualified new-generation researchers, and larger investments and funding.

This book, *Plant genomics*, aiming to provide updated discussions on current advances, challenges, and future perspectives of plant genomics research and application, has compiled 10 chapters from the plant researchers worldwide, including results of the first draft of full genome sequencing and assembly of a leaf chicory plant. Chapters also have reviewed and discussed strategies for plant genome sequencing, assembly, and its challenges, new-generation sequencing platforms for comparative genomics of cereal crops and non-model cactus plants, and the characterization of small RNA/micro-RNA world of plant genomes. Several chapters have covered the advances toward plant resistance genomics and molecular breeding of bacterial diseases in ryegrasses, RNA interference technology in plants, and some aspects of salinity tolerance genomics in agricultural crops. I trust that these chapter materials will be additional reading sources for scientist, students, and readers interested with the plant science development.

I greatly acknowledge the efforts of all authors of the book chapters for their timely response to our book project invitation, chapter proposal development, and writing and revising full chapters per my editorial requests. I also thank the InTech book department for giving me the opportunity to work on this book project, and Ms. Sandra Bakic and Ms. Iva Simcic, InTech's Publishing Process Managers, for their overwhelming efforts with publishing of this book.

Ibrokhim Y. Abdurakhmonov
Center of Genomics and Bioinformatics,
Academy of Sciences of Uzbekistan,
Tashkent,
Uzbekistan

Introduction to Plant Genomics

Genomics Era for Plants and Crop Species – Advances Made and Needed Tasks Ahead

Ibrokhim Y. Abdurakhmonov

Additional information is available at the end of the chapter

Abstract

Historically, unintentional plant selection and subsequent crop domestication, coupled with the need and desire to get more food and feed products, have resulted in the continuous development of plant breeding and genetics efforts. The progress made toward this goal elucidated plant genome compositions and led to decoding the full DNA sequences of plant genomes controlling the entire plant life. Plant genomics aims to develop high-throughput genome-wide-scale technologies, tools, and methodologies to elucidate the basics of genetic traits/characteristics, genetic diversities, and by-product production; to understand the phenotypic development throughout plant ontogenesis with genetic by environmental interactions; to map important loci in the genome; and to accelerate crop improvement. Plant genomics research efforts have continuously increased in the past 30 years due to the availability of cost-effective, high-throughput DNA sequencing platforms that resulted in fully sequenced 100 plant genomes with broad implications for every aspect of plant biology research and application. These technological advances, however, also have generated many unexpected challenges and grand tasks ahead. In this introductory chapter, I aimed briefly to summarize some advances made in plant genomics studies in the past three decades, plant genome sequencing efforts, current state-of-the-art technological developments of genomics era, and some of current grand challenges and needed tasks ahead in the genomics and post-genomics era. I also highlighted the related book chapters contributed by different authors in this book.

Keywords: Plant genome sequencing, genetical genomics, genomic selection, 1KP, 1001 plant genomes, GEEN

1. Introduction

The Plant Kingdom is a key of the food chain in our planet. Plant domestication by humankind occurred in early societal development, and subsequent agricultural practice and uninten-

tional and intentional plant breeding led to developing productive crop species that provided food and feed products for all living organisms, including humans [1, 2]. Plant species are very diverse and there are about 300,000 plant species in the world [3]. Humankind presently grows ~2000 plant species [4] in the agriculturally suitable land of 15.5 million square kilometers to fulfill the human diet. Crop domestication with subsequent breeding and farming has created 15 priority crop species, which provide more than 90% of food products [1, 5]. Besides feeding properties, plants supply clothing and housing materials, balance agrobiosenosis and earth ecology, provide medicines and treatment for many diseases, produce energy and biofuels, and have many other key properties and usages to understand life in our planet [6–10].

Plant domestication, coupled with the need and desire to get more food and feed products, has resulted in continuous development of breeding and genetics efforts [2, 4]. Early primitive selection attempts have subsequently developed the methods of shuffling traits/characteristics between plant genotypes via controlled sexual crosses that discovered the genetics of key characteristics of crops. Furthermore, the development of biological sciences and understanding of the Mendelian and quantitative genetics of phenotypic variations in plant genotypes, equipped with optimized, targeted, and efficient selection, phenotyping, and statistical methods as well as advanced agrochemical technologies of the past centuries, have revolutionized crop breeding efforts. These advances have resulted in the development of superior crop genotypes that have helped to increase agricultural production [11]. Thanks to the "Green Revolution" [11, 12], the efficient exploitation of plant genetic diversity and plant germplasm resources, novel cultivar development, and better and suitable agrochemical technologies for the past 50 years, the world average cereal crop yield has increased 2.6 times (1.35–3.51), whereas there was 5-fold increase in maize production [11]. There are many such examples of successful conventional breeding efforts. Despite this, food deficiency and human starvation still exist widely and will become even worse with an increase of global human population to ~9 billion by 2050 [13], whereby ~1 billion people may suffer hunger [14]. There is a desire and need to feed the increasing human population, sustain agricultural production, and overcome newly emerging biosecurity issues in the era of global climate change with ever worsening environmental conditions on earth, and societal globalization and technological advances [15, 16].

These prompted the plant research community to enrich and power the conventional plant breeding and genetics methods with precise tools beyond conventional hybridization, selection, and cultivation/farming practices. This is also dictated by the long duration of conventional breeding and crop improvement, impacted by the limitations in phenotypic evaluations, masking the effect of the environment, polygenic nature of many key traits with many unnoticed minor genetic components [11], negative genetic correlations between important agronomic traits [15, 17, 18], linkage drags, and distorted segregation issues in hybridization between diverse genotypes [15, 17–19].

To address all these, plant researchers have attempted to decipher the molecular basis of genetic diversities by cloning and sequencing the genes encoding the trait of interest and utilize them in plant breeding as tools in vertical or even via revolutionizing horizontal gene transfers [11]. Progress made toward this goal has elucidated plant genome composition and led to decoding the entire DNA sequences of plant genomes conditioning plant ontogenesis. Here comes "genomics" that was derived from the use of the term "genome"—a haploid set of

chromosomes—coined by Winkeler in 1920. First used in 1986, genomics defined "the enterprise that aimed to map and sequence the entire human genome" [20]. Similarly, "plant genomics" is a discipline of plant sciences targeting to decode, characterize, and study the genetic (DNA/RNA) compositions, structures, organizations, and functions as well as molecular genetic interactions/networks of a plant genome [20–29]. Plant genomics aims to develop large-scale high-throughput technologies and efficient tools and methodologies to elucidate the basics of genetic traits/characteristics, genetic diversities, and by-product production; to understand the phenotypic development throughout plant ontogenesis with genetic by environmental interactions; to map important loci throughout the genome; and to accelerate the crop breeding and selection in a genome-wide scale.

Plant genomics research efforts have continuously increased in the past 30 years. The numbers of scientific publications on plant genomics research have drastically increased and reached 17,210 scientific publications in 2015, as indexed in the PubMed database [30], with its first increase in 2000/2001, following a significant peak after 2010 (Figure 1). The first fully sequenced plant genome was the model plant Arabidopsis, which was published in 2000. Since then, almost 50 plant genomes were fully decoded by 2013 [31] and the plant sciences community has finished more than 100 plant genomes by 2015 [32]. Furthermore, the plant sciences community extendedly portrayed a sequencing vision of 1001 Arabidopsis accessions [33, 34] and sequencing 1000 plant species [35] that "will have broad implications for areas as diverse as evolutionary sciences, plant breeding and human genetics" while generating many unexpected challenges and grand tasks ahead.
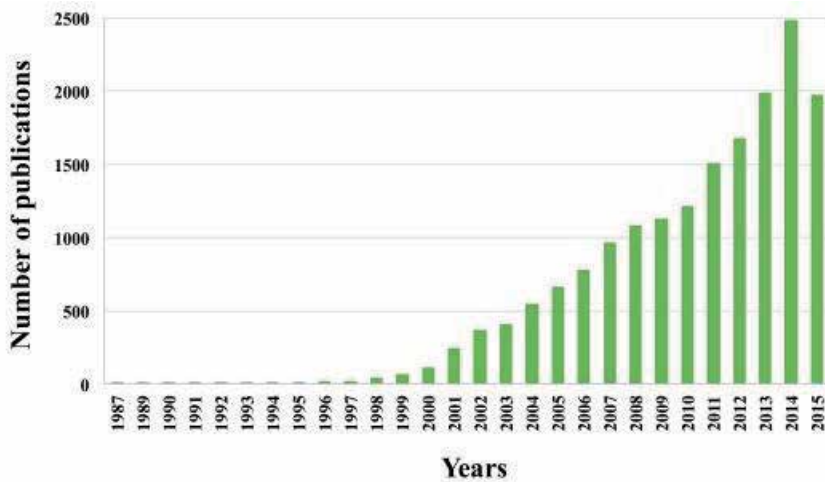
Figure 1. Dynamics of "plant genomics" keyword-retrieved scientific publications in the past three decades. Source: PubMed [30].

2. Genome of plants and crop species

2.1. Challenges and advantages

Compared to other eukaryotic systems, plant genomes are more complex, which create challenges to study its DNA compositions. First of all, the extraction of high-quality DNA from plant tissues, abundantly enriched with phenolic and other metabolic compounds with high affinity to DNA, is conventionally challenging. This interferes with efficient library preparation for whole-genome sequencing [1], although researchers have optimized methodologies to overcome existing issues [36].

Furthermore, plant genomes have widely different chromosome numbers, transposon/retrotransposon transcript retention property, and highly varied ploidy levels with many supergenes, pseudogenes, and repetitive elements including low-, medium-, and high-copy number DNA sequences such as transcribed genes, rRNA genes, and retro-elements or short repetitive sequences, respectively. As a result, plant genomes can be 100 times larger in sizes when compared to animal or other model eukaryotic genomes [1] and may contain many paralogous DNA sequences that make sequencing and genome assemblies difficult, which often will generate false-positive errors [37]. For instance, one of the largest examples of sequenced plant genomes, sugarcane (12 Gbs) and hexaploid wheat genome with 17 Gbs in size, represents 80% repetitive elements [1, 32].

Moreover, these massive repetitive "junk" DNA sequences, organized as a simple tandem repeat, repeat single-copy interspersion, inverted repeats, and compound tandem array arrangements, somewhat mask functionally vital single-copy genes, which create a challenge to characterize and clone important individual genes [32, 37].

Open pollinated, self-pollinated, and clonally propagated plant species have a high level of nucleotide diversity. This can be exemplified by the nucleotide diversity of maize, barley, and grape genomes, where maize genome, for instance, has 10-fold (up to 13%) more polymorphic sites between individual genotypes compared to humans with similar genome size [32, 37]. These polymorphism sites create a challenge in sequence assembly due to the higher rates of nucleotide mismatches to the reference genome.

Lastly, plants tend to have abundant copies of chloroplast genome with two inverted repeat organizations as well as large inversions in some plants with some exchanged regions between nuclear genomes. This creates another challenge in the assembly of repetitive and exchanged regions of chloroplast genomes [32]. The same issue exists in the case of mitochondrial genomes, although it is common for animal genomes as well. All these challenges and complications mentioned above may result in generating fragmented, isolated, and incorrect assemblies in the background of high-copy repeats and paralogous sequences.

However, some specific methodologies and bioinformatics tools have been developed to minimize these challenges. These include the optimized DNA isolation from difficult plant materials [36], use of high-density linkage maps, identification and sorting out of paralogous alleles using local patterns of linkage disequilibrium, and sequencing diploid relatives or

ancestor-like genomes of polyploid plants [37]. The use of laser capture microdissection techniques can isolate individual cell types or chromosome or its arm that could minimize the ploidy or paralogy complexities [27]. Moreover, the use of third-generation single-molecule sequencing approaches [1] and novel assembly methods such as optical mapping and long-range Hi-C interactions can also minimize some of challenging cases with the plant genomes mentioned here, which have been well addressed and covered in detail in a chapter by Deschamps and Llaca in this book.

At the same time, along with these challenges and complexities, plants also offer advantages [37] in genome analyses over other eukaryotic systems. This is due to the clonal propagation and indefinite seed storage properties, which create an opportunity for repeated collection of the same DNA samples for sequencing and studying its phenotype multiple times in many generations across replicated environments [37]. There are no ethical issues associated with the multiple use of plant materials, as it is a sensitive issue for animal cases. The possibility of self-pollination or forced crosses advantageously helps to create highly homozygous samples to reduce existing heterozygosity. There is an opportunity of obtaining double haploid plant genomes [37]. Plant genomes tend to have large chromosomal segments conserved across a large number of taxa in closely related plant species. The collinearity and synteny of plant genomes are very useful to use reference genomes of model species to study homologous and orthologous genes from yet unsequenced genomes [20].

2.2. Sequenced plant genomes

The ability to sequence DNA molecules, which was made possible in the 1970s with the introduction of the "plus and minus" sequencing technique of Sanger and Coulson [38] and Maxam and Gilbert [39], is generally considered to be the starting point of genomics sciences. Later, the simple, long-read chain-terminating dideoxynucleotide DNA sequencing method [40] has become a method of choice to decode genetic sequences. Its eventual automation [41] had extended the capacity and power of this chain-termination sequencing methods to decode the entire genome sequences of living organisms. Because of technological advances and automated sequencing instrumentations [27], a large-scale sequencing of cDNA libraries made it possible to perform serial analysis of gene expression (SAGE) and expressed sequences tags (ESTs). These were the first genomics technologies for all organisms, including plant genomes [42]. Furthermore, these advances powered by microarray tools routinely used by many individual laboratories worldwide have helped to identify the genome structures and functional and regulatory elements across genomes [27] and have facilitated to develop high-throughput reliable molecular markers for genome/trait mapping studies.

The development and generation of massively parallel sequencing technologies [44] provided cost-effective, new-generation sequencing (NGS) platforms that have helped to completely decode the entire genome of many different organisms within a short period. For instance, in plants, the first sequenced genome was a model plant *Arabidopsis thaliana* with 125 Mbs in size, 25,489 individual genes, and 14% repetitive elements, which was published in 2000 [5]. Further, more than 109 plant genomes have been fully sequenced by 2015 [32], including 21 monocots and 83 eudicots, 10 model and 15 non-model plant genomes, five non-flowering plants, and

69 crop species with 6 crop model genomes and 15 wild crop relatives [32]. Following the Arabidopsis model, several rice (*Oryza sativa*) genomes in 2002 to 2005, black cottonwood (*Populus trichocarpa*) genome in 2006, and grape (*Vitis vinifera*) genome in 2008 were fully sequenced. Sequencing whole plant genomes has increased in subsequent years, and 10 plant genomes had been sequenced in 2011. About 80% of sequenced genomes were accomplished in the past 3 years (2012–2014; Figure 2).
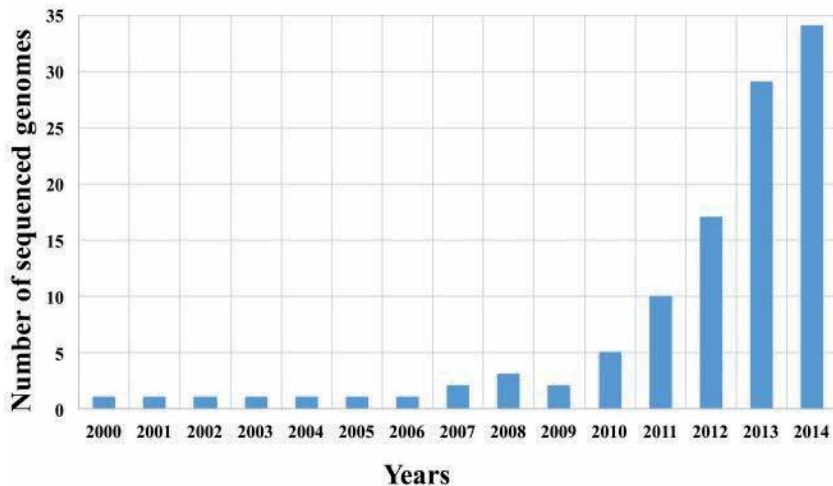
Figure 2. A number of sequenced plant genomes from 2000 to 2014. Source: Ref. [32].

The smallest plant genomes sequenced so far [32] are the two eudicot plants: corkscrew (*Genlisea aurea*) with 64 Mbs genome size and 17,755 genes [45] and bladderwort (*Utricularia gibba*) with 77 Mbs genome size and 28,500 genes [46]. In contrast, the largest genomes sequenced are from gymnosperm plants, including Norway spruce (19,600 Mbs) [47], white spruce (20,000 Mbs) [48], and loblolly pine (23,200 Mbs) [49]. The largest genome sequenced from crop species is the hexaploid wheat (*Triticum aestivum*) with a genome size of 17,000 Mbs [50]. An average size for all published plant genomes is 1850 Mbs. Per published plant genome data [32], the gene numbers of the smallest to largest genomes are within the range of 17,755 (corkscrew) to 124,201 (hexaploid wheat) with an average of 40,738 genes for all sequenced genomes. Repetitive elements are highly variable among published genomes that varied from 3% (bladderwort) to 85% (maize, *Zea mays*) with an average estimate of 46% per genome. These sequenced plant genomes not only provided an updated knowledge on structural compositions and complexities of plant genomes but also elucidated the evolution of gymnosperm and angiosperm plants and specific gene families contributing to the radiation of flowering plants. We learned some direct correlation between genome sizes and gene numbers/repetitive elements, although it does not strictly follow the rules, which was evidenced by several exceptions. For example, one of the largest genomes, Norway spruce, has ~28,000 genes, which is similar to the smallest genome bladderwort. Moreover, medium-sized maize genome (2300 Mbs) or wild tomato (1200 Mbs) contains

more or approximately the same (<80%) contents of repetitive elements compared to the largest sequenced genome of loblolly pine (23,200 Mbs) [32].

In this book, the chapter by Galla et al. (Section 2) presents the results of the first draft of the full genome sequence and assembly of a fresh salad plant leaf chicory (*Cichorium intybus* subsp. *intybus* var. *foliosum* L., 2*n*=2*x*=18, and 1.3 Gbs genome size), named as Radicchio in Italian. The results of decoding the full genome of leaf chicory will "extend the current knowledge of the genome organization and gene composition of leaf chicory, which is crucial for developing new tools and diagnostic markers useful for our breeding strategies in Radicchio" and will be an important addendum to the list of sequenced plant genomes.

2.3. Sequencing "1001 genotypes" and "1000 plant species"

The availability of a few whole reference genomes limits our full understanding of ecotypic variations that affect the function and adaptive evolution of plant species in various climatic conditions. It reduces the power of genome-wide tagging of biologically meaningful natural variations. In other words, the general perceptions are that "a single reference genome is not enough" for plant biology to explain and understand the existing natural variations in particular plant species and its populations [33, 34]. It also limits the development of efficient tools for genome analyses. To address this, as mentioned above, the Arabidopsis plant research community has developed a vision of sequencing a larger number of Arabidopsis genotype accessions, including various ecotypic and experimental population samples. As of today, the "1001 genome sequencing project of Arabidopsis accessions" has completed the full genome sequencing of 1100 Arabidopsis accessions [33, 34] "to record the genetic variation in the entire genome of many strains of the reference plant *Arabidopsis thaliana*" and with the future objective to develop efficient genome analysis tools and software [33].

To understand the tree of life of the Plant Kingdom and study its evolutionary aspects in comparison to other life forms, the international multi-disciplinary consortium of "The 1000 Plants (oneKP or 1KP) Initiative" has generated a large-scale gene sequencing data for more than 1000 various plant species [35]. Rather than concentrating on single species accessions as in the 1001 Arabidopsis whole-genome sequencing project [33, 34], the "1KP" project targeted 1000 distinct plant species with the objective of generating only functionally expressed (i.e., transcriptome) gene sequences. The plant species selected for the project had no restriction, and the samples were "chosen to represent every species known to science, across the Plant Kingdom, at some phylogenetically or taxonomically defensible levels" [35]. The 1KP sample list consists of 1328 entries [51] broadly grouped by phylogenetically (angiosperm, non-flowering, and green algae species) and by application (agriculture, medicine, biochemistry, and extremophytes). Most of these species have been sequenced for the first time (Table 1).

To date, an average of 2000 Mbs transcriptome sequence data have been generated for these 1KP plant species using 28 Illumina Genome Analyzer next-generation DNA sequencing machines at the Beijing Genomics Institute (BGI-Shenzhen, China) [35]. Ultimately, the obtained genomic sequence data will be used to analyze the phylogenetic, taxonomic, and evolutionary relationships of plant species, to study plant speciation, and to determine the timing of gene duplications during speciation events [35, 52]. However, the biggest limitation

is associated with sequencing only transcriptomes rather targeting the whole genome, which limits obtaining many non-coding and repetitive portions of genomes. The results of "1001" and "1KP" sequencing efforts will undoubtedly open a new paradigm for plant genomics and its above-mentioned sub-disciplines. The results should not only accelerate crop improvement and boost the agricultural and medicine production worldwide but also help to understand the basics of plant life, evolution, speciation, and plant adaptations to the extreme environments in the era of global climate change and technological advancements.

Plant species	*Number of samples
Phylogenetic groups	
Angiosperms	830
Angiosperms: Onagraceae samples	50
Non-flowering plants	257
Green algae	241
Application groups	
Medicinal samples	142
Medicine - Alkaloid samples	30
Medicine - Chemotherapeutic samples	12
Biochemistry - Lipid Biosynthesis samples	15
Agriculture - C3/C4 samples	93
Agriculture - Weeds	25
Extremophyte samples	31
Halophytes samples	18

*The number of samples overlaps among groups. Source: Ref. [51].

Table 1. Plant species samples chosen for the "1KP" plant genome sequencing project.

In this book, we have presented several chapters targeting to review and discuss the strategies for sequencing and assembly challenges (by Deschamps and Llaca), new-generation sequencing platforms for comparative genomics of cereal crops (by Sikhakhane et al.) and non-model cactus plant Nopal (*Opuntia* spp.; by Alonso-Herrada et al.), and characterization of small RNA world of plant genomes (Hernández-Salazar et al.). These chapters describe the current advances and future needs on these topics.

3. Crop improvement in the genomics and post-genomics era

3.1. Genomics-assisted selection or genomic selection

At present, the reference genomes for many agricultural plants including specialty crops have been sequenced, as reviewed by Michael and VanBuren [32], which created a new paradigm for modern crop breeding. Crop breeding, which is powered and enriched by molecular

markers, genetic linkage maps, QTL mapping, association mapping, and marker-assisted selection methods in the past century [37, 53], has now greatly accelerated and become ever productive and efficient in the plant genomics era [26]. This is due to the (1) availability of large-scale transcriptome and whole-genome reference sequences [32]; (2) high-throughput SNP marker collection and cost-effective, automated, and high-throughput genotyping platforms (HTP) and technologies (e.g., genotyping by sequencing or GBS), allowing breeders to screen multiple genotypes within a short time [23, 26]; (3) identification and use of expression QTLs (genetical genomics) in breeding [22]; and (4) opportunity to perform genome-wide selection (i.e., genomic selection) [26].

The biggest driving force for genomics-assisted crop breeding in the plant genomics era has been the inexpensive sequencing and re-sequencing opportunity for population individuals of genetic crosses and breeding lines. This helps to precisely identify and link genetic variations to the phenotypic expressions, taking into account the rare and private allelic variations that are abundant in crop line population or germplasm resources [26, 53, 54]. Furthermore, the availability of SNP marker collections and automated genotyping platforms provided a better genome converge to perform genome-wide genotype-to-phenotype associations (GWAS) [11, 37]. Also, when whole-genome sequences are not available and SNP markers are present in a limited number, the breeders using GBS and HTS platforms can readily genotype their mapping population and can provide genomic selections for the targeted crops of interest [23, 26, 54]. Although it was first applied for animal breeding [55], recently genomic selection has been successfully applied to a number of plant species [56–62], including studies using GBS in the context of genomic selection [26]. Most importantly, the application of available genomics tools and a large number of high-throughput DNA markers and new-generation genotyping platforms have made the "breeding by design" [63] possible and have developed "virtual breeding" approaches [64] for efficient crop improvement. Several chapters in this book have covered the advances toward plant resistance genomics and molecular breeding against bacterial diseases in ryegrasses (see the chapter by Dr. Takahashi) as well as biotic/abiotic stress tolerance in agriculture crops (see the chapters by Onaga and Wydra, and Rao et al.).

The availability of genome sequences and a large number of SNP marker collections also provided the analysis of copy number variations (CNVs) in crop genomes, and their links to the key traits have greatly enhanced the crop improvement programs [11, 22, 23, 26, 37]. Furthermore, although challenges are evident, the opportunity provided by post-genome sequencing advances has help to integrate and enrich genomic selection with key proteome and metabolome markers. This significantly fostered and powered up the breeding of complex traits [22] of crops. Consequently, the knowledge gained through plant genomics coupled with proteomic and metabolomic advances has facilitated the emergence of an innovative approach of "personalized" agriculture through the utilization of chemical genomics [21]. This requires the translation of knowledge and expertise of the pharmaceutical industry on the development of "personalized medicine" to treat each person based on its reaction to the medical drugs into the agriculture. Because of high-throughput genome analysis, it is possible to date that many plant compounds, including herbicides, growth regulators and phytohormones, elicitors, low molecular metabolites (e.g., salicylic acids), and/or synthetic hybrid chemicals, can be screened for genetic response of individual crop genotypes and to study their mechanism of actions

contributing to agricultural productivity. Once identified, highly genotype-specific chemical compounds can be developed that impact better than traditionally applied "fit for all" chemicals/growth stimulators and fertilizers. A combination of such chemical genomics approach, proteomics and metabolomics with genetic engineering, and genomic selection will further provide a way for "personalized" agriculture that sustains crop production (for detailed discussions, see a review by Stokes and McCourt [21]).

3.2. Novel transgenomics tools and biotech crops

Crop improvement is also greatly impacted by novel transgenomics and genome editing technologies developed as a result of plant genome characterization and understanding in the era of plant genomics. In the past two decades, a variety of novel transgenomics technologies have been developed to replace or enrich the traditional transgenesis-based genetic engineering and plant molecular biotechnology [65]. These novel technologies include antisense, RNA interference (RNAi), artificial microRNA expression (amiR), virus-induced gene silencing (VGS), zinc-finger nuclease (ZFN), transcription activator-like effects nucleases (TALENs), oligonucleotide-directed mutagenesis (ODM) of Cibus Rapid Trait Development System (RTDS), and clustered regularly interspaced short palindromic repeats/Cas9 (CRISPR/Cas9) technologies [65, 66]. These novel transgenomics technologies including genome-editing tools,the latter also referred to as genome editing with engineered nucleases (GEEN), are widely developed and utilized to investigate the gene function and apply to solve problems in medicine and agriculture. They are become methods of choice for major functional genomics and biotechnological studies [67]. RNA-mediated genome manipulation (RNAi) tools down-regulate the target genes due to gene silencing effects at transcriptional (TGS) or post-transcriptional (PTGS) levels, whereas GEEN systems help to insert, replace, or remove specific regions of DNA from a genome using artificially engineered nucleases that are referred to as "molecular scissors" [68–70]. For a detailed description of RNAi, readers are suggested to read a chapter by Ricano-Rodriguez et al. in this book as well as to the recently published "RNA Interference" book by InTech Open.

The potential application of RNA-mediated gene silencing methods for crop improvement, including RNAi in plant biotechnology, is huge and the technology has already generated many successful examples in a wide range of technical, food, and horticulture crops. For example, RNAi was used to improve crop yield, food/fiber quality [18, 71–75], resistance to pests, and biotic/abiotic stresses [76, 77], which are being considered for commercialization or are already in commercial production [78]. Employing ODM-mediated single nucleotide editing in Arabidopsis, targeting the BFP gene, has demonstrated a precise edition of CAC to TAC, converting histidine (H66) to tyrosine (Y66) in GFP protein that offered a non-transgenic breeding tool for crops [66]. Similarly, GEEN tools have also provided a new strategy for "trait stacking," whereby several desired traits are physically linked to ensure their co-segregation during the breeding processes [79]. The examples include *A. thaliana* [80–82] and *Z. mays* [83], where ZFN-assisted gene targeting has helped to heritably insert herbicide-resistant genes (SuRA/SuRB and PAT) into the targeted sites in the genome [83]. Although other GEEN technologies such as TALEN [84–92] and CRSPR/Cas9 [93] are just picking its application in plants, their utilization in Arabidopsis [84], maize [85], rice [86–88], potato [89, 90], wheat [65],

barley [91], and plum [92] holds a great promise and potential for non-transgenic crop genome modification and improvement [65, 94].

4. Grand tasks ahead

The revolutionizing advances made in the past three decades in plant genomics and its sub-disciplines provided a mass of novel opportunities with easy-solution applications and high-throughput, cost-effective, and time-effective technologies. Plant genomics era increased our understanding of the basis of complex life processes/traits in plants and crop species, and it paved a way for effective improvement of plants to fulfill our diet and other needs. However, it also piled up challenging grand tasks ahead for current genomics and post-genomics era. Several chapters of this book have discussed some aspects of these challenges, and I tried to briefly summarize some of them here.

As mentioned above, tremendous achievements have been made toward sequencing more than hundreds of plant genomes including major crop species and specialty, model/non-model, wild, vascular, flowering, and polypoid plants [31, 32]. There are ongoing and fasci-nating consortia projects of sequencing "1001 genotypes of Arabidopsis" and "1000 various plant species" [33–35, 51, 52]. However, the first current and future task ahead is to extend such large-scale, multiple accession genome sequencing initiatives for each priority agricul-tural and specialty crop species including their wild relatives and ancestor-like genome representatives. Although it sounds largely ambitious, this task will be mandatory and important for the next plant genome sequencing phase. This is to effectively use all variations existing among plant/crop germplasm resources and its ecotypic populations and to design efficient GWAS analysis and consequent genomic selections as well as tools/software pro-grams for better analyzing plant genomes and improving genome assembly issues [33–35]. This is especially needed for polyploidy crops [24, 32, 37] because the sequencing of many polyploids and their subgenomes would increase our understanding of the complexity of polypoidy, gene silencing, epigenetics, and biased retention and expression of genes after polyploidization [24, 95–97]. Furthermore, it also helps to discover all natural variations and lost genes during crop domestication that should be useful to restore the key agriculturally important traits in the future.

Sequencing the entire genome of 1KP samples, rather concentrating on only transcriptome/ exome, is also the necessary task ahead that would elucidate many important noncoding sequences from these plant species. Results would be useful for plant evolutionary, speciation and taxonomy studies. There are ongoing planning and targets toward this goal, and it should not cause much trouble in the land of experiences gained and inexpensive high-throughput sequencing technologies [1, 27, 32].

Although high-throughput DNA sequencing instrumentation exists and keeps evolving to better versions year-to-year, the consequent task is still to improve the sequence length that would solve many incorrect sequence sites and genome assembly challenges that plant genomics faces currently [1, 32]. Some of the currently ongoing efforts and possible solution with the advent of third-generation sequencing platforms and genome assembly tools and methodologies highlighted herein have been discussed by several book chapters in this book.

A consequent grand task and challenge with the completion of the above-highlighted tasks is the handling, organizing, systematizing, and visualizing a huge amount of plant genome sequencing ("Big Data") data that require urgent attention, effort, collaborative work, and investment. There is an urgent need to develop more efficient bioinformatics platforms to handle plant genome data due to challenges, specificities, complexities, and sizes of currently available and future sequenced plant genomes mentioned herein [1, 98]. Funding this aspect of plant genomics and bioinformatics research is a necessary key step [1] for future advances on this task ahead.

Furthermore, the most important current and future post-genomics grand task ahead is to link the sequence variation(s) with phenotype(s), trait expression, and epigenetic and adaptive features of plants to their living environment and extreme conditions. The successful completion of this task will require the combined approaches of genomics with bioinformatics, proteomics, metabolomics, phenomics, genomic selections, genetical genomics, reverse genomics, system biology, etc. [11, 21–29, 64, 65, 98]. In other words, there is a need to make sequenced genomes "functional" [31] and biologically meaningful [29, 37]. This also requires the integration of all available genomic and phenotypic data to identify key networks that also require downstream effort of integration of specific networks to networks of other systems in order to connect heterogeneous data [29]. There are suggested thoughts and tasks for plant genomics that should target to develop plant genome-specific "Encyclopedia of DNA Elements (ENCODE)" [31, 32], which will be an important achievement in the next phases of development. There is a need to use molecular phenotyping (i.e., using molecular process such as protein-RNA interactions, translation rates, etc.) in QTL mapping [23] that would help to precisely link the sequence variation(s) to its phenotype(s). There is a task for the development and translation of the concept "personalized agriculture" [21] that requires an attention as an unexplored area in crops with the availability of sequenced genomes and high-throughput genotype, proteome, metabolome, and phenotype profiling platforms and rapid crop line development tools such as genomic selection and new-generation genome-editing tools mentioned above. All these will help to minimize the current challenges with improved crop line development costs through efficient breeding [11, 22, 23, 26]. These particular grand tasks further highlight a need for extended effort and work on the development of inexpensive high-throughput plant phenotyping [25, 26] and plant proteome and metabolome profiling tools and instrumentation [27, 28] by utilizing small amount single-cell-derived samples [27–29].

A parallel grand task to the above-outlined needs is to have concentrated efforts on the timely application of novel transgenomics and genome-editing tools for all types of plants and to optimize it for routine large- and short-scale biotechnology industry usage. There are grandest tasks to (1) utilize the complex effects of plant developmental genes (e.g., core microRNA/ RNAi machinery) to simultaneously improve the key traits and overcome negative trait correlations [15, 18] and (2) optimize and better design novel transgenomics and genome-editing technologies for the key priority crops and plant by-product production. In addition, there are needs to (3) identify the appropriate choice of plant tissues for genome editing, (4) reduce or eliminate side effects and off-target toxicity and mutagenesis of application of novel genome modification technologies, and (5) develop reliable screens for the detection of edited genome samples [99]. The revolutionizing effects of these novel genome-editing/manipulation

technologies and genome-edited organisms (GEOs) as well as their safer nature compared to conventional transgenesis are evident. However, without objective or proper regulatory policies, providing understanding and removing confusion of regulatory agencies and stakeholders [94], "these technologies may not live up to their full potential" [64] if they are regulated as genetically modified organisms bearing foreign genes [64, 94]. Therefore, this is one of the most important grand tasks ahead in the front of plant sciences research community in the era of plant genomics and post-genomics.

Finally, the grandest task is a preparation of well-qualified next-generation scientists capable of continuing plant genomics tasks highlighted herein with the understanding of conventional plant biology, ecology, plant breeding, evolution, taxonomy, modern "omics" disciplines, and cross-related scientific disciplines (e.g., mathematics, computing, and modeling) [1, 98]. Importantly, they are required to have a capability to utilize modern computing and instrumentation platforms and bioinformatics knowledge [29]. For instance, there is a huge need for a new generation of molecular breeders [100] with full knowledge and appreciation of conventional plant breeding aspects including the understanding of agrotechnology methodologies, genetic diversity of crop germplasm, and randomized multi-environmental field trails. These breeders also need to have abilities to handle, work, and utilize the sequenced genomes, high-throughput genotyping, and phenotyping platforms. This is a bottleneck for plant genomics at present, which requires urgent awareness, attention, and investment.

5. Conclusions

Thus, in the past three decades, plant genomics has evolved from the enrichment and advances made in conventional genetics and breeding, molecular biology, molecular genetics, molecular breeding, and molecular biotechnology in the land of high-throughput DNA sequencing technologies powering the plant research community to sequence and understand the genetic compositions, structures, architectures, and functions of full plant genomes. The technological and instrumentation advancements as well as the desire and need to feed the increasing human population, overcome biosecurity issues, and sustain agricultural production in the era of global climate change, the societal globalization, and technological advancements have been the main driving forces for plant genomics development. These led to sequence and assemble entire plant genomes including very complex polyploid plants, annotate gene functions, link the sequence variation(s) to the phenotype(s), and exploit sequence variation(s) in plant/crop improvement in genome-wide scale or through targeted native modification of plant genomes in a highly sequence-specific manner.

To date, more than 100 plant genomes including a large number of crops as well as flowering, non-flowering, crop wild relative, model and non-model, and specialty plants have been fully sequenced. As a result, it expanded our knowledge and understanding of many aspects of plant biology, genetics, breeding, and crop evolution and domestication, which contributed to the development of analytical and breeding tools, resulting in accelerated crop improvement programs. To look even deeper scales, more than 1100 Arabidopsis accessions from various

eco-geographic origin and experimental populations have been fully sequenced, which will equip plant researchers with better analysis tools and help in tagging and exploiting the biologically meaningful variations. Furthermore, transcriptome profiling of 1000 distinct plant species with agricultural, medicinal, biochemical, and evolutionary utilization has a great value and will be "a gold mining" opportunity for plant biology to explain the evolution of tree of life and Plant Kingdom speciation. All of these successes have significantly accelerated crop improvement using novel genomic selections and new-generation genome-editing and manipulation technologies.

These advances, briefly highlighted herein, also have generated a number of grand challenges and mandatory tasks ahead in plant genomics and post-genomics era. There are many tasks ahead for the plant genomics community, which require more collaborations, integrated approaches, better computing capacity and analytical tools, accelerated training and education of well-qualified researchers, and larger investments. In this book, the authors tried to highlight some updates on current plant genomics efforts with future perspectives. We trust that the next phase of plant genomics efforts and development will be more exciting and help to solve current and future issues in front of humanity.

Acknowledgements

Author details

Ibrokhim Y. Abdurakhmonov[*]

*Address all correspondence to: ibrokhim.abdurakhmonov@genomics.uz; genomics@uzs-ci.net

Center of Genomics and Bioinformatics, Academy of Sciences of Uzbekistan, Tashkent, Uzbekistan

References

[1] Schatz MC, Witkowski J, McCombie WR. Current challenges in de novo plant ge‐ nome sequencing and assembly. Genome Biol. 2012;13:243. DOI:10.1186/ gb-2012-13-4-243

[2] Kossmann J. Grand challenges in plant biotechnology. Front Plant Sci. 2012;3:61. DOI: 10.3389/fpls.2012.00061

[3] Li X, Yang Y, Henry RJ, Rossetto M, Wang Y, Chen S. Plant DNA barcoding: from gene to genome. Biol Rev Camb Philos Soc. 2015;90:157-66. DOI:10.1111/brv.12104.

[4] Abdurakhmonov IY, editor. Plant breeding. Rijeka: InTech; 2012. 352 p. DOI: 10.5772/1389

[5] United Nations Food and Agriculture Organization. Dimensions of Need - An atlas of food and agriculture. Staple foods: What do people eat [Internet]. 2015. Available from: http://www.fao.org/docrep/u8480e/u8480e07.htm [Accessed: 2015-11-10]

[6] Falkowski P, Scholes RJ, Boyle E, Canadell J, Canfield D, Elser J, Gruber N, Hibbard K, Hogberg P, Linder S, Mackenzie FT, Moore B, Pedersen T, Rosenthal Y, Seitzinger S, Smetacek V, Steffen W. The global carbon cycle: a test of our knowledge of earth as a system. Science. 2000;290:291–296. DOI: 10.1126/science.290.5490.291

[7] Pimentel D, Harvey C, Resosudarmo P, Sinclair K, Kurz D, McNair M, Crist S, Shpritz L, Fitton L, Saffouri R, Blair R. Environmental and economic costs of soil ero‐ sion and conservation benefits. Science. 1995;267:1117–1123. DOI: 10.1126/science. 267.5201.1117

[8] Fairley P. Introduction: Next generation biofuels. Nature. 2011;474:S2–5. DOI: 10.1038/474S02a

[9] Mann J. Natural products in cancer chemotherapy: past, present and future. Nat Rev Cancer. 2002;2:143–148. doi: 10.1038/nrc723

[10] Lippman Z, Gendrel AV, Black M, Vaughn MW, Dedhia N, McCombie WR, Lavine K, Mittal V, May B, Kasschau KD, Carrington JC, Doerge RW, Colot V, Martienssen R. Role of transposable elements in heterochromatin and epigenetic control. Nature. 2004;430:471–476. DOI: 10.1038/nature02651

[11] Prohens J. Plant Breeding: A Success Story to be Continued Thanks to the Advances in Genomics. Front Plant Sci. 2011;2:51. DOI:10.3389/fpls.2011.00051

[12] Jain HK. The Green Revolution: History, Impact and Future. Houston: Studium Press LLC. 2010. 256 p.

[13] Hubert BM, Rosegrant van Boekel MAJS, Ortiz R. The Future of Food: Scenarios for 2050. Crop Sci. 2010;50:S-33–50. DOI:10.2135/cropsci2009.09.0530

[14] Higgins TJ. Time to modify the GM debate [Internet]. 2011. Available from: http://theconversation.edu.au/time-to-modify-the-gm-debate-210.   [Accessed   from 2015-11-10]

[15] Abdurakhmonov IY. Role of genomic studies in boosting yield. In: Proceedings of International Cotton Advisory Board (ICAC); 20 September-4 October 2013; Cartagena. Colombia. p. 7-22

[16] Sharma SB, You Y, Varaprasad RS. Biosecurity – an integral part of global food security strategy. Indian Journal of Plant Protection. 2008;2:165-172.

[17] Selection for increased lint yield and correlated responses in Upland cotton, *Gossypium hirsutum* L. Crop Sci. 1967;7:637–640. DOI:10.2135/cropsci1967.0011183X000700060024x

[18] Abdurakhmonov IY, Buriev ZT, Saha S, Jenkins JN, Abdukarimov A, Pepper AE. Cotton *PHYA1* RNAi enhances major fiber quality and agronomic traits of cotton (*Gossypium hirsutum* L). Nature Communications. 2014;4:3062; DOI:10. 1038/ ncomms4062

[19] Saha S, Stelly DM, Raska DA, Wu J, Jenkins JN, McCarty JC, Makamov A, Gotmare V, Abdurakhmonov IY, Campbell BT. Chromosome substitution lines: concept, development and utilization in the genetic improvement of Upland cotton. In Abdurakhmonov IY, editor. Plant Breeding. Rijeka: InTech; 2012. p. 107-128. DOI: 10.5772/35585

[20] Varma A, Shrivastava N. The role of plant genomics in biotechnology, In: Doelle, HW, DaSilva EJ, editors. Biotechnology. Encyclopedia of Life Support Systems(EOLSS), Developed under the Auspices of the UNESCO. Oxford: Eolss Publishers. Available from: http://www.eolss.net/sample-chapters/c17/e6-58-07-15.pdf [Accessed: 2015-11-11]

[21] Stokes ME, McCourt P. Towards personalized agriculture: what chemical genomics can bring to plant biotechnology. Front Plant Sci. 2014;5:344. DOI: 10.3389/fpls. 2014.00344

[22] Joosen RV, Ligterink W, Hilhorst HW, Keurentjes JJ. Advances in genetical genomics of plants. Curr Genomics. 2009;10:540-9. DOI: 10.2174/138920209789503914.

[23] Jimenez-Goimez JM. Next generation quantitative genetics in plants. Front Plant Sci. 2011;2:77. DOI:10.3389/fpls.2011.00077

[24] Song Q, Chen ZJ. Epigenetic and developmental regulation in plant polyploids. Curr Opin Plant Biol. 2015;24:101-9. DOI: 10.1016/j.pbi.2015.02.007.

[25] Fahlgren N, Gehan MA, Baxter I. Lights, camera, action: high-throughput plant phenotyping is ready for a close-up. Curr Opin Plant Biol. 2015;24:93-9. DOI:10.1016/ j.pbi.2015.02.006

[26] Poland J. Breeding-assisted genomics. Curr Opin Plant Biol. 2015;24:119-24. DOI: 10.1016/j.pbi.2015.02.009

[27] Deal RB. Grand Challenge: Accelerating Discovery through Technology Development. Front Plant Sci. 2011;2:41. DOI: 10.3389/fpls.2011.00041

[28] Heazlewood JL. The green proteome: challenges in plant proteomics. Front Plant Sci. 2011;2:6. DOI:10.3389/fpls.2011.00006

[29] Fernie AR. Grand Challenges in Plant Systems Biology:Closing the Circle(s). Front Plant Sci. 2012;3:35. DOI:10.3389/fpls.2012.00035

[30] PubMed database [Internet]. 2015. Available from: http://www.ncbi.nlm.nih.gov/ pubmed [Accessed from 2015-09-06].

[31] Michael TP, Jackson S. The First 50 Plant Genomes. The Plant Genome. 2013;6:1-7. DOI:10.3835/plantgenome2013.03.0001.

[32] Michael TP, VanBuren R. Progress, challenges and the future of crop genomes. Curr Opin Plant Biol. 2015;24:71-81. DOI: 10.1016/j.pbi.2015.02.002

[33] 1001 genomes: A Catalog of *Arabidopsis thaliana* Genetic Variation [Internet]. Available from: http://1001genomes.org/ [Accessed; 2015-11-10]

[34] Weigel D, Mott R. The 1001 genomes project for *Arabidopsis thaliana*. Genome Biol. 2009;10:107. DOI: 10.1186/gb-2009-10-5-107

[35] Leebens-Mack J. OneKp Capstone Wiki [Internet]. 2015. Available from: https:// pods.iplantcollaborative.org/wiki/display/iptol/OneKP±Capstone±Wiki   [Accessed: 2015-11-10]

[36] Healey A, Furtado A, Cooper T, Henry RJ. Protocol: a simple method for extracting next-generation sequencing quality genomic DNA from recalcitrant plant species. Plant Methods. 2014;10:21. DOI: 10.1186/1746-4811-10-21

[37] Morrell PL, Buckler ES, Ross-Ibarra J. Crop genomics: advances and applications. Nat Rev Genet. 2011;13:85-96. DOI: 10.1038/nrg3097

[38] Sanger F, Coulson AR. A rapid method for determining sequences in DNA byprimed synthesis with DNA polymerase. J Mol Biol. 1975;94:441-8. DOI: 10.1258/jrsm.99.2.62

[39] Maxam AM, Gilbert W. A new method for sequencing DNA. Proc Natl Acad Sci U SA. 1977;74:560-4. DOI:10.1073/pnas.74.2.560

[40] Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences USA. 1977;74:5463-7. DOI: 10.1073/pnas.74.12.5463

[41] Prober JM, Trainor GL, Dam RJ, Hobbs FW, Robertson CW, Zagursky RJ, Cocuzza AJ, Jensen MA, Baumeister K. A system for rapid DNA sequencing with fluorescent

chain-terminating dideoxynucleotides. Science. 1987;238:336-41. DOI:10.1126/science. 2443975

[42] Briggs SP. Plant genomics: more than food for thought. Proceedings of the National Academy of Sciences USA. 1998;95:1986-8. DOI:10.1073/pnas.95.5.1986

[43] Ren B, Robert F, Wyrick JJ, Aparicio O, Jennings EG, Simon I, Zeitlinger J,Schreiber J, Hannett N, Kanin E, Volkert TL, Wilson CJ, Bell SP, Young RA. Genome-wide location and function of DNA binding proteins. Science. 2000;290:2306-9. DOI:10.1126/science.290.5500.2306

[44] Morozova O, Marra MA. Applications of next-generation sequencing technologies in functional genomics. Genomics. 2008;92:255-64. DOI:10.1016/j.ygeno.2008.07.001

[45] Leushkin EV, Sutormin RA, Nabieva ER, Penin AA, Kondrashov AS, Logacheva MD. The miniature genome of a carnivorous plant Genlisea aurea contains a low number of genes and short non-coding sequences. BMC Genomics. 2013;14:476. DOI: 10.1186/1471-2164-14-476

[46] Ibarra-Laclette E, Lyons E, Hernández-Guzmán G, Pérez-Torres CA,Carretero-Paulet L, Chang TH, Lan T, Welch AJ, Juárez MJ, Simpson J, Fernández-Cortés A, Arteaga-Vázquez M, Góngora-Castillo E, Acevedo-Hernández G,Schuster SC, Himmelbauer H, Minoche AE, Xu S, Lynch M, Oropeza-Aburto A,Cervantes-Pérez SA, de Jesús Ortega-Estrada M, Cervantes-Luevano JI, Michael TP, Mockler T, Bryant D, Herrera-Estrella A, Albert VA, Herrera-Estrella L.Architecture and evolution of a minute plant genome. Nature. 2013;498:94-8. DOI: 10.1038/nature12132

[47] Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin YC, Scofield DG, Vezzi F, Delhomme N, Giacomello S, Alexeyenko A, Vicedomini R, Sahlin K, Sherwood E, Elfstrand M, Gramzow L, Holmberg K, Hällman J, Keech O, Klasson L, Koriabine M, Kucukoglu M, Käller M, Luthman J, Lysholm F, Niittylä T, Olson A, Rilakovic N, Ritland C, Rosselló JA, Sena J, Svensson T, Talavera-López C, Theißen G, Tuominen H, Vanneste K, Wu ZQ, Zhang B, Zerbe P, Arvestad L, Bhalerao R, Bohlmann J, Bousquet J, Garcia Gil R, Hvidsten TR, de Jong P, MacKay J, Morgante M, Ritland K, Sundberg B, Thompson SL, Van de Peer Y, Andersson B, Nilsson O, Ingvarsson PK, Lundeberg J, Jansson S. The Norway spruce genome sequence and conifer genome evolution. Nature. 2013;497:579-84. DOI: 10.1038/nature12211

[48] Birol I, Raymond A, Jackman SD, Pleasance S, Coope R, Taylor GA, Yuen MM,Keeling CI, Brand D, Vandervalk BP, Kirk H, Pandoh P, Moore RA, Zhao Y, Mungall AJ, Jaquish B, Yanchuk A, Ritland C, Boyle B, Bousquet J, Ritland K, Mackay J, Bohlmann J, Jones SJ. Assembling the 20 Gb white spruce (Picea glauca) genome from whole-genome shotgun sequencing data. Bioinformatics. 2013;29:1492-7. DOI: 10.1093/bioinformatics/btt178

[49] Neale DB, Wegrzyn JL, Stevens KA, Zimin AV, Puiu D, Crepeau MW, Cardeno C, Koriabine M, Holtz-Morris AE, Liechty JD, Martínez-García PJ, Vasquez-Gross HA, Lin BY, Zieve JJ, Dougherty WM, Fuentes-Soriano S, Wu LS, Gilbert D, Marçais G,

Roberts M, Holt C, Yandell M, Davis JM, Smith KE, Dean JF, Lorenz WW, Whetten RW, Sederoff R, Wheeler N, McGuire PE, Main D, Loopstra CA, Mockaitis K, deJong PJ, Yorke JA, Salzberg SL, Langley CH. Decoding the massive genome of loblolly pine using haploid DNA and novel assembly strategies. Genome Biol. 2014;15:R59. DOI: 10.1186/gb-2014-15-3-r59

[50] International Wheat Genome Sequencing Consortium (IWGSC). A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. Science. 2014;345:1251788. DOI: 10.1126/science.1251788

[51] List of sequenced plant samples [Internet]. 2015. Available from: http://www.onekp.com/samples/list.php [Accessed: 2015-11-10]

[52] Matasci N, Hung LH, Yan Z, Carpenter EJ, Wickett NJ, Mirarab S, Nguyen N, Warnow T, Ayyampalayam S, Barker M, Burleigh JG, Gitzendanner MA, Wafula E, Der JP, dePamphilis CW, Roure B, Philippe H, Ruhfel BR, Miles NW, Graham SW, Mathews S, Surek B, Melkonian M, Soltis DE, Soltis PS, Rothfels C, Pokorny L, Shaw JA, DeGironimo L, Stevenson DW, Villarreal JC, Chen T, Kutchan TM, Rolf M, Baucom RS, Deyholos MK, Samudrala R, Tian Z, Wu X, Sun X, Zhang Y, Wang J, Leebens-Mack J, Wong GK. Data access for the 1,000 Plants (1KP) project. Gigascience. 2014;3:17. DOI: 10.1186/2047-217X-3-17

[53] Abdurakhmonov IY, Abdukarimov A. Application of association mapping to understanding the genetic diversity of plant germplasm resources. International Journal of Plant Genomics. 2008:574927. DOI:10.1155/2008/574927

[54] Kumpatla SP, Buyyarapu R, Abdurakhmonov IY, Mammadov JA. Genomics-assisted plant breeding in the 21st century: technological advances and progress. In: Abdurakhmonov IY, editor. Plant Breeding. Rijeka: InTech; 2012. p 134-184. DOI: 10.5772/37458

[55] Meuwissen TH, Hayes BJ, Goddard ME. Prediction of total genetic value using genome-wide dense marker maps. Genetics. 2001;157:1819-29

[56] Cros D, Denis M, Bouvet JM, Sánchez L. Long-term genomic selection for heterosis without dominance in multiplicative traits: case study of bunch production in oil palm. BMC Genomics. 2015;16:651. DOI: 10.1186/s12864-015-1866-9

[57] Longin CF, Mi X, Würschum T. Genomic selection in wheat: optimum allocation of test resources and comparison of breeding strategies for line and hybrid breeding. Theor Appl Genet. 2015;128:1297-306. DOI:10.1007/s00122-015-2505-1

[58] Iwata H, Ebana K, Uga Y, Hayashi T. Genomic prediction of biological shape: elliptic Fourier analysis and kernel partial least squares (PLS) regression applied to grain shape prediction in rice (Oryza sativa L.). PLoS One. 2015;10:e0120610. DOI: 10.1371/journal.pone.0120610

[59] Spindel J, Begum H, Akdemir D, Virk P, Collard B, Redoña E, Atlin G, Jannink JL, McCouch SR. Genomic selection and association mapping in rice (Oryza sativa): ef-

fect of trait genetic architecture, training population composition, marker number and statistical model on accuracy of rice genomic selection in elite, tropical rice breeding lines. PLoS Genet. 2015;1:e1004982. DOI: 10.1371/journal.pgen.1004982

[60] Cros D, Denis M, Sánchez L, Cochard B, Flori A, Durand-Gasselin T, Nouy B, Omoré A, Pomiès V, Riou V, Suryana E, Bouvet JM. Genomic selection prediction accuracy in a perennial crop: case study of oil palm (Elaeis guineensis Jacq.). Theor Appl Genet. 2015;128:397 410. DOI: 10.1007/s00122-014-2439-z

[61] Beaulieu J, Doerksen TK, MacKay J, Rainville A, Bousquet J. Genomic selection accuracies within and between environments and small breeding groups in white spruce. BMC Genomics. 2014;15:1048. DOI: 10.1186/1471-2164-15-1048

[62] Lipka AE, Lu F, Cherney JH, Buckler ES, Casler MD, Costich DE. Accelerating the switchgrass (Panicum virgatum L.) breeding cycle using genomic selection approaches. PLoS One. 2014;9:e112227. DOI: 10.1371/journal.pone.0112227

[63] Peleman JD, van der Voort JR. Breeding by design. Trends Plant Sci. 2003;8:330-4. DOI:10.1016/S1360-1385(03)00134-1

[64] Andersen SB. Virtual breeding. In: Abdurakhmonov IY, editor. Plant Breeding. Rijeka: InTech; 2012. p. 107-128. DOI: 10.5772/29802

[65] Ricroch AE, Henard-Damave MC. Next biotech plants: new traits, crops, developers and technologies for addressing global challenges. Crit Rev Biotechnol. 2015: Early online:1-16. DOI:10.3109/07388551.2015.1004521

[66] Sauer NJ, Mozoruk J, Miller RB, Warburg ZJ, Walker KA, Beetham PR, Schöpke CR,Gocal GF. Oligonucleotide-directed mutagenesis for precision gene editing. Plant Biotechnol J. 2015 (in press). DOI: 10.1111/pbi.12496

[67] McMahon MA, Rahdar M, Porteus M. Gene editing: not just for translation anymore. Nat Meth. 2011;9:28-31. DOI:10.1038/nmeth.1811

[68] Esvelt, KM.; Wang, HH. Genome-scale engineering for systems and synthetic biology. Mol Syst Biol. 2014;9:641. DOI:10.1038/msb.2012.66

[69] Tan WS; Carlson DF; Walton MW, Fahrenkrug SC; Hackett PB. Precision editing of large animal genomes. Adv Genet 2012;80:37–97. DOI:10.1016/B978-0-12-404742-6.00002-8

[70] Puchta H, Fauser F. Gene targeting in plants: 25 years later. Int. J. Dev. Biol 2013;57:629–637. DOI:10.1387/ijdb.130194hp

[71] Saurabh S, Vidyarthi AS, Prasad D. RNA interference: concept to reality in crop improvement. Planta. 2014;239: 543–564. DOI:10.1007/s00425-013-2019-5

[72] Sunilkumar G, Campbell L, Puckhaber L; Stipanovic R, Rathore K. Engineering cottonseed for use in human nutrition by tissue-specific reduction of toxic gossypol. Proc Natl Acad Sci USA.2006:103: 18054–9. DOI:10.1073/pnas.0605389103

[73] Siritunga D, Sayre R. Generation of cyanogen-free transgenic cassava. Planta. 2003;217:367–73. DOI:10.1007/s00425-003-1005-8

[74] Le L, Lorenz Y, Scheurer, S, Fötisch K, Enrique E, Bartra J, Biemelt S, Vieths S, Sonnewald U. Design of tomato fruits with reduced allergenicity by dsRNAi-mediated inhibition of ns-LTP (Lyc e 3) expression. Plant Biotechnol J. 2006;4:231–42. DOI: 10.1111/j.1467-7652.2005.00175.x

[75] Niggeweg R, Michael A, Martin C.Engineering plants with increased levels of the antioxidant chlorogenic acid. Nat Biotechnol. 2004;22:746–54. DOI:10.1038/nbt966

[76] Zadeh A, Foster G. Transgenic resistance to tobacco ringspot virus. Acta Virol. 2004;48:145–152

[77] Mao YB, Cai WJ, Wang JW, Hong GJ, Tao XY, Wang LJ, Huang YP, Chen XY. Silencing a cotton bollworm P450 monooxygenase gene by plant-mediated RNAi impairs larval tolerance of gossypol. Nat Biotechnol. 2007;25:1307-1313. DOI: 10.1038/nbt1352

[78] Heinemann JA, Agapito-Tenfen SZ, Carman JA. A comparative evaluation of the regulation of GM crops or products containing dsRNA and suggested improvements to risk assessments. Environ Int. 2013;55:43-55. DOI: 10.1016/j.envint.2013.02.010

[79] Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD. Genome editing with engineered zinc finger nucleases. Nat Rev Genet. 2010;11:636-646. DOI:10.1038/nrg2842

[80] Townsend JA, Wright DA, Winfrey RJ, Fu F, Maeder ML, Joung JK, Voytas DF. High-frequency modification of plant genes using engineered zinc-finger nucleases. Nature. 2009;459:442-445. DOI:10.1038/nature07845

[81] Zhang F, Maeder ML, Unger-Wallace E, Hoshaw JP, Reyon D, Christian M, Li X, Pierick CJ, Dobbs D, Peterson T, Joung JK, Voytas DF. High frequency targeted mutagenesis in Arabidopsis thaliana using zinc finger nucleases. Proceedings of the National Academy of Sciences USA. 2010;107:12028-33. doi: 10.1073/pnas.0914991107

[82] Osakabe K, Osakabe Y, Toki S. Site-directed mutagenesis in Arabidopsis using custom-designed zinc finger nucleases. Proceedings of the National Academy of Sciences USA. 2010;107:12034-9. DOI: 10.1073/pnas

[83] Shukla VK, Doyon Y, Miller JC, DeKelver RC, Moehle EA, Worden SE, Mitchell JC,Arnold NL, Gopalan S, Meng X, Choi VM, Rock JM, Wu YY, Katibah GE, Zhifang G,McCaskill D, Simpson MA, Blakeslee B, Greenwalt SA, Butler HJ, Hinkley SJ, Zhang L, Rebar EJ, Gregory PD, Urnov FD. Precise genome modification in the crop species Zea mays using zinc-finger nucleases. Nature. 2009;459:437-41. DOI: 10.1038/nature07992

[84] Forner J, Pfeiffer A, Langenecker T, Manavella PA, Lohmann JU. Germline-transmitted genome editing in Arabidopsis thaliana Using TAL-effector-nucleases. PLoS One. 2015;10:e0121056. DOI: 10.1371/journal.pone.0121056

[85] Char SN, Unger-Wallace E, Frame B, Briggs SA, Main M, Spalding MH, Vollbrecht E, Wang K, Yang B. Heritable site-specific mutagenesis using TALENs in maize. Plant Biotechnol J. 2015;13:1002-10. DOI: 10.1111/pbi.12344

[86] Zhang H, Gou F, Zhang J, Liu W, Li Q, Mao Y, Botella JR, Zhu JK.TALEN-mediated targeted mutagenesis produces a large variety of heritable mutations in rice. Plant Biotechnol J. 2015 (in press). DOI: 10.1111/pbi.12372

[87] Wang M, Liu Y, Zhang C, Liu J, Liu X, Wang L, Wang W, Chen H, Wei C, Ye X, Li X, Tu J. Gene editing by co-transformation of TALEN and chimeric RNA/DNA oligonucleotides on the rice OsEPSPS gene and the inheritance of mutations. PLoS One. 2015;10:e0122755. DOI: 10.1371/journal.pone.0122755

[88] Shan Q, Zhang Y, Chen K, Zhang K, Gao C. Creation of fragrant rice by targeted knockout of the OsBADH2 gene using TALEN technology. Plant Biotechnol J. 2015;13:791-800. DOI: 10.1111/pbi.12312

[89] Nicolia A, Proux-Wéra E, Åhman I, Onkokesung N, Andersson M, Andreasson E, ZhuLH. Targeted gene mutation in tetraploid potato through transient TALEN expression in protoplasts. J Biotechnol. 2015;204:17-24. DOI:10.1016/j.jbiotec.2015.03.021

[90] Clasen BM, Stoddard TJ, Luo S, Demorest ZL, Li J, Cedrone F, Tibebu R, Davison S, Ray EE, Daulhac A, Coffman A, Yabandith A, Retterath A, Haun W, Baltes NJ, Mathis L, Voytas DF, Zhang F. Improving cold storage and processing traits in potato through targeted gene knockout. Plant Biotechnol J. 2015 (in press). DOI: 10.1111/pbi.12370

[91] Budhagatapalli N, Rutten T, Gurushidze M, Kumlehn J, Hensel G. Targeted Modification of Gene Function Exploiting Homology-Directed Repair of TALEN-Mediated Double-Strand Breaks in Barley. G3 (Bethesda). 2015;5:1857-63. DOI: 10.1534/g3.115.018762

[92] Ilardi V, Tavazza M. Biotechnological strategies and tools for Plum pox virus resistance: trans-, intra-, cis-genesis, and beyond. Front Plant Sci. 2015;6:379. DOI: 10.3389/fpls.2015.00379

[93] Jain M. Function genomics of abiotic stress tolerance in plants: a CRISPR approach. Front Plant Sci. 2015;6:375. DOI: 10.3389/fpls.2015.00375

[94] Wolt JD, Wang K, Yang B. The Regulatory Status of Genome-edited Crops. Plant Biotechnol J. 2015 (in press). DOI: 10.1111/pbi.12444

[95] Chaudhary B, Flagel L, Stupar RM, Udall JA, Verma N, Springer NM, Wendel JF. Reciprocal silencing, transcriptional bias and functional divergence of homeologs in polyploid cotton (*Gossypium*). Genetics. 2009;182:503-17. DOI:10.1534/genetics.109.102608

[96] Renny-Byfield S, Gallagher JP, Grover CE, Szadkowski E, Page JT, Udall JA, Wang X, Paterson AH, Wendel JF. Ancient gene duplicates in Gossypium (cotton) exhibit

near-complete expression divergence. Genome Biol Evol. 2014;6:559-71. DOI: 10.1093/gbe/evu037

[97] Yoo MJ, Szadkowski E, Wendel JF. Homoeolog expression bias and expression level dominance in allopolyploid cotton. Heredity (Edinb). 2013;110:171-80. DOI: 10.1038/hdy.2012.94

[98] Sinha NR. Plant developmental biology in the post-genomic era. Front Plant Sci. 2011;2:11. DOI:10.3389/fpls.2011.00011

[99] Puchta, H. & Hohn, B., Breaking news: Plants mutate right on target. Proc Natl Acad Sci U S A. 2010;107:11657-8. DOI: 10.1073/pnas

[100] Moose SP, Mumm RH. Molecular plant breeding as the foundation for 21st century-crop improvement. Plant Physiol. 2008;147:969-77. DOI:10.1104/pp.108.118232

Sequencing and Assembling Plant Genomes

Integration of Next-generation Sequencing Technologies with Comparative Genomics in Cereals

Thandeka N. Sikhakhane, Sandiswa Figlan, Learnmore Mwadzingeni, Rodomiro Ortiz and Toi J. Tsilo

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/61763

Abstract

Cereals are the major sources of calories worldwide. Their production should be high to achieve food security, despite the projected increase in global population. Genomics research may enhance cereal productivity. Genomics immensely benefits from robust next-generation sequencing (NGS) techniques, which produce vast amounts of sequence data in a time and cost-efficient way. Research has demonstrated that gene sequences among closely related species that share common ancestry have remained well conserved over millions of years of evolution. Comparative genomics allows for comparison of genome sequences across different species, with the implication that genomes with large sizes can be investigated using closely related species with smaller genomes. This offers prospects of studying genes in a single species and, in turn, gaining information on their functions in other related species. Comparative genomics is expected to provide invaluable information on the control of gene function in complex cereal genomes, and also in designing molecular markers across related species. This chapter discusses advances in sequencing technologies, their application in cereal genomics and their potential contribution to the understanding of the relationships between the different cereal genomes and their phenotypes.

Keywords: Bioinformatics, Cereals, Comparative genomics, Next-generation sequencing, Synteny

1. Introduction

Significant limitations to cereal crop production and productivity pose a threat to global food security since these crops are the main sources of calories that support the ever-growing human population. Despite the significant progress that has been made in the improvement of edible

yield through classical breeding techniques, the current rates of increase in grain yield in several major cereal crops are still too slow to catch up with the increasing demand of the growing population [1, 2]. This is likely to get worse according to the projected climate change scenarios [3], as it also affects biotic stresses such as pests, diseases and weeds, and abiotic stresses including drought, extreme temperatures, salinity and nutrient deficiencies [4-6]. Although there are various strategies to cope with these constraints, Kole [7] suggested the use of genomics-assisted breeding as an effective and economic strategy.

Despite the sustainability of breeding resilient crops, there are still several genomic constraints to genome-based selection and stress resistance improvement, particularly for multigenic traits. A poor understanding of the genetic basis and the regulatory mechanisms of various stresses is among the major challenges for successful genetic manipulation through gene introgression, gene pyramiding, gene stacking or gene silencing. Additionally, more diagnostic genetic markers are necessary to improve the current limited success in marker application in both foreground and background selection. These challenges are related to the fact that genomes of some cereal crops are not yet fully sequenced and annotated, either because the crops have been under-researched or the genomes are huge and structurally complex. For instance, the hexaploid wheat (*Triticum aestivum*) genome is the largest (about 17 billion nucleotides) among cultivated cereals, and is multifaceted by repetitive DNA sequences [8]. Furthermore, dissection of the genetic and regulatory mechanisms of host plant resistance is complicated because most traits of interest are multigenic and thus influenced by several genes with additive and nonadditive gene effects. Hence, tools that detect the genetic variation at the genome sequence level allow all genes controlling particular traits to be investigated for various genetic applications to realize phenotypic gains from genetic manipulation.

Enhanced application of next-generation sequencing (NGS) techniques in cereal crops is revolutionizing and speeding up plant breeding. The advances that have been made so far in the use of NGS, particularly with the human genome in the field of medicine, and on various model crops through plant biotechnology, envisions the following in cereals and other crops: first, complete sequencing of small and less complex plant genomes is increasingly becoming possible as costs have dropped significantly and more sequences are being generated in a shorter time than before. Secondly, the genetic mechanisms of particular traits in huge and complex plant genomes can now be investigated using small and less complex genomes of related plants sharing conserved regions through comparative genomics. This will potentially identify genes or quantitative trait loci (QTL) and putative single nucleotide polymorphism (SNP) markers for genome-wide association mapping and annotation of genomes. This chapter discusses the advances made in improving sequencing technologies and how these advances can assist in generating complete sequences for the improvement of genome-aided selection. This will also assist in identifying the unique sequences responsible for the major differences existing among cereals.

2. The need for high-throughput genome and transcriptome sequencing

Since the discovery of the DNA molecule by Friedrich Miescher in 1869 [9], and the subsequent exposition of its double-helical structure by Watson and Crick in 1953, significant knowledge

has been gained on the flow of genetic information. Understanding how this genetic informa-tion influences the phenotype (trait) of interest has, however, remained a challenge. This is mainly because the overall instruction contributing to the phenotype is not restricted to the coding region but is also influenced by some posttranscriptional modifications controlled by noncoding DNA [10-12]. Also, multigenic traits are influenced by complex interactions of alleles at different loci, having major or minor influence [13]. These, together with differential genotype-by-environment interactions, add to the structural and functional complexity of most cereal genomes that are multifaceted by repetitive DNA sequences, transposable elements and polyploid genomes, as in the case of wheat and finger millet (*Eleusine coracana*) [8, 14]. Whole genome and transcriptome sequencing therefore become a necessity so that all the genomic and transcriptomic variation can be detected. NGS and various 'omic' technolo-gies, including genomics, transcriptomics, proteomics, metabolomics and phenomics, offer prospects towards whole-genome annotation; particularly in cereals that have small and less complex genomes. This will simplify comparative genomics and evolutionary genetic re-search, which will enhance the manipulation and exploitation of important genes for cereal improvement.

NGS technologies are one of the available tools that can produce complete sequences for diverse research at the DNA and RNA level within and across species. Firstly, this will make it easy to obtain the entire DNA, coding and noncoding regions. Secondly, this will simplify studies on the whole transcriptome, including RNAs involved in protein synthesis such as the messenger, ribosomal, signal recognition particle, transfer and transfer-messenger RNAs and other RNAs involved in posttranscriptional modifications, such as small RNAs [15]. Quanti-fication of such transcripts through NGS under various stress conditions will precisely determine the levels of gene expression within and across different species.

3. Advances in sequencing technologies

Since the pioneering of genome sequencing through technologies such as Sanger sequencing [16], significant advances have been made to resolve the limitations of the early technologies. This has seen the development of more sophisticated sequencing technologies that allow *de novo* genome sequencing, generating vast amounts of data in a short period at low costs. Table 1 summarizes the advances made in sequencing technology development, from the advent of the chain termination sequencing [16], to prominent NGS technologies including Roche/454 sequencing [17], Illumina (Solexa) sequencing [18], sequencing by oligonucleotide ligation and detection (SOLiD) [19], the single molecule sequence pioneered by Helicos Biosciences [20] and Ion Torrent sequencing [21]. These technological advances are instrumental in whole-genome research and are expected to simplify comparative genomics within species and across distantly related cereals and grasses. Several modifications are available for each of these technologies and fine-tuned protocols are constantly being developed to address some of the current limitations.

Although NGS technologies have enormous prospective benefits, they come with their own limitations that need to be addressed to realize their full potential. Key among these drawbacks are the bioinformatic and computational challenges related to storage, image analysis, base

calling and integration of the large amounts of data that are generated in several terabytes per day. Apparently, the large amount of sequence data that is being generated on a daily basis in cereal genomics cannot be transformed into information that is useful for the detection of important genomic variants within and among species or in identifying genes that are differentially expressed under particular stress conditions. Hence, investment in computational and high-throughput bioinformatic equipment and human resources and combining the various NGS technologies will allow the data generated using different NGS techniques by various laboratories to be related and used to build onto each other. Unlike traditional marker technologies, NGS is currently dissociated from phenomics, yet it should be complementary to high-throughput phenotyping in order to relate sequence variations to traits of interest for progressive discoveries through genome-wide association mapping, particularly for multigenic traits like adaptation to drought in complex cereal genomes [22]. Additionally, NGS technologies are still associated with high error rates [23] and short read lengths that limit data analysis accuracy. This further confuses detection and distinction of sequence variations including large amounts of duplications, deletions, inversions and chromosomal rearrangements that characterize cereal genomes.

Technologies (Developer)	Year	Sequencing chemistry	Throughput	Read length	References
Sanger sequencing (Frederick Sanger and team)	1977	Involves DNA polymerase based selective amplicon-termination of in vitro DNA replication by radioactively or fluorescently labeled di-deoxynucleotide triphosphates, followed by electrophoresis and UV or X-ray spectra detection of DNA sequences. Major limitations of the Sanger technique Since the technique relies on cloning vectors, there is potential for a mix up of the target sequences with some DNA portions from the clonal vector. Additionally, it requires a lot of labor and space since multiplexing is not possible.	Up to 84 Kb per about 3 hr run	Up to 1000 bp	[16]
Roche/454 (Life Sciences)	2004	First NGS technique This is a sequencing by synthesis (SBS) technique where DNA fragments attached to adapters annealed to beads are PCR amplified using adapter specific primers. Addition of each dNTP is associated with the release of a pyrophosphate, which is converted to ATP energy used to produce an optical signal (light). The light allows reading of the beads to which the dNTP is added, hence deducing the sequence (Pyrosequencing).	700 Mb of sequence data per 23 hr run	Up to 1,000 bp	[17]

Technologies (Developer)	Year	Sequencing chemistry	Throughput	Read length	References
Illumina/ Solexa (Illumina -Inc)	2006	This sequencing by synthesis technology uses dNTPs with reversible dye-terminators ensuring that DNA polymerase enzyme adds only one base to a growing DNA strand. The terminators are removed after the images of the four dNTPs added to the growing sequence are recorded, and the cycle is repeated.	20 to 130 Gb in 15 to 30 hr run time	Up to 3000 bp	[18]
ABI SOLiD (Life Technologies)	2006	Clonal bead populations prepared from a library of same species of DNA fragments each with a universal P1 adapters are attached on the surface of magnetic beads. On the universal adapters, primers are hybridized, on which a set of four dye-labeled di-base probes compete for ligation. A series of ligation cycles is followed by cleavage of the extension product then the template is reset for the next cycle of ligation by annealing a primer complimentary to the next adapter.	At least 20 Gb in about 3.5 days run time	Up to 50 bp	[19]
Helicos (Helicos Biosciences)	2009	This is a single molecule fluorescent sequencing technique which achieves direct DNA or RNA sequencing without amplification through imaging light emitting single molecules corresponding to each nucleotide base. DNA sequencing is achieved through an imaging system. The technology identifies the exact sequence of a piece of DNA and does not require PCR amplification, thus have reduced amplification bias.	21 to 35 Gb per 8 days run	35 bp average	[20]
Ion Torrent (Life Technologies)	2010	This sequencing by synthesis technique involves detection of a pH change caused by hydrogen ions released when a dNTP complementary to the leading unpaired template nucleotide is added to the growing strand. The electrical pulses transmitted to a computer in this process are subsequently translated into a DNA sequence	Up to 2 Gb per 2.3 to 7.3 hr run	35 to 400 bp (average: 200 bp)	[21]

Table 1. Evolution of next-generation sequencing technologies.

4. Application of next-generation sequencing in cereal biotechnology

Among the major cereals, the relatively small rice (*Oryza sativa*) genome (~389 Mb) has long been fully sequenced by the International Rice Genome Project [24]. Kawahara [25] recently demonstrated, however, the robustness of NGS technologies by revising the rice genome using the Illumina and Roche 454 pyrosequencing platforms. Their study noted some errors in the initial assembly. This research provides sufficient evidence that high quality and validated reference genomes can be produced among most cereals through resequencing using NGS technologies. Also, a recent whole genome-wide study of the hexaploid wheat genome (~17 Gb) using the Roche/454 pyrosequencing technology reviewed the capacity of NGS technologies to resequence huge and complex genomes and to identify SNPs for dissection of quantitative traits [26]. Similarly, Illumina sequencing was recently used to quantify the transposable element (TE) content in the complex maize *(Zea mays)* genome (~2.3 Gb) [27] and to estimate their potential contribution to the genome size differences between the cultivated species and its close relative, *Zea luxurians* [28]. The latter also reported high proportions of conserved TE families between the two species, revealing the potential of NGS technologies to enhance evolutionary and comparative genomic studies. Other major cereals whose genomes have been sequenced and are expected to further benefit from NGS technologies include barley *(Hordeum vulgare)* (~5.1 Gb) [29] and sorghum *(Sorghum bicolor)* (~730 Mb) [30].

Minor and under-researched cereals such as the allotetraploid finger millet *(Eleusine coracana)* —which has a genome size of about 1.76 Gb [31]—and the diploids, pearl millet *(Pennisetum glaucum)*—with a genome size of about 4.6 Gb [32]—and tef *(Eragrostis tef)*—with a 714 to 733 Mb genome [33]—have not received much benefit from NGS technologies. However, these crops are expected to benefit from the African Orphan Crops Consortium that has the mandate to use the latest scientific equipment and techniques to sequence, assemble and annotate genomes of under-researched crops [34]. These minor crops are renowned for their adaptation to various biotic and abiotic stresses, particularly drought. Thus, sequencing or resequencing their genomes will potentially expose huge amounts of relevant genetic information for cereal improvement. NGS technologies will have great application in comparing genomic features of cereal crops through comparative genomic research.

5. Comparative genomics in cereal crops

Core questions unanswered with traditional cereal biotechnology approaches include: (1) What are the genetic foundations that underlie the similarities between different grass species or individuals within a species? (2) What are the genetic variations responsible for the detected phenotypic differences? Comparative genomics is the branch of biology in which DNA sequence information from genomes of different life forms are compared in an effort to directly answer these questions. It was founded mainly on various ideas. Firstly, comprehensive analysis and comparison of whole genomes can uncover the essentially conserved and the important variable components of any set of genomes [35]. Secondly, differences in genome sequence (genotype) contribute to differences in genome function and therefore explain differences between phenotypic traits [36]. The application of comparative genomic informa-

tion on various plants including cereals has, however, been a challenge previously because of the large genome sizes of most species, which are complicated by high rates of structural rearrangements mainly due to transposable elements, duplications and inversions [35], as listed in Table 2.

Species	Clade (Subfamily, Tribe)	Ploidy level	Genome size	Repetitive DNA and retrotransposon content	References
Rice (*Oryza sativa*)	Ehrhartoideae, Oryzeae	$2n=2x=24$	420 to 460 Mb	~35%	[41]
Sorghum (*Sorghum bicolor*)	Panicoideae, Andropogoneae	$2n=2x=20$	~730 Mb	~61%	[30]
Pearl millet (*Pennisetum glaucum*)	Panicoideae, Paniceae	$2n=2x=14$	~4.6 Gb		[32]
Finger millet (*Eleusine coracana*)	Chloridoideae, Eragrostideae	$2n=4x=40$	3.3 to 3.8 Gb		[31]
Maize (*Zea mays*)	Panicoideae, Andropogoneae	$2n=2x=20$	~2.3 Gb	~78%	[27]
Barley (*Hordeum vulgare*)	Pooideae, Triticeae	$2n=2x=14$	~5.1 Gb	~76%	[29]
Rye (*Secale cereale*)	Pooideae, Triticeae	$2n=2x=14$	~7.9 Mb	~92%	No reference available
Bread wheat (*Triticum aestivum*)	Pooideae, Triticeae	$2n=6x=42$	~17 Gb	~80%	[8]
Oat (*Avena sativa*)	Pooideae, Triticeae	$2n=6x=42$	~11.3 Gb		No reference available
Tef (*Eragrostis tef*)	Chloridoideae,	$2n=4x=40$	~672 Mb		[33]

Table 2. Genome size, structure and genomic resources of major cereal species.

The application of comparative genomics for crop improvement has evolved over time. In the grass family, significant research provided remarkable and comprehensive datasets demonstrating high degree of collinearity or synteny among genomes at chromosome (macro) and gene (micro) levels [37, 38]. *Synteny*, from the Greek *syn* (together with) and *taenia* (ribbon), refers to loci contained within the same chromosome. Collinearity, on the other hand, refers to some degree of conservation of gene order between chromosomes of different species or between nonhomologous chromosomes of a single species [39]. A large number of sequences within the grass family has remained considerably conserved at the genome level over millions of years of evolution, irrespective of the differences in ploidy level, chromosome number and haploid DNA content [37]. This conservation of gene content and order at the megabase level makes it easy to use species with small genome sizes such as *Arabidopsis* and rice as model species for studying similar gene contents in other related species. Their applications include allele discovery, positional cloning, and comparative studies in related species [40]. There is, however, limited synteny and gene homology between *Arabidopsis* and rice, but an extensive collinearity between the latter and other grasses, thereby suggesting that rice is an appropriate

grass model species for cereal comparative genomics [41]. In this case, rice and purple false brome (*Brachypodium distachyon)* (genome size ~355 Mb), both of which are from the grass family, serve as functional model species for cereal comparative genomics owing to their small and fully sequenced genomes. Moreover, *Brachypodium* showed conservation of gene content and family structure with rice and sorghum [42]. A phylogenetic study carried out on seven grass species also revealed a close evolutionary relationship of *Brachypodium* with maize, barley and wheat based on 335 commonly shared sequences [43].

Microcollinearity has numerous interesting applications in cereal genome analysis including the transfer of genetic markers between species and the identification of candidate genes across species borders [44]. It is possible, due to such advances, to intensively study, decipher and understand the genetic makeup of the cereal genomes including those of rice, maize, wheat, barley and sorghum [30, 45-47]. Comparing the gene sequences of these cereal crops is the initial step towards understanding their morphological and functional similarities and differences. Comparative analysis research has been extended to the DNA sequence (micro) level, to allow the investigation of conservation of coding and noncoding regions as well as characterization of molecular mechanisms of genome evolution [38].

6. Several examples of macro- and microcollinearity in cereal crops

The advent of molecular markers and molecular mapping allowed researchers to conduct comparative mapping research, comparing gene orders and content of genes and markers along chromosomes of related species. The first research of large-scale restriction fragment length polymorphism (RFLP) mapping in several economically important crop genomes included the genomes of wheat, rice, maize, oat and barley. They are benchmarks for the discovery of collinearity in the grass family [44]. Hence, in the past, exploiting RFLPs to compare genomes was a valuable method as the markers made it possible to map, for the first time, a huge number of randomly distributed polymorphic loci in a single population and provided the foundation for efficient, whole-genome studies at the molecular level [48]. The application of RFLP technology in comparative genome analysis studies revealed that an extensive commonality in gene content and arrangement was a basic chromosomal property, thus prompting the idea that the genetic map could be used to tie all grasses into a single model system. This led to the construction of a consensus grass map based on 25 rice linkage blocks [37, 38]. The resolution of the genetic maps, however, proved to be very low with an average of one marker in every 5 to 10 centimorgans (cM), allowing the detection of only large rearrangements. The RFLP markers used to construct the maps were also low-copy, therefore limiting the detection of small deletions, inversions and whole or partial genome duplication events [49]. The use of RFLP markers for comparative mapping also had difficulty to assess orthologous (derived from a common ancestor by speciation) and paralogous (derived by duplication within one genome) relationships in gene families. Having these challenges associated with traditional genotyping, the NGS techniques discussed above are expected to advance comparative genomics because they provide actual DNA sequences that allow interspecies or intergeneric comparisons.

Traditional genome analyses have provided sufficient evidence that cereal genomes share conserved regions at either macro or micro levels. For example, a comparative genomics study on rice and maize indicated high levels of collinearity between the two genomes with some chromosomes or their arms—accounting for at least 67% of the two genomes—having almost similar gene order and sequences [46]. Similarly, large proportions of conserved regions between rice and wheat chromosomes were identified with major differences arising from chromosomal rearrangements [40, 45]. Conservation of about 24% of grass-specific gene orders have been reported in sorghum [30], including high collinearity with rice [50]. Thus, sorghum can also serve as a model species for cereal genomic studies due to its relatively small genome size and wide adaptability. High levels of microcollinearity have been demonstrated between chromosome 6 of rice and the telomeric regions of barley chromosome 1P, which further confirm the usefulness of mapping the small rice genome for map-based cloning of important genes in complex genomes [47]. Figures 1 and 2 illustrate the conservation of synteny and collinearity among different cereals by revealing the syntenic relationships between chromosomes of cereal crops. Furthermore, Figure 2B reveals that the 10 maize progenitor chromosomes and the 10 linkage groups of sorghum appear to be similar, thus exposing their evolutionary divergence from rice that could be their common ancestor before speciation [51]. The study of such evolutionary relationships and changes that occurred after cereals diverged from their progenitors will further be enhanced through comparative genomics integrated with NGS and next-next or third-generation sequencing techniques, which can generate more resolute physical maps. Availability of updated genome sequences will expose the multiple breaks in collinearity occurring in the genome compositions due to structural rearrangements caused by transposable elements, inversions, deletions and duplications. The macro- and microcollinearities described in this section are exposed by the observed phenotypic similarities that exist among different cereal species.
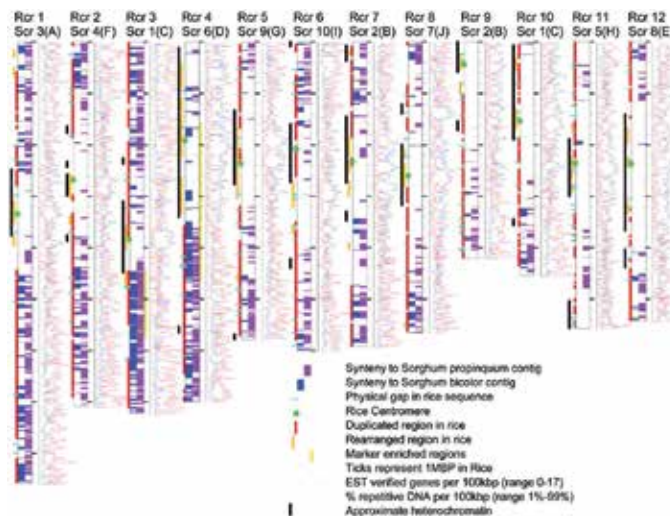
Source: Bowers, et al. [50].

Figure 1. Microsynteny conservation between sorghum and rice.
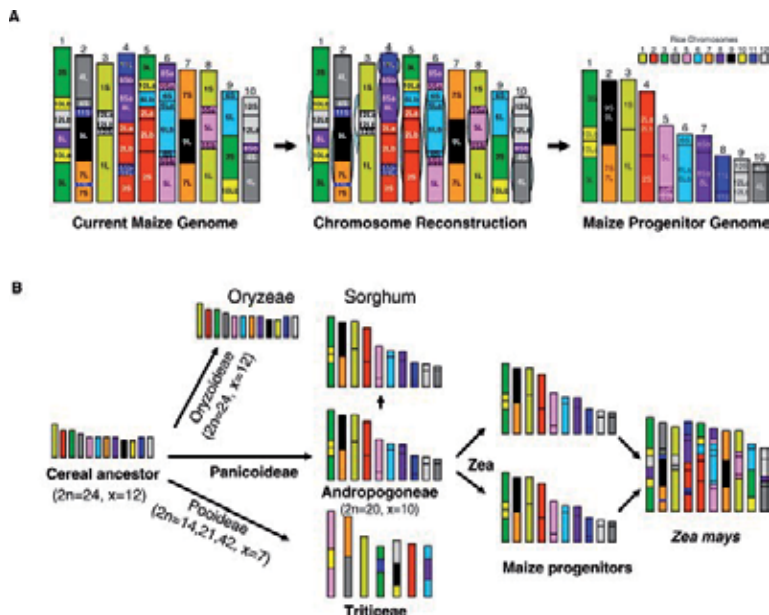
Figure 2. Conservation and changes in rice, maize, sorghum and wheat chromosomes during cereal speciation.

7. Phenotypic commonality in cereals

The conservation of synteny and collinearity of genes among cereals is highly attributed to the common phenotypic features or characteristics that are evidence that they share common ancestry, while their differences mainly stem from chromosomal rearrangements and polyploidization as shown in Figure 2. Their morphological similarity (Figure 3) also shows evidence that they share common ancestry. Based on phenotype alone, most also share similar rooting system, leaf venation, flowering habits, tillering, inflorescences, physiological behavior such as vernalization requirements, and adaptation to biotic and abiotic stresses. For example, some cereals are hosts of common diseases, as in the case of maize streak virus (MSV), wheat streak mosaic virus (WSMV) and rusts [52, 53], while others are nonhosts, as in the case of rice to rusts. The differences in phenotype and genome structure among all these species could be due to mutations, breaks in collinearity and loss of synteny that occurred in their genomes over millions of years. Such differences can be traced through comparative genomic analysis, particularly with the aid of high-throughput sequencing techniques. Likewise, the similarity in phenotype and genome structure could be due to sharing a common ancestry (Figures 2 and 3). This finding therefore reveals some phenotypes along with gene orders and sequences that have been conserved over millions of years.
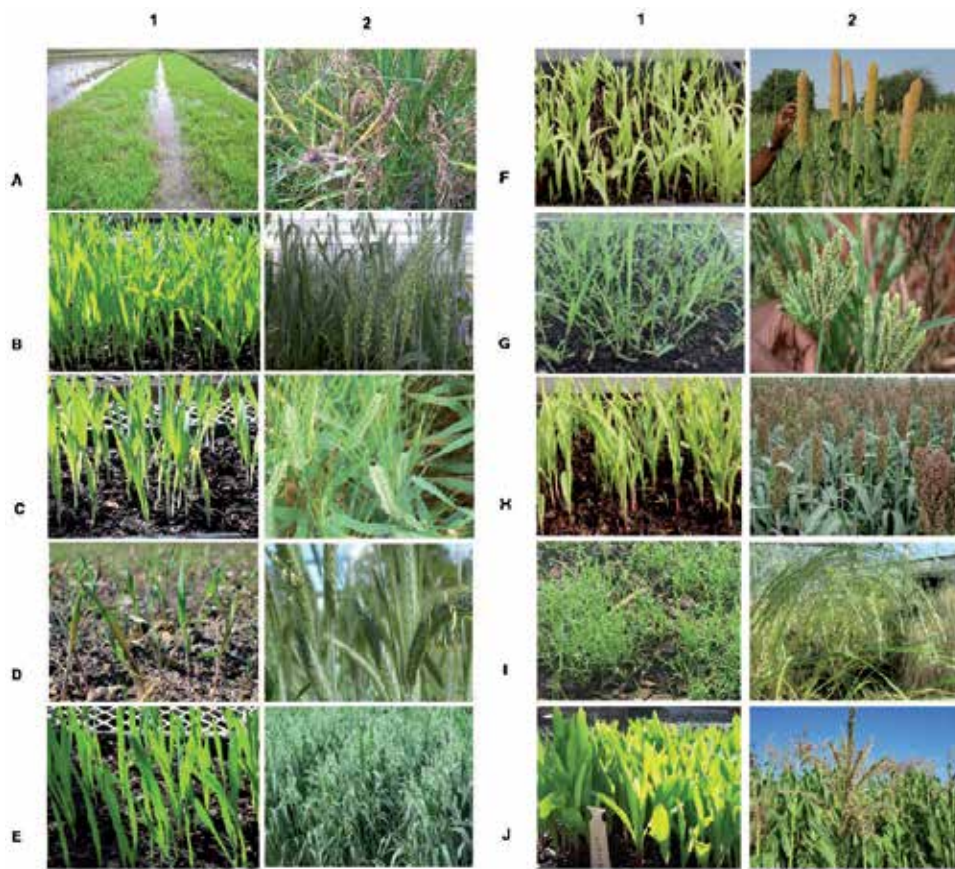
Figure 3. Phenotypic commonality in cereals. Similar seedlings: (A1) rice seedlings, (B1) wheat seedlings, (C1) barley seedlings, (D1) rye seedlings, (E1) oat seedlings, (F1) pearl millet seedlings, (G1) finger millet seedlings, (H1) sorghum seedlings, (I1) tef seedlings, (J1) maize seedlings. Similar heads: (A2) rice heads, (B2) wheat heads, (C2) barley heads, (D2) rye heads, (E2) oat heads, (F2) pearl millet heads, (G2) finger millet heads, (H2) sorghum heads, (I2) tef heads, (J2) maize heads.

8. Outlook

Plant species have highly conserved regions at DNA sequence level, whereas the bulk of the large genomes consist of repetitive DNA sequences, most of which are species-specific. Comparative genomics have opened new avenues for map-based positional cloning of genes encoding important traits on large and intricate genomes through investigating small and less complex genomes. In grasses, rice and *Brachypodium* have been identified as model species for such research since they have small and stable genomes. This, however, requires the integration of NGS techniques so that all the conserved and nonconserved regions can be fully sequenced and annotated with the aid of other "omic" technologies. Hence, the future of

comparative genomics studies in cereals will largely rely on cost-effective sequencing technologies along with computational systems that handle large numbers of sequences, thus allowing effective sequence comparisons across species of interest. The substantial evidence regarding a common ancestry of cereals—based on genome and morphological structures—led to the successful use of the genome sequence of one species to share a light on the function of that sequence in other related species. A wide adoption of this approach across different cereals will speed up gains and generate useful databases and datasets for effective cereal breeding. Furthermore, researchers will be able to use other widely adapted cereals like sorghum and some of the under-researched cereals as models for sequencing genes and alleles responsible for unique traits such as wide adaptation to stress-prone environments due to increased sequencing throughput. There is, however, a need to invest in advanced computational and bioinformatics tools to handle and analyze huge datasets that will be generated through these technology advances.

Acknowledgements

Author details

Thandeka N. Sikhakhane[1,2], Sandiswa Figlan[1], Learnmore Mwadzingeni[1], Rodomiro Ortiz[3] and Toi J. Tsilo[1,2,4*]

*Address all correspondence to: tsilot@arc.agric.za

1 Agricultural Research Council, Small Grain Institute, Bethlehem, South Africa

2 Department of Life and Consumer Sciences, University of South Africa, Pretoria, South Africa

3 Department of Plant Breeding, Swedish University of Agricultural Sciences, Alnarp, Sweden

4 Department of Plant Production, University of Venda, Thohoyandou, South Africa

References

[1] Ray DK, et al. Yield trends are insufficient to double global crop production by 2050. 2013. DOI: 10.1371/journal.pone.0066428

[2] Lobell DB and CB Field. Global scale climate–crop yield relationships and the impacts of recent warming. Environmental research letters, 2007. 2(1): p. 014002. DOI: 10.1088/1748-9326/2/1/014002

[3] Rosenzweig C, et al. Climate change and extreme weather events; implications for food production, plant diseases, and pests. Global change & human health, 2001. 2(2): p. 90-104.

[4] Waddington SR, et al. Getting the focus right: production constraints for six major food crops in Asian and African farming systems. Food security, 2010. 2(1): p. 27-48. DOI: 10.1007/s12571-010-0053-8

[5] Chakraborty S and AC Newton. Climate change, plant diseases and food security: an overview. Plant Pathology, 2011. 60(1): p. 2-14. DOI: 10.1111/j.1365-3059.2010.02411.x

[6] Duveiller E, RP Singh, and JM Nicol. The challenges of maintaining wheat productivity: pests, diseases, and potential epidemics. Euphytica, 2007. 157(3): p. 417-430. DOI: 10.1007/s10681-007-9380-z

[7] Kole C, et al. Application of genomics-assisted breeding for generation of climate resilient crops: progress and prospects. Frontiers in Plant Science, 2015. 6. DOI: 10.3389/fpls.2015.00563

[8] Paux E, et al. Characterizing the composition and evolution of homoeologous genomes in hexaploid wheat through BAC-end sequencing on chromosome 3B. The Plant Journal, 2006. 48(3): p. 463-474. DOI: 10.1111/j.1365-313X.2006.02891.x

[9] Dahm R. Friedrich Miescher and the discovery of DNA. Developmental biology, 2005. 278(2): p. 274-288. DOI: 10.1016/j.ydbio.2004.11.028

[10] Cogoni C and G Macino. Post-transcriptional gene silencing across kingdoms. Current opinion in genetics & development, 2000. 10(6): p. 638-643. DOI: 10.1016/S0959-437X(00)00134-9

[11] Depicker A and M Van Montagu. Post-transcriptional gene silencing in plants. Current opinion in cell biology, 1997. 9(3): p. 373-382. DOI: 0955067400900373

[12] Saze H, et al. DNA methylation in plants: relationship to small RNAs and histone modifications, and functions in transposon inactivation. Plant and Cell Physiology, 2012. 53(5): p. 766-784. DOI: 10.1093/pcp/pcs008

[13] Alonso-Blanco C and B Méndez-Vigo. Genetic architecture of naturally occurring quantitative traits in plants: an updated synthesis. Current opinion in plant biology, 2014. 18: p. 37-43. DOI: 10.1016/j.pbi.2014.01.002

[14]   Edwards D, J Batley, and RJ Snowdon. Accessing complex crop genomes with next-generation sequencing. Theoretical and Applied Genetics, 2013. 126(1): p. 1-11. DOI: 10.1007/s00122-012-1964-x

[15]   Taft RJ, et al. Non-coding RNAs: regulators of disease. The Journal of pathology, 2010. 220(2): p. 126-139. DOI: 10.1002/path.2638

[16]   Sanger F, S Nicklen, and AR Coulson. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences, 1977. 74(12): p. 5463-5467.

[17]   Roche. GS FLX+ System. Sanger-like read lengths - the power of next-gen throughput. Roche Diagnostics GmbH; Available from: http://454.com/downloads/GSFLXApplicationFlyer_FINALv2.pdf [Accessed: 2015-09-05].

[18]   Illumina. Next-Generation Sequencing 2015; Available from: http://www.illumina.com/ [Accessed: 2015-09-05].

[19]   TFS. SOLiD® Next-Generation Sequencing 2015; Available from: https://www.life-technologies.com/za/en/home/life-science/sequencing/next-generation-sequencing/solid-next-generation-sequencing.html [Accessed: 2015-09-05].

[20]   SeqLL. Helicos Genetic Analysis Aool 2014; Available from: http://seqll.com/ [Accessed: 210-09-05].

[21]   TFS. Ion Personal Genome Machine® (PGM™) System 2015; Available from: http://www.lifetechnologies.com/order/catalog/product/4462921 [Accessed: 2015-09-05].

[22]   Mwadzingeni L, et al. Breeding wheat for drought tolerance: Progress and technologies. Journal of Integrative Agriculture, 2015. DOI: 10.1016/s2095-3119(15)61102-9

[23]   Harismendy O, et al. Evaluation of next generation sequencing platforms for population targeted sequencing studies. Genome biol, 2009. 10(3): p. R32.

[24]   IRGSP. The map-based sequence of the rice genome. Nature, 2005. 436(7052): p. 793-800. DOI: 10.1038/nature03895

[25]   Kawahara Y, et al. Improvement of the Oryza sativa Nipponbare reference genome using next generation sequence and optical map data. Rice, 2013. 6(1): p. 4.

[26]   Brenchley R, et al. Analysis of the bread wheat genome using whole-genome shotgun sequencing. Nature, 2012. 491(7426): p. 705-710. DOI: 10.1038/nature11650

[27]   Schnable PS, et al. The B73 maize genome: complexity, diversity, and dynamics. science, 2009. 326(5956): p. 1112-1115. DOI: 10.1126/science.1178534

[28]   Tenaillon MI, et al. Genome size and transposable element content as determined by high-throughput sequencing in maize and Zea luxurians. Genome biology and evolution, 2011. 3: p. 219-229. DOI: 10.1093/gbe/evr008

[29]   IBGSC. A physical, genetic and functional sequence assembly of the barley genome. Nature, 2012. 491(7426): p. 711-716. DOI: 10.1038/nature11543

[30] Paterson AH, et al. The Sorghum bicolor genome and the diversification of grasses. Nature, 2009. 457(7229): p. 551-556. DOI: doi:10.1038/nature07723

[31] Mysore KS and V Baird. Nuclear DNA content in species of Eleusine (Gramineae): A critical re-evaluation using laser flow cytometry. Plant Systematics and Evolution, 1997. 207(1-2): p. 1-11. DOI: 10.1007/BF00985206

[32] Martel E, et al. Genome size variation and basic chromosome number in pearl millet and fourteen related Pennisetum species. Journal of Heredity, 1997. 88(2): p. 139-143.

[33] Ayele M, et al. Flow cytometric analysis of nuclear genome of the Ethiopian cereal tef [Eragrostis tef (Zucc.) Trotter]. Genetica, 1996. 98(2): p. 211-215.

[34] AOCC. Improving nutrition through African Orphan Crops. 2013; Available from: http://www.mars.com/global/african-orphan-crops.aspx [Accessed: 2015-09-05].

[35] Bennetzen JL and M Chen. Grass genomic synteny illuminates plant genome function and evolution. Rice, 2008. 1(2): p. 109-118.

[36] Schuit F. Comparative genomics: beyond the horizon of the next research grant. Diabetologia, 2015: p. 1-5. DOI: 10.1007/s00125-015-3620-x

[37] Devos KM and MD Gale. Genome relationships: the grass model in current research. The Plant Cell, 2000. 12(5): p. 637-646.

[38] Feuillet C and B Keller. Comparative genomics in the grass family: molecular characterization of grass genome structure and evolution. Annals of botany, 2002. 89(1): p. 3-10. DOI: 10.1093/aob.2002.mcafv0a0il8a

[39] Eckardt NA. Everything in Its Place: Conservation of Gene Order among Distantly Related Plant Species. The Plant Cell, 2001. 13(4): p. 723-725. DOI: 10.2307/3871334

[40] Sorrells ME, et al. Comparative DNA sequence analysis of wheat and rice genomes. Genome research, 2003. 13(8): p. 1818-1827.

[41] Goff SA, et al. A draft sequence of the rice genome (Oryza sativa L. ssp. japonica). Science, 2002. 296(5565): p. 92-100.

[42] Vogel JP, et al. Genome sequencing and analysis of the model grass Brachypodium distachyon. Nature, 2010. 463(7282): p. 763-768. DOI: 10.1038/nature08747

[43] Huo N, et al. The nuclear genome of Brachypodium distachyon: analysis of BAC end sequences. Functional & integrative genomics, 2008. 8(2): p. 135-147. DOI: 10.1007/s10142-007-0062-7

[44] Smilde WD, et al. New evidence for the synteny of rice chromosome 1 and barley chromosome 3H from rice expressed sequence tags. Genome, 2001. 44(3): p. 361-367. DOI: 10.1139/gen-44-3-361

[45] Ahn S, et al. Homoeologous relationships of rice, wheat and maize chromosomes. Molecular and General Genetics MGG, 1993. 241(5-6): p. 483-490.

[46] Ahn S and SD Tanksley. Comparative linkage maps of the rice and maize genomes. Proceedings of the National Academy of Sciences, 1993. 90(17): p. 7980-7984. http://www.pnas.org/content/90/17/7980.abstract

[47] Kilian A, et al. Rice-barley synteny and its application to saturation mapping of the barley Rpg1 region. Nucleic acids research, 1995. 23(14): p. 2729-2733. DOI: 10.1093/nar/23.14.2729

[48] McCouch SR. Genomics and synteny. Plant Physiology, 2001. 125(1): p. 152-155.

[49] Paterson AH, et al. Comparative genomics of plant chromosomes. The Plant Cell, 2000. 12(9): p. 1523-1539.

[50] Bowers JE, et al. Comparative physical mapping links conservation of microsynteny to chromosome structure and recombination in grasses. Proceedings of the National Academy of Sciences of the United States of America, 2005. 102(37): p. 13206-13211. DOI: 10.1073pnas.0502365102

[51] Wei F, et al. Physical and genetic structure of the maize genome reflects its complex evolutionary history. PLoS Genet, 2007. 3(7): p. e123. DOI: 10.1371/journal.pgen.0030123

[52] Cummins GB, The rust fungi of cereals, grasses and bamboos. 1971: Berlin, Springer-Verlag.

[53] Brakke M. Wheat streak mosaic virus. CMI/AAB descriptions of plant viruses, 1971. 48: p. 1-4.

Strategies for Sequence Assembly of Plant Genomes

Stéphane Deschamps and Victor Llaca

Additional information is available at the end of the chapter

Abstract

The field of plant genome assembly has greatly benefited from the development and widespread adoption of next-generation DNA sequencing platforms. Very high sequencing throughputs and low costs per nucleotide have considerably reduced the technical and budgetary constraints associated with early assembly projects done primarily with a traditional Sanger-based approach. Those improvements led to a sharp increase in the number of plant genomes being sequenced, including large and complex genomes of economically important crops. Although next-generation DNA sequencing has considerably improved our understanding of the overall structure and dynamics of many plant genomes, severe limitations still remain because next-generation DNA sequencing reads typically are shorter than Sanger reads. In addition, the software tools used to *de novo* assemble sequences are not necessarily designed to optimize the use of short reads. These cause challenges, common to many plant species with large genome sizes, high repeat contents, polyploidy and genome-wide duplications. This chapter provides an overview of historical and current methods used to sequence and assemble plant genomes, along with new solutions offered by the emergence of technologies such as single molecule sequencing and optical mapping to address the limitations of current sequence assemblies.

Keywords: Sequencing, Plant, Genome, Assembly

1. Introduction

Genome sequencing, assembling and annotation have been major priorities in plant genetics research during the past 20 years. The release of draft reference genomes have typically constituted major milestones and have proven to be invaluable for the analysis and characterization of genome architecture, genes and their expression, diversity and evolution [1–5]. The expansion of sequence information in a growing number of taxa has contributed to comparative studies and the implementation of molecular breeding and biotechnology approaches for crop improvement [6, 7]. The construction of the first plant genomes was made

possible by applying considerable resources, coordination and effort to enabling automated Sanger-based sequencing technologies and computational algorithms. Starting in 2005, a series of technological revolutions in DNA sequencing, driven in large part by the goal of affordable personalized genome sequencing, radically changed the sequencing model. First, new technologies drastically increased throughput while reducing costs and times in data collection. Additional technologies then enabled long single-molecule reads and algorithms that were more suitable to resolve complex genomes [8, 9].

In addition to these advances, the genomics community has benefited from the development and implementation of complementary mapping technologies and methods that have facilitated the scaffolding of sequences and integration to genetic maps. This review provides a historical and technical perspective of methods and technologies applied to genome reference assembly in plants as well as current advances and future directions.

2. The development of Sanger sequencing for *de novo* assembly of plant genomes

The construction of reference genomes was initially enabled by technological advances in sequencing using the Sanger method [10]. During the 1980s and 1990s, the introduction of thermal cycle sequencing, single-tube reactions and fluorescence-tagged terminator chemistry [11] facilitated the development of high-capacity sequencing platforms. Additional improvements in parallelization, base quality assessment, read length and cost-effectiveness were achieved by the development of automatic base-calling and capillary electrophoresis [12, 13]. With no major modifications made in the past years, automated high-throughput Sanger sequencing is performed by parallel reactions that include a mixture of the DNA template, primer, DNA polymerase, and deoxynucleotides (dNTPs). A proportion of dideoxynucleotide terminators (ddNTP) are included in the reaction, each labelled with a different fluorescent dye. DNA molecules are extended from templates using a thermal cycling reaction and terminated by random incorporation of the labelled ddNTPs, which are detected by laser excitation of the fluorescent labels after capillary-based electrophoresis. The differences in dye excitation profiles are recorded and translated by a computer to generate the sequence. Primary analysis software then calls nucleotides from the raw sequences, assigning a corresponding quality score at each position [6, 14].

The complete sequencing of the first bacterial genomes [15,16] as well as the creation of initiatives aimed at sequencing the genomes of *Sacharomyces cerevisae*, *Caenorhabditis elegans*, *Drosophila melanogaster* and *Homo sapiens* provided the technical and technological framework for the initial sequencing of genomes in plants [17–21]. These projects validated the idea of applying a scaled-up form of shotgun sequencing [22]. Shotgun sequencing relied on computer algorithms to enable *in silico* assembly of overlapping sequencing reads derived from randomly-generated subclones. The development of software suites such as Phred, Phrap and Consed [23] allowed calling bases, setting individual base quality, assembling overlapping reads, assigning assembly quality scores, viewing final assemblies and extracting consensus sequences. Two major genomic shotgun sequencing strategies were defined at that time: (1) whole-genome shotgun sequencing (WGS) and (2) clone-by-clone, also referred to as BAC-by-

BAC sequencing. In WGS, genomic DNA is randomly sheared and the ends of the cloned fragments are directly sequenced and assembled. This strategy is the simplest, and it was initially used in small bacterial and yeast genomes. Later, it was also used in *D. melanogaster* and one of two initiatives aimed at sequencing and assembling the reference human genome [19, 21]. Major improvements to *de novo* WGS assembly came from using strategies that relied on paired-end reads from multiple libraries with different average insert sizes and the optimization of software with algorithms that use end-sequence distance information from these libraries.

The second Sanger sequence assembly strategy, clone-by-clone, was successfully deployed in projects aimed at complex eukaryotic genomes. In clone-by-clone genome assembly, shotgun sequencing is performed in libraries derived from individual genomic large-insert clones, selected in a minimum tile path according to physical and genetic map information [24, 25]. The most common type of large-insert clone is the bacterial artificial chromosome (BAC), which can stably carry genomic inserts ranging from 100 to 300 kb and is relatively easy to maintain and purify. Accordingly, this method is usually referred to as BAC-by-BAC, although additional vector systems have been used in assembly projects, including yeast artificial chromosomes (YACs), P1 artificial chromosomes (PAC), transformation-competent artificial chromosomes (TACs), cosmids and fosmids. The two major genomic shotgun-sequencing approaches, WGS and BAC-by-BAC, had advantages and disadvantages when applied to Sanger-based sequencing platforms, depending on the genome of interest. The clone-by-clone approach benefited from working in small units, effectively reducing complexity and computational requirements. This approach minimized problems associated with the misassembly of highly repetitive DNA and therefore provided a better, more complete assembly in plants and other complex eukaryotic genomes. WGS projects were computationally intensive and were less effective bridging repetitive regions in complex genomes but benefited from considerably lower cost, time and logistics [14].

The first completed reference plant genome was from the model system *Arabidopsis thaliana*, accession Columbia [26]. At that time, it was only the third multicellular eukaryotic genome to be published, after *C. elegans* and *D. melanogaster*. The nuclear genome of Arabidopsis is distributed in five chromosomes, and it is only approximately 4% the size of the human genome. The *A. thaliana* genome initiative used multiple types of available large-insert libraries including cosmids, BACs, PACs and TACs. Shotgun clones were constructed and then mapped by restriction fragment fingerprinting as well as screening with hybridization or polymerase chain reaction (PCR) markers. End sequences for 47,788 BAC clones were further used to anchor clones, integrate contigs and help select a minimum tiling path. Each of 1,569 clones in a minimum tiling path were selected, sequenced bidirectionally and assembled at estimated error rates of less than 1 in 10,000 bases. Direct PCR products were used to close some gaps and YACs allowed the characterization of telomere sequences. As initially published, the total length of sequenced regions was 115.4 Mb, in addition to an estimated 10 Mb nonsequenced centromeric and rDNA repeat regions. Since the original publication, the Arabidopsis genome sequence reference has been subjected to several rounds of improvements, each time reducing gaps and extending the sequence towards the centromeric regions [27].

The second published plant genome reference was rice (*Oryza sativa*). While the rice genome is more than 2-fold the size of Arabidopsis, approximately 390 Mb, it is one of the smallest

genomes of any major crop, less than 15% the size of the human genome. Like Arabidopsis, the rice genome was completed using a Sanger-only clone-by-clone approach [28] that required the initial construction, fingerprinting and physical mapping of a large number of random BACs and PACs. In total, 3,401 mapped clones in a minimum tiling path were selected from the physical map, randomly sheared and individually end-sequenced to approximately 10-fold coverage. Clone sequences were assembled and low-quality regions were finished using targeted sequencing. Gaps were closed and low-quality regions resolved by sequencing PCR fragments, plasmids and fosmids.

The draft reference genome of Maize, one of the most important crops in the world, is considered the last major published plant genome project based primarily on a Sanger BAC-by-BAC strategy [29]. At 2.3 Gb and spanning 10 chromosomes, the nuclear genome of maize is considerably larger than that of rice and Arabidopsis, approximately 3/4 the size of the human genome. A set of 16,848 minimally overlapping BAC clones, derived from an integrated physical and genetic map, were selected and end-sequenced. The assembly was performed after adding additional data derived from cDNA sequences and sequences from subtractive libraries with methyl-filtered DNA and high C_0t techniques, resulting in a whole-genome assembly (B73 RefGen_v1) made of 2,048 Mb in 125,325 sequence contigs and 61,161 scaffolds [29]. Unlike the completed genomes of rice and Arabidopsis, most sequenced BACs in the first version of the maize draft genome are unfinished. Gaps and low-quality regions in BACs were not systematically closed by PCR sequencing or other target approaches. Therefore, while the BACs used in the minimum tiling path were mapped, the order and orientation of individual contigs within a single BAC could be incorrect. Subsequent versions of the genome have been improved by targeting gaps and adding alternative sequencing strategies described later in this review.

Finally, it is important to mention that a significant number of plant genome sequencing initiatives have used WGS strategies, which provide a considerable reduction in time and cost associated with cloning, construction, mapping and selection. Sanger WGS genome projects included those of poplar tree, grape, and papaya [30–32]. Later refinements to the process enabled the sequencing of *Brachypodium distachyon* [33] as well as the larger genomes of *Sorghum bicolor* (~730 Mb) [34] and soybean, an ancestral tetraploid (1.1 Mb) [35]. It should be noticed that, as demonstrated by the Maize genome project, the two Sanger shotgun assembly approaches, as well as later sequence technologies, are not mutually exclusive and may be complementary to increase quality and coverage.

The high cost and logistics of plant projects based on clone-by-clone Sanger sequencing required extensive funding, the creation of large collaborative consortia and several years of fingerprinting and sequencing work. The cost of the project by the Arabidopsis Genome Initiative has been estimated at US$70 million [36]. The International Rice Genome Sequencing Project (IRGSP), which included groups from 11 different nations, took over 5 years to complete. During its early stages, IRGSP had estimated that the project would take 10 years and cost a staggering US$200 million [37]. The Maize draft genome was accomplished by multiple laboratories at an estimated cost of tens of millions in a joint NSF/DOE/USDA program. It is worth noticing that, while the cost and time required to accomplish Sanger WGS projects are in fact lower than those based on a clone-by-clone approach, they are still consid-erable for today's standards. The sequencing of the 1.1-Gb soybean genome, the largest

published plant genome based on a Sanger WGS approach, provides an example of such a cost. It was completed in less than two years although it took a group of 18 institutions several million dollars to generate and assemble more than 15 million Sanger reads from multiple libraries with average sizes ranging from 3.3 kb to 135 kb [35].

Besides cost and time considerations, these early Sanger-only projects posed considerable technical challenges. Despite the extensive resources deployed towards the sequencing of the Arabidopsis and rice genomes, which are usually considered as finished, as well as other projects mentioned in this review, they all have representation gaps. A considerable number of gaps correspond to regions that are "unclonable" under the conditions used to prepare BAC and other genomic libraries. Although many of these regions correspond to tandem repeats such as telomeric sequences and other repetitive regions, it may also include gene space [29]. Moreover, the maximum length of quality Sanger reads, usually 800–900 bp, as well as technical issues associated with the sequencing of DNA stretches with strong secondary structures or extensive homopolymers, create conditions for additional sequencing gaps, even in regions with physical coverage.

Finally, most plant genomes are characterized by elevated proportions of highly repetitive DNA and by the presence of segmental duplications or full genome duplications due to polyploidization events [38], which can be problematic during assembly. The 1C genome content in Maize, for example, is smaller than in humans but it consists of higher proportions and larger tracks of high-copy elements such as retrotransposable elements [29, 38]. At least some of the differences between the assembled and estimated genomes of the Maize line B73 could be attributed to the assembly-based collapse of highly similar long terminal repeats (LTRs) at the end of retrotransposons. It is important to note that all the Sanger-only initiatives corresponded to plant species with genomes that were considerably smaller than the average 5.8-Gb plant genome. Plant genomes have a considerably wider size range than in mammals, and in some important crops (e.g. wheat), nuclear genomes can be more than 15 Gb long, well beyond the practical realm of Sanger sequencing. Although BAC-by-BAC approaches can reduce complexity by more than 10,000 fold, Sanger-based assembly remains difficult and prohibitively expensive in plant genomes of moderate or large size. The WGS approach is even more sensitive to the complexity of plant genomes as it increases the potential for assembly artefacts due to haplotype and homeolog collapse in regions with high identity. Reductions in time and cost in WGS projects are achieved at the expense of assembly fidelity in repetitive regions and expanded need for computational resources.

3. Next-generation sequencing technologies applied to *de novo* assembly of plant genomes

3.1. Second-generation sequencing technologies

As indicated above, successful whole-genome sequencing projects have been achieved with the use of Sanger technology. However, such projects require dealing with several complicating factors, including high costs and relatively long turnaround times to completion. The emergence of next-generation sequencing (NGS) technologies has changed this paradigm, both

by reducing costs and increasing sequencing throughputs, while at the same time introducing complexity related to the relative short reads of NGS reads. Several NGS technologies have emerged in the past 7 to 8 years [for reviews, see refs. 39–41]. All follow a relatively uniform approach to library construction and sequencing. To complete sequencing: (1) universal adapters are ligated at the end of single DNA molecule templates; (2) adapter-ligated DNA templates are amplified via PCR to create a cluster of identical isoforms and (3) clusters are loaded on sequencers and nucleotide incorporations occur in parallel on millions of clusters. These generate an amplified signal that is recognized by the platform and translated into a base call.

The most widely used NGS technology nowadays is the one commercialized by Illumina [42], whose high-throughput instrument, the HiSeq4000, can produce up to 1.5 Tb of sequencing data in approximately 3.5 days.. In the Illumina sequencing platform, sequencing templates generated during library construction are immobilized on a solid surface, and a "bridge PCR" approach allows for the localized amplification of millions of single DNA molecules, thus generating millions of clusters, each containing thousands of copies of the original DNA molecules [43]. Sequencing then is performed using a sequencing-by-synthesis approach where single-base extension allows the incorporation of a fluorescently labelled nucleotide (a blocking chemical moiety at the 3' hydroxyl end allows the incorporation of one base only). Once incorporated, the label is detected and the resulting signal subsequently translated into a base call. Finally, the fluorescent dye and the blocking 3' agent are cleaved, allowing the next single base incorporation event to occur. Through the use of alternating cycles of base incorporation, image capture and dye cleavage, the Illumina sequencing technology can produce reads that are up to 300 bp in length. The relatively high error rate (~0.1% or 10 times higher than Sanger sequencing) [39] can be compensated by very high sequencing coverage, thus allowing random errors at any given base position to be ignored below a certain frequency threshold. The relative short read of Illumina sequencing reads can be explained by several noise factors accumulating after each cycle, including phasing, where imperfect single-base incorporation and imperfect cleavage of the dye and 3' hydroxyl blocking moiety lead to the accumulation of copies of various lengths within a cluster, and the subsequent increase of signal-to-noise ratio after each cycle [44].

3.2. Third-generation sequencing technologies

De novo assemblies of plant genomes have been performed with NGS reads only, either with reads generated on the Illumina platform alone or with reads generated with the Illumina platform combined with reads generated on the Roche 454 second-generation sequencing platform [45]. However, those assemblies generally are fragmented, resulting in low N50 values and a high number of contigs, mostly because of the overall short read length, the complexity of the genome and the presence of conserved regions whose length exceeds the length of NGS reads and thus cannot be extended during the *de novo* assembly process. The emergence of third-generation sequencing technologies [46, 47] has started to address some of the inherent limitations of sequencing and assembling large and complex plant genomes. Those technologies are characterized by the parallel sequencing of single molecules of DNA (rather than "clusters"), thus avoiding phasing issues, and the resulting sequences tend to be in the kb range, offering the opportunity to assemble genomes and generating longer contigs

by encompassing complex and conserved genomic regions and allowing relatively high-confidence assemblies of overlapping reads. However, single sequencing reads tend to exhibit relatively high error rates (~15%–25% on average). Deep sequencing coverage or repeated sequencing of the same DNA fragments therefore are required to offset the presence of a high number of sequencing errors [48, 49]. As of today, two companies have developed and commercialized third-generation sequencing technologies, namely, Pacific Biosciences [e.g., 50] (Menlo Park, CA) and Oxford Nanopore Technologies [e.g., 51] (Oxford, UK). Each company uses vastly different approaches to sequencing. The Pacific Biosciences (PacBio) RS II system uses a sequencing-by-synthesis approach to offer up to ~40-kb reads, where base incorporation is monitored in a real-time fashion. Nanoscale holes, described as Zero Mode Waveguides ("ZMW") are located on a chip, where individual polymerases are covalently attached to the surface of each ZMW. Individual nucleotides with a fluorescent label attached to the phosphate chain are incorporated to the elongating strand and the excited dye emits a signal that is captured before diffusion of the released pyrophosphate, and translated into a specific base call. DNA fragments used as template are ligated to "bell-shaped" adapters at both ends, thus facilitating the sequencing of DNA fragments through multiple passes and the creation of a more accurate consensus sequence. The overall stability and activity of the polymerase remain limited by photo damage and the progressive dissociation of the poly-merase/template complex from the surface of the ZMW. It is therefore expected that reads generated from smaller DNA fragments will exhibit higher consensus accuracy than reads from larger DNA fragments. Oxford Nanopore Technologies released the MinION sequencing device in early access mode in 2014. Like the PacBio RS II system, the MinION delivers long reads in a real-time fashion, from single molecules of DNA. In that particular case, however, sequencing is performed by measuring the change in ionic current when a single DNA strand translocates through a protein nanopore located in an insulated membrane. The resulting signal is measured and translated into a base call. Because no enzyme is involved in the DNA sequencing process, it is expected that read length will be driven mostly by the physical length of the DNA strand being sequenced. Library construction involves the ligation of two types of adapters to DNA fragment, one "Y-shaped" adapter with a bound protein that unwinds the double-stranded DNA and facilitates the translocation of a single strand through the pore, and one "bell-shaped" adapter at the other end that allows the translocation, and sequencing, of both the sense and antisense strands. Sequencing reads then are generated by aligning base calls from the two strands and producing a higher quality consensus sequence.

3.3. Challenges in assembling plant genomes

De novo assembly of genomes has closely mimicked the trends and improvements in sequencing technologies and accompanying sequencing assembly software over the years [45]. The emergence of next-generation sequencing technologies has allowed a much larger number of plant genomes to be sequenced and assembled than what would have been deemed possible with Sanger sequencing alone, mostly because of the costs and labor involved in such projects. However, the complexity of the majority of those genomes still makes it a challenge to resolve them with short reads alone [52, 53]. As a result, most plant genome assemblies are highly fragmented, with large number of contigs and conserved regions of the genome in an unfin-

ished state [54]. The presence of highly conserved repeats often exceeding 10 kb in length represents a major challenge in assembling plant genomes. The most common types of repeats in plants are type II long-terminal repeat (LTR) retrotransposons and their proliferation within a genome often explains most of the structural variations between strains [55]. Their movement also results in genome expansion, where repeats represent, in some instances, more than 80–90% of the structural content of a particular genome [29]. Repeat expansion also can lead to very large genome sizes. While NGS technologies can generate enough raw data to cover an entire genome in a relatively cost-effective manner, assembling such a large amount of data often represents a major computational challenge. For example, the assembly of the loblolly pine genome (~22 Gb), which represents the largest genome assembled to date, could be solved only using condensed sets and read pooling prior to assembly [56]. Assembling large and repeat-rich genomes can also be facilitated by using supplemental layers of information, such as the physical distance between "paired" reads (end-sequences generated at both ends of a particular DNA fragment) in mate-pair libraries. Another challenge for *de novo* assembly of plant genome is the issue of polyploidy [57]. Polyploidy is an important force in plant genome evolution and it is estimated that ~80% of all living plants are polyploids [58], while close to 100% of all plant lineages have a paleo-polyploidy event in their history. As a consequence, some plants species, including economically important crop species like soybean [35], have entire gene families consisting of highly similar paralogs. Those gene families are the direct result of paleo-polyploidization events where the merger of genomes has been followed by extensive structural rearrangements, including gene loss, and the modification of gene expression for paralogs within a particular gene family. The diploid genomes of progenitor species can be used to determine the origin and structure of contigs when assembling large polyploid genomes [59]. Finally, heterozygosity may represent another important challenge when assembling plant genomes. Outcrossing species like grape, for instance, exhibit up to 13% sequence divergence between alleles, and the existence of such variation will impact contig assembly when both alleles are sequenced in a whole-genome assembly project [31].

3.4. Examples of plant genome assemblies

According to Michael and Van Buren [45], over 100 plants genomes have been sequenced since 2000, out of which 63% are genomes from various crop species. As indicated above, different Sanger sequencing strategies have been applied with varying degrees of success on several plant genomes. However, the most successful Sanger-based genome assemblies have been obtained from relatively small genomes (Arabidopsis, rice), while *de novo* assemblies for larger and complex genomes, such as maize, remains partial and unfinished (manual improvements of the maize genome were limited to nonrepetitive regions only). In addition, due to the high costs and labor associated with such approaches, and the need for (in most cases) an international consortium to complete such projects, a vast majority of the most recent genomes have been sequenced using either a hybrid approach, complementing Sanger sequencing with NGS data, or using NGS data alone, from various NGS platforms. Such platforms include Illumina, 454/Roche, and more recently, Pacific Biosciences.

The domesticated tomato genome [60] represents an example of Sanger/NGS hybrid genome assembly. A total of 30,800 BAC clones from three different BAC libraries were shotgun-sequenced and end-sequenced, generating a total of 3.3 Gb of Sanger reads. In addition, 454/Roche shotgun and mate-pair sequencing was performed, both on BAC pools and whole-genome DNA preparation, using different insert sizes and generating a total of 21 Gb of NGS data. The *de novo* assembly of Sanger and 454 data was performed using the Newbler assembly software [61] and other sequence assembly and alignment tools. Further scaffolding and polishing of the assembly were performed when integrating BAC end-sequence data and additional high-coverage Illumina and ABI/SOLiD data. Taken together, the *de novo* assembly resulted in 3,761 scaffolds totalling to 782 Mb, with 95% of the assembled scaffold sequences present in 225 scaffolds. The predicted tomato genome size is approximately 900 Mb. The correctness and integrity of the assembly were validated through different means including the alignment of clone end-sequences, publicly available tomato EST sequences, and alignment of BAC contigs from a sequence-based physical BAC map. Interestingly comparison of the tomato, potato and grape genomes supported the existence of two successive whole-genome triplication events in common ancestors that added new gene family members that mediate important fruit functions, such as enzymes involved in ethylene biosynthesis (examples of whole genome duplication or triplication events abound among plant genomes that have been sequenced to date).

Because of the relatively cheap costs involved, a large number of plant genomes have been sequenced and assembled using NGS technologies alone. This includes the assembly of the complex tetraploid genome of cultivated cotton (*Gossypium arboreum*) [62]. The tetraploid cultivated cotton genome has a genome size of approximately 1.7 Gb and is thought to have appeared 1–2 million years ago through interspecific hybridization between diploid A (*Gossypium arboretum*) and D (*Gossypium raimondii*) subgenome progenitors. A total of 371.5 Gb of shotgun Illumina data was generated with various insert sizes ranging from 180 bp to 40 kb and complemented with 33,454 BAC end sequences. The assembly was performed with SOAPdenovo [63], which resulted in 40,381 contigs, anchored and oriented in 7,914 scaffolds, ranging in length from 140 kb to 5.9 Mb with 90% of the contigs included in 3,740 scaffolds.

An example of a smaller, relatively less complex genome assembly is that of the crop species *Brassica rapa* [64]. An estimated 72× sequencing coverage of the genome was generated, corresponding to Illumina shotgun paired-end data from NGS libraries with insert sizes ranging from 200 bp to 10 kb, and assembled using SOAPdenovo [63]. The resulting assembly was made of 14,207 contigs larger than 2 kb, further assembled into 794 scaffolds, totalling approximately 283.8 Mb and estimated to cover more than 98% of the gene space, based on alignments of 214,425 *B. rapa* public EST sequences and 52,712 unigenes from the BrGP database [65]. Further assessment of the integrity of the assembly was performed by aligning BAC clone Sanger sequences reported in previous studies.

While a large number of genomes have been sequenced with NGS technologies alone, the relatively short reads of the major NGS platforms that have been used in those assembly projects, combined with the general complexity of most of those genomes, generally require

the use of alternative methods to facilitate the assembly or confirm its integrity. These methods rely on the use of various types of NGS libraries, such as mate-pair large inserts, or the use of Sanger-derived sequencing data such as EST or BAC-based shotgun reads. However, scaffolding of NGS contigs, based on using pairing information between NGS reads originating from the same DNA fragment, generally leads to unresolved gaps between contigs, often due to the presence of large repeat regions whose size exceed the length and resolution of short NGS reads. As a result, significant portions of any given scaffold contain large batches of unknown sequences, and of unknown length. To address these issues and improve plant genome assemblies, researchers have developed a series of multifaceted solutions, combining alignment to known public data, such as ESTs or BAC ends, or, when available, reference genomes from related species, integration of physical and genetic map data, or new technologies. Some of these approaches have been described in the next chapter.

4. Complementary approaches to *de novo* assembly of plant genomes

4.1. Long-read assembly

NGS assembly strategies based on the use of short reads cannot solve long and identical transposable elements abundantly present in most plant genomes. The use of long reads is expected to address some of those shortcomings and improve the overall quality of *de novo* assembly by ordering contigs, closing gaps, and improving scaffolding. As a consequence, researchers have started to adopt the single-molecule long-read sequencing technology from Pacific Biosciences in plant genome assembling projects. Spinach is an example of such genome assembly efforts. Spinach is a diploid species with a genome size estimated at 989 Mb. Van Deynze *et al*. [66] sequenced and assembled the Spinach genome using large fragment libraries of Pacific Biosciences sequence reads. They generated a 60× coverage of the genome, with 20% of the reads larger than 20 kb. Data were assembled using PacBio's Hierarchical Genome Assembly Process (HGAP) [67], which showed that long-read assemblies exhibited a 63-fold improvement in contig size over an Illumina-only assembly, derived from multiple Illumina libraries.

A distinct strategy to long-read assembly, namely, the Illumina TruSeq Synthetic Long-Read (SLR) strategy [68], is also expected to improve the quality of assemblies generated with short reads only. In SLR libraries, genomic DNA is fragmented to ~10 kb and individual indexed Illumina libraries are generated in parallel from highly diluted pools of sheared DNA fragments. After Illumina sequencing and data deconvolution, the original ~10 kb fragments can be reassembled, effectively reducing the complexity level of the assembly and generating very-high quality synthetic long reads that can subsequently be assembled together or used for haplotype resolution.

The use of long reads in *de novo* assembly is bound to become more prevalent in the near future, reducing the number of scaffolds while at the same time increasing their average length. The use of PacBio in smaller genomes, such as microbial genomes, has already demonstrated that

the assemblies often result in contigs corresponding in most cases to individual chromosomes or plasmids present in the microbial cells. Likewise, it is likely that future plant studies will include such long reads, either alone or in combination with short-read NGS data to improve assembly and coverage in questionable regions, and to confirm the integrity of the assembly in a manner similar to Sanger data with current NGS assemblies.

4.2. Genetic anchoring

The emergence of NGS technologies has rapidly led researchers to develop methods and assays for variant discovery in various plant genomes. Some studies have shown that Single nucleotide polymorphisms (SNPs) can be discovered in parental inbred lines using next-generation sequencing [69]. Entire mapping populations also have been simultaneously sequenced and genotyped, in a process known as "genotyping-by-sequencing" (GBS) [70, 71], discovering in the process extensive lists of segregating markers within the mapped population [72, 73], that can be completed by using known reference maps or sequences to impute missing marker data from individual haplotypes. Various reduced-representation methods have been employed for NGS-derived SNP discovery in plant species where whole-genome shotgun sequencing still remains too expensive for sequencing more than a few individuals [71]. These methods include the use of restriction enzyme digestion–based assays with methyl-sensitive restriction endonucleases [74, 75], or methods based on sequence capture approaches [76], to sequence and map gene-rich portions of a genome, and allowing the anchoring of SNPs in a relatively unambiguous manner.

More recently, ultradense linkage maps have been created from genotyping by whole genome sequencing of a genetic mapping population. It has been used to place whole-genome sequencing contigs into a map, thus anchoring, and ordering, sequencing of contigs [77]. Such an approach requires using a genetic linkage map as a framework, into which SNPs derived from the whole genome sequencing assembly can be integrated into a genetic framework derived from low coverage whole-genome sequencing data from a segregating population. The genetic position of the sequence-derived SNPs can then be used to assign chromosomal locations to the contigs harboring them. Such an approach has been used in the context of a whole-genome assembly project in barley where genetic anchoring was applied to a whole-genome assembly [78]. SNPs discovered by sequencing individuals from two mapping populations at low coverage (~1×) were placed into genetic maps that had been previously constructed through different means, including SNP array data and GBS, or made from the whole-genome shotgun sequencing data of the population. Their genetic positions then were used to assign chromosomal locations, and integrate into the combined physical and genetic genome framework, approximately two-thirds of all whole-genome shotgun sequencing contigs. While highly effective in plants, where mapping populations are often readily available, it must be noted that such an approach is limited by the overall recombination landscape, and the subsequent relationship between physical and genetic distance within a particular region of the genome [76]. Recombination events in plants often occur in distal regions of the chromosomes, and peri-centromeric regions may require very large mapping populations to improve their resolution. In addition, recent studies have suggested that

specific features of the genome, such as chromosomal inversion, translocation and duplication varying between the two parents used to generate the mapping population, may lead to errors and potentially confound genome assemblies.

4.3. BAC pool sequencing in gene-rich regions

A large number of genome assemblies have been generated with the help of physical maps and the use of a BAC-by-BAC sequencing approach. While laborious and costly, this approach still remains relevant as it offers multiple advantages over a whole-genome sequencing approach, especially in terms of assembling sequencing reads conserved in the context of a whole-genome assembly but mapping exclusively to a defined portion of a genome in the context of an individual clone assembly. Lonardi *et al*. [80] proposed a modified version of clone sequencing to take advantage of the massive sequencing capacity offered by NGS platforms. In that study, subsets of overlapping genome-tiling BAC clones were selected and pooled according to a multidimensional grid design. Each pool then was sequenced on an Illumina HiSeq2000 instrument. The resulting paired-end reads were deconvoluted by determining, for each read the intersection between the pool it originates from and the individual BAC clone(s) within that same pool covering the portion of the genome the read corresponds to, based on physical map information. Once deconvolution is achieved, reads can be assembled using an NGS assembler (Velvet) [81], to recreate the sequence of the original BAC clone. Such an approach was successfully tested in barley BAC clones selected based on BAC-unigene associations described in that same study, thus suggesting that BAC pool sequencing can be used in correlation with existing physical maps to complement or correct whole-genome sequencing assemblies, offering in the process the likelihood of higher quality contig sequence assemblies in gene-rich regions of complex plant genomes.

4.4. Optical mapping

Optical mapping is a single-molecule approach that produces fingerprints using ordered restriction maps [82] or specific nick sites [83]. After enzymatic treatment and subsequent incorporation of fluorescent labels, the DNA molecules are stretched on a glass surface or in a nanochannel array and directly imaged to locate regions corresponding to the restriction sites or nick sites within the molecule. Distances between those sites are then inferred to produce an optical map of the DNA molecule. Two commercial platforms currently are available, namely, the Opgen Argus [84] and the BioNano Genomics Irys [85] systems. Using such techniques, very large DNA molecules, in the Mb range, can be interrogated for the presence and location of short recognition sites (whose sequence will vary with the enzyme being used to treat the DNA). Consensus optical maps then can be created by determining the overlap, under highly redundant conditions, between optical maps of single DNA molecules. Such consensus maps have to take into account the possibility of errors inherent to this type of technology, including star activity and false enzyme cuts, or the possibility of chimeric maps when joining, for example, optically mapped molecules containing paralogous genomic regions.

Optical maps can be used for multiple applications, including comparative genomics and structural variation detection, as well as the development of optical map-guided genome assemblies, where the optical map is aligned and compared to *in silico* digested contig sequences. Optical map-guided genome assemblies can assist in building high-quality genome assemblies by providing evidence of the ordering of adjacent contigs and scaffolds, or by assessing the overall sequence accuracy of contigs and suggesting potential errors in an assembly, such as inversions, translocations or chimeric contig or scaffold sequences. The addition of optical maps to a genome assembly often results in a significant increase in the scaffold N50 value. For example, Hastie *et al.* [86] used the mapping of tiling BAC clones in a 2.1 Mb highly repetitive region of *Aegilops tauschii* (the D-genome donor of hexaploid wheat) to correct several misassemblies and improve the assembly from 75% to 95% complete. In another study [87], a high-resolution optical map, spanning 91% of the maize genome, was built, and used to characterize gaps within contigs, the maize genetic-physical (FPC) map and the reference pseudomolecules. Results also suggested that the placement of 12 FPC contigs on the maize genetic-physical map required re-evaluation.

4.5. Long-range Hi-C interactions

High-throughput Chromosome capture (Hi-C) is a method that uses cross-linking of DNA-binding protein to DNA followed by restriction digestion and self-ligation of protein-bound DNA fragments, to probe genome-wide three-dimensional chromatin interactions between chromosomal regions bound to the same proteins (such as enhancer and promoter regions) [88]. There is a statistically higher probability that those regions are located on the same chromosome rather than on different chromosomes, as expected within the context of chromosomes located in distinct three-dimensional spaces within the nucleus. As a result, a vast majority of Hi-C read pairs (where each paired reads correspond to reads that may be millions of bases apart from each other on the same chromosome) can be used to determine what two contigs can be linked together on the same chromosome, based on the Hi-C paired reads they each contain.

Burton *et al.* [89] evaluated the use of Hi-C datasets for long-range scaffolding of *de novo* whole-genome assemblies. This approach works, first, by aligning Hi-C reads to *de novo* assembly contig sequences and indexing each contig to their respective chromosomes, ordering contigs within each respective chromosome group by using higher Hi-C interaction densities expected between closely located contigs, and orienting ordered contigs using the location and orientation of Hi-C reads within each contig. The approach tested on existing human and mouse contig datasets generated from next-generation shotgun and mate-pair sequencing reads showed that a vast majority of the contigs could be grouped (98.2% and 98% of all sequences, in human and mouse, respectively) and ordered (94.4% and 86.7% of all grouped sequences, in human and mouse, respectively) within individual chromosomes when combined with Hi-C sequencing reads. Similar studies, where Hi-C datasets were used to complement *de novo* assembly generated with next-generation sequencing reads have been performed in human and mouse by Kaplan and Dekker [90] and Selvaraj *et al.* [91].

4.6. Long-range scaffolding

Two companies, namely, 10X Genomics [92] (Pleasanton, CA) and Dovetail Genomics [93] (Santa Cruz, CA), recently presented new ways to assemble short reads delivered by the Illumina technology. The GemCode instrument from 10X Genomics is a microfluidic device used to partition very long DNA molecules (typically 50 kb or more) into oil-based droplets and to prepare Illumina-compatible libraries in combination with "gel beads", each containing a unique 14-bp indexing barcode. Once sequencing is performed, in-house software deconvolutes the barcodes and reconstructs the sequence of the original DNA subfragments as to where they originate from on the original long DNA molecule. In contrast to 10X Genomics, Dovetail Genomics approach does not necessarily require an instrument but requires larger amount of starting material for preparing samples. Dovetail's approach works essentially by *in vitro* making a Hi-C library from chromatin-free purified DNA, thus recreating intramolecular interactions while reducing intermolecular ones. The resulting fragments can then be selected for mate-pair sets capturing long-range intramolecular interactions for genome scaffolding. While not yet applied on plant genome assemblies, it is presumed that the strategies and technologies highlighted above could potentially assist in grouping and ordering contigs and scaffolds from gene-rich regions of diploid plant genomes.

5. Conclusion

Reference genomes are now available for a significant number of plant species. The emergence of NGS technologies has made it possible to sequence genomes not only from economically important crop species but also from nonstandard model and special plants whose genomes otherwise might not have been sequenced due to the requirements for large funds, instrumentation and personnel that was witnessed in earlier pre-NGS days. While great progress has been made, assembling such genomes still remains challenging due to their inherent complexity and the relative absence of long-range connectivity, lost during DNA fragmentation and short-read sequencing. As a result, plant genome assemblies tend to be highly fragmented, and focused essentially on unique "gene-rich" regions, while large fractions of the genomes, namely, complex repeat and conserved regions, remain unassembled. Researchers have come up with creative ways to address those shortcomings, including the use of mate-pair NGS libraries, the complementation of physical assemblies with genetic maps, or the use of new technologies for sequencing, physical mapping or scaffolding. It is hoped that the routine use of such novel approaches will help in elucidating the biological aspects of genomes by allowing true comparative and structural analysis between species, strains, tissue or environment.

Acknowledgements

Author details

Stéphane Deschamps[*] and Victor Llaca

*Address all correspondence to: stephane.deschamps@cgr.dupont.com

DuPont Pioneer, Wilmington, Delaware, USA

References

[1] Varshney RK, Navak SN, May GD, Jackson SA. Next-generation sequencing technologies and their implications for crop genetics and breeding.Trends Biotechnol. 2009;27:522–530. DOI: 10.1016/j.tibtech.2009.05.006

[2] Messing J, Llaca V.Importance of anchor genomes for any plant genome project. Proc. Natl. Acad. U.S.A.1998;95:2017–2020. DOI: 10.1073/pnas.95.5.2017

[3] Edwards D, Batley J.Plant genome sequencing: applications for crop improvement. Plant Biotechnol. J.2010;8:2–9. DOI: 10.1111/j.1467-7652.2009.00459.x

[4] Jackson SA, Iwata A, Lee SH, Schmutz J, Shoemaker R. Sequencing crop genomes: approaches and applications. New Phytol.2011;191:915–925. DOI: 10.1111/j. 1469-8137.2011.03804.x

[5] Hou H, Atlihan N, Lu ZX. New biotechnology enhances the application of cisgenesis in plant breeding. Front. Plant Sci. 2014;11:389. DOI: 10.3389/fpls.2014.00389

[6] Green ED. Strategies for the systematic sequencing of complex genomes. Nat. Rev. Genet. 2001;2:573–583. DOI: 10.1038/35084503

[7] Liu, W, Yuan, JS, Stewart, CN. Advanced genetic tools for plant biotechnology. Nat Rev Genet. 2013;14:781–793. DOI: 10.1038/nrg3583

[8] Shendure, J, Aiden, EL.The expanding scope of DNA sequencing. Nature Biotechnol. 2012;30:1084–1094.DOI:10.1038/nbt.2421

[9] Mardis ER.A decade's perspective on DNA sequencing technology. Nature. 2011;470:198–203. DOI: 10.1038/nature09796

[10] Sanger F, Coulson AR. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase.J. Mol. Biol. 1975;94:441–448. DOI: 10.1016/0022-2836(75)90213-2

[11] Trainor GL. DNA sequencing, automation and the human genome. Anal. Chem. 1990;62:418–426. DOI: 10.1021/ac00204a001

[12] Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res.1998;8:175–185. DOI: 10.1101/gr. 8.3.175

[13] Karger BL, Guttman A. DNA sequencing by CE. Electrophoresis. 2009;30:S196–202. DOI: 10.1002/elps.200900218

[14] Llaca V.Sequencing technologies and their use in plant biotechnology and breeding. In: Dr. Anjana Munshi, editor. DNA Sequencing—Methods and Applications. 2012. p. 35–60. DOI: 10.5772/37918

[15] Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, *et al*.Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science. 1995;269:496–512. DOI: 10.1126/science.7542800

[16] Himmelreich R, Hilbert H, Plagens H, Pirkl E, Li BC, Herrmann R.Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*. Nucleic Acids Res. 1996;24:4420–4449. DOI: 10.1093/nar/24.22.4420

[17] Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, *et al*.Life with 6000 genes. Science. 1996;274:563–567. DOI: 10.1126/science.274.5287.546

[18] *C. elegans* Sequencing Consortium. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. Science. 1998;282:2012–2018. DOI: 10.1126/science. 282.5396.2012

[19] Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, *et al*.The genome sequence of *Drosophila melanogaster*. Science. 2000;287:2185–2195. DOI: 10.1126/science.287.5461.2185

[20] International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature. 2001;409:860–921. DOI: 10.1038/35057062

[21] Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, *et al*.The sequence of the human genome. Science. 2001;291:1304–1351. DOI: 10.1126/science.1058040

[22] Staden R.A strategy of DNA sequencing employing computer programs. Nucleic Acids Res. 1979;6:2601–2610. DOI: 10.1093/nar/6.7.2601

[23] Gordon D, Abajian C, Green P. Consed: a graphical tool for sequence finishing. Genome Res.1998;8:195–202. DOI: 10.1101/gr.8.3.195

[24] Soderlund C, Longden I, Mott R. FPC: a system for building contigs from restriction fingerprinted clones. Comput. Appl. Biosci. 1997;13:523–535. DOI: 10.1093/bioinformatics/13.5.523

[25] Ding Y, Johnson MD, Chen WQ, Wong D, Chen YJ, Benson SC, *et al*.Five-color-based high-information-content fingerprinting of bacterial artificial chromosome clones using type IIS restriction endonucleases. Genomics. 2001;74:142–154. DOI: 10.1006/geno.2001.6547

[26] Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature. 2000;408:796–815. DOI: 10.1038/35048692

[27] Hosouchi, T, Kumekawa N, Tsuruoka H, Kotani H. Physical map-based sizes of the centromeric regions of *Arabidopsis thaliana* chromosomes 1, 2, and 3. DNA Res. 2002;9:117–121. DOI: 10.1093/dnares/9.4.117

[28] International Rice Genome Sequencing Project. The map-based sequence of the rice genome. Nature. 2005;436:793–800. DOI: 10.1038/nature03895

[29] Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, *et al*.The B73 maize genome: complexity, diversity, and dynamics. Science. 2009;326:1112–1115. DOI: 10.1126/science.1178534

[30] Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, *et al*.The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science. 2006;313:1596–1604. DOI: 10.1126/science.1128691

[31] Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, et al.The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature. 2007;449:463–467. DOI: 10.1038/nature06148

[32] Ming R, Hou S, Feng Y, Yu Q, Fionne-Laporte A, Saw JH, *et al*.The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). Nature. 2008;452:991–996. DOI: 10.1038/nature06856

[33] International Brachypodium Initiative. Genome sequencing and analysis of the model grass *Brachypodium distachyon*. Nature. 2010;463:763–768. DOI: 10.1038/nature08747

[34] Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, *et al*.The *Sorghum bicolor* genome and the diversification of grasses. Nature. 2009;457:551–556. DOI: 10.1038/nature07723

[35] Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, *et al*.Genome sequence of the palaeopolyploid soybean. Nature. 2010;463:178–183. DOI: 10.1038/nature08670

[36] Feuillet C, Leach JE, Rogers J, Schnable PS, Eversole K. Crop genome sequencing: lessons and rationales. Trends Plant Sci. 2011;16:77–88. DOI: 10.1016/j.tplants.2010.10.005

[37] Saegusa A. US firm's bid to sequence rice genome causes stir in Japan. Nature. 1999;398:545. DOI: 10.1038/19123

[38] Llaca V, Deschamps S, Campbell M. Genome diversity in maize. J. Botany. 2011;104172. DOI: 10.1155/2011/104172

[39] Shendure J, Ji H.Next-generation DNA sequencing. Nat. Biotechnol.2008;28:1135–1145. DOI: 10.1038/nbt1486

[40] Ansorge WJ.Next-generation DNA sequencing techniques. N. Biotechnol. 2009;25:195–203. DOI: 10.1016/j.nbt.2008.12.009

[41] Niedringhaus TP, Milanova D, Kerby MB, Snyder MP, Barron AE.Landscape of next-generation sequencing technologies. Anal. Chem.2011;83:4327–4341. DOI: 10.1021/ac2010857

[42] Bentley DR, Balasubramanjan S, Swerdlow HP, Smith GP, Milton J, Brown CG, *et al*.Accurate whole human genome sequencing using reversible terminator chemistry. Nature. 2008;456:53–59. DOI: 10.1038/nature07517

[43] Fedurco M, Romieu A, Williams S, Lawrence I, Turcatti G.BTA, a novel reagent for DNA attachment on glass and efficient generation of solid-phase amplified DNA colonies. Nucleic Acids Res. 2006;34:e22. DOI: 10.1093/nar/gnj023

[44] Erlich Y, Mitra PP, de la Bastide M, McCombie WR, Hannon GJ. Alta-cyclic: a self-optimizing base caller for next-generation sequencing. Nat. Methods. 2008;5:679–682. DOI: 10.1038/nmeth.1230

[45] Michael TP, Van Buren R. Progress, challenges and the future of crop genomes. Curr. Opin. Plant Biol.2015;24:71–81. DOI: 10.1016/j.pbi.2015.02.002

[46] Schadt EE, Turner S, Kasarskis A. A window into third-generation sequencing. Hum. Mol. Genet.2010;19:R227–240. DOI: 10.1093/hmg/ddq416

[47] Feng Y, Zhang Y, Ying C, Wang D, Du C. Nanopore-based fourth-generation DNA sequencing technology. Genomics Proteomics Bioinformatics. 2015;13:4–16. DOI: 10.1016/j.gpb.2015.01.009

[48] Anton BP, Mongodin EF, Agrawal S, Fomenkov A, Byrd DR, Roberts RJ, *et al*.Complete genome sequence of ER2796, a DNA methyltransferase-deficient strain of *Escherichia coli* K-12. PLoS One. 2015;10:e0127446. DOI: 10.1371/journal.pone.0127446

[49] Loman NJ, Quick J, Simpson JT. A complete bacterial genome assembled de novo using only nanopore sequencing data. Nat. Methods. 2015;12:733–735. DOI: 10.1038/nmeth.3444

[50] Plant and animal whole genome sequencing [Internet]. 2015. Available from: http://www.pacb.com/applications/whole-genome-sequencing/plant-animal/   [Accessed 2015-10-14]

[51] DNA: nanopore sequencing [Internet]. 2015. Available from: https://nanopore-tech.com/applications/dna-nanopore-sequencing [Accessed 2015-10-14]

[52] Bennetzen JL, Ma J, Devos KM. Mechanisms of recent genome size variation in flowering plants. Ann. Bot.2005;95:127–132. DOI: 10.1093/aob/mci008

[53] Rostoks N, Park YJ, Ramakrishna W, Ma J, Druka A, Shiloff BA, *et al*.Genomic sequencing reveals gene content, genomic organization, and recombination relation-

ships in barley. Funct. Integr. Genomics. 2002;2:51–59. DOI: 10.1007/s10142-002-0055-5

[54] Claros MG, Bautista R, Guererro-Fernandez D, Benzerki H, Seoane P, Fernandez-Pozo N. Why assembling plant genome sequences is so challenging. Biology (Basel). 2012;1:439–459. DOI: 10.3390/biology1020439

[55] Proost S, Pattyn P, Gerats T, Van de Peer Y. Journey through the past: 150 million years of plant genome evolution. Plant J.2011;66:58–65. DOI: 10.1111/j.1365-313X.2011.04521.x

[56] Zimin A, Stevens KA, Crepeau MW, Holtz-Morris A, Koriabine M, Marcais G, *et al*. Sequencing and assembly of the 22-Gb loblolly pine genome. Genetics. 2014;196:875–890. DOI: 10.1534/genetics.113.159715

[57] Comai L.The advantages and disadvantages of being polyploid. Nat. Rev. Genet. 2005;6:836–846. DOI: 10.1038/nrg1711

[58] Meyers LA, Levin DA. On the abundance of polyploids in flowering plants. Evolution. 2006;60:1198–1206. DOI: 10.1554/05-629.1

[59] Ling HQ, Zhao S, Liu D, Wang J, Sun H, Zhang C, *et al*.Draft genome of the wheat A-genome progenitor *Triticum urartu*. Nature. 2013;496:87–90. DOI: 10.1038/nature11997

[60] Tomato Genome Consortium. The tomato genome sequence provides insights into fleshy fruit evolution. Nature. 2012;485:635–641. DOI: 10.1038/nature11119

[61] Nederbragt AJ.On the middle ground between open source and commercial software—the case of the Newbler program. Genome Biol. 2014;15:113. DOI: 10.1186/gb4173

[62] Li F, Fan G, Wang K, Sun F, Yuan Y, Song G, *et al*.Genome sequence of the cultivated cotton *Gossypium arboreum*. Nat. Genet.2014;46:567–572. DOI: 10.1038/ng.2987

[63] Li R, Zhu H, Ruan J, Qian W, Fang X, Shi Z, *et al*.De novo assembly of human genomes with massively parallel short read sequencing. Genome Res.2010;20:265–272. DOI: 10.1101/gr.097261.109

[64] Wang X, Wang H, Wang J, Sun R, Wu J, Liu S, *et al*.The genome of the mesopolyploid crop species *Brassica rapa*. Nat. Genet.2011;43:1035–1039. DOI: 10.1038/ng.919

[65] Brassica database [Internet]. 2010. Available from: http://brassicadb.org/brad/index.php [Accessed 2015-10-14]

[66] Van Deynze A, Ashrafi H, Hickey L, Peluso P, Rank D, Chin J, *et al*.Using spinach to compare technologies for whole genome assemblies. In: Plant & Animal Genome XXIII; 10–14 January 2015; San Diego, CA.

[67] Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, *et al*.Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat. Methods. 2013;10:563–569. DOI: 10.1038/nmeth.2474

[68] McCoy RC, Taylor RW, Blauwkamp TA, Kelley JL, Kertesz M, Pushkarev D, *et al*. Illumina TruSeq synthetic long-reads empower de novo assembly and resolve complex, highly-repetitive transposable elements. PLoS One. 2014;9:e106689. DOI: 10.1371/journal.pone.0106689.

[69] Deschamps S, Campbell M. Utilization of next-generation sequencing platforms in plant genomics and genetic variant discovery. Mol. Breed.2010;25:553–570. DOI: 10.1007/s11032-009-9357-9

[70] Elshire RJ, Glaubitz JC, Suri Q, Poland JA, Kawamoto K, Buckler ES, *et al*.A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One. 2011;6:e19379. DOI: 10.1371/journal.pone.0019379

[71] Deschamps S, Llaca V, May GD. Genotyping-by-sequencing in plants. Biology (Basel). 2012;1:460–483. DOI: 10.3390/biology1030460

[72] Huang X, Feng Q, Qian Q, Zhao Q, Wang L, Wang A, *et al*.High-throughput genotyping by whole-genome resequencing. Genome Res.2009;19:1068–1076. DOI: 10.1101/gr.089516.108

[73] Trick M, Adamski NM, Mugford SG, Jiang CC, Febrer M, Uauy C. Combining SNP discovery from next-generation sequencing data with bulked segregant analysis (BSA) to fine-map genes in polyploidy wheat. BMC Plant Biol.2012;12. DOI: 10.1186/1471-2229-12-14

[74] Fellers JP. Genome filtering using methylation-sensitive restriction enzymes with six base pair recognition sites. The Plant Genome. 2008;1:146–152. DOI: 10.3835/plantgenome2008.05.0245

[75] Gore MA, Wright MH, Ersoz ES, Bouffard P, Szekeres ES, Jarvie TP, *et al*.Large-scale discovery of gene-enriched SNPs. The Plant Genome. 2009;2:121–133. DOI: doi: 10.3835/plantgenome2009.01.0002

[76] Muraya MM, Schmutzer T, Ulpinnis C, Scholz U, Altmann T. Targeted sequencing reveals large-scale sequence polymorphism in maize candidate genes for biomass production and composition. PLoS One. 2015;10:e0132120. DOI: 10.1371/journal.pone.0132120

[77] Mascher M, Muehlbauer GJ, Rokhsar DS, Chapman J, Schmutz J, Barry K, *et al*.Anchoring and ordering NGS contig assemblies by population sequencing (POPSEQ). Plant J. 2013;76:718–727. DOI: 10.1111/tpj.12319

[78] Ariyadasa R, Mascher M, Nussbaumer T, Schulte D, Frenkel Z, Poursarebani N, *et al*.A sequence-ready physical map of barley anchored genetically by two million sin-

gle-nucleotide polymorphisms. Plant Physiol.2014;164:412–423. DOI: 10.1104/pp. 113.228213

[79] Mascher M, Stein N. Genetic anchoring of whole-genome shotgun assemblies. Front. Genet.2014;5:208. DOI: 10.3389/fgene.2014.00208

[80] Lonardi S, Duma D, Alpert M, Cordero F, Beccuti M, Bhat PR, *et al*.Combinatorial pooling enables selective sequencing of the barley gene space. PLoS Comput. Biol. 2013;9:e1003010. DOI: 10.1371/journal.pcbi.1003010

[81] Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res.2008;18:821–829. DOI: 10.1101/gr.074492.107

[82] Schwartz DC, Li X, Hernandez LI, Ramnarain SP, Huff EJ, Wang YK. Ordered restriction maps of Saccharomyces cerevisiae chromosomes constructed by optical mapping. Science. 1993;262:110–114. DOI: 10.1126/science.8211116

[83] Lam ET, Hastie A, Lin C, Ehrlich D, Das SK, Austin MD, *et al*.Genome mapping on nanochannel arrays for structural variation analysis and sequence assembly. Nat. Biotechnol. 2012;30:771–776. DOI: 10.1038/nbt.2303

[84] Argus system [Internet]. 2015. Available from: http://opgen.com/genomic-services/ argus-system [Accessed 2015-10-14]

[85] Irys system [Internet]. 2015. Available from: http://www.bionanogenomics.com/ products/ [Accessed 2015-10-14]

[86] Hastie AR, Dong L, Smith A, Finklestein J, Lam ET, Huo N, *et al*.Rapid genome mapping in nanochannel arrays for highly complete and accurate de novo sequence assembly of the complex *Aegilops tauschii* genome. PLoS One. 2013;8:e55864. DOI: 10.1371/journal.pone.0055864

[87] Zhou S, Wei F, Nguyen J, Bechner M, Potamousis K, Goldstein S, *et al*.A single molecule scaffold for the maize genome. PLoS Genet.2009;5:e1000711. DOI: 10.1371/journal.pgen.1000711

[88] Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, *et al*.Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science. 2009;326:289–293. DOI: 10.1126/science.1181369

[89] Burton JN, Adey A, Patwardhan RP, Qiu R, Kitzman JO, Shendure J. Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. Nat. Biotechnol.2013;31:1119–1125. DOI: 10.1038/nbt.2727

[90] Kaplan N, Dekker J. High-throughput genome scaffolding from in vivo DNA interaction frequency. Nat. Biotechnol.2013;31:1143–1147. DOI: 10.1038/nbt.2768

[91] Selvaraj, S, R Dixon J, Bansai V, Ren B. Whole-genome haplotype reconstruction using proximity-ligation and shotgun sequencing. Nat. Biotechnol.2013;31:1111–1118. DOI: 10.1038/nbt.2728

[92] Megabase scale phasing from 10× Genomics [Internet]. 2015. Available from: http://10xgenomics.com/sites/default/files/downloads/10x-0009_app_note_phasing_press_ready.pdf [Accessed 2015-10-14]

[93] Chromosome-scale shotgun assembly using an in vitro method for long-range linkage [Internet]. 2015. Available from: http://arxiv.org/abs/1502.05331 [Accessed 2015-10-14]

Toward a First High-quality Genome Draft for Marker-assisted Breeding in Leaf Chicory, Radicchio (*Cichorium intybus* L.)

Giulio Galla, Andrea Ghedina, Silvano C. Tiozzo and Gianni Barcaccia

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/61747

Abstract

Radicchio (*Cichorium intybus* subsp. *intybus* var. *foliosum* L.) is one of the most important leaf chicories, used mainly as a component for fresh salads. Recently, we sequenced and annotated the first draft of the leaf chicory genome, as we believe it will have an extraordinary impact from both scientific and economic points of view. Indeed, the availability of the first genome sequence for this plant species will provide a powerful tool to be exploited in the identification of markers associated with or genes responsible for relevant agronomic traits, influencing crop productivity and product quality. The plant material used for the sequencing of the leaf chicory genome belongs to the Radicchio of the Chioggia type. Genomic DNA was used for library preparation with the TruSeq DNA Sample Preparation chemistry (Illumina). Sequencing reactions were performed with the Illumina platforms HiSeq and MySeq, and sequence reads were then assembled and annotated. We are confident that our efforts will extend the current knowledge of the genome organization and gene composition of leaf chicory, which is crucial for developing new tools and diagnostic markers useful for our breeding strategies in Radicchio.

Keywords: Genome draft, marker-assisted breeding, gene prediction, SSR markers, SNP calling

1. Introduction

The common Italian name of Radicchio was adopted in recent years by all the most internationally used languages and indicates a highly differentiated group of chicories, with red or variegated leaves. Radicchio (*Cichorium intybus* subsp. *intybus* var. *foliosum* L.) is currently one of the most important leaf chicories, used mainly as a component for fresh salads but also very

often cooked and prepared differently according to local traditions and alimentary habits [1]. This plant species belongs to the Asteraceae family and includes several cultivar groups whose commercial food products are the leaves, namely Witloof, Pain de sucre, and Catalogne, as well as several types of Radicchio.

From the reproductive point of view, Radicchio is prevalently allogamous, due to an efficient sporophytic self-incompatibility system, proterandry and gametophytic competition favoring allo-pollen grains and tubes [1]. Probably known by the Egyptians and used as food and/or medicinal plants by the ancient Greeks and Romans, this species gradually underwent a process of naturalization and domestication in Europe during the past few centuries. This plant has become part of both natural and agricultural environments of Italy. Currently, among the different biotypes of leaf chicories, the so-called Radicchio of Chioggia, native to and very extensively grown in northeastern Italy, is the Radicchio cultivar acquiring more and more commercial interest worldwide. In Italy, the Radicchio of Chioggia is cultivated on a total area of approximately 16–18,000 ha, half of which is in the Veneto region, with a total production of approximately 270,000 tons (more than 60% obtained using professional seeds), reaching an overall turnover of approximately € 10,000,000 per year.

Grown plant materials are usually represented by landraces or their directly derived synthetics that are known to possess a high variation and adaptation to the natural and anthropological environment where they originated from and are still cultivated. These populations are characterized by high-quality traits and have been maintained or even improved over the years by local farmers through phenotypical selection according to their own criteria and more recently by seed companies through genotypical selection following intercross or polycross schemes combined with progeny tests to obtain populations showing superior DUS scores for both agronomic and commercial traits. The breeding programs currently underway by local firms and regional institutions exploit the best landraces and aim to isolate individuals amenable for use as parents for the constitution of narrow genetic base synthetic varieties and/ or to select inbred lines suitable for the production of heterotic F1 hybrids [2]. In recent years, phenotypic evaluation trials are increasingly assisted by genotypic selection procedures through the use of molecular markers scattered throughout the genome. In fact, marker-assisted breeding allows the identification of the parental individuals or the inbred lines showing the best general or specific combining ability in order to breed synthetics and hybrids, respectively.

Radicchio, like the other leaf chicories, is diploid ($2n=2x=18$) and is characterized by an estimated haploid genome size of approximately 1.3 Gb. In recent years, three distinct saturated molecular linkage maps were constructed for leaf chicories, covering approximately 1,200 cM [3-5]. Its linkage groups were mainly based on neutral SSR markers, but many EST-derived SNP markers were also mapped. A method for genotyping elite breeding stocks of Radicchio, both local and modern varieties, assaying mapped SSR marker loci possibly linked to EST-rich regions and scoring PIC>0.5, was recently developed using multiplex PCRs [6]. Here, we are dealing with a research and development project aimed at sequencing and annotating the first draft of the leaf chicory genome as we believe it will have an extraordinary impact from both scientific and economic points of view. Indeed, the availability of the first

genome sequence for this plant species will provide a powerful tool to be exploited in the identification of markers associated with or genes responsible for relevant agronomic traits, influencing crop productivity and product quality. As an example, data and knowhow produced in this research project will be useful for detailed studies of the genetic control of male-sterility and self-incompatibility in this species.

The plant material that we used for the sequencing of the leaf chicory genome belongs to the Radicchio of Chioggia type, specifically to the male fertile inbred line named SEG111. This type was chosen as the most suitable accession based on the following criteria: i) the commercial relevance of the variety of origin; ii) the availability of clonal materials; iii) robust phenotypic and genotypic characterization; iv) a high degree of homozygosity (80%); and v) high breeding value as pollen parent of F1 hybrids. Sequencing reactions of the genomic DNA library were performed with Illumina HiSeq and MySeq platforms to combine the high number of reads originated by the former with the longer sequences produced by the latter. Here, we report original data from the bioinformatic assembly of the first genome draft of Radicchio, along with the most relevant findings that emerged from an extensive *de novo* gene prediction and *in silico* functional annotation of more than 18,000 unigenes. Analyses were performed according to established computational biology protocols by taking advantage of the publically available reference transcriptome data for *Cichorium intybus* [7]. The main preliminary findings on the genome organization and gene composition of Radicchio are presented, and the potentials of newly annotated expressed sequences and diagnostic microsatellite markers in breeding programs are critically discussed.

2. Materials and methods

2.1. Plant materials

Plant materials used for the sequencing belong to a variety of commercial relevance of the Radicchio of Chioggia type. The clone chosen derives from the inbred line SEG111 and shows a degree of homozygosity equal to 80% [6]. In particular, this clone was obtained by several cycles of selfing from plants yearly selected on the basis of a robust phenotypic and genotypic characterization, being also characterized by high-quality agronomic traits on farm and the ability to be easily cloned in vitro.

2.2. DNA isolation and sequencing

DNA was isolated from 150 mg of fresh leaf tissue using a CTAB-based protocol [8]. The eventual contamination of RNA was avoided with an RNase A (Sigma-Aldrich) treatment. DNA samples were eluted in 80–100 μL of 0.1× TE buffer (100 mM Tris-HCl 1, 0.1 mM EDTA, pH=8). The integrity of the extracted DNA samples was estimated through electrophoresis in 0.8% agarose/1× TAE gels containing 1× SYBR Safe DNA Gel Stain (Life Technologies, USA). The purity and quantity of the DNA extracts were assessed with a NanoDrop spectrophotometer (Thermo Scientific, USA). Then, 1 μg of high-quality DNA

was used for library preparation with the TruSeq DNA Sample Preparation chemistry (Illumina). Sequencing reactions were performed with the Illumina platforms: HiSeq (1 lane, 2 × 100 bp) and MySeq (1 lane, 2 × 300 bp).

2.3. *De novo* assembly and annotation

All high-quality reads generated from the two sequencing reactions were assembled in a single reference genome. Assemblies were attempted with three pieces of software: i) Velvet [9]; ii) SPAdes [10]; and iii) CLC Genomics Workbench 6.5 (Qiagen). The average coverage was estimated for the run HiSeq by calculating the frequency distribution of 25-mers [11].

To annotate all assembled contigs, a BLASTX-based approach was used to compare the *C. intybus* sequences to a subset of the NR protein collection that was made by focusing on the clade pentapetalae [12]. Moreover, the GI identifiers of the best BLASTX hits, having E-value ≤1.0E-15 and similarity ≥70%, were mapped to the UniprotKB protein database [13] to extract Gene Ontology annotations [14] and KEGG terms [15] for functional annotations. Further enrichment of enzyme annotations was made with the BLAST2GO software v1.3.3 using the function "direct GO to Enzyme annotation". The BLAST2GO software v1.3.3 [16, 17] was used to reduce the complexity of the data and perform basic statistics on ontological annotations, as reported by Galla *et al.* [18].

SSRs were detected among the 522.301 contigs via MISA [19]. The parameters were adjusted to identify perfect and complex mono-, di-, tri-, tetra-, penta-, and hexanucleotide motifs with a minimum of 49, 13, 9, 8, 8, and 8 repeats, respectively. Repeated elements were detected with a BLASTN-based approach using a PGSB Repeat Element Database in all blast searches [20]. The parameters set for the identification of Transposable Elements (TEs) were: reward 1, penalty 1, gap_open 2, gap_extend 2, word_size 9, dust no. An E-value cutoff of 1.0E-9 was adopted to filter the BLAST results.

Two public *C. intybus* transcriptomes CHI-2418 and CHI-Witloof originally developed from plant seedlings [7] corresponding to a wild accession of leaf chicory and a cultivated variety of witloof, respectively, were mapped to the reference genome using the CLC Genomics Workbench V7.02 (Qiagen). Mappings were performed with default mapping parameters, including mismatch cost: 2; insertion cost: 3; deletion cost: 3; length fraction: 0.5; and similarity fraction: 0.8. Non-specific matches were ignored and not included in the annotation tracks. For nucleotide variant analysis, the appropriate reference masking options were used to map transcriptome reads selectively over the sequences annotated as CDS or TEs. The variant detection analysis was done by using the Basic Variant Detection tool of the CLC Genomics Workbench V7.02 (Qiagen) with default parameters. As general filters, positions with coverage above 100,000 were not considered. Base quality filters were turned on and set to default parameters. All variants included in homopolymer regions with minimum length of 3nt, and with frequency below 0.8 were also removed from the dataset. As coverage and count filters, all variants with a minimum count lower than 20 were discarded.

3. Results

3.1. Genome assembly statistics

To obtain the first genome draft of leaf chicory, a single genomic library produced from the inbred line SEG111 was sequenced using the Illumina MySeq and HiSeq platforms. Here, we report the genome assembly results derived from the CLC Genomic Workbench assembly output. Figure 1 describes the frequency distribution of 25-mers in the HiSeq data.
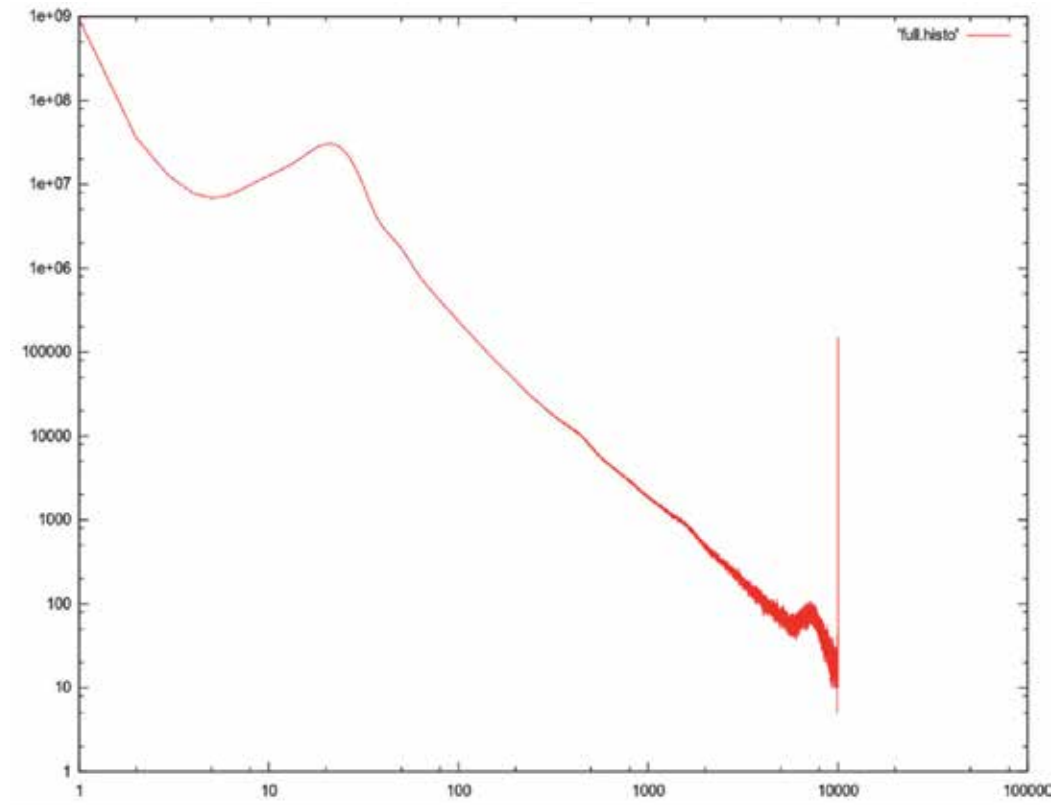
Figure 1. Frequency distribution of 25-mers in the HiSeq data (logarithmic scale for both axes)

The data shown suggest that the average coverage in the HiSeq run is approximately 21×. Additionally, the curve indicates that a certain number of sequences are present with a relatively high frequency within the genome. This might indicate that repeated elements are relatively abundant within the genome. As a consequence, the estimated size of the assembled genome draft is 760 Mb.

We obtained 58,392,530 and 389,385,400 raw reads through the MySeq and HiSeq platforms, respectively. The *de novo* assembly of the two datasets in a unique reference genome draft assembled 724,009,424 nucleotides into 522,301 contigs (Table 1). The maximum contig length

was equal to 379,698 bp, whereas the minimum contig length was set to 200 bp, with an average contig length of 1,386 bp. Overall statistics are summarized in Table 1.

Total number of contigs	522,301
Total No. of assembled nucleotides (nt)	724,009,424
GC percentage	34.8%
Average contig length (bp)	1,386
Minimum contig length (bp)	200
Maximum contig length (bp)	379,698
N75	1,051
N50	3,131

Table 1. Summary statistics of the sequence assembly generated from *Cichorium intybus*.

The length distribution of the contig size, expressed in base pairs, is reported in Figure 2.
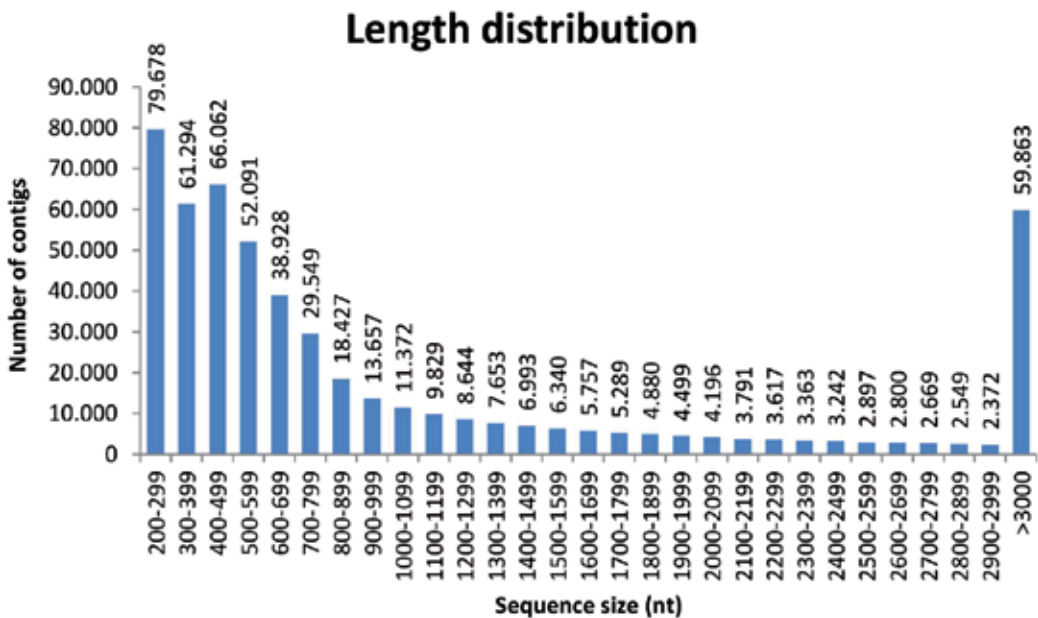
Figure 2. Distribution of length of contigs for leaf chicory

As much as 68.9% of the recovered sequences are contained within a length spanning from 200 nt to 999 nt. The interval length ranging between 1,000 nt and 2,999 nt is represented by 19.7% of the assembled contigs, whereas the proportion of contigs whose length is higher or equal to 3,000 nt corresponds to 11.5%.

We searched the genome sequence assembly for TEs and estimated their abundance using a BLASTN strategy. The proportion of base pairs annotated as TEs out of the total amount of assembled nucleotides was equal to 6.3% (Table 2).

Key	Classification	Number	Abundance (%)	Length (bp)	Percentage over the assembled genome
02.01	Class I retroelement	273	0.19	85,241	0.012%
02.01.01	LTR Retrotransposon	82,260	56.55	19,658,874	2.715%
02.01.01.05	Ty1/copia	35,802	24.61	17,519,102	2.420%
02.01.01.10	Ty3/gypsy	23,651	16.26	7,121,605	0.984%
02.01.02	non-LTR Retrotransposon	354	0.24	106,259	0.015%
02.05	Class II: DNA Transposon	1,976	1.36	713,119	0.098%
02	Unclassified mobile element	861	0.59	199,301	0.028%
10 / 90 / 99	High Copy Number Genes and additional attributes	283	0.19	51,577	0.007%
Total		145,462	100.0	45,455,078	6.278%

Table 2. Classification statistics of transposable elements (TEs) in Radicchio genome draft assembly.

The retroelements were the most abundant elements (>97% of the total). Within the major class of retroelements, Long Terminal Repeat (LTR) retrotransposons proved to be the dominant class (56.55%) in the leaf chicory genome. Moreover, the Copia-type (24.61%) and the Gypsy-type (16.26%) appeared to be the most abundant LTR retrotransposons. A total of 273 (0.2%) elements were annotated as retroelements, but they lacked the assignation to a specific class based on sequence similarity and conservation. Non-LTR retrotransposons were detected to a very low extent (0.24%). Less than 2% of the total repeat elements were annotated as DNA transposons.

3.2. Discovery of SSR loci

Overall, we identified 66,785 SSR containing regions. As many as 52,186 and 11,501 sequences proved to contain one or more microsatellites, respectively. These numbers included 1,226 mononucleotide SSR motifs (which were no longer taken into account for further computations). We found a total number of di- or multinucleotide SSR motifs equaling 65,559.

The most common SSR elements were those showing a dinucleotide motif (89.0%), followed by trinucleotide (7.1%) and tetranucleotide (3.0%) ones. Microsatellites revealing a pentanucleotide and hexanucleotide motif were less than 1.0% of the total. Overall data are summarized in Table 3.

Type of motif	Range of repeat numbers				Total No.	Percentage (%)
	8-12	**13-17**	**18-22**	**>22**		
Di-nucleotide	0	8,333	7,100	42,913	58,346	89.0
Tri-nucleotide	1,822	1,769	762	321	4,674	7.1
Tetra-nucleotide	1,114	475	205	202	1,996	3.0
Penta-nucleotide	69	23	0	2	94	0.1
Hexa-nucleotide	359	80	8	2	449	0.7
Total	3,364	10,680	8,075	43,440		
Percentage (%)	5.1	16.3	12.3	66.3		

Table 3. Number of SSRs detected in the Radicchio genome draft assembly. For each type of motif, the number of SSRs identified in the range of repeated numbers is reported. Albeit present in the genome, mono-nucleotide SSRs were not considered in this analysis.

3.3. Functional annotation of contig sequences

Functional annotation of the assembled contigs was performed with a BLASTX approach, according to which all contig sequences were used to query different public protein databases (Table 4).

Public database	No. of Hits (gene models)	No. of *C. intybus* contigs
NR	38,782	80,862
Arabidopsis	16,689	50,417
GO	14,073	45,381
KEGG	4,512	22,273

Table 4. Summary statistics of functional annotations for leaf chicory genome sequences in public protein databases. As for the NR database, only the protein sequences from the clade pentapetalae of eudicots were considered. The Arabidopsis proteome used in all BLAST analysis was TAIR10.

The database enclosing all public protein sequences belonging to the pentapetalae clade of the eudicots, which includes the sub-clades of rosids and asterids to which leaf chicory belongs, provided a total of 38,782 hits. The proteome of *Arabidopsis thaliana* alone scored 16,689 hits when an E-value cutoff of 1.0E-15 was applied for the screening of the most reliable BLASTX hits.

Two public *C. intybus* transcriptomes originally developed from plant seedlings and provided by UC DAVIS, the Compositae Genome Project (CHI-2418 and CHI-Witloof) [7] were mapped to the reference genome using the appropriate mapping function of the CLC Genomics Workbench.

By doing so, we were able to map 76.5% and 78.0% of the sequences, respectively. Data derived from the mapping of two *C. intybus* transcriptomes were used to integrate the annotation of the assembled contigs. BLAST and mapping data integration increased the BLAST-based annotation with an additional set of 1,995 contigs.

Arabidopsis matches were used to retrieve both GO and KEGG annotations from public databases. We could finally assign one or multiple GO terms to 45,381 leaf chicory genome contigs. The analysis performed against the GO illustrate 14,073 genes annotated with terms belonging to one or multiple vocabularies. Of these, 24,634 contigs were annotated for their putative biological process, 39,118 contigs were related to a molecular function, and 37,561 contigs were associated to a specific cellular component. Figure 3 shows the fine distribution of the 14,073 hits caught by our Radicchio contigs from the TAIR database according to the aforementioned three GO categories.
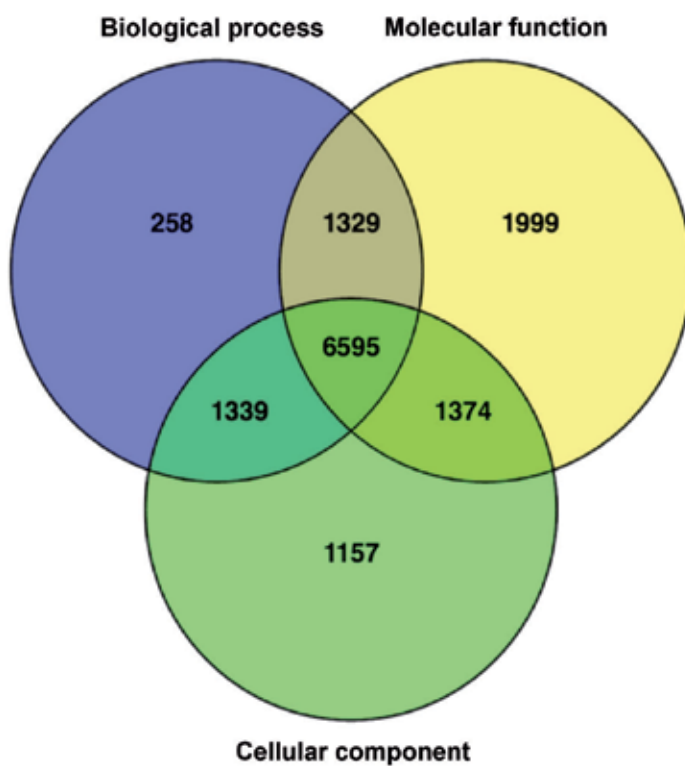
Figure 3. Venn diagram showing the fine distribution according to GO terms of the 14,073 *A. thaliana* hits matching our leaf chicory contigs

Among all the terms underlined by the GO vocabulary for the biological process, our investigations were focused on terms related to the response to biotic and abiotic stresses (Figure 4), hormonal responses (Figure 5), and flower and seed development (Figure 6). Of the 15 most

interesting processes for molecular breeding in leaf chicory, 7 and 8 were linked to biotic and abiotic stresses, respectively (see Figure 4). The ontological terms were assigned to 2,388 and 3,844 genome contigs, respectively.
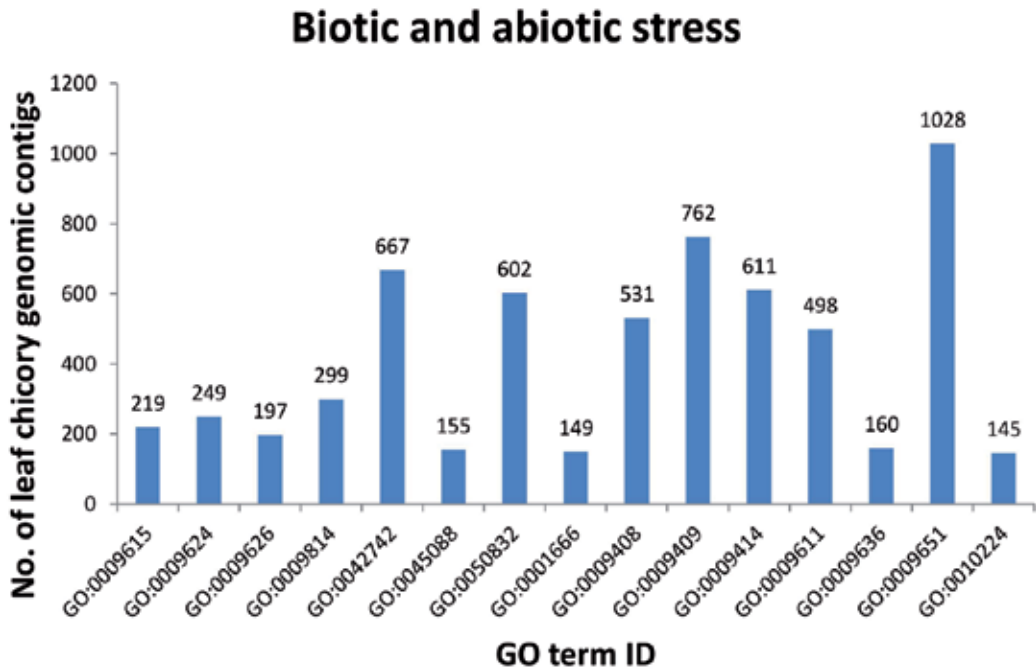
Figure 4. Number of *C. intybus* genomic contigs for response to biotic (the first 7) and abiotic (the last 8) stress

The computational analysis for the identification of SSR elements within these contigs unveiled 495 motifs linked to biotic stresses and 841 motifs associated with abiotic stresses. Among the biotic stresses, the most abundant gene ontology (GO) term was GO:0042742, which corresponds to the "defense response to bacterium" and shows a match with 667 genome contigs containing 135 microsatellites. Concerning the abiotic stresses, the GO term assigned with the higher frequency was GO:0009651, which accounts for processes related to "response to salt stress" and matches 1,028 genome contigs containing 249 microsatellites.

Data of hormonal responses and processes of flower and seed development are reported in Figures 5 and 6. The analysis for hormonal responses noted nine different GO terms, for a total of 3,344 genome contigs, and 833 SSR elements linked to these sequences and terms. In particular, the term "response to jasmonic acid stimulus" (GO:0009753) was the most represented, with 478 matches with different genome contigs, including 118 SSR motifs (Figure 5).

Results of the GO term annotation of genome contigs according to the flower and seed developmental processes are reported in Figure 6.

Hormonal responses
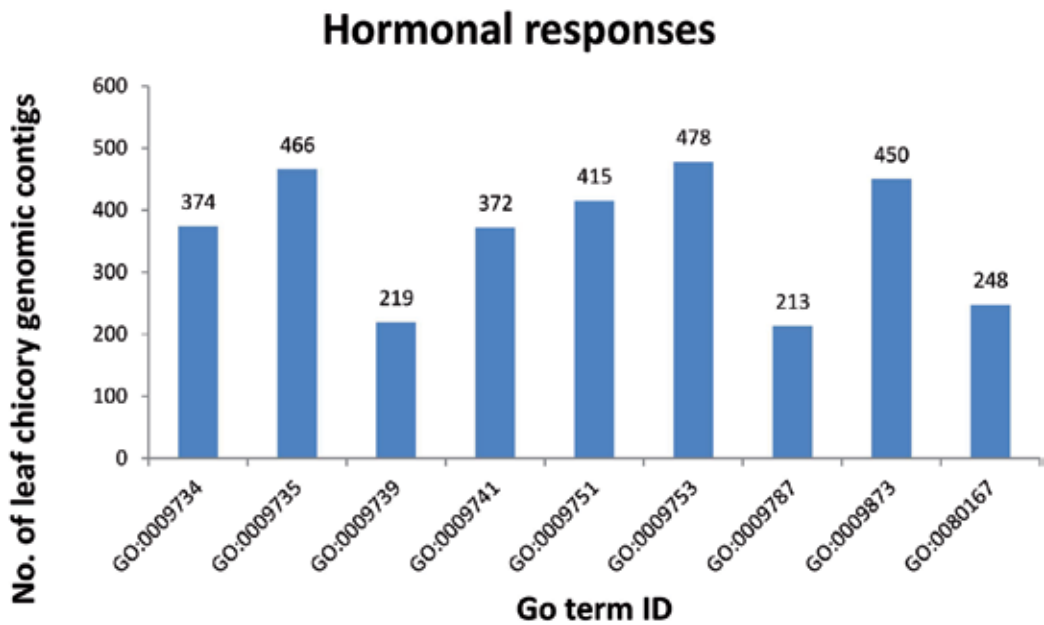
Figure 5. Number of *C. intybus* genomic contigs for hormonal response

Flower and seed developement
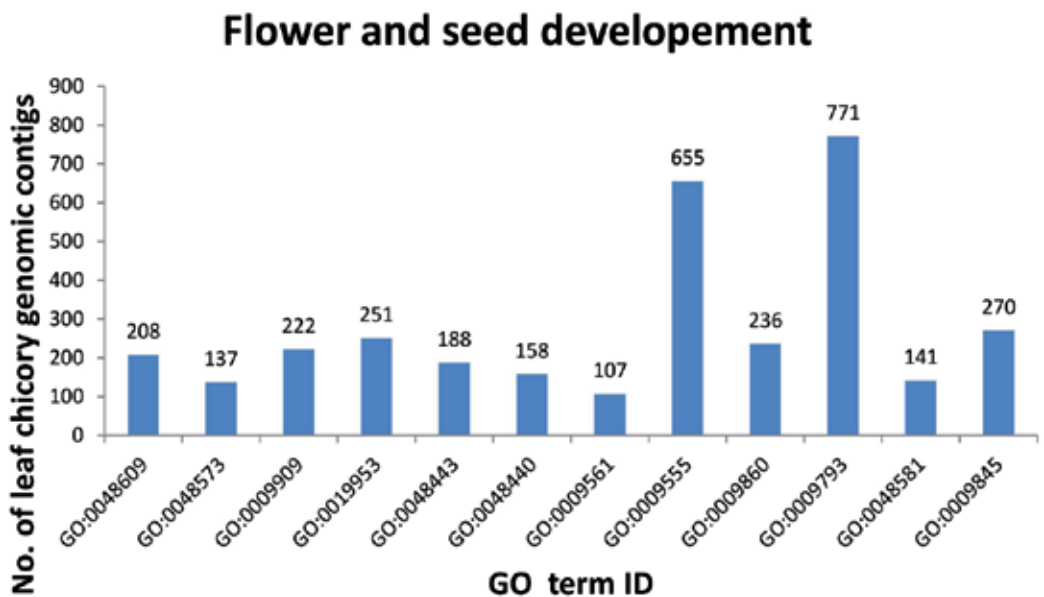
Figure 6. Number of *C. intybus* genomic contigs for flower and seed (only the last 3) development

The flower development process was embraced by selecting nine ontological terms, whereas three terms were assigned to seed development and seed germination. A total of 2,162 contigs

were annotated with GO terms related to flower development; 496 of these were also annotated for the presence of one or multiple SSRs. In particular, the term "pollen development" (GO: 0009555) was the most abundant, with 655 contigs containing 153 SSR motifs.

As far as the seed development process is concerned, we annotated 1,182 contigs linked to this GO term, 273 of which co-localized with one or multiple SSRs. Among these, the most abundant ontological term was "embryo development ending in seed dormancy" (GO: 0009793) as it is assigned to 771 contigs, co-localizing with 171 SSR elements.

Using the Kyoto Encyclopaedia of Genes and Genomes database (http://www.genome.jp/kegg/), a total of 22,273 contigs enabled the mapping of 795 enzymes to 157 metabolic pathways. Among the metabolic pathways with the highest number of mapped reads, we found fructose and mannose metabolism (418 gene models matched), phenylpropanoid biosynthesis (415 gene models matched) and tryptophan metabolism (380 gene models matched). The biosynthetic pathway of flavonoid biosynthesis, described in map:00941, is relevant as the biosynthesis of flavonoid is directly connected to the synthesis of anthocyanin (Figure 7), whose accumulation contributes to the pigmentation of leaf chicories. This map includes 236 gene models that were assigned to 14 unique enzymes, including CHS (CHALCONE SYNTHASE), CHI (CHALCONE ISOMERASE), and ANS (ANTHOCYANIDIN SYNTHASE), among others.
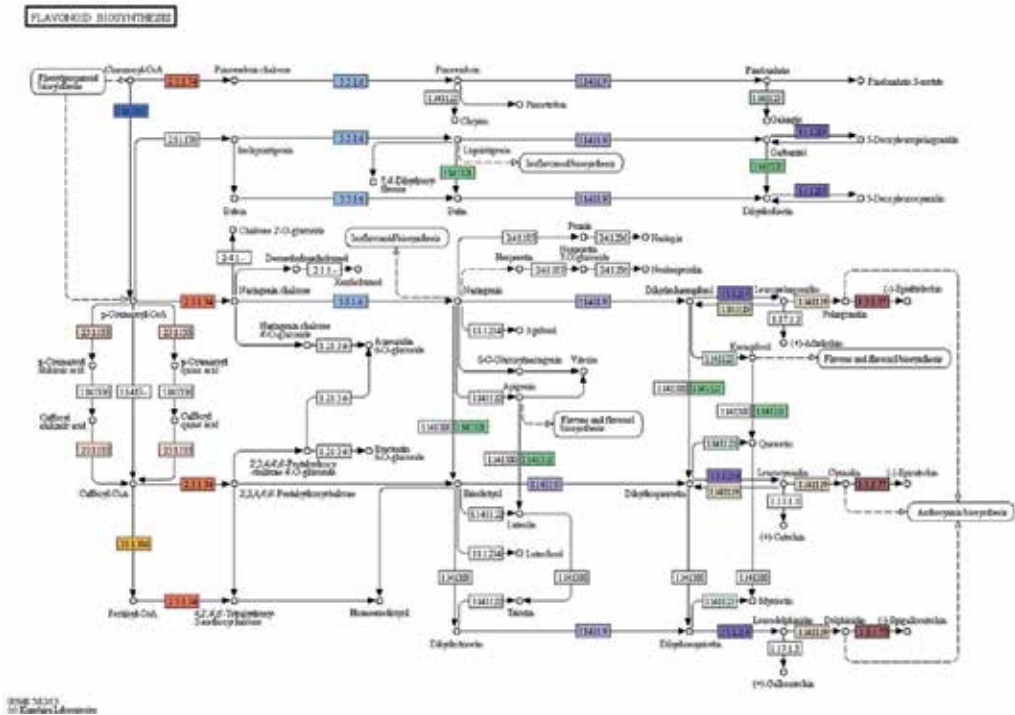
Figure 7. KEGG pathway for flavonoid biosynthesis (Map:00941)

KEGG data related to a number of selected metabolic pathways were exploited to find SSR regions potentially associated with highly valuable phenotypes in this plant species. The number of SSRs putatively linked to the most interesting phenotypic traits with breeding values in leaf chicory is displayed in Table 5.

KEGG map ID	Metabolic pathway	Characteristic	No. of SSRs
map00909	Sesquiterpenoid and triterpenoid biosynthesis	Bitter taste	107
map00053	Ascorbate and aldarate metabolism	Vitamin C content	172
map00940	Phenylpropanoid biosynthesis	Leaf color	281
map00941	Flavonoid biosynthesis	Leaf color	173
map00942	Anthocyanin biosynthesis	Leaf color	180
map00943	Isoflavonoid biosynthesis	Leaf color	5
map00944	Flavone and flavonol biosynthesis	Leaf color	128
map00040	Pentose and glucuronate interconversions	Response to cold	96
map00051	Fructose and mannose metabolism	Response to cold	259
map00052	Galactose metabolism	Response to cold	31
map00061	Fatty acid biosynthesis	Response to cold	39
map00260	Glycine, serine and threonine metabolism	Response to cold	60
map00290	Valine, leucine and isoleucine biosynthesis	Response to cold	13
map00330	Arginine and proline metabolism	Response to cold	55
map00410	beta-Alanine metabolism	Response to cold	16
map00480	Glutathione metabolism	Response to cold	48
map00500	Starch and sucrosa metabolism	Response to cold	164
map00561	Glycerolipid metabolism	Response to cold	159
map00564	Glycerophospholipid metabolism	Response to cold	124
map00592	alpha-Linolenic acid metabolism	Response to cold	66
map00710	Calvin cycle	Response to cold	28
map00780	Biotin metabolism	Response to cold	18
map00960	Tropane, piperidine and pyridine alkaloid biosynthesis	Response to cold	97

Table 5. Number of SSRs located in contig sequences annotated for the presence of proteins with known enzymatic activity in relevant metabolic pathways for the breeding of leaf chicory.

Considering the overall grouping of selected metabolic pathways, we identified many microsatellite sequences putatively linked to important traits, according to their potential effect on plant characteristics. For instance, 107 SSRs were linked to bitter taste, 172 SSRs were associated with vitamin C biosynthesis and metabolism, and 767 SSRs located in sequence contigs encoding enzymes of the flavonoid and anthocyanin biosynthetic pathways, thus potentially associated with the leaf color. The most represented characteristic is the response to cold. For this trait, we analyzed 16 different metabolic pathways that altogether led to the selection of 1,273 microsatellites potentially associated with one or multiple genes actively involved in the plant response to cold eventually, but not exclusively, through the accumulation of sugar.

We also performed the calling of nucleotide variants. Stringent quality criteria were used for discriminating sequence variations from sequencing errors and mutations introduced during cDNA synthesis. Only sequence variations with mapping quality scores over the established thresholds were annotated, leading to the identification of 123,943 and 121,086 variants that were present only in the leaf chicory transcriptome CHI-2418 (wild type) or the Witloof transcriptome CHI-Witloof (cultivated type), respectively. A total of 119,729 variants were shared by both *C. intybus* transcriptomes. The average number of variants per contig ranged from 9.5 to 10.5 in the two assemblies (Table 6), yielding one single variation per 100 bp in both cases.

	Radicchio CDS – 29,175 contigs			Radicchio TEs – 122,745 contigs		
	CHI-2418	**CHI-Witloof**	**Shared**	**CHI-2418**	**CHI-Witloof**	**Shared**
No. contigs	12,725	12,739	11,419	2,016	1,924	1,554
No. variants	123,843	121,086	119,729	10,662	10,651	10,246
No. variants/contigs	9.75	9.52	10.52	5.29	5.54	6.61
No. variants/100 bp	0.99 (1.14)	0.98 (1.14)	1.14 (2.05)	3.26 (3.64)	3.16 (3.50)	5.42 (8.88)
SNVs	115,678	113,049	107,255	9,532	9,605	9,006
MNVs	5,367	5,439	8,475	507	441	261
Insertions	2,044	2,036	2,166	556	552	714
Deletions	754	562	1,833	67	53	265

Table 6. Summary statistics of nucleotide variants restricted to gemonic regions of Radicchio annotated as CDS and Transposable Elements. Nucleotide variants were detected by using the transcriptomes CHI-2418 (wild type leaf chicory) and CHI-Witloof (cultivated Witloof type). For each transcriptome, the number of contigs displaying one or multiple variants, the number of variants and the number of variants per contigs are indicated. The number of variants per 100bp is also reported. Variants present in both transcriptomes are indicated as shared.

The vast majority of variants were Single Nucleotide Variants (SNVs), whereas Multi Nucleotide Variants (MNVs), Insertions, and Deletions were found to a considerably lower extent (Table 6). On average, the proportion of SNVs and MNVs was comparable in the CDS and TE contigs and equal to about 90% and 5%, respectively.

Among all contigs annotated as TEs, those characterized by the presence of one or multiple variants were 10,662 and 10,651 for the two transcriptomes (Table 6). The average number of variants per contig was equal to 5.3 and 5.5. Despite the relatively low abundance of polymorphic residues in these regions, the average number of variants per 100 bp was equal to 3.3 and 3.2. Single Nucleotide Polymorphisms (SNPs) were by far the most abundant type of variants in TEs as well as in CDS regions (Table 6). In particular, transversions and transitions were on average 37% (ranging from 35.6% and 37.8%) and 63% (ranging from 62.2% and 64.4%) of the point mutations, respectively. The total number of nonsynonymous SNPs calculated with the reference transcriptomes was equal to 13,559 (10.9%) and 11,197 (9.2%) for wild-type leaf chicory and cultivated Witloof accessions, respectively.

4. Discussion

Here, we report the uncovering of the first draft of the Radicchio genome. This highly relevant discovery was achieved by combining the recent advancement of next-generation sequencing technologies on the public side with the significant investment of financial resources in research and development on the private side.

Currently, conventional agronomic-based selection methods are supported by molecular marker-assisted breeding schemes. In recent years, we have demonstrated that the constitution of F1 hybrids is not only feasible in a small experimental scheme but also realizable and profitable on a large commercial scale (*e.g.*, registered CPVO varieties TT4070/F1, TT5010/F1, TT5070/F1, and TT4010/F1 in progress). F1 hybrids are varieties manifesting heterosis, or hybrid vigor, which refers to the phenomenon in which highly heterozygous progeny plants obtained by crossing genetically divergent inbred or pure lines exhibit greater biomass, faster speed of development, higher resistance to pests and better adaptation to environmental stresses than the two homozygous parents. Critical steps of an applicative breeding program are the production of parental inbreds. Two highly relevant factors in this context are the selection of self-compatible genotypes, to be used as pollen donors, and the identification of male-sterile genotypes, to be used as seed parents in large-scale crosses [21, 22].

It is worth mentioning that there are several reasons why the constitution of F1 hybrids is a strategic choice for a seed company. First, the crop yield of modern F1 hybrid varieties is usually much higher than that of traditional OP or synthetic varieties. Second, the uniformity of F1 populations and the way to legally protect their parental lines allow a seed company to adopt a plant breeder's rights, promoting genetic research and development programs that are very expensive and require many years. Finally, the need for breeding hybrid varieties also promotes the preservation of local varieties because the selection of appropriate inbred or pure lines as parents in pairwise cross-combinations requires the exploration and exploitation of germplasm resources. Our expectation is that F1 hybrid varieties will be bred and adopted with increasing frequency in Radicchio. Consequently, we invested in the sequencing and annotation of the first draft of the leaf chicory genome as it will have an extraordinary impact from both scientific and economic points of view. Indeed, the availability of the first genome

sequence for this plant species will provide a powerful tool to be exploited in the identification of markers associated with or genes responsible for relevant agronomic traits, influencing crop productivity and product quality. As an example, data and knowhow produced in this research project will be capitalized on in subsequent years to plan and develop basic studies and applied research on male-sterility and self-incompatibility in this species.

The availability of high-quality sequencing platforms (*i.e.*, Illumina) on the one hand, and specific and high-performing software for genome data assembly and gene set analysis on the other, made this project feasible. High-quality genomic DNA libraries were used for sequencing reactions performed with the Illumina platforms HiSeq and MySeq, originating a total of 197 million (mln) short reads and 29 mln longer sequences passing quality filters, respectively, which were then bioinformatically assembled to obtain the first genome draft. On the basis of this strategy, the genome draft of leaf chicory is composed of approximately 500,000 contigs, forming approximately 720 Mb. Based on the distribution of 25-mer frequencies, we estimated that the genome coverage is close to 25X. The same distribution also indicates that a significant part of the genome might be composed of highly repeated elements, as indicated by the number of k-mers that appears to be present with high frequency.

Nucleotide variant calling for the Radicchio genome showed comparable number of polymorphisms in the pairwise comparisons with the two publically available transcriptomes, originally developed from seedlings of two leaf chicory accessions (*i.e.*, wild and cultivated types). The total number of variants discovered in the CDS regions was shown to be approximately 10 times higher than the ones found in the TEs. This result might be a consequence of low expression, or silencing, of numerous transposable elements at the level of plant seedlings, as indicated by the finding that the mapping of the two transcriptomes to the reference genome failed to align sequences to about 98% of the contigs annotated as TEs. Noteworthy, the number of variations per 100 base pairs was significantly higher in the TEs than in the CSD sequences. This result might be explained by the accumulation of mutations in noncoding sequences, as most of the TEs are.

Overall, Single Nucleotide Variants (SNVs) were the most common variants compared with In/Del mutations. Since SNP mutations very often result in silent mutations, their high proportion in the CDS regions was an expected result. In/Del mutations that usually occur in silenced or functionally disrupted genes, along with noncoding regions, were found at a low rate in CDS regions.

TEs were found to occur, at least in one copy, in the 23.50% of the 522,301 contigs that constitute our chicory genome draft assembly. Retrotransposons proved to be the most abundant elements in the Radicchio genome. This finding is in agreement with data from previous studies [23-26]. It is worth mentioning that Copia-type elements were more abundant than Gypsy-type elements, forming the predominant subclass of LTR retrotransposons.

Although the amount of TEs of the totally assembled sequences was much lower than that reported for other species, the class ratio of the TE types corresponds to that found in previous studies [23-26]. Our estimate of TEs in leaf chicory is equal to 6.28% of the contigs length, which is much lower than amounts reported for soybean (59%), pigeonpea (52%), alfalfa (27%), trefoil

(34%), and chickpea (40%) [25, 27-30]. One of the reasons could be that our BLAST strategy chosen to find repeated elements in the genome was less efficient than specific software (*e.g.*, RepeatScout and RepeatMasker [31, 32]). Another reason could be the lack of TEs in the assembled portion of the Radicchio genome due to the low complexity of these repeated DNA regions.

The BLAST strategy with the nonredundant (NR) pentapetalae protein database produced the best output in terms of similarity with our contigs. This is undoubtedly due to the availability of large collections of sequences from species taxonomically related to leaf chicory, such as *Beta vulgaris*, *Helianthus annuus,* and *Lactuca sativa*, among others. Unfortunately, the depth of annotation of these recently sequenced genomes is frequently not comparable to that of the long-studied *Arabidopsis thaliana*. Although BLAST results obtained by querying the NR database proved to be highly informative in terms of the number of hits producing alignments with significant e-value, the annotation of the leaf chicory assembled contigs was more successful when the *A. thaliana* database was used alone. Therefore, a possible alternative for future enrichment of the current annotation state would imply the use of software (*e.g.*, Blast2GO) that could extract the annotation codes from multiple BLAST hits, provide the appropriate specificity cutoff, and assign the mapped GO terms to the original query.

Our choice to use the TAIR10 database to annotate our sequence contigs led to the annotation of a large number of assembled sequences and provided precious information concerning the putative process, or eventually, the metabolic pathways in which genes are putatively active.

The ability to annotate a certain number of sequences is not only exclusively dictated by the length and quality of the query sequences but also by their match with orthologous sequences that need to be annotated in depth.

This would be the case of annotations for metabolic pathways not actively studied or present in *A. thaliana* and for processes whose study is hampered by biological or physical circum-stances. This might explain some discrepancies in annotations for male and female gameto-genesis (Figure 6). From the graph, it is easy to understand the large discrepancy between the number of contigs presented for the term "Megagametogenesis" (GO:0009561), just 107, and the term "Pollen development" (GO:0009555), cited in the results as the most prevalent (more than six times that of megagametogenesis). We can suppose that this difference might not be due to a real difference in the number of genes involved in these two reproductive processes but rather to the lower number of genes known to be involved in female sporogenesis and gametogenesis.

Similarly, enzymes involved in the biosynthesis of germacren-type sesquiterpenoids, such as the germacrene-A synthase (EC:4.2.3.23), which are responsible for the biosynthesis of lactones associated with bitter taste in leaf chicory, are not known or properly characterized in *A. thaliana*.

Another fundamental finding of our study is the large number of SSR markers that were found in the assembled contigs. We can affirm that the leaf chicory genome shows an unexpected number and distribution of repeated sequences. Submitting our Radicchio draft to MISA software, we were able to reveal such a number of potential SSR markers. It is

therefore interesting that we were able to link a reasonably large number of microsatellites to each item here presented for both GO terms and KEGG maps. In the results, we presented only a small selection of important characteristics that could be utilized in marker-assisted selection and breeding programs in Radicchio. Together with SSRs, thousands of sequences that could be used in Single Nucleotide Polymorphism (SNP) analysis were associated to fundamental biosynthetic pathways or metabolism enzymes. This is a crucial starting point for modern breeding in leaf chicories.

It is noteworthy that further studies must be conducted to determine whether and how these potential markers could be exploited in molecular breeding programs. As a final step, gene prediction and annotation were also performed according to established computational biology protocols by taking advantage of the reference transcriptome data publically available for *Cichorium intybus* L. These sequences allowed us to learn the number, sequence, and role of the ~25.000 genes of the Radicchio's genome. This finding represents an important achievement for Italian agriculture genetics as a whole and opens new perspectives in both basic and applied research programs in Radicchio. It will have great impacts, potentials, and advantages in terms of breeding methods and tools useful for the constitution and protection of new varieties. Information obtained by the sequencing of the genome will be exploitable to detect and dissect the chromosomal regions where the genetic factors that control the expression of important agronomic and qualitative traits are located in Radicchio.

Modern marker-assisted breeding (MAB) technology based on traditional methods using molecular markers such as SSRs and SNPs, without relations to genetic modification (GM) techniques, will now be planned and adopted for breeding of vigorous and uniform F1 hybrids combining quality, uniformity, and productivity traits in the same genotypes.

In conclusion, our study will contribute to increase and reinforce the reliability of Italian seed firms and local activities of the Veneto region associated with the cultivation and commercialization of Radicchio plant varieties and food products; the seed market of this species will have the chance to become highly professional and more competitive at the national and international levels. To uncover the sequence of a given genome means to gain a robust scientific background and technological knowhow, which in short time can play a crucial role in addressing and solving issues related to the cultivation and protection of modern Radicchio varieties. In fact, we are confident that our efforts will extend the current knowledge of the genome organization and gene composition of leaf chicories, which is crucial in the development of new tools and diagnostic markers useful for our breeding strategies, and allow researchers for more focused studies on chromosome regions controlling relevant agronomic traits of Radicchio. In addition, conducting novel research programs for the preservation and valorization of the biodiversity, still present in the Radicchio germplasm of the Veneto region, is very important and accomplished through the genetic characterization of the most locally dominant and historically important landraces using sequenced genome information of Radicchio presented in this work.

Author details

Giulio Galla[1], Andrea Ghedina[1], Silvano C. Tiozzo[2] and Gianni Barcaccia[1*]

*Address all correspondence to: gianni.barcaccia@unipd.it

1 DAFNAE, Laboratory of Plant Genetics and Breeding, University of Padua, Legnaro, PD, Italy

2 T&T Produce, Sant'Anna di Chioggia, VE, Italy

References

[1] Lucchin M, Varotto S, Barcaccia G, Parrini P. Chicory and Endive. In: Springer Science, editor. *Handbook of Plant Breeding.* New York: 2008. p. 1-46. DOI: 10.1007/978-0-387-30443-4_1

[2] Barcaccia G, Pallottini L, Soattin M. Genomic DNA fingerprints as a tool for identifying cultivated types of radicchio (*Cichorium intybus* L.) from Veneto, Italy. *Plant Breed*. 2003;122:178-183. DOI: 10.1046/j.1439-0523.2003.00786.x

[3] Cadalen T, Morchen M, Blassiau C, Clabaut A, Scheer I, Hilbert JL *et al*. Development of SSR markers and construction of a consensus genetic map for chicory (*Cichorium intybus* L.). *Mol Breed*. 2010;25:699-722. DOI: 10.1007/s11032-009-9369-5

[4] Gonthier L, Blassiau C, Mörchen M, Cadalen T, Poiret M, Hendriks T, Quillet MC. High-density genetic maps for loci involved in nuclear male sterility (NMS1) and sporophytic self-incompatibility (S-locus) in chicory (*Cichorium intybus* L., Asteraceae). *Theor Appl Genet*. 2013;126(8):2103-2121. DOI: 10.1007/s00122-013-2122-9

[5] Muys C, Thienpont CN, Dauchot N, Maudoux O, Draye X, Van Cutsem P. Integration of AFLPs, SSRs and SNPs markers into a new genetic map of industrial chicory (*Cichorium intybus* L. var. *sativum*). *Plant Breed*. 2014;133(1):130-137. DOI: 10.1111/pbr.12113

[6] Ghedina A, Galla G, Cadalen T, Hilbert JL, Tiozzo CS, Barcaccia G. A method for genotyping elite breeding stocks of leaf chicory (*Cichorium intybus* L.). *BMC Research Notes*. Forthcoming.

[7] UC Davis. The Compositae Genome Project [Internet]. Available from: http://compgenomics.ucdavis.edu/cgp_wd_assemblies.php#13427 [Accessed: June 2015]

[8] Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull*. 1987;19:11-15.

[9]   R. Zerbino DR, Birney E. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res*. 2008;18(5):821-829. DOI: 10.1101/gr.074492.107

[10] Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 2012;19(5):455-477. DOI: 10.1089/cmb.2012.0021

[11] Cornell University Library - Binghang Liu, Yujian Shi, Jianying Yuan, Xuesong Hu, Hao Zhang, Nan Li, Zhenyu Li, Yanxiang Chen, Desheng Mu, Wei Fan. Estimation of genomic characteristics by analyzing kmer [Internet]. 2013. Available from: http://arxiv.org/abs/1308.2012

[12] NCBI [Internet]. Available from: http://www.ncbi.nlm.nih.gov/

[13] UniProt [Internet]. Available from: http://www.uniprot.org/

[14] Gene Ontology Consortium [Internet]. [Updated: http://geneontology.org/].

[15] KEGG: Kyoto Encyclopedia of Genes and Genomes [Internet]. Available from: http://www.genome.jp/kegg/

[16] Blast2GO [Internet]. Available from: https://www.blast2go.com/

[17] Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*. 2005;21(18):3674-3676. DOI: 10.1093/bioinformatics/bti610

[18] Galla G, Barcaccia G, Ramina A, Collani S, Alagna F, Baldoni L, Cultrera N. GM, Martinelli F, Sebastiani L, Tonutti P. Computational annotation of genes differentially expressed along olive fruit development. *BMC Plant Biol*. 2009;9(128). DOI: 10.1186/1471-2229-9-128

[19] MISA - MIcroSAtellite identification tool [Internet]. [Updated: 5/14/02]. Available from: http://pgrc.ipk-gatersleben.de/misa/

[20] Helmholtz Zentrum München - Deutsches Forschungszentrum für Gesundheit und Umwelt (GmbH). PGSB Plant Genome and Systems Biology [Internet]. Available from: http://pgsb.helmholtz-muenchen.de/plant/recat/

[21] Barcaccia G, Tiozzo CS. New male sterile Cichorium spp. mutant, parts or derivatives, where male sterility is due to a recessive nuclear mutation linked to a polymorphic molecular marker, useful for producing F1 hybrids of Cichorium spp. EU PATENT No. WO2012163389-A1. 2012;DOI: 1013140/2.1.4749.5044

[22] Barcaccia G, Tiozzo CS. New male sterile mutant of leaf chicory, including Radicchio, used to produce chicory plants and seeds with traits such as male sterility exhibiting cytological phenotype with shapeless, small and shrunken microspores in dehiscent anthers. USA PATENT No. US20140157448-A1. 2014;DOI: 10.13140/2.1.3176.6400

[23] Chan AP, Crabtree J, Zhao Q, Lorenzi H, Orvis J, Puiu D, Melake-Berhan A, Jones KM, Redman J, Chen G, Cahoon EB, Gedil M, Stanke M, Haas BJ, Wortman JR, Fraser-Liggett CM, Ravel J, Rabinowicz PD. Draft genome sequence of the oilseed species *Ricinus communis*. *Natur Biotechnol*. 2010;28:951-956. DOI: doi:10.1038/nbt.1674

[24] Rahman AYA, Usharraj AO, Misra BB, Thottathil GP, Jayasekaran K, Feng Y, Hou S, Ong SY, Ng FL, Lee LS, Tan HS, Sakaff MKLM, Teh BS, Khoo BF, Badai SS, Aziz NA, Yuryev A, Knudsen B, Dionne-Laporte A, Mchunu NP, Yu Q, Langston BJ, Freitas TAK, Young AG, Chen R, Wang L, Najimudin N, Saito JA, Alam M. Draft genome sequence of the rubber tree *Hevea brasiliensis*. *BMC Genomics*. 2013;14(75). DOI: 10.1186/1471-2164-14-75

[25] Jain M, Misra G, Patel RG, Priya P, Jhanwar S, Khan AW, Shah N, Singh VK, Garg R, Jeena G, Yadav M, Kant C, Sharma P, Yadav G, Bhatia S, Tyagi AK, Chattopadhyay D. A draft genome sequence of the pulse crop chickpea (*Cicer arietinum* L.). *Plant J*. 2013;74:715-729. DOI: 10.1111/tpj.12173

[26] He N, Zhang C, Qi X, Zhao S, Tao Y, Yang G, Lee T-H, Wang X, Cai Q, Li D, Lu M, Liao S, Luo G,He R, Tan X, Xu Y, Li T, Zhao A, Jia L, Fu Q, Zeng Q, Gao C, Ma B, Liang J, Wang X, Shang J, Song P, Wu H, Fan L, Wang Q, Shuai Q, Zhu J, Wei C, Zhu-Salzman K, Jin D, Wang J, Liu T, Yu M, Tang C, Wang Z, Dai F, Chen J, Liu J, Zhao S, Lin T, Zhang S, Wang J, Wang J, Yang H, Yang G, Wang J, Paterson A.H, Xia Q, Ji Q, Xiangc Z. Draft genome sequence of the mulberry tree *Morus notabilis*. *Natur Commun*. 2013;4. DOI: 10.1038/ncomms3445

[27] Sato S, Nakamura Y, Kaneko T *et al*. Genome structure of the legume, *Lotus japonicus*. *DNA Res*. 2008;15(4):227-239. DOI: 10.1093/dnares/dsn008

[28] Schmutz J, Cannon SB, Schlueter J *et al*. Genome sequence of the palaeopolyploid soybean. *Nature*. 2010;463:178-183. DOI: 10.1038/nature08670

[29] Varshney RK, Chen W, Li Y *et al*. Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. *Natur Biotechnol*. 2012;30:83-89. DOI: 10.1038/nbt.2022

[30] Young ND, Debelle F, Oldroyd GE *et al*. The *Medicago* genome provides insight into the evolution of rhizobial symbioses. *Nature*. 2011;480:520-524. DOI: 10.1038/nature10625

[31] Price AL, Jones NC, Pevzner PA. De novo identification of repeat families in large genomes. *Bioinformatics*. 2005;21:351-358. DOI: 10.1093/bioinformatics/bti1018

[32] Smit AFA, Hubley R, Green P. Repeat Masker Open-4.0 [Internet]. 2013-2015. Available from: http://www.repeatmasker.org

Sequencing of Non-model Plants for Understanding the Physiological Responses in Plants

Jannette Alonso-Herrada, Ismael Urrutia, Tania Escobar-Feregrino,
Porfirio Gutiérrez-Martínez, Ana Angélica Feregrino-Pérez,
Irineo Torres-Pacheco, Ramón G. Guevara González,
Sergio Casas-Flores and Andrés Cruz-Hernández

Additional information is available at the end of the chapter

Abstract

From a genomic point of view, plants are complex organisms. Plants adapt to the environment, by developing different physiological and genetic properties, changing their genomic and expression profiles of adaptive factors, as exemplified by polyploidy studies. These characteristics along with the presence of duplicated genes/genomes make sequencing with early low-throughput DNA sequencing technologies in plants a challenging task. With the development of new technologies for molecular analysis, including transcriptome, proteome or microarray profiling, a new perspective in the genomic analysis was open, making possible to programs in species without genomic maps. The opportunity to extend molecular studies from laboratory model scale toward naturally occurring plant populations made it possible to precisely answer the longstanding important ecological and evolutionary questions. Some plant species have unique properties that could help to understand their adaptability to environment, crop production, pest protection or other biological processes. Molecular studies on non-model plants, including algae, mosses, ferns and plants with very specific characteristics are ongoing.

Keywords: Genome size, NGS, polyploidy, transcriptome, wild materials

1. Introduction

The first wave of plant genome sequencing has passed, and now new era has started in plant genomics research with new-generation sequence (NGS) strategies, which require a mixture of economic and scientific needs. Until now, several crops have been sequenced and some

others crop´s sequencing is underway, which will greatly help to elucidate unknown biological processes and the phylogenetic relationship among crop plants. Furthermore, the genomic data analysis and its integration of the biological systems will help to establish fundamental models to understand the evolution, development, and adaptability of the plants.

A genomic sequence is an important information for the basic research and understanding of plant evolution and development. It serves as a tool for engineering new genotypes [1]. Different plant species have different amounts of DNA [2]. The DNA of most plants includes from 100 million to the largest example of 150 billion base pairs (designed as alphabets), organized into 20,000 to 50,000 genes [1]. The most important contribution in the field of plant genome analysis is the discovery that many higher plants share a blueprint gene content. As distantly related plant taxa, monocots and dicots, which diverged from a common ancestor about 200 million years ago, retain some common gene order along the genome [1].

The development of new strategies and technologies for genome sequencing, can lead to development of programs to get a partial (transcriptome) or a complete DNA sequence (whole genome sequencing (WGS)) for non-model plants. The costs for these projects are now accessible. The first plant genome sequencing project, represented an effort of several years and millions of dollars. Now, the costs for the sequencing are in the order of thousands of dollars and there are new bioinformatics tools available for the analysis of generated sequence data.

The world's population depends on a few crops such as rice, wheat, maize, and potato for their food. In the following decades the world will face the tremendous challenge of feeding the global population [3]. The study of plant genomics in non-model plants will help to reveal the genetic factors and biochemical pathways involved in many processes such as flowering, nutrition, disease, and pest resistance, as well as tolerance of plants to abiotic stresses.

Model organisms are important for biological and agricultural approaches. The research in model organisms has generated a huge amount of important information on different molecular factors that contribute to plant growth and development, however it has some limitations [4]. The study of wild plants will help to overcome these limitations. Wild plants are well adapted to extreme conditions, and resistant to plant pathogens.

Model organisms, have a limited number of uniform narrow-based genotypic samples or variability limited to a number of specific plants. The study on how some plants survive in extreme conditions may also provide some clues about the mechanisms of plant response to biotic or abiotic stresses. Some non-model organisms are an extraordinary source of plant secondary metabolites.

The wild plant genotypes sometimes does not look attractive for breeding programs because of their morphology; however, they are the repository of ancestral genes and very important sources for the rescue of specific traits.

In Mexico, there are a large number of wild plant populations fit for breeding or sequencing programs are yet to be identified. Wild plant populations are genetically diverse and are source of genes that encode proteins potentially used for health, industrial, or ecological purposes.

2. Non model plants as model for environmental adaptation

The cactus plants of Cactaceae family are an example for a plant that can adapt to several environmental conditions. One of these plants is Nopal (*Opuntia* spp.), which belongs to the genera Opuntia and Nopalea [5]. This is an endemic plant found in semiarid areas in Mexico, but it grows along the American continent, from Canada to La Patagonia in Argentina, where environmental conditions are different from each area. Recently, Nopal plants have become the world´s interesting alternative fruit and forage crop. Only few varieties of nopal fruit originated from the Mexican nopal germplasm, are available in the market.

The history of first Nopal use in Mexico dates back to the ancient Mesoamerican civilizations; people used to collect cladodes and fruits from wild materials, for their nutritional qualities and medicinal purposes. The Spanish conquerors spread Nopal in America and Europe; now it is cultivated in Italy, Morocco, Tunisia, Greece, Israel, India, Philippines, China, Australia, South Africa, Brazil, Argentina, Colombia and the United States [6-8].

Although nopal is propagated asexually for commercial purposes, seed propagation is essential for breeding. Nopal apomixis makes the screening of individual crops obtained from crosses difficult and complicates the genetic studies [5]. Although no genomic map exists for this multipurpose plant, several efforts have been made to get some genomic approaches, and it has been included in the 1000 genomes sequencing program. To date, extensive efforts on cDNA microarrays, microRNAs (miRNAs) microarrays, mRNA deep sequencing and molecular markers studies have been made.

To study the genes associated with crassulacean acid metabolism (CAM), an expressed sequence tag (EST) database of different developmental stages of various tissues was created [9]. Sequences were assembled and compared with the available plant and genetic databases; genes involved in circadian regulation and CAM were identified in plants grown under a long day regime. Three kinds of expression profiles were found: transcripts oscillated with a 24-h periodicity; transcripts of the light-active genes adapted to cycles of 12-h periodicity; and

arrhythmic accumulation patterns. Some genes were scored best to a 12-h rhythm, suggesting a difference with Arabidopsis at level of circadian clock gene interactions. The results indicate that changes at the CAM metabolism are the result of modified circadian regulation at the transcriptional and posttranscriptional levels [9].

In addition, the gene regulation trough miRNAs has been explored [10]. miRNAs are a class of small non-coding RNAs that regulate gene expression. A combination of Northern blot and tissue print hybridization was used, to identify conserved miRNAs expressed during nopal (*Opuntia ficus indica*) fruit development. A comparative analysis detected 34 miRNAs expressed differentially. These miRNA were clustered different groups and associated with the different phases of fruit development. Gradual expression of several miRNAs was observed during fruit development. The work provided the evidence of miRNA expression in the cactus fruit and the basis for future research on miRNAs in Opuntia [10].

One transcendental work is related to the analysis of genomic content in 23 Opuntia species by flow cytometry [11]. A main interest on Opuntia genomes was related to the DNA content

because; almost all the genotypes have a ploidy level of 4x, 8x or 12x; of their genetic complexity. In four different ploidy levels having 2C-DNA amounts, DNA content ranged from 3.75 Giga base pairs (Gb) (*Opuntia incarnadilla* Griffiths) to 5.87 Giga base pairs (Gb) (*Opuntia heliabravoana* Scheinvar) among the samples analyzed.

The 2C DNA content when compared with other species; such as maize, shows that genome of Opuntia is less complex than that of maize (Table 1), which makes Opuntia suitable for a genomic sequencing program.

Common name	Scientific name	Family	Ploidy level	2C genome size (pg)	2C genome size (Gb)
Arabidopsis	*Arabidopsis thaliana*	Brassicaceae	2x	0.30	0.29
Soy bean	*Glycine max*	Leguminosae	2x	2.31	2.25
Maíze	*Zea mays*	Gramineae	2x	13.49	13.19
Tuna charola	*Opuntia streptacantha*	Cactaceae	8x	4.64	4.53
Tuna	*Opuntia megacantha*	Cactaceae	8x	5.01	4.89
Tuna blanca	*Opuntia ficus indica*	Cactaceae	8x	4.90	4.79
Tuna robusta	*Opuntia robusta*	Cactaceae	8x	4.98	4.87
Xoconoxtle	*Opuntia joconostle*	Cactaceae	8x	4.7	4.59

Table 1. Opuntia DNA content and other plant species

The nopal and its products need a more deeply analysis to maximize the real value of this crop. It is a multipurpose plant that is very important in the life of the people because it impact on the economy, nutrition, medicinal practices, and fuel production. Two main aspects are now in the focus for increasing its crop value:

1. Some crops have been sequenced and some others are in progress, however the nopal is waiting to be sequenced. Once sequenced, it would help to understand several mechanisms of plant adaptation to different environments, and will give us clues about controlling the process for adaptation to extreme conditions in other plants [12].

2. A new important aspect involves miRNAs, which are thought to be fine-tuning mechanisms in gene regulation [10]. Wrong expression of miRNAs can produce pleiotropic effects on development. It would be no surprise to discover that several events related to plant adaptation were under the control of miRNA expression. In the future, the expression of miRNA and siRNA will serve as tools for the generation of new *Opuntia* phenotypes. In these experiments the role of different molecules or pathways involved in seed formation, ripening delay or fruit development could be revealed.

3. Non model plants as source of industrial solutions

Development of modern society has led to an increased emission of pollutants into the environment, from industrial and domestic activities, as well as from mining, agriculture and crafting [13]. These compounds are a threat to all the organisms; therefore, numerous methods have been developed to reduce the impact caused by pollution. Conventional methods for the removal of pollutants in soil and water are often costly and can irreversibly affect the properties of the soil, as well as the organisms that inhabit those places [14].

Bioremediation is a tool used to clean pollutants in soil and water, and it is referred to the chemical transformation of pollutants through the use of microorganisms and plants [15]. The genomic content of plants for remediation has been calculated by different methods, and the sizes are included in Table 2. As shown in Table 2, there are no genomic complex organisms and some have been sequenced already.

It is important to consider that there is a great diversity of plants grow under different climates, which belong to different families. This allows researchers to have a wide variety of candidate plants that fits the scientific needs. Some plants of the Asteraceae, Brassicaceae and Solanaceae families have been found as tolerant to different pollutants. According to Lopez et al. [15], plants use a mechanism to alleviate the environmental stresses, by the following three phases:

(1) Absorption, excretion and detoxification of pollutants; (2) the distribution of pollutants throughout the plant and their excretion via volatilization; and (3) detoxification of pollutants by phytoremediation, by any one of the following processes: phytoextraction, rhizofiltration, phytoestimulation, phytostabilization, phytovolatilization or phytodegradation [15,16].

Phytostabilization allow to reduce the bioavailability and mobility of contaminants, avoiding underground transport layers or the atmosphere [15,16]. This process is less expensive than other methods, is easy to apply and aesthetically pleasing.

Phytodegradation is the transformation of organic pollutants in simpler molecules. In certain instances, degradation products will serve to accelerate plant growth, and other cases the contaminants are biotransformed. The phytodegradation has been employed for the removal of explosives, such as TNT, halogenated hydrocarbons, Bis-phenol, PAHs and organochlorine and organophosphorus pesticides [14].

In phytovolatilization, plants absorb water along with the soluble organic and inorganic pollutants (As, Se and Hg). Some of the contaminants can reach the leaves and get evaporated or volatilized into the atmosphere. Plants such as *Bigelovii Salicornia*, *Brassica juncea*, *Astragalus bisulcatus* and *Chara canescens* have been used for bioremediation of Se pollution and *Arabidopsis thaliana* has been used for bioremediation of Hg [14].

Rhizofiltration uses plants to remove contaminants from water environment through the root. In rhizofiltration, these plants are grown in hydroponic way. When the root system is well developed, the plants are introduced into polluted water with metals, where the roots absorb and accumulate. Numerous aquatic plants have the ability to accumulate pollutants, and some examples of these are as follows: *Scirpus lacustris*, *Lemna gibba*, *Azolla caroliniana*, *Elatine*

trianda, Wolffia papulifera, Polygonum punctatum, Myriophylhum aquaticum, and *Mentha palustris* (for Al, As, Au, Cd, Cr, Cu, Cr, Fe, Hg, Mg, Mn, Ni, Pb, Se, Sr, Zn,) [14,15].

Function	Species	Family	1C (Gb)	1C (pg)	Sequencing year	Reference
Phytoestabilization Pb, Zn, Cd, As, Cu, Mn	*Hordeum vulgare*	Gramineae	5.1	5.5	2012	International. Barley Genome Consortium [17]
Mercury Phytovolatilization	*Arabidopsis thaliana*	Cruciferae	0.125	0.16	2000	Arabidopsis Genome Initiative [18]
Phytoextraction Cd, Zn, Pb, Ni, Ag, Cr, Cu, Hg,	*Brassica juncea*	Cruciferae	1.49	1.092	--	Johnston et al. [19].
	Helianthus annuus	Compositae	3.5	2.43	2012	Staton et al. [20]
	Brassica napus	Cruciferae	1.12	1.15	2014	Boulos et al. [21]
Petroleum Contaminants Degradation	*Sorghum bicolor*	Gramineae	1.68	0.835	2009	Paterson et al. [22]
	Medicago sativa	Leguminosae	1.75	0.86	2011	Young et al. [23]
Elimination Cd, Pd, Zn, Cu, Ni, Cr	*Brassica nigra*	Cruciferae	0.632	0.647	--	Johnston et al. [19]
	Helianthus annuus	Compositae	3.5	2.43	2012	Staton et al. [20]
Insecticide Accumulation	*Cucumis sp*	Cucurbitaceae	0.68	0.66	2009	Huang et al [24]
	Cucurbita sp	Cucurbitaceae	0.34	0.33	--	Šisko et al. [25]
Phytoextraction Zn, HgNO$_3$	*Helianthus annuus*	Compositae	3.5	2.43	2012	Staton et al. [20]

Table 2. DNA content in Plants used for bioremediation

Phytoextraction or absorption is carried out by the plant roots and accumulation of polluting metals in the stems and leaves. Some plants used for this approach are: *Thlaspi caerulescens*; *Sedum alfredii, Viola* and *Vertiveria baoshanensis*; *Alyssum murale, Trifolium nigriscens, Psychotria douarrei, Pruinosa geissois, Homalium guillainii, Hybanthus floribundus, Sebertia acuminata, Stackhousia tryonii, Pimelea leptospermoides, Aeollanthus biformifolius*; *Haumaniastrum robertii*; *Brassica juncea, Helianthus annuus, Sesbania drummondii* and *Brassica napus* (for Ag; Cd, Cr, Cu, Hg, Ni, Pb, and Zn) [14,15].

Phytodegradation in plants and microorganisms is associated with, degradation of organic pollutants into harmless products and, mineralization into CO_2 and H_2O. Plants such as *Populus* spp. are introduced to absorb the contaminants in soil pores and prevent leaking to other soil layers [15,26].

4. Perspectives for genome sequencing and genome information from non-model plants to plant breeding

Next-generation sequencing (NGS), include several and different technologies which has its own set of characteristics. NGS generates huge amounts of sequence data in a very cost-effective way [27].

The increased number of WGS projects means that more organisms, are becoming important genetic models [28]. At the same time, many molecular studies are focusing on natural variation and adaptation in classical genetic model species, or close relatives of these, such as *Arabidopsis*, thus closing the gap between model and non-model organisms.

For example, the assembled genomes could be used as a reference sequence for further transcriptome analysis or re-sequencing and surveys of genetic variation. They may also be used to develop other genomic tools, such as proteomics and microarrays hybridization [29].

After the novel transcriptome has been annotated using a genomic reference species, it can be used as a starting point for more detailed functional characterizations of desired organisms, using gene ontology databases.

With RNA-seq protocols, or longer sequence reads will also improve applications because large haplotype blocks including several linked polymorphisms will become available. Wherein hundreds of genes are analyzed simultaneously. Some of these may be involved in important phenotypic variation, and this is relevant from the conservation point of view because such variation may be important to maintain within the population.

In the future, the bottleneck is more likely to be at the bioinformatics rather than in producing the sequences [30] because a huge number of biologists are trying to order

the genomic data with biological sense. New approaches for data storage and processing will be needed, because currently available databases might be unable to cope up with the rapid generation of new sequencing data [31].

5. Conclusions and prospects

Plants provide food for all living organisms, and just 15 crop plants provide 90% of the world's food intake [32]. Plant species are responsible for maintaining the balance of the carbon cycle, for developing and maintaining soil from erosion, and plant products are used as human medicines [33, 34]. For these reasons, there is great interest in sequencing plant genomes, but

so far relatively few plant species have been sequenced compared with the hundreds of thousands of species around the world.

Non-model plants are becoming very attractive sources for different purposes, for their ability to adaptation to extreme environments and to produce specific metabolites that can be used for food and medicinal purposes. The materials must be characterized at molecular level to develop any strategy for the generation of genetic data, that is molecular markers, cDNA sequencing, and cDNA microarrays, to have reference data to compare with model organisms.

Large complex plant genomes remain a particularly difficult challenge for *de novo* assembly for various biological, bioinformatics, and biomolecular reasons. Plant genomes can be nearly 100 times larger than the sequenced mammalian genomes [35]. The next frontier for plant genomics is to characterize the diversity of genomic variations across large populations, deeply annotate their functional elements, and develop predictive quantitative models relating genotype to phenotype.

Acknowledgements

Author details

Jannette Alonso-Herrada[1], Ismael Urrutia[1], Tania Escobar-Feregrino[2], Porfirio Gutiérrez-Martínez[4], Ana Angélica Feregrino-Pérez[1], Irineo Torres-Pacheco[1], Ramón G. Guevara González[1], Sergio Casas-Flores[5] and Andrés Cruz-Hernández[1,3*]

*Address all correspondence to: andrex1998@hotmail.com

1 Engineering Faculty, Autonomus University of Querétaro, Circuito Universitario, Cerro de las Campanas s/n, C.P. Santiago de Querétaro, Querétaro, México

2 Natural Sciences Faculty, Autonomus University of Querétaro, Circuito Universitario, Cerro de las Campanas s/n, C.P. Santiago de Querétaro, Querétaro, México

3 Chemistry Faculty, Autonomus University of Querétaro, Circuito Universitario, Cerro de las Campanas s/n, C.P. Santiago de Querétaro, Querétaro, México

4 Technological Institute of Tepic. Food Science Postgrade, Integral Laboratory on Food Science and Biotechnology Research. C.P. Tepic, Nayarit, México

5 IPICYT, Molecular Biology Division, C.P. 78216, San Luis Potosí, México

References

[1] Paterson A, Richardson C. Toward a cactus genome project: utilizing information from well studied model plants to map *Opuntia* chromosomes. J Prof Assoc Cactus Dev.1996;1:106–109.

[2] Aramuganathan K, Earle ED. Nuclear DNA content of some important plant species. Plant Mol Biol Rep.1991;9:208–218. DOI: doi.org/10.1007/BF02672069

[3] Food and Agriculture Organization. Our priorities: The FAO strategic objectives. Version 1.1–4. Food and Agriculture Organization of the United Nations. Rome. 2013. www.fao.org/docrep/018/mi317e/mi317e.pdf

[4] de Magalháes JP. The big, the bad and the ugly. EMBO Rep. 2015;16:771-776. DOI: 10.15252/embr.201540606

[5] Mondragón–Jacobo, C. Verification of the apomictic origin of cactus pear (*Opuntia* spp. *Cactaceae*) seedlings of open pollinated and crosses from central Mexico. J Prof Assoc Cactus Dev. 2001;3:49–56. DOI: 10.1016/S0140-1963(5)80016-3

[6] Flores–Valdés C, de Luna JM, Ramírez P.P. Nopalito world market. ASERCA–UACh–CIESTAAM. Chapingo, México. 1995; pp. 16–75.

[7] Barbera G, Inglese P, Pimienta–Barrios E. Agroecology, Cultivation and Uses of Cactus Pear. FAO Plant Production and Protection. Food and Agriculture Organization of the United Nations. Rome. 1995; Paper 132.

[8] Pimienta–Barrios E. Prickly pear (*Opuntia* spp.): a valuable fruit crop for the semiarid lands of México. J Arid Environ. 1994;28:1–11. DOI: 10.1016/S0140-1963(05)80016-3

[9] Mallona I, Egea-Cortines M, Weiss J. Conserved and divergent rhythms of crassulacean acid metabolism-related and core clock gene expression in the cactus *Opuntia ficus-indica*. Plant Physiol; 2011;156:1978–1989. DOI: 10.1104/pp.111.179275

[10] Rosas-Cárdenas FF, Caballero-Pérez J, Gutiérrez-Ramos X, Marsch-Martínez N, Cruz-Hernández A, de Folter S. miRNA expression during prickly pear cactus fruit development. Planta. 2014. DOI: 10.1007/s00425-014-2193-0

[11] Segura S, Scheinvar L, Olalde G, Leblanc O, Filardo S, Muratalla A, Gallegos C, Flores C. Genome sizes and ploidy levels in Mexican cactus pear species *Opuntia* (Tourn.) Mill. series *Streptacanthae* Britton et Rose, *Leucotrichae* DC., *Heliabravoanae* Scheinvar and *Robustae* Britton et Rose. Genet Res Crop Evol. 2007;54:1033–1041. DOI: 10.1007/s10722-006-9196-z

[12] Jackson S, Rounsley S, Purugganan M. Comparative sequencing of plant genomes: choices to make. Plant Cell. 2006;18:1100–1104. DOI: 10,1105/tpc.106,042192

[13] Wei S, Zhou Q, Saha UK. Hyperaccumulative characteristics of weed species to heavy metals. Water Air Soil Pollut. 2008;192:173–181. DOI: 10.1007/s11270-008-9644-9

[14] Delgadillo LAE, González RCA, Prieto GF, Villagómez IJR, Acevedo SO. Fitorremediation: an alternative to eliminate contamination. Tropical Subtropical Agroecosyst. 2011;14:597- 612

[15] López MS, Gallegos MM, Pérez FL, Gutiérrez RM. Mechanisms of fitorremediation of contaminated soils with organic xenobiotic molecules. Rev Int Contam Ambient. 2005; 2:91-100.

[16] Cartman EP, Crossman TL. *In situ* Treatment Technology. Chapter 9: Phytoremediation. Second Edition. Lewis Publishers. USA. 2001

[17] The International Barley Genome Sequencing Consortium. A physical, genetic and functional sequence assembly of the barley genome. Nature. 2012:491:711-717. DOI: 10.1038/nature11543

[18] The Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis Thaliana*. Nature. 2000;408:796-815. DOI: 10.1038/35048692

[19] Johnston JS, Pepper AE, Hall AE, Chen ZJ, Hodnet G, Drabek J, Lopez, Price J. Evolution of Genome Size in Brassicaceae. Ann Bot. 2005;95:229-235. DOI: 10.1093/aob/mci016

[20] Staton SE, Bakken BH, Blackman BK, Chapman MA, Kane NC, Tang SU, Mark C, Knapp SJ, Rieseberg LH, Burke JM. The sunflower (*Helianthus annuus* L.) genome reflects a recent history of biased accumulation of transposable elements. Plant J. 2012;72:142–153. DOI: 10.1111/j.1365-313X.2012.05072.x

[21] Boulos C, Denoeud F, Liu S, Parkin I, Tang H, Wang X, Chiquet J, Belcram H, Tong C, Samans B, Corréa M, Da Silva C, Just J, Falentin C, Koh C, Le Clainche I, Bernard M, Bento P, Noel B, Labadie K, Alberti A, Charles M, Arnaud D, Guo H, Daviaud C, Alamery S, Jabbari K, Zhao M, Edger P, Chelaifa H, Tack D, Lassalle G, Mestiri I, Schnel N, Le Paslier M, Fan G, Renault V, Bayer P, Golicz A, Manoli S, Lee T, Thi V, Chalabi S, Hu Q, Fan C, Tollenaere R, Lu Y, Battail C, Shen J, Sidebottom C, Wang X, Canaguier A, Chauveau A, Bérard A, Guan M, Liu Z, Sun F, Lim Y, Lyons E, Town E, Bancroft I, Wang X, Meng J, Ma J, Pires J, King G, Brunel D, Delourme R, Renard M, Aury J, Adams K, Batley J, Snowdon R, Tost J, Edwards D, Zhou Y, Hua W, Sharpe A, Paterson A, Guan C, Wincker P. Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. Science. 2014;345:950-953. DOI: 10.1126/science.1253435

[22] Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A, Schmutz J, Spannagl M, Tang H, Wang X, Wicker T, Bharti AK, Chapman J, Feltus FA, Gowik U, Grigoriev IV, Lyons E, Mahe CA, Martis M, Narechania A, Otillar RP, Penning BW, Salamov AA, Wang Y, Zhang

L, Carpita NC, Freeling M, Gingle AR, Hash CT, Keller B, Klein P, Kresovich S, McCann MC, Ming R, Peterson DG, Mehboob-ur-Rahman, Ware D, Westhoff P, Maye K, Messing J, Rokhsar DS. The *Sorghum bicolor* genome and the diversification of grasses. Nature. 2009;457:551-556. DOI:10.1038/nature07723

[23] Young N, Debelle F, Oldroyd GE, Geurts R, Cannon S, Udvardi M, Benedito V, May-er K, Gouzy J, Schoof H, Van de Peer Y, Proost S, Cook D, Meyers B, Spannagl M, Cheung F, De Mita S, Krishnakumar V, Gundlach H, Zhou S, Mudge J, Bharti A, Murray J, Naoumkina M, Rosen B, Silverstein K, Tang H, Rombauts S, Zhao P, Zhou P, Barbe V, Bardou P, Bechner M, Bellec A, Berger A, Berges H, Bidwell S, Bisseling T, Choisne N, Couloux A, Denny R, Deshpande S, Dai X, Doyle J, Dudez A, Farmer A, Fouteau S, Franken C, Gibelin C, Gish J, Goldstein S, González A, Green P, Hallab A, Hartog M, Hua A, Humphray S, Jeong D, Jing Y, Jocker A, Kenton S, Kim D, Klee K, Lai H, Lang C, Lin S, Macmil S, Magdelenat G, Matthews L, McCorrison J, Mona-ghan E, Jeong-HwanMun, Najar F, Nicholson C, Noirot C, O'Bleness M, Paule C, Poulain J, Prion F, Qin B, Qu C, Retzel E, Riddle C, Sallet E, Samain S, Samson N, Sanders I, Saurat O, Scarpelli C, Schiex T, Segurens B, Severin A, Sherrier D, Shi R, Sims S, Singer S, Sinharoy S, Sterck L, Viollet A, Wang B, Wang K, Wang M, Wang X, Warfsmann J, Weissenbach J, White D, White J, Wiley G, Wincker P, Xing Y, Limei Y, Yao Z, Ying F, Zhai J, Zhou L, Zuber A, Denarie J, Dixon R, May G, Schwartz D, Rog-ers J, Quetier F, Town C, Roe B. Medicago genome provides insight into the evolu-tion of rhizobial symbioses. Nature. 2011;480: 796-810. DOI: 10.1038/nature10625

[24] Huang S, Li R, Zhang Z, Li L, Gu X, Fan W, Lucas W, Wang X, Xie B, Ni P, Ren Y, Zhu H, Li J, Lin K, Jin W, Fei Z, Li G, Staub J, Kilian A, van der Vossen E, Wu Y, Guo J, He J, Jia Z, Ren Y, Tian G, Lu Y, Ruan J, Qian W, Wang M, Huang Q, Li B, Xuan Z, Cao J, Asan, Wu Z, Zhang J, Cai Q, Bai Y, Zhao B, Han Y, Li Y, Li X, Wang S, Shi Q, Liu S, Cho W, Kim J, Xu Y, Heller-Uszynska K, Mia H, Cheng Z, Zhang S, Wu J, Yang Y, Kang H, Li M, Liang H, Ren X, Shi Z, Wen M, Jian M, Yang H, Zhang G, Yang Z, Chen R, Liu S, Li J, Ma L, Liu H, Zhou Y, Zhao J, Fang X, Li G, Fang L, Li Y, Liu D, Zheng H, Zhang Y, Qin N, Li Z, Yang G, Yang S, Bolund L, Kristiansen K, Zheng H, Li S, Zhang X, Yang H, Wang J, Sun R, Zhang B, Jiang S, Wang J, Du Y, Li S. The genome of the cucumber, *Cucumis sativus* L. Nature Gen. 2009; 41:1275-1283. DOI: 10.1038/ng.475

[25] Šisko M, Ivancic A, Bohanec B. Genome size analysis in the genus *Cucurbita* and its use for determination of interspecific hybrids obtained using the embryo-rescue tech-nique. Plant Sci. 2003;165: 663–669. DOI: 10.1016/S0168-9452(03)00256-5

[26] Jankaitė A, Vasarevičius S. Use of *Poaceae f*. species to decontaminate soil from heavy metals. Ekologija. 2007;53:84–89.

[27] Gilad Y, Pritchard JK, Thornton K. Characterizing natural variation using next-gener-ation sequencing technologies. Trends Genet. 2009;25:463–471. DOI: 10.1016/j.tig. 2009.09.003

[28] Mitchell-Olds T, Feder M, Wray G. Evolutionary and ecological functional genomics. Heredity. 2008;100:101–102. DOI: 10.1038/sj.hdy.6801015

[29] Harr B, Turner LM. Genome-wide analysis of alternative splicing evolution among Mus subspecies. Mol Ecol. 2010;19(s1):228–239.DOI: 10.1111/j.1365-294X.2009.04490.x

[30] Schuster SC. Next-generation sequencing transforms todays biology. Nat Meth. 2008;5:16–18.DOI: 10.1038/nmeth1156

[31] Ekblom R, Galindo J. Applications of next generation sequencing in molecular ecology of non-model organisms. Heredity. 2011;107:1–15. DOI: 10.1038/hdy.2010.152

[32] United Nations Food and Agriculture Organization: Dimensions of Need - An atlas of food and agriculture. Staple foods: What do people eat. Available from http://www.fao.org/docrep/u8480e/u8480e07.htm

[33] Pimentel D, Harvey C, Resosudarmo P, Sinclair K, Kurz D, McNair M, Crist S, Shpritz L, Fitton L, Saffouri R, Blair R: Environmental and economic costs of soil erosion and conservation benefits. Science.1995;267:1117-1123.DOI: 10.1126/science.267.5201.1117

[34] Mann J. Natural products in cancer chemotherapy: past, present and future. Nat Rev Cancer. 2002;2:143-148. DOI: 10.1038/nrc723

[35] Schatz S, Witkowski J, McCombie R. Current challenges in *de novo* plant genome sequencing and assembly. Genome Biol. 2012;13:243-249. DOI: 10.1186/gb-2012-13-4-243

Regulation of Plant Genes and Genomes by Small RNAs

MicroRNAs Sequencing for Understanding the Genetic Regulation of Plant Genomes

Christopher A. Cedillo-Jiménez, Marcelo Hernández–Salazar, Tania Escobar-Feregrino, Juan Caballero-Pérez, Mario Arteaga-Vázquez, Alfredo Cruz-Ramírez, Irineo Torres-Pacheco, Ramón Guevara-González and Andrés Cruz-Hernández

Additional information is available at the end of the chapter

Abstract

MicroRNAs (miRNAs) are endogenous non-coding RNAs that play important regulatory roles in animals and plants by targeting mRNAs for cleavage or translational repression. Small RNAs are classified into different types by their biogenesis and mode of action, such as miRNAs, siRNAs, piRNAs, and snoRNAs. In the case of miRNAs, this specific type regulates gene expression in plants and animals by targeting mRNAs for cleavage and translational repression, respectively. Diverse miRNAs regulate plant development, metabolism, and responses to biotic and abiotic stresses. The identification of miRNAs has been accomplished in diverse species, organs and developmental or diverse biotic and abiotic stress conditions. Novel massive sequencing techniques and further bioinformatics analysis have allowed the identification of hundreds of miRNAs in *Arabidopsis thaliana*, *Oryza sativa*, *Malus domestica*, *Zea mays*, *Solanum lycopersicum*, and other plants. Functional characterization of a given miRNA in a specific biological context has shown their role in the fine-tuning mechanisms of posttranscriptional gene regulation. In this chapter, besides making a summary of genome-wide miRNA profiling in plants, we describe how gain and loss of function approaches influence plant phenotypes that affect development, physiology or stress responses, pointing to miRNAs as effective tools for the generation of new plant phenotypes that improve plant productivity and conservation.

Keywords: Gene expression, Plant development, mRNA targeting, miRNA, sRNAs

1. Introduction

MicroRNAs (miRNAs) are a class of non-coding endogenous small RNAs (sRNAs) that have attracted a huge interest from scientists. Experimental and computational approaches have demostrated that miRNAs play crucial roles during plant growth and development. Expression of miRNAs is highly regulated at both the transcriptional and postranscriptional level.

The development of new sequencing technologies has been crucial for the identify of novel miRNAs and to understand their function in specific process including the adaptation of plants to extreme environments. Plant development and of course, in the improvement of plants for human consumption. This book chapter highlights research progress on plant miRNAs and their various functions on plant growth, development, and stress responses.

2. Plant sRNAs´classification

Regulatory sRNAs are ubiquitous components of endogenous plant transcriptomes, as well as common responses to exogenous viral infections and introduced double-stranded RNA molecules (dsRNA). They range from 20 to 24 nucleotides in length. Endogenous sRNAs are processed from dsRNA precursors sRNAs can be classified based on their origin and function. In plants, those derived from single stranded precursors capable of adquiring an imperfect extensive nearly perfect dsRNA precursors hairpin are called hairpin small RNAs (hpRNAS) and those derived from dsRNA are referred to as small interfering RNAs (siRNAs). Frequently, siRNAs fall into one of three additional groups: heterochromatic siRNAs, secondary siRNAs, and natural antisense transcript siRNAS (NAT-siRNAs; Figure 1) [1].

2.1. Micro-RNAs (miRNAs)

miRNAs are a well-studied subset of hpRNAs defined by the highly precise excision of one or sometimes a few functional products, which correspond to the mature miRNAs (Figure 1) [1, 2]. miRNAs have a defined set of miRNA targets [1, 3], and individual miRNA families can be evolutionary conserved [1, 4]. Most plant miRNAs require a member of the DCL1-clade for their biogenesis, and a member of the AGO1-clade to exert their function, although some exceptions have been described [1, 5].

Several miRNA families are conserved in plant species, and some are conserved from mosses to flowering plants [1, 6]. Conserved miRNAs have homologous target mRNAs in several species, showing that miRNA/target relationships are very stable during plant evolution [1, 3]. However, some relationships between plant miRNAs and their targets can be novel; for example, miR159 is a highly conserved miRNA that targets MYB transcription factors in several plant species. Buxdorf *et al*. [7]; found that miR159 in tomato, also targets SGN-U567133, a non-MYB mRNA. Expression of SGN-U567133 causes developmental defects, suggesting that regulation of gene expression trough miR159 of this non-canonical target has a functional consequence [1].

However, not all plant miRNAs are conserved; some of the miRNAs present in any given plant species seem to be unique to that species, and some other miRNAs are conserved only be-

tween closely related species [1]. The lineage-specific miRNAs are different in some ways from the more conserved miRNAs. These miRNA have more heterogeneous processing from their hairpin precursors, and have low abundance, and they are generally encoded by single genes instead of multiple paralogs [1, 8]. These differences suggest that some lineage-specific miRNAs could be transient, nonfunctional entities, and categorize them as a distinct subset of miRNAs (Figure 1) [1].
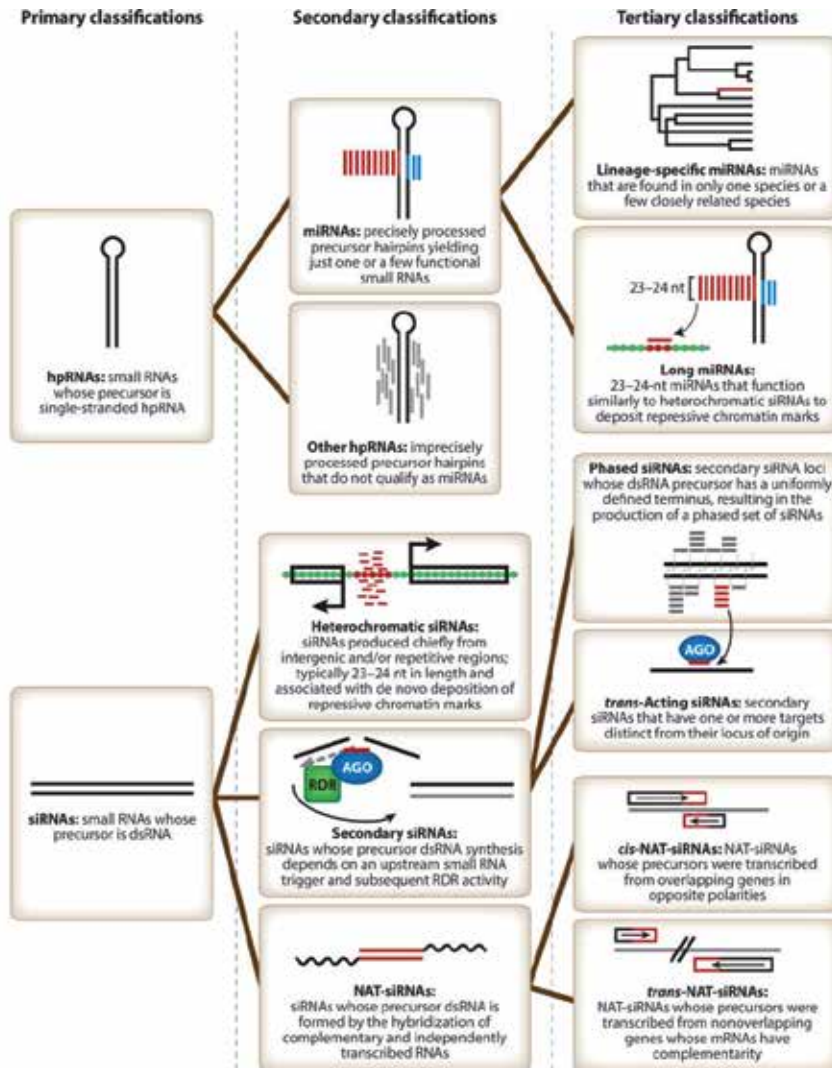
Figure 1. Hierarchical classification system for endogenous plant small RNAs. Thick black lines indicate hierarchical relationships. Abbreviations: double-stranded RNA (dsRNA); hairpin RNA (hpRNA); microRNA (miRNA); natural antisense transcript small interfering RNA (NAT-siRNA); small interfering RNA (siRNA) [1].

3. Plant miRNAs´ biogenesis

The mode of action of miRNAs between plant and animal kingdoms is different [9]. The miRNA biogenesis in plants has been documented using *Arabidopsis thaliana* as a model plant [10]. Primary miRNAs (pri-miRNAs) are transcribed by the RNA polymerase II (Figure 2), from regions in between coding genes [11]. The pri-miRNA adquires a hairpin secondary structure, and its length ranges from approximately 70 to many hundreds of bases [10]. The protein Dawdle (DDL) is an RNA-binding protein that stabilizes a subset of pri-miRNAs for the subsequent export to D-bodies [12]. D-bodies are compartments where miRNAs are processed by the joint action of Serrate (SE) and Hyponastic leaves 1 (HYL1) that together with Dicer-like 1 (process the pri-miRNA into a precursor-miRNA (pre-miRNA). Susequently, DCL1 releases the miRNA duplex containing the mature miRNA and the passenger miRNA* (Figure 2) [13].
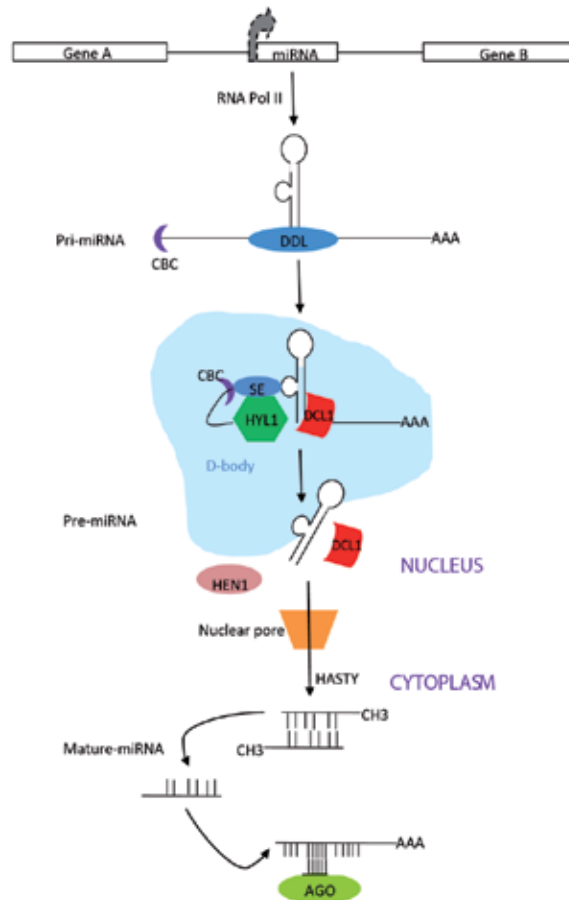
Figure 2. miRNA biosynthesis in plants. miRNA biosynthesis complex in nucleus and cytoplasm Dawdle protein (DDL), binding complex (CDC), Serrate protein (SE), Hyponastic leaves 1 (HYL1), Dicer-like 1 (DCL1), exportin-5 ortholog (HASTY), methylation protein (HEN1), and Argonaute (AGO)

The miRNA released by DCL1 is exported out of the nucleus through the exportin - 5´ortholog (HASTY). The pre-miRNA is converted into a mature miRNA, but the double strand persist, so one of the two strands is the guide strand and the other is the passenger (that gets degraded), the mature guide miRNA is methylated by HEN1 to prevent the degradation by sRNA degrading nuclease [14].

The next step is the recruiting of the miRNA by an Argonaute (AGO) protein. In *Arabidopsis thaliana* 10 different AGO proteins were are present [15]. the AGO/mature miRNA complex its finds target mRNA based on the complementarity between the guide sequence and the mRNA, then it cleaves its target the target thanks to AGO's slicer activity (Figure 2).

4. Plant miRNAs´ function and plant development

A miRNA is a 21–24 nucleotide RNA product of a non-protein-coding gene. Plants, like animals, have a large number of miRNA-encoding genes in their genomes. Plant miRNAs have been predicted or confirmed to regulate a variety of processes, such as development, metabolism, and stress responses (Table 1). Plant miRNAs have been predicted or confirmed to regulate genes encoding various types of proteins. A major category of miRNA target genes consists of transcription factors or other regulatory proteins that function in plant development or signal transduction [16]. The first evidence that sRNAs play roles in plant development, came from mutants with impaired sRNA biogenesis or function [17].

4.1. Auxin signaling

Auxins are critical for plant development and their interaction with the environment. Local concentration of auxins, as established by polar auxin transport, provides the growth pattern on the axis of the embryo from root stem cells, and controls primordia outgrowth from meristems, as well as initiate, lateral root formation, and gravitropic responses.

A number of genes in auxin signaling are confirmed or predicted targets of miRNAs. The TIR1 auxin receptor is a predicted target of miR393 [17] (Table 1). Several auxin response factors, such as ARF10, ARF16, and ARF17, contain potential binding sites for miR160 [18] and ARF6; ARF8 have sites for miR167 [19]. In *in vivo* assays, miR160 can guide the cleavage of ARF10 and ARF16, and miR167 guide the cleavage of ARF8 mRNA [20]. The expression of a miR160-resistant version of ARF17 (5mARF17) leads to developmental defects with abnormalities, such as leaf serration, leaf curling, early flowering, altered floral morphology, and reduced fertility [20]. This indicates that regulation of ARF17 mediated by miR160 is crucial for different aspects of plant development (Table 1).

4.2. Plant organ boundary formation

Three members of the NAC gene family, *CUP-SHAPED COTYLEDON* (*CUC*) *1*, *2*, and *3*, have overlapping functions on organ boundary formation and in the initiation of shoot apical meristem (SAM). These three genes are expressed first in the boundaries of embryo cotyledons and later in the boundaries of floral organs [21].

miRNA	TARGET FAMILY	TARGETS	FUNCTION IN PLANTS
mir156	SBP	*SPL2, SPL3, SPL4, SPL10*	Apical dominance
mir159	MYB TCP	*MYB33, MYB65,* *TCP2, TCP3, TCP4,* *TCP10, TCP24*	Male sterility
mir160	ARF	*ARF10,ARF16, ARF17*	Root development
mir164	NAC	*CUC1, CUC2, NAC1,* *At5g07680, At5g61430*	Aging induced cell death. Senescence
mir166	HD-ZIPIII	*PHB, PHV, REV, ATHB-8,* *ATHB-15*	Female sterility Organ polarity
mir167	ARF	*ARF6, ARF8*	Auxin signaling
mir169	HAP2	*At1g17590, At1g72830,* *At1g54160, At3g05690*	Root architecture Stress response
mir171	SCL	*SCL6-III, SCL6-IV*	Developmental patterning
mir172	AP2	*AP2, TOE1, TOE2, TOE3*	Flower development
mir393	bZIP	*At1g27340* *arf10, arf16 and arf17*	Auxin signaling Root development
mir396	GRF	*GRL1, GRL2, GRL3, GRL7,* *GRL8, GRL9*	Cell proliferation Leaf development
mir444	MADS	*Os02g49840*	Defense response

Table 1. miRNA targets and their associated function

miR164 targets CUC1 and CUC2 but not CUC3 [17]. Assays of miR164 overexpression in wild-type plants (under the control of CaMV35S promoter) lead to floral organ fusion [20] and cotyledon fusion, in a lower frequency [22]. Expressing a CUC2 resistan version to miR164 can restore sepal separation in miR164 overexpressing lines [22]. A wild-type plant expression of CUC1 resistant to miR164 results in changes in sepal (reduced) and petal number (increased) and broadened leaves [20].

Expression of the CUC2 mir164-resistan version also increased the width of the boundary domain between sepals [22]. A similar effect was observed in miRNA mutants such as *dcl1*, *hen1*, and *hyl1* [22]. The sepal boundary expansion could explain the phenotypes related to narrow sepals in *dcl, hen1*, and *hyl1* mutants and reduced sepal number in CUC genes. miR164 is potentially encoded by a gene family of three members (MIR164 family), where MIR164c was identified as a regulator of petal number in flowers [23].

4.3. Polarity at the leaves and floral organs

Leaves and floral organs, are initiated as primordia on the SAM or floral meristems. These lateral organs have polar structures that develop in the adaxial side in the primordium, and

they differ from the side that faces away from the meristem (called the abaxial side) [16]. The polarity of lateral organs is established through antagonistic interactions between two groups of genes: the class III homeodomain leucine zipper (HD-zip) including *PHABULOSA* (*PHB*), *PHAVOLUTA* (*PHV*), and *REVOLUTA* (*REV*); the other group includes the KANADI family (*KAN1*, *2*, and *3*) [16, 24]. The HD-zip genes are expressed in the adaxial domain, and the *KAN* genes are expressed in the abaxial domain in a differential way. Mutations in *PHB*, *PHV*, and *REV* genes result in adaxialized leaves and floral organs, and also contribute to the polarity of the vasculature [16].

Analysis of a particular set of mutants in PHB, PHV and REV revealed their regulation by miR165/miR166 the wild types, and shown that mutations in the binding sites of miRNA165/166 affecting the regulation of these genes at the miRNA level rather than the activity of the proteins [16, 18]. In a gain-of-function assays with the *phb-d* allele, the expression domain of the gene expands into the abaxial region [16, 25]. It shows that a regulation mediated by miRNA inhibited the PHB expression in the adaxial domain. The mechanism would be that degradation of the HD-zip mRNAs by miR165/166 in the abaxial domain clears the mRNAs from this domain. Also, it was found that miR165/166 causes DNA methylation of the PHB and PHV genes [26]. The miR165/166-mediated regulation of HD-zip genes is highly conserved, and the miR165/166-binding site is very similar among angiosperms, gymnosperms, ferns, and mosses [27].

4.4. Organ identity in flowers

Floral organs are organized in whorls at the floral meristem. The identities of the floral primordia are directed by the activities of floral homeotic genes known as the A, B, and C genes, and the regulation is known as the ABC model [16]. The A and C genes specify the identities of the perianth and reproductive organs, respectively. Mutations in AGAMOUS (class C gene) replace the reproductive organs by perianth organs, and mutations in APETALA2 (class A gene) lead to the loss of perianth. It suggests that A and C genes are antagonists to their activity within the floral meristem [16].

AP2 contains a binding site for miR172 and is regulated by miR172 *in vivo* (Table 1). Overexpression of miR172 under the control of CaMV35S promoter causes a reduction in the levels of AP2 protein, and the phenotypes at the floral structure are similar to those in ap2 mutants [28]. The overexpression of a miR172 resistant form of AP2 cDNA replaces the reproductive organs by perianth organs [28]. The expression assays using the AP2 promoter result in severe floral patterning defects [16, 28]. These results indicate the importance of miR172 in repressing AP2 in the inner two whorls in floral initiation. As in the HD-zip/KAN/miR165/166 system, miRNA serves as a negative regulator of one of the two antagonist functions.

Another miRNA, miR159 plays a role in reproductive development (Table 1) by regulating two MYB-domain transcription factor genes, MYB33 and MYB65. These two genes act redundantly to prevent the hypertrophy of the tapetum during anther development [29]. miR159 reduces the MYB33 and MYB65 expression to anthers. Transformed plants with MYB33 (a miR159-resistant version) under its own promoter arrested plant growth at various stages, suggesting that MYB33 expression by miR159 is critical for plant development [29].

4.5. Developmental transitions

The SAM generates leaves during the vegetative phase and flowers during the reproductive phase. The types of leaves produced at vegetative stage that are put out earlier (juvenile leaves) differ from the ones produced later (adult leaves). These changes between the developmental states are regulated by sRNAs[16].

miR172 regulates some other AP2-like genes, such as *TOE1*, *TOE2*, *TOE3*, *SMZ*, and *SNZ*, in addition to AP2. The *toe1-1* mutation results in an early flowering phenotype. The *toe2-1* mutant does not show a phenotype in flowering time, although the *toe1-1* and *toe2-1* double mutant flowers much earlier than the wild type, this indicates that TOE1 and TOE2 are redundant repressors of the transition from vegetative to reproductive stages [28]. The overexpression of *TOE1* (toe1-1D) provokes a delayed flowering time phenotype. A similar late flowering phenotype is caused by SMZ and SNZ overexpression [30]. When miR172 is overexpressed in TOE overexpressed lines, it results in early flowering phenotype, changing the late flowering phenotype of *toe1-1D* [28]. Hence, miR172 regulates the transition from vegetative to reproductive stage through the *TOE* genes regulation, via translational inhibition. However, overexpression of miR172 does not lead to a decrease of *TOE1* mRNA level, suggesting that this miRNA regulates different targets with different mechanisms [30].

miR156 overexpression also affects flowering time. 35S::MIR156 plants show a late flowering phenotype [31]. miR156 targets *Squamosa* promoter binding protein-like (SPL) transcription factor genes [18]; the role of *spl* genes in floral transition needs to be studied. Overexpression of miR319/Jaw results also in delayed flowering. miR319/Jaw targets to TCP transcription factors, and the participation of these factors in flowering is unknown [32]. miR159 overexpression leads to delayed flowering under short day conditions [33].

A pronounced vegetative phase change is observed in maize; an *AP2-like* gene *glossy*15 promotes juvenile leaf identity and its RNA is only found in juvenile leaves [34]. *glossy*15 contains a miR172-binding site suggesting that *glossy*15 is a target of miR172 as demonstrate *in vivo*. miR172 expression is correlated with the specification of adult leaf characteristics. It seems that miR172 clears glossy15 mRNA in adult leaves to promote the vegetative phase change [35].

4.6. Leaf development

The *CINCINNATA* (*CIN*) gene from snapdragon is required for differential regulation of cell division in leaf morphogenesis, where the resultant phenotype is a flat leaf [36]. *CIN* is a member of the TCP family of transcription factors. The overexpression of miR319/Jaw in Arabidopsis results in the reduction of the mRNA levels in five *TCP* genes containing miR319/Jaw-binding sites [32]. Overexpression of a resistant version of *TCP2 to* miR319/Jaw restores phenotype of miR319/Jaw overexpression [32].

4.7. sRNA metabolism

DCL1 contains a binding site for miR162 and miR162-guided cleavage products of DCL1 mRNA are detected *in vivo* [37]. *DCL1* mRNA abundance is augmented in mutants defective

in miRNA biogenesis (such as *dcl1* or *hen1*) [37]. The *AGO1* gene is targeted by miR168. Overexpression of miR168-resistant version of *AGO1* affects miRNA function, the phenotype is a miRNA that targets overaccumulation, and the plants show phenotypes similar to miRNA biogenesis mutants, such as *dcl1*, *hen1*, and *hyl1* [38]. AGO2, an argonaute gene, contains a binding site for miR403 in its 3′-UTR [39]. The regulation of genes involved in sRNA metabolism or function by miRNAs involves a feedback mechanism to ensure an adequate level of activity for the different the sRNA pathways.

5. miRNAs sequencing and prediction

Since the early beginning of noncoding RNA findings in developmental patterning, researchers have emphasized the bioinformatic challenges [40-42], not only in the miRNA discovery but also in the target prediction in order to better understand the expression, processing, and mechanism of regulation through base pairing recognition [12, 43, 44].

Genetic screening and direct cloning approaches work for simple miRNA candidate per event; therefore, these technologies have been replaced with Northern blotting, qRT-PCR and miRNA array assays. However, since the availability of next generation sequencing, this kind of technologies has been more frequently used as an efficient strategy for detailed research on plant miRNA of a wide variety of species because of the generation of millions of sequences per run. It can not only identify miRNA but also generate expression profiles. Moreover, big data analyses require a support of computational tools in order to extract relevant and refined information [45, 46].

Parallel to the development of deep sequencing strategies, sophisticated computational tools, and refined databases have played a major role in the effort to obtain a genome-wide profiling of miRNAs. Because of such effort, MiRBase, TargetScan, Plant Non-coding RNA Database (PNRD), miRNEST 2.0, and miRDeepFinder have been developed as free-access tools available for the study of miRNAs [46, 47].

miRBase Database [48] is a public repository first established in 2002. Nowadays, it is managed by Griffiths-Jones lab at the University of Manchester. In one of its latest version (released v. 21, June 2014) there are a total of 28,645 hairpin precursors and 35,828 mature products all over across 223 species. Besides, the miRBase includes a functional miRNA information connected with Wikipedia resource [48].

Tha Plant miRNA Database (PMRD, http://bioinformatics.cau.edu.cn/PMRD/) was created in 2009 by Prof. Zhen Su´s lab to integrate only plant miRNA data from public databases to keep together sequence information, secondary structure, target genes, and expression profiles [49]. New findings for ncRNA were identified, such as epigenetic regulators. The central focus of miRNA on the PMRD became limited for regulatory repository of data. PNRD (http://structuralbiology.cau.edu.cn/PNRD/index.php) is the updated version on PMRD released with improvements in functional analysis and service [50].

miRNESt was developed in 2012 by the Laboratory of Functional Genomics as a comprehensive repository for plant, animal and virus miRNAs. In consulted version (miRNEST2.0, http://

rhesus.amu.edu.pl/mirnest/copy/home.php) are included 522 miRNA from animal and plant with prediction data cross to 15 external databases, predicted targets for plant candidates supported by experimental validation, miRNA gene structure, and degradome data [51].

miRDeepFinder is a software developed in order to identify miRNA cross their target from deep sequencing. This package also provides to analyze miRNA functionality and it is a specific tool for plant species. From biogenesis, gene regulation as well as target recognition, abundance miRNA/miRNA* analysis, miRDeepFinder is capable to analyze RNA deep sequencing.

There are some computational approaches available for miRNA expression as well as miRNA target determination. Moreover, the false positive data are estimated for around 24-70%,therefore the experimental validation is required to characterize miRNA function [52].

There are methodologies for miRNA function validation at different levels: protein level by Western blot and mRNA level by qRT-PCR [52]. qRT-PCR is one of the most used techniques for detection of miRNA expression due for its high sensitivity to miRNA detection, the capability to identify single nucleotide changes [53]. miR-RACE (PCR-based) is an effective method to determine the precise sequence of miRNA at their 5′ and 3′ ends, which can distinguish between members of a miRNA family, and they can determine expression patterns at different family member levels [53, 54].

Parallel analysis of RNA ends (PARE) is a modification of miR-RACE, deep sequencing, and bioinformatic analysis for high-throughput sequencing. PARE is also known as degradome or genome-wide mapping of uncapped transcripts. PARE is used to plant systems manly to identify large subsets of miRNA targets to direct cleavage [55].

6. Conclusions and perspectives

Genetic, biochemical, and genomic studies have revealed a diverse array of endogenous sRNAs in plants, and resulted in the identification of several distinct classes of sRNAs. The study of these sRNAs has increased our knowledge regarding the function of these gene regulatory molecules.

Traditional computational approaches have made great progress in predicting new miRNAs in combination with molecular analysis. Most of the miRNAs are likely to be non conserved and/or species specific; this makes it hard to adapt the current approaches to predict non-conserved miRNA genes, and how to identify non-conserved miRNAs in non-model species where a reference genome is not available is an area of intense research.

Studies on miRNA target identification represents a big challenge beyond the identification of miRNA genes; total number of miRNA targets per miRNA family is still unknown and a large number of predicted miRNAs have not been validated experimentally. The identification of miRNA targets will improve our understanding of miRNA-mediated regulation of plant growth and development. miRNAs regulate gene expression by cleaving mRNA or by repressing mRNA translation; thus, now it is possible to design artificial miRNAs to suppress

the expression of a target gene in order to study gene function, similar to the use of alternative molecular tools (i.e. antisense mRNA) used for studying gene functions.

The use of miRNA to improve plant yields, quality, or resistance to various environmental stresses including insect and pathogen infection will come with the understanding of miRNA regulation over specific processes. Future study of miRNAs will provide us with tools for improving crop growth and quality.

Acknowledgements

Author details

Christopher A. Cedillo-Jiménez[1], Marcelo Hernández–Salazar[1], Tania Escobar-Feregrino[2], Juan Caballero-Pérez[3], Mario Arteaga-Vázquez[4], Alfredo Cruz-Ramírez[5], Irineo Torres-Pacheco[1], Ramón Guevara-González[1] and Andrés Cruz-Hernández[1*]

*Address all correspondence to: andrex1998@hotmail.com

1 Engineering Faculty, Autonomous University of Querétaro, Circuito Universitario, Cerro de las Campanas s/n, C.P. Santiago de Querétaro, Querétaro, México

2 Natural Sciences Faculty, Autonomous University of Querétaro, Circuito Universitario, Cerro de las Campanas s/n, C.P. Santiago de Querétaro, Querétaro, México

3 Chemical Sciences Faculty, Autonomous University of Querétaro, Circuito Universitario, Cerro de las Campanas s/n, C.P. Santiago de Querétaro, Querétaro, México

4 INBIOTECA, Universidad Veracruzana, Jalapa, Veracruz, México

5 LANGEBIO-CINVESTAV, Irapuato, Gto, México

References

[1] Axtell M J. Classification and comparison of small RNAs from plants. Ann. Rev. Plant Biol. 2013;64:137-159. DOI:10.1146/annurev-arplant-050312-120043

[2] Meyers BC, Axtell MJ, Bartel B, Bartel DP, Baulcombe D, Bowman J.L., Cao X, Carrington JC, Chen¹ X, Green PJ, Griffiths-Jones S, Jacobsen S, Mallory AC, Martienssen RA, Poethig RS, Qi Y, Vaucheret¹ H, Voinnet O, Watanabe Y, Weigel D, Jian-Kang Z. Criteria for annotation of plant microRNAs. Plant Cell. 2008;20:3186–3190. DOI:10. 1105/tpc.108.064311

[3] Jones-Rhoades MW. Conservation and divergence in plant microRNAs. Plant Mol. Biol. 2012;80:3–16. DOI:10.1007/s11103-011-9829-2

[4] Cuperus JT, Carbonell A, Fahlgren N, Garcia-Ruiz H, Burke RT, Takeda A, Sullivan CM, Gilbert SD, Montgomery TA, Carrington JC. Unique functionality of 22 nt miRNAs in triggering RDR6-dependent siRNA biogenesis from target transcripts in Arabidopsis. Nat. Struct. Mol. Biol. 2010;17:997–1003. DOI:10.1038/nsmb.1866

[5] Wu L, Zhou H, Zhang Q, Zhang J, Ni F, Chang L, Yijun Q. DNA methylation mediated by a microRNA pathway. Mol. Cell. 2010;38:465–475. DOI:10.1016/j.molcel. 2010.03.008

[6] Cuperus JT, Fahlgren N, Carrington JC. Evolution and functional diversification of miRNAgenes. Plant Cell. 2011;23:431–42. DOI:10.1105/tpc.110.082784

[7] Buxdorf K, Hendelman A, Stav R, Lapidot M, Ori N, Arazi T. Identification and characterization of a novel miR159 target not related to MYB in tomato. Planta. 2010; 232:1009–1022. DOI:10.1007/s00425-010-1231-9

[8] Ma Z, Coruh C, Axtell MJ. *Arabidopsis lyrata* small RNAs: Transient MIRNA and small interfering RNA loci within the *Arabidopsis* genus. Plant Cell. 2010;22:1090–1103. DOI:10.1105/tpc.110.073882

[9] Pogue AI, Clement C, Hill JM, Lukiw WJ. Evolution of microRNA (miRNA) Structure and Function in Plants and Animals: Relevance to Aging and Disease. J Aging Sci. 2014;2: 119: 1-5. DOI:10.4172/2329-8847.1000119

[10] Axtell MJ, Westholm JO, Lai EC. Vive la difference: Biogenesis and evolution of microRNAs in plants and animals. Genome Biol. 2011;12: 221. DOI:10.1186/gb-2011-12-4-221

[11] Ameres SL, Zamore, PD. Diversifying microRNA sequence and function. Nat. Rev. Mol. Cell Biol. 2013;14:475- 488. DOI:10.1038/nrm3611

[12] Voinnet O. Origin, biogenesis, and activity of plant microRNAs. Cell. 2009;136:669-687. DOI:10.1016/j.cell.2009.01.046

[13] Vaucheret H. Gene silencing: Mode of miRNA biogenesis matters. Nat. Plants. 2015;1:475-483. DOI:10.1038/NPLANTS.2015.19

[14] Jover-Gil S, Candela H, Robles P, Aguilera V, Barrero JM, Micol JL, Ponce M R. The microRNA pathway genes AGO1, HEN1 and HYL1 participate in leaf proximal–distal, venation and stomatal patterning in Arabidopsis. Plant Cell Physiol. 2012;53:1322-1333. DOI:10.1093/pcp/pcs077

[15] Zhang H, Xia R, Meyers B C, Walbot V. Evolution, functions, and mysteries of plant ARGONAUTE proteins. Curr. Opin. Plant Biol. 2015;27:84-90. DOI:10.1016/j.pbi. 2015.06.011

[16] Chen X. A microRNA as a translational repressor of APETALA2 in Arabidopsis flower development. Science. 2004;303:2022– 2025. DOI:10.1126/science.1088060

[17] Jones-Rhoades MW, Bartel DP. Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. Mol. Cell. 2004;14:787–799. DOI:10.1016/j.molcel.2004.05.027

[18] Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP. Prediction of plant microRNA targets. Cell. 2002;110:513–520. DOI:10.1016/S0092-8674(02)00863-2

[19] Mallory AC, Reinhart BJ, Jones-Rhoades MW, Tang G, Zamore PD, Barton MK, Bartel DP. MicroRNA control of PHABULOSA in leaf development: Importance of pairing to the microRNA 5′-region. EMBO J. 2004;23:3356–3364. DOI:10.1038/sj.emboj. 7600340

[20] Mallory AC, Dugas DV, Bartel DP, Bartel B. MicroRNA regulation of NAC-domain targets is required for proper formation and separation of adjacent embryonic, vegetative, and floral organs. Curr. Biol. 2004;14:1035–1046. DOI:10.1016/j.cub.2004.06.022

[21] Vroemen CW, Mordhorst AP, Albrecht C, Kwaaitaal MA, de Vries SC. The CUP-SHAPED COTYLEDON3 gene is required for boundary and shoot meristem formation in Arabidopsis. Plant Cell. 2003;15:1563–1577. DOI: 10.1105/tpc.012203

[22] Laufs P, Peaucelle A, Morin H, Traas J. MicroRNA regulation of the CUC genes is required for boundary size control in *Arabidopsis* meristems. Development. 2005;131:4311–4322. DOI:10.1242/dev.01320

[23] Baker CC, Sieber P, Wellmer F, Meyerowitz EM. The early extra petals1 mutant uncovers a role for microRNA miR164c in regulating petal number in *Arabidopsis*. Curr. Biol. 2005;15:303–315. DOI:10.1016/j.cub.2005.02.017

[24] Otsuga D, DeGuzman B, Prigge MJ, Drews GN, Clark SE. REVOLUTA regulates meristem initiation at lateral positions. Plant J. 2001;25:223–236. DOI:10.1111/j. 1365-313X.2001.00959.x

[25] McConnell JR, Emery J, Eshed Y, Bao N, Bowman J, Barton MK. Role of PHABULOSA and PHAVOLUTA in determining radial patterning in shoots. Nature. 2001;411:709–713. DOI:10.1038/35079635

[26] Bao N, Lye KW, Barton MK. MicroRNA binding sites in *Arabidopsis* class III HD-ZIP mRNAs are required for methylation of the template chromosome. Dev. Cell. 2004;7:653–662. DOI:10.1016/j.devcel.2004.10.003

[27] Floyd SK, Bowman JL. Gene regulation: Ancient microRNA target sequences in plants. Nature. 2004;428:485–486. DOI:10.1038/428485a

[28] Aukerman MJ, Sakai H. Regulation of flowering time and floral organ identity by a microRNA and its APETALA2-like target genes. Plant Cell. 2003;15:2730–2741. DOI: 10.1105/tpc.016238

[29] Millar AA, Gubler F. The Arabidopsis GAMYB-like genes, MYB33 and MYB65, are microRNA-regulated genes that redundantly facilitate anther development. Plant Cell. 2005;17:705–721. DOI:10.1105/tpc.104.027920

[30] Schmid M, Uhlenhaut NH, Godard F, Demar M, Bressan R, Weigel D, Lohmann JU. Dissection of floral induction pathways using global expression analysis. Development. 2003;130:6001–6012. DOI:10.1242/dev.00842

[31] Schwab R, Palatnik JF, Riester M, Schommer C, Schmid M, Weigel D. Specific effects of microRNAs on the plant transcriptome. Dev. Cell. 2006;8:517–527. DOI:10.1016/j.devcel.2005.01.018

[32] Palatnik JF, Allen E, Wu X, Schommer C, Schwab R, Carrington JC, Weigel D. Control of leaf morphogenesis by microRNAs. Nature. 2003;425:257–263. DOI:10.1038/nature01958

[33] Achard P, Herr A, Baulcombe DC, Harber NP. Modulation of floral development by a gibberellin regulated microRNA. Development. 2004;131:3357–3365. DOI:10.1242/dev.01206

[34] Kerstetter RA, Poethig RS. The specification of leaf identity during shoot development. Annu. Rev. Cell. Dev. Biol. 1998;14:373–398. DOI:10.1146/annurev.cellbio.14.1.373

[35] Lauter N, Kampani A, Carlson S, Goebel M, Moose SP. MicroRNA172 down-regulates glossy15 to promote vegetative phase change in maize. Proc. Natl. Acad. Sci. USA. 2005;102:9412- 9417. DOI:10.1073/pnas.0503927102

[36] Nath U, Crawford BC, Carpenter R, Coen E. Genetic control of surface curvature. Science. 2003;299:1404–1407. DOI:10.1126/science.1079354

[37] Xie Z, Kasschau KD, Carrington JC. Negative feedback regulation of Dicer-like1 in *Arabidopsis* by microRNA-guided mRNA degradation. Curr. Biol. 2003;13:784–789. DOI:10.1016/S0960-9822(03)00281-1

[38] Vaucheret H, Vazquez F, Crete P, Bartel DP. The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. Genes Dev. 2004;18:1187–1197. DOI:10.1101/gad.1201404

[39] Allen E, Xie Z, Gustafson AM, Carrington JC. MicroRNA-directed phasing during trans-acting siRNA biogenesis in plants. Cell. 2005;121:207–221. DOI:10.1016/j.cell.2005.04.004

[40] Lee RC, Feinbaum RL, Ambros V. The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to Iin-14. Cell. 1993;75:843–854. DOI: 10.1016/0092-8674(93)90529-Y

[41] Ambros V. MicroRNAs: Tiny regulators with great potential. Cell. 2001;107:823–826. DOI:10.1016/S0092-8674(01)00616-X

[42] Lee RC, Ambros V. An extensive class of small RNAs in *Caenorhabditis elegans*. Science. 2001;294:862–864. DOI:10.1126/science.1065329

[43] Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T. Identification of novel genes coding for small expressed RNAs. Science. 2001;294:853–858. DOI:10.1126/science. 1064921

[44] Zhang B, Chen Y, Hu X, Guo J. Affiliated with The Institute of Tropical Bioscience and Biotechnology, Chinese Academy of Tropical Agricultural Sciences (CATAS). Computational identification of microRNAs and their targets in apple. *Comput.* Biol Chem, 2006;30:395–407. DOI:10.1007/s13258-013-0070-z

[45] Dai Q, Smibert P, Lai EC. Exploiting Drosophila genetics to understand microRNA function and regulation. In E. Hornstein, ed. *MicroRNAs in Development*. Academic Press (Elsevier), 2012. p. 222.

[46] Zhang B, Wang Q. MicroRNA-based biotechnology for plant improvement. J Cell Physiol. 2015;230:1–15. DOI:10.1002/jcp.24685

[47] Neilsen CT, Goodall GJ, Bracken CP. IsomiRs-The overlooked repertoire in the dynamic microRNAome. Trends Genet. 2012;11:544–549. DOI:10.1016/j.tig.2012.07.005

[48] Kozomara A, Griffiths-Jones S. MiRBase: Annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res. 2014;42:68–73. DOI:10.1093/nar/gkt1181

[49] Zhang, Z, Yu J, Li D, Zhang Z, Liu F, Zhou X, Wang T, Ling Y, Su Z. PMRD: Plant microRNA database. Nucleic Acids Res. 2009;38:806–813. DOI:10.1093/nar/gkp818

[50] Yi X, Zhang Z, Ling Y, Xu⸍ W, Su Z. PNRD: A plant non-coding RNA database. Nucleic Acids Res. 2014;43: D982–D989. DOI:10.1093/nar/gku1162.

[51] Szcześniak MW, Deorowicz S, Gapski J, Kaczyński Ł, Makalowska I. miRNEST database: An integrative approach in microRNA search and annotation. Nucleic Acids Res. 2012;40:198–204. DOI:10.1093/nar/gkr1159

[52] Thomson DW, Bracken CP, Goodall GJ. Experimental strategies for microRNA target identification. Nucleic Acids Res. 2011;39:6845–6853. DOI:10.1093/nar/gkr330

[53] Leng X, Leng X, Fang J, Pervaiz T, Li Y, Wang X, Liu D, Zhu X, Fang J. Characterization of expression patterns of grapevine microRNA family members using microRNA rapid amplification of complementary DNA ends. Plant Genome. 2015;8:1–9. DOI:10.3835/ plantgenome2014.10.0069

[54] Wang C, Han J, Liu C, Kibet KN, Kayesh E, Shangguan L, Li X, Fang J. Identification of microRNAs from Amur grape (*Vitis amurensis* Rupr.) by deep sequencing and analysis

of microRNA variations with bioinformatics. BMC Genomics. 2012;13:122. DOI: 10.1186/1471-2164-13-122

[55] German M, Luo S, Schroth G, Meyers BC, Green PJ. Construction of Parallel Analysis of RNA Ends (PARE) libraries for the study of cleaved miRNA targets and the RNA degradome. Nat. Protoc. 2009;4:356–362. DOI:10.1038/nprot.2009.8

The Extraordinary Nature of RNA Interference in Understanding Gene Downregulation Mechanism in Plants

Jorge Ricaño-Rodríguez, Jacel Adame-García, Silvia Portilla-Vázquez, José M. Ramos-Prado and Enrique Hipólito-Romero

Additional information is available at the end of the chapter

Abstract

Gene silencing (also known as ribonucleic acid [RNA] interference [RNAi] or interfering RNA) was first recognized in plants and is considered one of the most significant discoveries in molecular biology in the last several years. These short-chain ribonucleic acid molecules regulate eukaryotic gene expression. The phenomenon involves a process that promotes RNA transcripts degradation through complementarity between RNA molecules and RNAi transcripts, resulting in the reduction of their translation levels. There are two principal classes of regulatory RNA molecules: small interfering RNAs (siRNA) and microRNAs (miRNA). Both are generated from the cleavage of double-stranded self-complementary RNA hairpins by a DICER enzyme that belongs to the RNase III family. Small RNAs (of about 21–24 nucleotides in size) guide specific effector Argonaute protein to a target nucleotide sequence by complementary base pairing. Thereby, the effector protein complex downregulates the expression of RNA or DNA targets. In plants, *cis*-regulatory RNAi sequences are involved in defense mechanisms against antagonistic organisms and transposition events, while *trans*-regulatory sequences participate in growth-related gene expression. siRNA also performs neutral antiviral defense mechanisms and adaptive stress responses. This document is an attempt to scrutinize the RNAi nature in understanding gene downregulation mechanism in plants and some technical applications.

Keywords: Plant gene silencing, RNAi biosafety, RNA-directed DNA methylation, RNA interference, small interfering RNA

1. Introduction

The discovery of ribonucleic acid (RNA) interference is undoubtedly one of the most important scientific events of the last decades. The beginning of this fascinating story takes place for the

first time in the early 1990s, when a few scientists attempted to increase the color in petunia flowers (*Petunia* sp.), through the addition of target gene copies involved in pigment biosynthesis pathways that were joined to very strong promoters and inserted into the petunia genome. Although respective results showed a decrease in floral color, those expected should be just the opposite. This meant that some transgenic plant lines used in the experiments exhibited suppression or co-suppression (gene silencing) that may be coordinated of both the transgene and the homologous endogenous plant gene. Therefore, it was concluded that plant tissues exhibiting gene suppression (co-suppression) had showed strong evidence of reduced steady-state levels of transgene and homologous messenger RNA (mRNA) [1–2].

Plant RNA silencing is divided into transcriptional gene silencing (TGS) and posttranscriptional gene silencing (PTGS) based on its action target. Although the molecular mechanism behind this phenomenon was unrecognized, shortly before, the results of co-suppression assays related to the production of tobacco etch virus (TEV)-resistant plants using transgenic lines that express the TEV coat protein were published [3–5].

Gene silencing was also referred to gene quelling in plants and fungi and later RNAi in animals. It is considered a conserved regulatory mechanism of gene expression and has been mostly characterized in eukaryotic cells. As far as we know, RNA silencing leads to a specific nucleotide sequencing process in plants that induces mRNA degradation or translation inhibition at the posttranscriptional level. On the other hand, in plants, it sometimes can cause epigenetic modifications at the transcriptional level, which depend on a process called RNA-directed DNA methylation (RdDM) [6–7]. In addition, siRNA-mediated RNA silencing also serves as natural antiviral defense mechanism (*e.g.,* virus-induced gene silencing [VIGS]) [8].

Since miRNA-mediated gene silencing pathway has emerged as a key regulatory mechanism for controlling gene expression, recent discoveries have shown that this pathway is composed of a series of different important components. Among others, it starts with a double-stranded RNA (dsRNA) trigger, followed by an intermediary processor called DICER (Argonaute protein) or a DICER-like protein (DCL). This peptide is a member of the endoribonucleases RNase III family that specifically cleaves dsRNA. The processor product, which consists of small RNAs (siRNAs or miRNAs) of about 21–24 nucleotides (nt) in size, activates an effector complex called RISC (RNA-induced silencing complex), where the Argonaute protein (AGO) (*i.e.,* essential catalytic component) works as a key player to initiate gene expression regulation. Posteriorly, RNA-dependent RNA polymerase (RDR) amplifies the dsRNA target (siRNAs-guided AGO) and cleaves the target RNA. These molecular interactions stabilize the dsRNA substrate to produce secondary siRNAs and maximize the silencing process. The entire complex is considered a gene silencing suppressor (GSS) [7].

Due to its effectiveness and relative ease of use, gene silencing technique has become a potential tool in both basic and applied research. Given the fact that phytopathogenic microorganisms are a major cause of plant diseases, RNA silencing-based resistance proves to be an effective biotechnological alternative to engineer resistant crops, among other applications. In either case, it is necessary to generate dsRNA trigger molecules before using RNAi to silence target genes that help to metabolic engineering of transgenic plants and generation of pest-resistant crops by inserting into plants a transgene that will produce homologous miRNA sequences. Finally, the recent discovery of dsRNA in unicellular eukaryotes implies that miRNAs have a

deep evolutionary history. The last indicates dsRNAs have evolved independently within eukaryotes through exaptation of their shared and inherited RNAi machinery [9].

2. RNAi machinery: Brief overview of its biogenesis

It is noteworthy that some authors believe that RNAi was first discovered in plants as "co-suppression" [1–2], but not in worms as PTGS [10]. For creating transgenic plants, several attempts have been made to engineer more desirable characteristics [11]. This is how the "co-suppression" concept was coined to explain the ability of exogenous elements to modify gene expression. Currently, the general comprehension that we have about RNAi emerges from an evolutionarily conserved gene regulatory mechanism in higher organisms.

It is known that some other molecules related to siRNA (*i.e.*, trans-acting siRNA and repeat-associated small interfering RNA commonly abbreviated "TAS," "ta-siRNA" or "tasiRNA" and "rasiRNA") repress gene expression through PTGS in plants. All of these molecules are a class of small RNAs involved in the RNAi pathway [11]. Many times, RNAi is considered a quelling process because it is the result of overexpression or suppression of specific transgenes.

According to some authors [12–13], dsRNA was characterized in detail after injecting anti-sense-stranded RNA into an organism that was an effective way to inhibit gene function. This was the first attempt to use an antisense RNA approach to inactivate a *Caenorhabditis elegans* gene [14]. Due to the above results and thanks to further investigations, it was concluded that the active molecules that triggered this phenomenon could be considerable amounts of dsRNA that interfered in vitro RNA transcripts. dsRNA injection into the nematode acted systematically to cause posttranscriptional depletion of homologous RNA. This methodology offers a way of specific and potent inactivation of gene function. It is also known that RNAi acts systemically when injected into the animal´s tissue, inhibiting gene function.

Through a variety of experiments, it has been suggested that RNAi destabilizes cleaved RNA after its processing. The nature of RNAi inspired Timmons and Fire [15] to perform a simple but efficient experiment that produced an astonishing result. Several nematodes were fed with bacteria that had been engineered to express dsRNA corresponding to *C. elegans unc*-22 gene. The organisms showed a similar phenotype (dependent on their food source) to that of *unc*-22 mutants. The ability to expose a vast number of samples with dsRNA established the basis for the development of a versatile tool to select RNAi-defective *C. elegans* mutants as well as target genes [16]. Small RNA molecules have been described according to their origin and function (*i.e.*, siRNAs, rasiRNAs, and miRNAs). RNA polymerization may produce dsRNA in nature (*e.g.*, viruses).

Although it is very common to observe transcript overlapping from repetitive sequences such as transposons and transgene arrays, dsRNA is rapidly processed into short RNA duplexes of about 21–28 nucleotides in length. A clear example of the natural function of these molecules is mRNAs or viral genomic/antigenomic RNAs that are recognized and split to several particles (translationally repressed). In addition, short RNAs are implicated in guiding chromatin modification [7]. RNA silencing mechanisms have been also recognized as antiviral defense against exogenous RNA viruses and random integration of transposable element transcripts.

The general role of gene silencing only became clear when it was realized that specific genes in plants and animals encode short forms of fold-back dsRNA5 (precursor molecules of miRNAs) [17]. There are three different metabolic pathways that induce RNAi and share a common molecular mechanism. These are currently known as miRNA, siRNA, and Piwi-associated RNA (RNAi that prevents transposons mobility through the genome), although the last one has been only found in animals [18]. Gene silencing is part of an miRNA or siRNA complex that works as splicing pattern to identify nucleotide sequences ready for degradation via RISC machinery.

The RISC complex is the result of several enzyme couplings involved in RNAi mechanism, that mediate target mRNA silencing through degradation or translational inhibition. miRNA production starts from a pre-miRNA (primary miRNA) transcript whose length sequence is about of 1000 nucleotides and create complementary loops, either single or double, as well as complementary sequences (5′–3′) [19]. Since this mechanism involves both endogenous and exogenous microsequences, their precursors produce dsRNA molecules of appropriate size in order to be linked to an effector protein. This phenomenon is mediated by an endoribonuclease enzyme (class III; DICER) with different structural domains, although the most important are those called PAZ (Piwi, Argonaute, and Zwelli) and helicase (*i.e.*, specific amino acid sequence responsible for unpacking genes). After an intensive search for the enzymatic mechanisms of gene silencing, DICER enzymes were first recognized as responsible for processing dsRNA to siRNA in *Drosophila* [20]. These enzymes contain a helicase and a couple of dimerized RNase and PAZ domains, although variability among organisms can be observed.

Helicase domains are RNAi precursors, which are perfectly aligned with dsRNA. Moreover, helicase metabolizes ATP (adenosine triphosphate) to translocate enzymes in order to generate a large number of sequences [21]. In plant genera such as *Arabidopsis*, DICER DCL1 (DICER-like1) proteins converge sequentially with pre-miRNAs for synthesizing loops and posteriorly with dsRNA of about 21 nucleotides in length. Through partial sequence alterations of RNA helicase domains caused by point mutations, it has been observed a reduction phenomenon of the amount of mature miRNA sequences. It is now known that plant DLC1 proteins are essential for a proper embryonic development [22].

In DICER proteins, PAZ domains have been extensively studied. Structurally, they have similarities to oligonucleotide–oligosaccharide structures, and theoretically, PAZ domains recognize the 3′ end of RNA substrates. On the other hand, recent studies have shown that they link not only the 3′- but also their 5′-phosphorylated substrates, where cleavage positions are recognized at a distance of 22 nucleotides [23–24]. In the conventional RNAi model, DICER enzymes interact in the cytoplasm to degrade their substrates prior to the RISC complex linkage.

DICER enzymes are important siRNA and miRNA intermediary pathways and generate dsRNA molecules as imperative substrates for Argonaute proteins. DICER are also considered common effectors of ribonucleoproteinic complexes linked to a single RNA sequence of 20–30 nucleotides complemented to target genes and conduct, at the same time, mRNA degradation [25]. Argonaute proteins contain four domains: terminally-N, PAZ, middle (MID), and Piwi terminally-C. The latter is typical of such complexes [26].

Many organisms express multiple members of this superfamily of proteins. For example, *Homo sapiens*, *Drosophila melanogaster*, and *Arabidopsis thaliana* express up to 8, 5, and 10 peptides, respectively. Individual members of each family are highly specialized in carrying out gene silencing process [23]. One of the most prominent roles of this class is its relationship with pre-ribosomal RNA synthesis (pre-rRNA) [27]. During the miRNA formation, HASTY proteins (exporter miRNA proteins) translocate their precursor into the cytoplasm. Subsequently, double-stranded precursor is dissociated and miRNA guide sequence is incorporated into a containing AUG protein complex, usually to form a specific RISC complex (miRISC) [28]. AGO1 PAZ domain complex links to miRNA and helps to incorporate miRISC. miRISC–miRNA complex prevents target gene expression, by either mRNA cleavage or translation inhibition [29]. In miRNA processing, introns among pre-miRNA sequences are removed through RNA splicing (posttranscriptional RNA maturation).
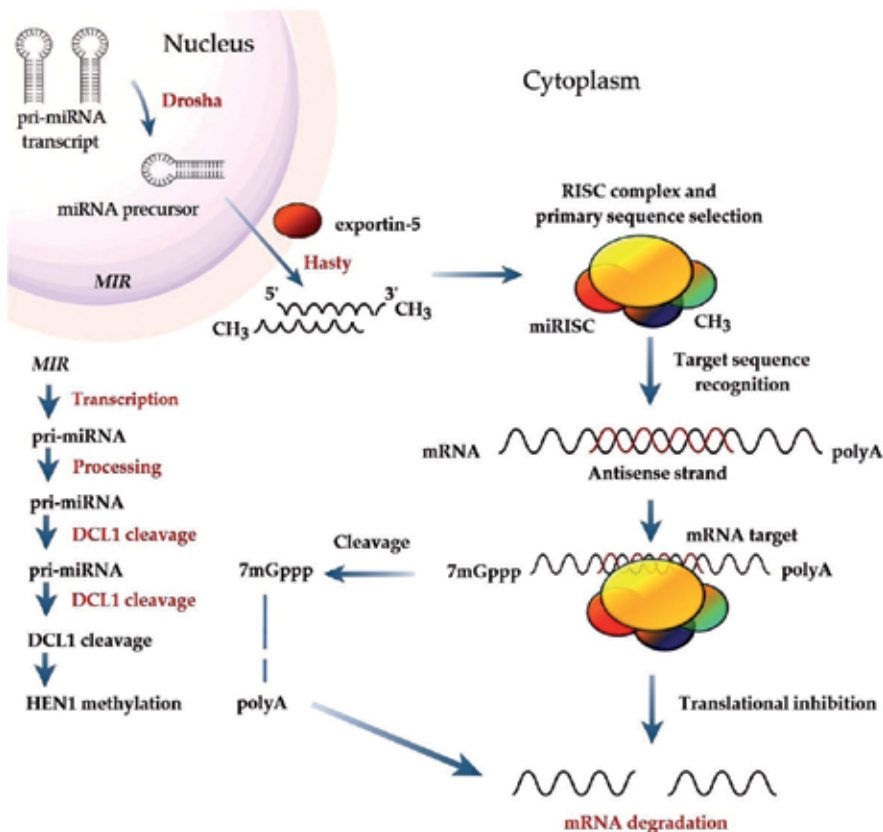
Figure 1. Pathway of siRNA-guided posttranscriptional regulation of gene expression. RNA polymerase II is mediating miRNA genes (*miR*) transcription that generates primary miRNA. DICER (DCL1) processing takes place in the nucleus through cap-binding complexes (CBC), DAWDLE (DDL), dsRNA-binding RNA protein hypnotastic leaves 1 (HYL1), and Hua enhancer 1 (Hen1) protein interactions. The HASTY (HST1; nucleocytoplasmic transporter activity) ortholog transports methylated miRNA to the cytoplasm and miRNA is coupled to RISC complex. miRNA guides miR–RISC complex in order to silence target mRNA by either excision or translational inhibition [99].

It has been recently discovered that there are ribonucleotide structures at the intermediate stage of the metabolic complex that allow the synthesis of specific molecules known as noncoding RNAs (ncRNAs), which are also considered regulatory RNA molecules (of 200 nucleotides) that are not translated into proteins [30]. They are intermediaries of target mRNA degradation that is finally identified by RISC complex, whose function is defined by different protein interactions [25]. Endoribonuclease RNase III DICER enzyme is the majorly involved key in RNAi and miRNA pathways. It plays an important role in assembling the RISC complex in addition to its catalytic function over microsequences [31].

RNase III DICER family enzymes are important intermediaries for siRNA and miRNA pathways. These peptides generate dsRNAs that will be linked to an Argonaute protein. Bacterial RNase III class I enzymes form DICER's active site, and it comprises a terminally-C RNase III domain [18]. In addition, prokaryotic enzymes are capable to dimerize and achieve a cleavage of both strands of dsRNA. DICER enzymes use RNase III pseudodimer domains of a single polypeptide with a single double-stranded RNA-binding domain (dsDRBD) to accomplish a similar dsRNA cleavage [32]. PAZ domain of these paired active sites has a terminal-N domain, and it recognizes the dsRNA end that is characteristic of RNAi intermediaries.

DICER proteins complexity can be attributed to multiple domain levels, ranging from several combinations of catalytic RNase III as well as the number of differently expressed proteins in single organism. In a generic RNAi model, DICER enzymes function in the cytoplasm, where they cleave their substrates before loading into RISC complex [23]. In recent years, DICER enzymes have been receiving much attention because they are capable of playing an important role in transcriptional gene silencing. Limited evidence suggests that DICER may also be found and functional in mammal cells. Among all DICER non-catalytic domains, PAZ has been one of the most intensively studied domains because of its presence in AGO proteins recognizing 3'-nucleotides of siRNAs [33].

3. Role of miRNAs in plant immunity

Eukaryotic cells are capable of modulating the stability of their miRNAs in response to environmental and endogenous stimuli and/or to regulate mRNA transcription levels (regulating mRNA transcript level). Such alterations in reducing mRNA levels are mediated by RNAi *cis* regulator and by RNA-binding proteins [34–35].

miRNA sequences are often related to the regulation of various biological processes such as stress mitigation [36]. *Arabidopsis* has two miR393a and miR393b genes that are processed almost identically when they mature and subsequently become miR393 sequences. This miRNA has been considered a nonfunctional sequence [37]. However, later studies showed the involvement of these molecules in plant immunity because of their interaction with AGO proteins during bacterial infections [38]. The sequence has a target gene called *MEMB12*, which encodes a structural protein of Golgi apparatus involved in vesicular secretion processes.

Plants respond to either biotic or abiotic environmental stresses by differential gene expression and miRNA sequences regulation. In several plant species, increased expression of miR160, miR167, and miR393 have been observed during drought conditions. It is known that miR393 blocks the expression of a gene encoding auxin receptors, while miR167 and miR160 interfere with the expression of some genes related to stress responses [39]. In addition, plant *miR-*sequences play important regulatory roles in many other processes (refer to Table 1 for some detailed examples).

Description	Annotation	Mature sequence	Reference
Arabidopsis thaliana miR156a stem-loop	Regulatory roles through complementary to mRNA	ath-miR156a-5′ (21-40 nt) ath-miR156a-3′ (83-104 nt)	[79]
Arabidopsis thaliana miR167a stem-loop	Target of mRNAs coding for auxin response factors, DNA binding proteins related to control transcription in response to the phytohormone auxin	ath-miR167a-5′ (19-39 nt) ath-miR167a-3′ (101-121 nt)	[80]
Arabidopsis thaliana miR168a stem-loop	Target of mRNAs coding for Argonaute (AGO1) proteins	ath-miR168a-5′ (18-38 nt) ath-miR168a-3′ (103-123 nt)	[79]
Arabidopsis thaliana miR169a stem-loop	Target of mRNA coding for CCAAT binding factor (CBF)-HAP2-like proteins	ath-miR169a-5′ (18-38 nt) ath-miR169a-3′ (190-209 nt)	[81]
Arabidopsis thaliana miR170a stem-loop	Target of mRNAs coding for GRAS domain (family of transcription factors whose members have been implicated in radial patterning in roots, signaling by gibberellin and light signaling	ath-miR170a-5′ (18-38 nt) ath-miR170a-3′ (190-209 nt)	[82]
Arabidopsis thaliana miR172a stem-loop	Target of mRNAs coding for APETALA2-like transcription factors	ath-miR172a (78-98 nt)	[81]
Nicotiana tabacum miR6020b stem-loop	Regulatory roles through complementary to mRNA	nta-miR6020b (21-41 nt)	[83]
Oryza sativa miR156a stem-loop	Regulatory roles through complementary to mRNA	osa-miR172a (7-26 nt)	[80]
Physcomitrella patens miR1049 stem-loop	Regulatory roles through complementary to mRNA	ppt-miR1049 (89-109 nt)	[84]
Populus trichocarpa miR156d stem-loop	Family of plant non-coding RNA	ptc-miR156d (11-30 nt)	[85]
Ricinus communis miR156a stem-loop	Target of mRNAs coding for Argonaute (AGO1) proteins	rco-miR156a (6-26 nt)	[86]

Description	Annotation	Mature sequence	Reference
Saccharum officinarum miR408c stem-loop	Regulatory roles through complementary to mRNA	sof-miR408c (247-267nt)	[87]
Selaginella moellendorffii miR156 stem-loop	Regulatory roles through complementary to mRNA	smo-miR156c (11-31 nt)	[84]
Solanum tuberosum miR6022-stem-loop	Regulatory roles through complementary to mRNA	stu-miR6022 (197-217 nt)	[83]
Zea mays miR156b stem-loop	Regulatory roles through complementary to mRNA	zma-miR156b-5′ (21-40 nt) zma-miR156b-3′ (86-106 nt)	[88]

Table 1. Examples of representative microsequences and their role in plant physiology

Plants require at least 14 essential minerals coming from the soil for proper development; therefore, RNAi is involved in both regulation and homeostasis of nutrients [40]. It is worth mentioning that constructions of genomic libraries have proved to be very valuable for studies of miRNAs associated with these metabolic processes [41]. Thereby, biotechnological applications of miRNAs might require microarray studies helping to discover important miRNA-associated metabolic responses to water, heat, salt, biotic stress, and UV radiation, as well as stress-mediated hormonal regulation and nutrient homeostasis, and resulting in future creations of "biotech" lines resistant to adverse environmental conditions.
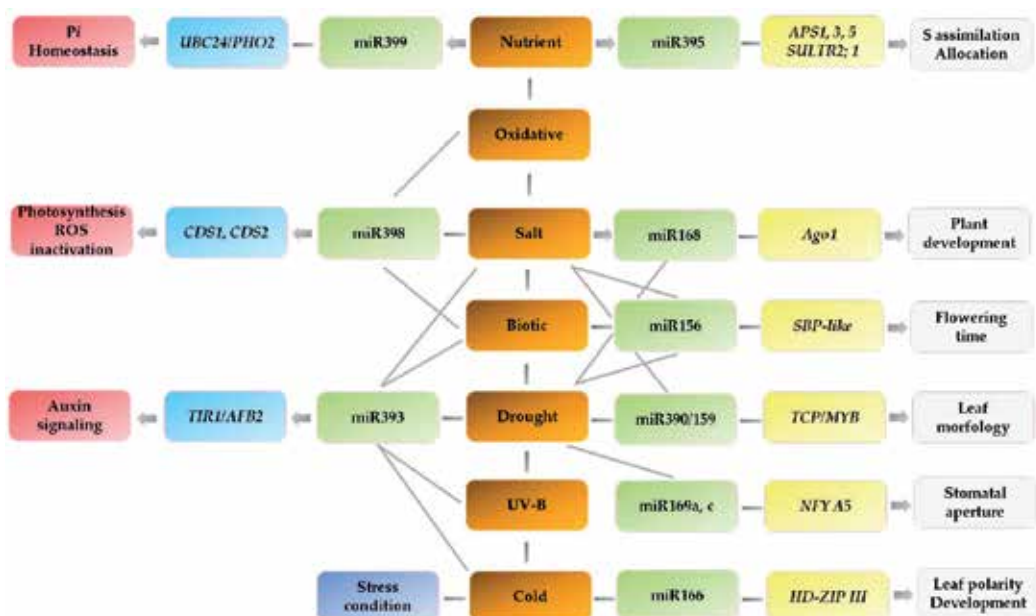
Figure 2. Summary of representative plant miRNAs involved in response to stresses. Modified from Kruszka et al., 2012 [40].

4. RNAi in crop protection against pest insects

As mentioned above, one of the first researches showing that RNAi could degrade specific mRNA sequences, resulting in blocking of the expression of certain insect genes, was conducted in *C. elegans*, a rhabditoid nematode [14]. The responsible researchers behind the project shared the Nobel Prize for Medicine in 2006 for what they called "a fundamental mechanism for controlling the flow of genetic information." To date, functional genomics using RNAi technology has been studied in several insect species, including orders such as Diptera, Coleoptera, Lepidoptera, Isoptera, Orthoptera, Hymenoptera, and Hemiptera, among others [42].

The functional approach of this tool has been successful in characterizing genes related to different physiological processes, including development, reproduction, behavior, and immune systems [43–44]. A viable biological control strategy based on RNAi application should target a gene that is vital for a proper physiological process as well as require an efficient delivery method for RNAi triggers. Recent research in insects has shown the in vitro micro-injection effect of synthetic double-stranded sequences in embryos [45]. However, although this delivery method provides a tool for understanding gene function, dsRNA microinjection may not be feasible for pest control due to its high cost. RNAi potential as biotechnological tool for controlling insect populations was first demonstrated after oral introduction of dsRNA into insect body [46]. The study was conducted using *Rhodnius prolixus* larvae, which were fed with a dsRNA developed from the genes sequence coding protein called nitrophorin 2 (anticoagulant encoding transcripts), after which a significant decrease of anticoagulant activity levels on insect´s salivary glands was observed.

In the same year, a research that involved *Epiphyas postvittana*, a lepidopteran that is capable of attacking up to 123 different species of dicotyledonous, was conducted [47]. Oral introduction of dsRNA target encoding intestinal proteins as well as intermediary pheromone-like protein synthesis in adult antennas decreased mRNA transcript levels in both tissues. In addition, assays related to *Aedes aegypti* showed that RNAi may be induced in insects through topical application [48]. In this study, dsRNA diluted in acetone caused *AaeIAP1* gene transcription that encodes an inhibitor protein of programmed cell death (apoptosis) in adult females, remained blocked. Thus, a significant increase in insect's mortality was observed.

Posteriorly, topical application of such molecules in borer moth larvae *Ostrinia furnacalis* showed similar effect. It was observed when RNAi inducer was introduced into larvae by direct spray of an aqueous solution containing double-stranded ribonucleotides, after which insects showed stunted growth as well as early death. Moreover, a significant reduction in egg hatchability compared to controls was observed. Besides, fluorescently labeled dsRNA molecules persisted in larval stages once they reached the intestine and hemocytes [49].

As mentioned above, artificial in vitro RNAi is expensive. Alternatively, a construction of a target gene-specific dsRNA vectors, its insertion into insect genomes and subsequent in vivo expression could be economically beneficial approach. Several recent investigations have allowed obtaining silencing vectors in bacteria host plants and plant viruses, which have been successfully implemented to study the expression of specific insect genes [50–53].

In addition, one way to generate genetically modified nematode-resistant plants is to produce copies (repeated and inverted) of target gene sequences in the plant tissue so that worms eating dsRNA-bearing plant material suffer from rapidly induced and triggered RNAi of important insect gene (s) under target. Although the results of RNAi potential to control insect pests as well as beneficial insects from parasites and diseases are encouraging, more research is necessary to understand the barriers and an efficient application. In the last several years, technical problems were uncovered, although a lot of concerns still remain. Future scientific efforts will help to solve current obstacles, which should allow this technology to be applied for integrated pest management (IPM) strategies as a novel way of action [54–57].

5. Gene silencing and viral immunity

Although there is little scientific background related to RNAi potential against various types of viruses capable of infecting animal cells (*e.g.,* dengue virus and *Drosophila*) [58–59], some studies suggest RNAi involvement in plant pathogenicity. Silencing viral suppressors affect the accumulation and function of siRNAs, including *trans*RNAi-mediated posttranscriptional gene silencing process that was recently discovered (tasiRNA; trans-acting siRNA). As a result, abnormal development of host organisms is often triggered [60–61]. As mentioned above, it can be considered that the effectiveness of RNAi technology was first demonstrated in 1998 [14]. In past decades, RNAi application was a successful tool for controlling various "difficult-to-eradicate" viral strains causing different pathologies in the wide range of economically important crops [62].

Plant gene silencing induced by viral agents (*i.e.,* VIGS; virus-induced gene silencing in plants) is one of the most common techniques that involves RNAi as immune mediator [63]. This technology allows implementing a system that releases dsRNA sequences in order to identify target viral genes, which generate multiple resistance mechanisms. In stable transgenic plants, this manipulation may require sequential processing or cross-linking among dsRNA sequences for considerable periods of time [64].

In addition, using RNAi has resulted in increasing immune resistance against viruses in different plant species, for example, (1) bean golden mosaic geminivirus (BGMV) [65], (2) rice dwarf virus (RDV) [66], (3) white leaf disease of rice (RHBV) [67], (4) rice tungro baciliform virus (RTBV) [68], (5) African cassava mosaic virus (ACMV) [69], (6) tobacco rattle virus (TRV) [70], and (7) citrus tristeza virus (CTV) [71], among others.

Functional approach of VIGS tool proves to be successful in characterizations of various physiological processes, including gene expression, development, reproduction, behavior, and immune system [43]. Presence of gene expression inhibitors in development of such diseases has to be consistent with the fact that inhibitors usually determine pathogenicity [72–73]. However, RNAi interaction in host metabolic pathways may not be the leading cause of infection symptoms because most of viral suppressors show no affection to plant metabolism [74].

In the conventional RNAi-mediated pathogenicity models, short ribonucleotide sequences are derived from infectious viruses, and host subviral RNA-induced gene silencing is carried out through random sequence complementarities. For example, transcribed gene expression related to self-complementary RNA hairpins (self-complementary hairpin RNA) encoding potato spindle tuber viroid sequences (PSTVd) is also capable of inducing viral symptoms in tomato (*Solanum lycopersicon*) [75]. Furthermore, RNAi-mediated pathogenicity models have shown that a darkening effect of tobacco plant, associated to the tobacco mosaic virus (TMV), is caused by a satellite RNA (*i.e.*, pathogenic RNA molecule). It is strongly inhibited by a silencing suppressor called P1/HC-Pro. Such wilt symptoms are due to a silencing effect on the chlorophyll biosynthetic-encoding *(CHLI)* gene [76–77].

RNAi-mediated gene silencing could be considered a general mechanism for pathogenicity of subviral RNA because such infective molecules may conduct gene silencing in various ways. siRNAs have high sequence identity degree with host´s promoter regions, and it may induce cytosine methylation by RNA-directed DNA methylation (RdDm), leading to transcriptional inactivation [78–82] as well as gene downregulation [83–87].
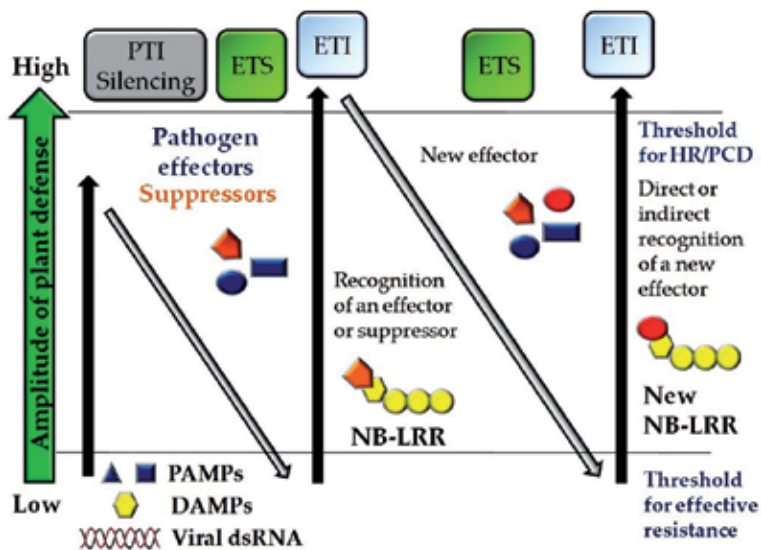
Figure 3. Zigzag model for evolution of innate immunity and silencing-based plant defense against viral and non-viral pathogens. Modified from Jones and Dangl, 2006 [88]. Susceptibility is proportional to PTI + silencing + ETS + ETI. Plants detect pathogen-associated molecular patterns (PAMPs) as well as host danger-associated molecular patterns (DAMPs) via pattern recognition receptors (PRRs) to induce pattern-triggered immunity (PTI). Plants additionally detect viral dsRNA to trigger RNA silencing. Effector-triggered susceptibility (ETS) results from the recognition of NB–LRR protein activating effector-triggered immunity (ETI; amplified version of PTI that passes a threshold for induction of hypersensitive response (HR) and programmed death cell (PDC). Pathogens that have lost the specifically recognized effector/suppressor are selected to help isolates to suppress ETI. NB–LRR plan alleles have evolved and selected to recognize newly acquired effectors resulting in ETI.

The zigzag model proposed by Jones and Dangl [88] shows the initial perception of pathogen-associated molecular patterns (PAMPs) as triggered immunity (TI)-based defense response

(*i.e.*, PAMP-TI) that regulates pathogenic growth and subsequent host infection. However, successful pathogens promote effector/virulence factors through PTI suppression. As a specific counteract action for pathogenic effectors, plants have evolved effector-triggered immunity (ETI), which is considered multiple rounds of effector-triggered susceptibility (ETS) followed by ETI.

On the basis of the above background, Zvereva and Pooggin [89] considered to extend this model to plant–virus interactions. On the other hand, because RNA silencing is an evolutionary conserved mechanism that defends organisms against transgenes and viruses, zigzag model may be related to specific *miR*-gene expression linked at the same time to plant innate immunity.

6. Human health approaches in gene silencing: biosafety and final considerations

The convention of biological diversity is intended to protect species from potential risks of genetic modified organisms (GMO), which are the result of applying modern biotechnological tools. On January 2000, Cartagena Protocol on Biosafety was signed by most of the developed countries. According to the Article 1 of this document, primary aim is to ensure a proper protection level in the field of safe transfer and handling of living modified organisms that may show adverse effects on conservation and sustainable use of biological diversity, considering also risks to human health, and specifically focusing on migration of species.

It is known that plant small RNAs help regulate several physiological processes such as growth and stress responses by attaching target mRNAs to modify their translation. Most people in the earth live on plant-based diets, and their food contains small RNAs from 19–24 nucleotides in size, among other bioactive molecules. Due to this fact, it is common that scientific community may ask the following: are plant small RNAs capable of regulating gene expression into the consumer's genome? [90–91]. Before giving our opinion, some cases of small RNAs/miRNAs application for customized human gene therapy as well as RNAi relationship to food security and environmental biosafety will be discussed.

Over 800 human miRNAs have been discovered to date, and exploiting new platforms for controlling their expression are of urgent need. For example, nanotechnology and biomaterial synthesis have developed solid knowledge of sensing treatments using miRNAs against cancer. It is important to understand that human systemic administration using optimized delivery systems of interfering molecules is critical for proper functioning of *miRs*. Thereby, liposome-based nano-vehicles are capable of efficient transporting of miRNAs and antisense RNA helping to accumulate them easier in the liver, spleen, and kidney [92–94].

If plant-implemented glyco-engineering techniques based on RNAi silencing could reduce target glycosyltransferases transcripts, virus-like particles (VLPs) production in transgenic plants may be a reliable path to develop CHIKV (chikungunya) vaccines, for example [95]. Transgenic rice seeds as bioreactor for molecular pharming systems show great promise for

producing and processing recombinant proteins. Some of the advantages over conventional plant host or animal bioreactors are the following: (1) high capacity to obtain considerable expression levels, (2) production cost is lower than that of conventional fermentation, and (3) high capacity of seed reproduction [96–97].

About two years passed since it was demonstrated the ability of dietary miRNAs to regulate an animal gene in the liver [98]; however, while a few opinions suggested this was a possible way of cross-kingdom gene regulation, majority of data suggest gastrointestinal uptake of dietary plant miRNAs is not possible due to fast acid digestion [99]. On the other hand, measured tissue and blood dietary miRNA levels reported are so few that their dietary impact is insignificant.

Since plants can be modified by engineering RNAi pathways to alternatively generate small RNA molecules, RNAi could generate new crop lines for providing protection against pest insects (including nematodes), without cross-linking new protein varieties into food. Due to this fact, credible ecological risk assessments (ERAs) that are primordial tasks for stakeholders should be constructed. ERAs will allow the characterization of exposure pathways and potential hazards for RNAi crops (*e.g.,* off- and nontarget effects, genetic mutations, and polymorphism) [100]. Risks are also associated with genome direct changes in plants for human consumption, commonly related to newly expressed proteins that eventually show toxicity and allergenicity. However, when aversely a target gene decreased its expression, safety implications in particular cases such as when a silenced enzyme substrate accumulates to toxic levels may be observed [101]. Currently, optimal threshold doses for most food allergens remain unknown, thereby oral challenges test capable of evaluating the effects of RNAi consumption should be carried out in the future [102].

Another major concern about using RNAi-transformed plants for improving crops selection is the use of antibiotic resistance markers because antibiotic resistance genes could raise environmental risks as these genes may trigger horizontal transfer. In that sense, gene horizontal transfer will lead to generating antibiotic resistant microorganisms [103]. On the other hand, transgenic lines such as siRNA-mediated virus-resistant plants may provide a solution to reduce the indiscriminate use of toxic pesticides [97]. It is worth mentioning that during an international scientific workshop (June 2014) organized by the European Food Safety Authority (EFSA), some of the selected key outcomes suggested that bioinformatic analyses will play an imperative role in the identification of possible human and environmental risk assessments of RNAi-based plants [104].

According to Yang and colleagues [90], summary of evidence regarding dietary miRNAs uptake and functionality in mammalian consumers may be divided into two parts: (1) **evidences against:** inconsistent exogenous levels in serum typically low, various feeding studies failed to show absorption of dietary microRNAs, target suppression is shown only in the initial study, in silico analyses suggest that crossed contamination may be the main cause of plant microRNA reads in animal tissue; (2) **evidences for:** oral uptake of miRNAs is well characterized in nematodes and insects (indirect evidence), detection of RNAi from different kingdoms (including mammalian organisms), detection of *miR*-sequences in mice fed with

cabbages, microRNAs absorbed by humans and mice fed milk, tumor suppression observed when miRNAs were orally delivered into mice.

7. Conclusions

The general understanding about RNAi nature is an evolutionary conserved gene regulatory mechanism on superior organisms with several interspecific variations, which allows the survival of species through the reduction of the number of homologous RNA silencing proteins.

RNAi molecular bases that are implemented for fighting several diseases caused by biological agents or extreme abiotic conditions are vital for sustainable agriculture. It has been found that the existence of several virulence factors caused by phytopathogens related to blocking recognition patterns and signaling in immune responses. However, despite knowing the outcome of these physiological processes, it was not entirely clear which could be the molecular mechanisms that trigger such phenomena. Just a few years ago, the principal pathway was discovered and now we know that gene silencing is caused by RNAi, whereby it may regulate gene expression in eukaryote organisms.

It is true that plant metabolic pathways regulate their gene expression through a silencing phenomenon that emerges from siRNA, miRNA, and tasiRNA; however, all these interfering molecules share common elements in their biogenesis and structural characteristics, as well as in action mechanisms involved in common cellular components. Although miRNAs discovery has delved into the role that RNAi plays in plant gene regulation, more questions arise about its nature; for example, how exactly trans-acting elements repress gene expression and how RNA interference is completely involved in the model for evolution of innate immunity and silencing-based plant defense against viral and nonviral pathogens proposed by Jones and Dangl? [88]. Likewise, it would be highly interesting to understand why some similar nature microsequences block the expression of genes encoding auxin receptors while others interfere stress responses (*e.g.*, miR393, miR167, and miR160, respectively) [39].

Small RNAi-directed gene regulation mechanism was independently discovered in plants, fungi, worms, and mammalian cells, and scientific attention has been focused mainly on the regulation of development, biotic and abiotic stress responses, as well as genome stability through controlling plant gene expression. In addition, the siRNA-mediated RNA silencing also functions as a neutral antiviral defense mechanism.

Some authors consider the future possibility of having a better approach on the exact location of target genes from agricultural interest organisms (*i.e.*, crops and insects) by means of artificial microRNAs generation (amiRNA) [105]. Such projections could improve research in crop plants and metabolic engineering through developing better predictable and artificial manipulable microsequences. miRNAs are also being exploited recently as new platforms for developing solid knowledge in different science fields such as medicine, nanotechnology, and integrated pest management. Thus, synthesis of RNAi in plant-based biofactories could be effective in several disciplines involved in forthcoming experiments.

Recent advances have shown the potential of RNAi for its future role in transgenic plants against pest insects in the environment [100]. Perhaps the most relevant application will be in modifying crop–pest interactions so that transgenic lines are capable of producing secondary metabolites against nematodes and some other pathogens. In fact, some researchers have proposed to extend this approach for controlling mammalian diseases.

The recent discovery of some of the most important RNAi molecular mechanisms is useful to discuss future applications in agricultural biotechnology, and attending the resulting food security concerns emerged from the *in situ* application of such tool must be imperative. As a result of a couple of studies on human effects of the consumption of plant foods with high levels of interfering microsequences, considerable uncertainties become noticeable, for example, the effect of these microarrays on the metabolism of those who directly consume engineered plant foods [90].

So far, limited reports related to food security as well as environmental risks involving RNAi are available, since RNAi biotechnological approaches are very difficult to scrutinize and, consequently, proofs of concept are difficult to obtain. In the future, potential and limitations of engineered plants, including alternative strategies for generating low allergic supplies like low weight proteins, should be studied by using bioinformatic tools followed by the respective studies (*i.e.*, physiological characterization of transgenic plants, toxicity and allergenicity of expressed proteins, as well as metabolites production and nutritional characteristics) [102].

Author details

Jorge Ricaño-Rodríguez[1*], Jacel Adame-García[2], Silvia Portilla-Vázquez[3],
José M. Ramos-Prado[1] and Enrique Hipólito-Romero[1]

*Address all correspondence to: jorgericano@gmail.com

1 Center for Eco-Literacy and Dialogue of Knowledge. University of Veracruz. Campus USBI. Col Emiliano Zapata. Xalapa, Veracruz, México

2 Cell Biology Laboratory. Technological Institute of Úrsulo Galván. Úrsulo Galván, Veracruz, México

3 Food Research and Development Unit (UNIDA). Laboratory of Genetics. Technological Institute of Veracruz, Veracruz, México

References

[1] Napoli C, Lemieux C, Jorgensen R. Introduction of a chimeric chalcone synthase gene into *Petunia* results in reversible cosuppression of homologous genes in trans. Plant Cell. 1990;2:279-289. DOI: 10.1105/tpc.2.4.279

[2]   van der Krol AR, Mur LA, Beld M, Mol JN, Stuitje AR. Flavonoid genes in petunia: Addition of a limited number of gene copies may lead to a suppression of gene expression. Plant Cell. 1990;2:291-299. DOI: 10.1105/tpc.2.4.291

[3]   Lindbo JA, Dougherty WG. Pathogen derived resistance to a potyvirus: Immune and resistant phenotypes in transgenic tobacco expressing altered forms of a potyvirus coat protein nucleotide sequence. Molecular Plant-Microbe Interactions. 1992;2:144-153. DOI: 10.1094/MPMI-5-144

[4]   Lindbo JA, Dougherty WG. Untranslatable transcripts of the tobacco etch virus coat protein gene sequence can interfere with tobacco etch virus replication in transgenic plants and protoplasts. Virology. 1992;189:725-733. DOI: 10.1016/0042-6822(92)90595-G

[5]   Lindbo JA, Siolva-Rosales L, Proebsting WM, Dougherty WG. Induction of a highly specific antiviral state in transgenic plants: Implications for regulation of gene expression and virus resistance. Plant Cell. 1993;5:1749-1759. DOI: 10.1105/tpc.5.12.1749

[6]   Prins M, Laimer M, Noris E, Schubert J, Wassenegger M, Tepfer M. Strategies for antiviral resistance in transgenic plants. Molecular Plant Pathology. 2008;9:73-83. DOI: 10.1111/j.1364-3703.2007.00447.x

[7]   Peragine A, Yoshikawa M, Wu G, Albrecht HL, Poethig RS. SGS3 and SGS2/ SDE1/ RDR6 are required for juvenile development and the production of trans-acting siRNAs in *Arabidopsis*. Genes & Development. 2004;18:2368-2379. DOI: 10.1101/gad.1231804

[8]   Ding SW. RNA-based antiviral immunity. Nature Reviews Immunology. 2010;10:632-644. DOI: 10.1038/nri2824

[9]   Tarver JE, Donoghue PCJ, Peterson KJ. Do miRNAs have a deep evolutionary history? Bioessays. 2012;34:857-866. DOI: 10.1002/bies.201200055

[10] Meister G, Tuschl T. Mechanisms of gene silencing by double-stranded RNA. Nature. 2004;431:343-349. DOI: 10.1038/nature02873

[11] Jorgensen RA, Cluster PD, English J, Que Q, Napoli CA. Chalcone synthase cosuppression phenotypes in petunia flowers: Comparison of sense vs. antisense constructs and single-copy vs. complex T-DNA sequences. Plant Molecular Biology. 1996;31:957-973. DOI: 10.1007/BF00040715

[12] Hunter GP. Gene silencing: Shrinking the black box of RNAi. Current Biology. 2000;10:R137-R149. DOI: 10.1016/S0960-9822(00)00325-0

[13] Duan CG, Wang CH, Guo HS. Application of RNA silencing to plant disease resistance. Silence. 2012;3:5. DOI: 10.1186/1758-907X-3-5

[14] Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature. 1998;391:806-811. DOI: 10.1038/35888

[15] Timmons L, Fire A. Specific interference by ingested dsRNA. Nature. 1998;395:854. DOI: 10.1038/27579

[16] Tabara H, Sarkissian M, Kelly WG, Fleenor J, Grishok A, Timmons L, Fire A, Mello CC. The *rde-1* gene, RNA interference, and transposon silencing in *C. elegans*. Cell. 1999;99:123-132. DOI: 10.1016/S0092-8674(00)81644-X

[17] Bartel DP. MicroRNAs: Genomics, biogenesis, mechanism, and function. Cell. 2004;116:281-297. DOI: 10.1016/S0092-8674(04)00045-5

[18] Shabalina SA, Koonin EV. Origins and evolution of eukaryotic RNA interference. Trends in Ecology and Evolution. 2008;23:578-587. DOI: 10.1016/j.tree.2008.06.005

[19] Saini HK, Griffiths-Jones S, Enright AJ. Genomic analysis of human microRNA transcripts. Proceedings of the National Academy of Sciences of the United States of America. 2007;104:19-24. DOI: 10.1073/pnas.0703890104

[20] Bernstein E, Caudy AA, Hammond SM, Hannon GJ. Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature. 2001;409:363-366. DOI: 10.1038/35053110

[21] Cenik ES, Fukunaga R, Lu G, Dutcher R, Wang Y. Phosphate and R2D2 restrict the substrate specificity of Dicer-2, an ATP-driven ribonuclease. Molecular Cell. 2011;42:172-184. DOI: 10.1016/j.molcel.2011.03.002

[22] Kasschau KD, Xie Z, Allen E, Llave C, Chapman EJ, Krizan KA, Carrington JC. P1/HC-Pro, a viral suppressor of RNA silencing, interferes with *Arabidopsis* development and miRNA function. Developmental Cell. 2003;4:205-217. DOI: 10.1016/S1534-5807(03)00025-X

[23] Doyle M, Jaskiewicz L, Filipowicz W. Dicer proteins and their role in gene silencing pathways. In: Guo F, Tamanoi F, editors. The Enzymes: Eukaryotic RNAses and their partners in RNA degradation and biogenesis.1st ed. Oxford: Academic Press; 2012. p. 1-28. DOI: 10.1016/B798-0-012-404741-9.00001-5

[24] Park JE, Heo I, Tian Y, Simanshu DK, Chang H, Jee D, Patel DJ, Kim VN. Dicer recognizes the 5′ end of RNA for efficient and accurate processing. Nature. 2011;475:201-205. DOI: 10.1038/nature10198

[25] Wilson RC, Doudna JA. Molecular Mechanism of RNA Interference. Annual Review of Biophysics. 2013;42:217-239. DOI: 10.1146/annurev-biophys-083012-130404

[26] Tolia NH, Joshua-Tor L. Slicer and the argonautes. Nature Chemical Biology. 2007;3:36-46. DOI: 10.1038/nchembio848

[27] Nicholson AW. Function, mechanism and regulation of bacterial ribonucleases. FEMS Microbiology Reviews. 1999;23:371-390. DOI: 10.1111/j. 1574-6976.1999.tb00405.x

[28] Voinnet O. Origin, biogenesis, and activity of plant microRNAs. Cell. 2009;136:669-687. DOI: 10.1016/j.cell.2004.06.017

[29] Meng Y, Shao C, Wang H, Chen M. The regulatory activities of plant microRNAs: A more dynamic perspective. Plant Physiology. 2011;157:1583-1595. DOI: 10.1016/j.cell. 2004.06.017

[30] Perkel JM. Visiting noncodarnia. Biotechniques. 2013;6:301-304.DOI: 10.1016/j.cell. 2004.06.017

[31] Pattanayak D, Solanke AU, Kumar PA. Plant RNA interference pathways: Diversity in function similarity in action. Plant Molecular Biology Reporter. 2013;31:493-506. DOI: 10.1016/j.sbi.2006.12.002

[32] MacRae IJ, Doudna JA. Ribonuclease revisited: Structural insights into ribonuclease III family enzymes. Current Opinion in Structural Biology. 2007;17:138-45. DOI: 10.1016/j.sbi.2006.12.002

[33] Zhang H, Kolb FA, Jaskiewicz L, Westhof E, Filipowicz W. Single processing center models for human DICER and Bacterial RNase III. Cell. 2004;118:57-68.DOI: 10.1016/j.cell.2004.06.017

[34] Staiger D, Korneli C, Lummer M, Navarro L. Emerging role for RNA-based regulation in plant immunity. New Phytologists. 2013;197:394-404. DOI: 10.1111/nph.12022

[35] Streitner C, Simpson CG, Shaw P, Danisman S, Brown JWS, Staiger D. Small changes in ambient temperature affect alternative splicing in *Arabidopsis thaliana*. Plant Signaling and Behavior. 2013;8:e24638.1-e24638.7. DOI: 10.4161/psb.24638

[36] Sunkar R, Li YF, Jagadeeswaran G. Functions of microRNAs in plant stress responses. Trends Plant Science. 2012;17:196-203. DOI: 10.1016/j.tplants.2012.01.010

[37] Jones-Rhoades MW, Bartel DP, Bartel B. MicroRNAs and their regulatory roles in plants. Annual Review of Plant Biology. 2006;57:19-53. DOI: 10.1146/annurev.arplant. 57.032905.105218

[38] Zhang X, Zhao H, Gao S, Wang WC, Katiyar-Agarwal S, Huang HD, Raikhel N, Jin H. *Arabidopsis* Argonaute 2 regulates innate immunity via miRNA393- mediated silencing of a Golgi localized SNARE gene, MEMB12. Molecular Cell. 2011;42:356-366. DOI: 10.1016/j.molcel.2011.04.010

[39] Sunkar R, Zhu JK. Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. Plant Cell. 2004;16:2001- 2019. DOI: 10.1105/tpc.104.022830

[40] Kruszka K, Pieczynski M, Windels D, Bielewicz D, Jarmolowski A, Szweykowska-Kulinska Z, Vazquez F. Role of microRNAs and other sRNAs of plants in their

changing environments. Journal of Plant Physiology. 2012;169:1664-1672.DOI: 10.1016/j.jplph.2012.03.009

[41] Wang L, Zheng J, Luo Y, Xu T, Zhang Q, Zhang L, Xu M, Wan J, Wang MB, Zhang C, Fan F. Construction of a genome wide RNAi mutant library in rice. Plant Biotechnology Journal. 2013;11:997-1005. DOI: 10.1111/pbi.12093

[42] Burand JP, Hunter WB. RNAi: Future in insect management. Journal of Invertebrate Pathology. 2013;112:568-574. DOI: 10.1016/j.jip.2012.07.012

[43] Belles X. Beyond Drosophila: RNAi *in vivo* and functional genomics in insects. Annual Review of Entomology. 2010;55:111-128. DOI: 10.1146/annurev-ento-112408-085301

[44] Roether S, Meister G. Small RNAs derived from longer non-coding RNAs. Biochimie. 2011;93:1905-1915. DOI: 10.1016/j.biochi.2011.07.032

[45] Gu L, Knipple DC. Recent advances in RNA interference research in insects: Implications for future insect pest management strategies. Crop Protection. 2013;45:36-40. DOI: 10.1016/j.cropro.2012.10.004

[46] Araujo RN, Santos A, Pinto FS, Gontijo NF, Lehane MJ, Pereira MH. RNA interference of the salivary gland nitrophorin 2 in the triatomine bug *Rhodnius prolixus* (Hemiptera: Reduviidae) by dsRNA ingestion or injection. Insect Biochemistry and Molecular Biology. 2006;36:683-693. DOI: 10.1016/j.ibmb.2006.05.012

[47] Turner CT, Davy MW, MacDiarmid RM, Plummer KM, Birch NP. RNA interference in the light brown apple moth, *Epiphyas postvittana* (Walker) induced by double-stranded RNA feeding. Insect Molecular Biology. 2006;15:383-391. DOI: 10.1111/j.1365-2583.2006.00656.x

[48] Pridgeon JW, Zhao L, Becnel JJ, Strickman DA, Clark GG, Linthicum KJ. Topically applied AaeIAP1 double-stranded RNA kills female adults of *Aedes aegypti*. Journal of Medical Entomology. 2008;45:414-420. DOI: 10.1603/0022-2585(2008)45[414:TAADRK]2.0.CO;2

[49] Wang Y, Zhang H, Li H, Miao X. Second-generation sequencing supply an effective way to screen RNAi targets in large scale for potential application in pest insect control. PLoSOne. 2011;6:e18644. DOI: 10.1371/journal.pone.0018644

[50] Li X, Zhuo K, Luo M, Sun L, Liao J. Molecular cloning and characterization of a calreticulin cDNA from the pinewood nematode *Bursaphelenchusxylophilus*. Experimental Parasitology. 2011;128:121-126. DOI: 10.1016/j.exppara.2011.02.017

[51] Zhu F, Xu J, Palli R, Ferguson J, Palli SR. Ingested RNA interference for managing the populations of the Colorado potato beetle, *Leptinotarsa decemlineata*. Pest Management Science. 2011;67:175-182. DOI: 10.1002/ps.2048

[52] Zha W, Peng X, Chen R, Du B, Zhu L, He G. Knockdown of midgut genes by dsRNA-transgenic plant-mediated RNA interference in the Hemipteran insect *Nila-parvatalugens*. PLoSOne. 2011;6:e20504.DOI: 10.1371/journal.pone.0020504

[53] Kumar P, Pandit SS. Baldwin IT. Tobacco rattle virus vector: A rapid and transient means of silencing *Manduca sexta* genes by plant mediated RNA interference. PLoSONE. 2012;7:e31347.DOI: 10.1371/journal.pone.0031347

[54] Fairbairn DJ, Cavallaro AS, Bernard M, Mahalinga-Iyer J, Graham MW, Botella JR. Host-delivered RNAi: An effective strategy to silence genes in plant parasitic nematodes. Planta. 2007;226:1525-1533.DOI: 10.1007/s00425-007-0588-x

[55] Yadav BC, Veluthami K, Subramaniam K. Host-generated double-stranded RNA induces RNAi in plant-parasitic nematodes and protects the host from infection. Molecular and Biochemical Parasitology. 2006;148:219-222. DOI: 10.1016/j.molbiopara. 2006.03.013

[56] Huang G, Allen R, Davis EL, Baum TJ, Hussay RS. Engineering broad root-knot resistance in transgenic plants by RNAi silencing of a conserved and essential root-knot nematode parasitism gene. Proceedings of the National Academy of Sciences of the United States of America. 2006;103:14302-14306. DOI: 10.1073/pnas.0604698103

[57] Steeves RM, de Bruin R, Kenter S, van der Hoorn R, van Blokland R. Transgenic soybeans expressing siRNAs specific to a major sperm protein gene suppress *Heterodera glycines* reproduction. Functional Plant Biology. 2006;33:991-999.DOI: 10.1071/FP06130

[58] Kakumani PK, Ponia SS, Rajgokul KS, Chinnappan M, Banerjea AC, Medigeshi GR, Malhotra P, Mukherjee SK, Bhatnagar RK. Role of RNAi in dengue viral replication and identification of NS4B as a RNAi suppressor. Journal of Virology. 2013;87:8870-8883. DOI: 10.1128/JVI.02774-12.

[59] Karlikow M, Goic B, Saleh MC. RNAi and antiviral defense in *Drosophila*: Setting up a systemic immune response. Developmental and Comparative Immunology. 2014;42:85-92. DOI: 10.1016/j.dci.2013.05.004

[60] Shivaprasad PV, Rajeswaran R, Blevins T, Schoelz J, Meins F, Hohn T, Pooggin MM. The CaMV transactivator/viroplasmin interferes with RDR6-dependent *trans*-acting and secondary siRNA pathways in *Arabidopsis.* Nucleic Acids Research. 2008;36:5896-5909. DOI: 10.1093/nar/gkn590

[61] Kasschau KD, Xie Z, Allen E, Llave C, Chapman EJ, Krizan KA, Carrington JC. P1/HC-Pro, a viral suppressor of RNA silencing, interferes with *Arabidopsis* development and miRNA function. Developmental Cell. 2003;4:205-217. DOI: 10.1016/S1534-5807(03)00025-X

[62] Cakir C, Tor M. Factors influencing barley stripe mosaic virus-mediated gene silencing in wheat. Physiological and Molecular Plant Pathology. 2010;74:246-253. DOI: 10.1016/j.pmpp.2010.04.001

[63] Senthil-Kumar M, Mysore KS. Virus-induced gene silencing can persist for more than two years and also be transmitted to progeny seedlings in *Nicotiana benthamiana* and tomato. Plant Biotechnology Journal. 2011;9:797-806. DOI: 10.1111/j. 1467-7652.2011.00589.x.

[64] Fernández LC, Donaire L, Livia D, Cisar L. RNA silencing and the interplay between plants and viruses. Recent Advances in Plant Virology. 2011;6:121-136. ISBN: 978-1-904455-75-2

[65] Bernstein E, Caudy AA. Hammond SM, Hannon GJ. Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature. 2001;409: 363-366. DOI: 10.1038/35053110;

[66] Takumi S, Motoyasu Y, Taiyun W, Hirohiko H, Toshihiro O. Silencing by RNAi of the gene for Pns12, a viroplasm matrix protein of Rice dwarf virus, results in strong resistance of transgenic rice plants to the virus. Plant Biotechnology Journal. 2008;7:24-32. DOI: 10.1111/j.1467-7652.2008.00366.x

[67] Hemmes H,Kaaij LD, Prins M, Goldbach R, Schnettler E. Binding of small interfering RNA molecules is crucial for RNA interference suppressor activity of rice hoja blanca virus NS3 in plants. Journal of General Virology. 2009;90:1762-1766. DOI: 10.1099/vir. 0.010488-0

[68] Tyagi H, Rajasubramaniam S, Rajam MV, Dasgupta I. RNA-interference in rice against Rice tungro bacilliform virus results in its decreased accumulation in inoculated rice plants. Transgenic Research. 2008;17:897-904. DOI:10.1007/s11248-008-9174-7

[69] Vanderschuren H, Alder A, Zhang P, Gruissem W. Dose-dependent RNAi-mediated geminivirus resistance in the tropical root crop cassava. Plant Molecular Biology. 2009;70:265-272. DOI: 10.1007/s11103-009-9472-3

[70] Velásquez AC, Chakravarthy S, Martin GB. Virus-induced gene silencing (VIGS) in *Nicotiana benthamiana* and tomato. Journal of Visualized Experiments. 2009;28:1292-1301. DOI: 10.3791/1292

[71] Soler S, Fagoaga C, Chiibi S, López C, Moreno P, Navarro L, Flores R, Peña L. RNAi mediated protection against citrus tristeza virus in transgenic citrus plants. In: Erdmann VA, Barciszewski J, editors. Non coding RNAs in plants. 1st ed. Berlin: Heidelberg; 2011. p. 447-460. DOI: 10.1007/978-3-642-19454-2_27

[72] Li HW, Lucy AP, Guo HS, Li WX, Ji LH, Wong SM, Ding SW. Strong host resistance targeted against a viral suppressor of the plant gene silencing defense mechanism.

The European Molecular Biology Organization Journal. 1999;18:2683-2691. DOI: 10.1093/emboj/18.10.2683

[73]   Qiu W, Park JW, Scholthof HB. Tombusvirus P19-mediated suppression of virus-induced gene silencing is controlled by genetic and dosage features that influence pathogenicity. Molecular Plant-Microbe Interactions. 2002;15:269-280. DOI: 10.1094/MPMI.2002.15.3.269

[74]   Chen J, Li WX, Xie D, Peng JR, Ding SW. Viral virulence protein suppresses RNA silencing-mediated defense but upregulates the role of microRNA in host gene expression. Plant Cell. 2004;16:1302-1313. DOI: 10.1105/tpc.018986

[75]   Wang MB, Bian XY, Wu LM, Liu LX, Smith NA, Isenegger D. On the role of RNA silencing in the pathogenicity and evolution of viroids and viral satellites. Proceedings of the National Academy of Sciences of the United States of America. 2004;101:3275-3280. DOI: 10.1073/pnas.0400104101

[76]   Smith NA, Eamens AL, Wang MB. Viral small interfering RNAs target host genes to mediate disease symptoms in plants. PloS Pathogens. 2011;7:e1002022. DOI: 10.1371/journal.ppat.1002022

[77]   Shimura H, Pantaleo V, Ishihara T, Myojo N, Inaba J, Sueda K, Burgyán J, Masuta C. A viral satellite RNA induces yellow symptoms on tobacco by targeting a gene involved in chlorophyll biosynthesis using the RNA silencing machinery. PLoS Pathogens. 2011;7:1002021. DOI: 10.1371/journal.ppat.1002021

[78]   Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP. Prediction of plant microRNA targets. Cell. 2002;110:513-520. DOI: 10.1016/S0092-8674(02)00863-2

[79]   Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP. MicroRNAs in plants. Genes & Development. 2002;16:1616-1626. DOI:10.1101/gad.1004402

[80]   Xie Z, Allen E, Fahlgren N, Calamar A, Givan SA, Carrington JC. Expression of *Arabidopsis* miRNA genes. Plant Physiology. 2005;138:2145-2154. DOI: 10.1104

[81]   Rajagopalan R, Vaucheret H, Trejo J, Bartel DP. A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*. Genes Development. 2006;20:3407-3425. DOI: 10.1101/gad.1476406

[82]   Li F, Pignatta D, Bendix C, Brunkard JO, Cohn MM, Tung J, Sun H, Kumar P, Baker B. MicroRNA regulation of plant innate immune receptors. Proceedings of the National Academy of Sciences of the United States of America. 2012;109:1790-1795. DOI: 10.1073/pnas.1118282109

[83]   Axtell MJ, Snyder JA, Bartel DP. Common functions for diverse small RNAs of land plants. Plant Cell. 2007;19:1750-1769. DOI: 10.1105/tpc.107.051706

[84] Lu S, Sun YH, Shi R, Clark C, Li L, Chiang VL. Novel and Mechanical Stress-responsive MicroRNAs in *Populus trichocarpa* that are absent from *Arabidopsis*. Plant Cell. 2005;17:2186:2203. DOI: 10.1105/tpc.105.033456

[85] Zeng C, Wang W, Zheng Y, Chen X, Bo W, Song S, Zhang W, Peng M. Conservation and divergence of microRNAs and their functions in Euphorbiaceous plants. Nucleic Acids Research. 2010;38:981-995. DOI: 10.1093/nar/gkp1035

[86] Dezulian T, Palatnik JF, Huson DH, Weigel D. Conservation and divergence of microRNA families in plants. Genome Biology. 2005;6:13-18.DOI: 10.1186/gb-2005-6-11-p13

[87] Zhang L, Chia JM, Kumari S, Stein JC, Liu Z, Narechania A, Maher CA, Guill K, McMullen MD, Ware D. A Genome-wide characterization of microRNA genes in maize. PLoS Genetics. 2009;5:e1000716. DOI: 10.1371/journal.pgen.1000716

[88] Jones JD, Dangl JL. The plant immune system. Nature. 2006;444:323–329. DOI: 10.1038/nature05286

[89] Zvereva AS, Pooggin MM. Silencing and innate Immunity in plant defense against viral and non-viral pathogens. Viruses. 2012;4:2578-2597. DOI: 10.3390/v4112578

[90] Yang J, Hirschi KD, Farmer LM. Dietary RNAs: New Stories Regarding Oral Delivery. Nutrients. 2015;7:3184-3199. DOI: 10.3390/nu7053184

[91] Bartel DP. MicroRNAs: Target Recognition and Regulatory Functions.Cell.2009;136: 215-233. DOI: 10.1016/j.cell.2009.01.002

[92] Conde J, Edelman ER, Artzi N. Target-responsive DNA/RNA nanomaterials for microRNA sensing and inhibition: The jack-of-all-trades in cancer nanotheranostics? Advanced Drug Delivery Reviews. 2015;81:169-183. DOI: 10.1016/j.addr.2014.09.003

[93] Strasser R, Stadlmann J, Svoboda B, Altmann F, Glossl J, Mach L. Molecular basis of N-acetylglucosaminyltransferase I deficiency in Arabidopsis thaliana plants lacking complex N-glycans. Biochemistry Journal. 2005;387:385-91. DOI: 10.1042/BJ20041686

[94] Schähs M, Strasser R, Stadlmann J, Kunert R, Rademacher T, Steinkellner H. Production of a monoclonal antibody in plants with a humanized N-glycosylation pattern. Plant Biotechnology Journal. 2007;5:657-63. DOI: 10.1111/j.1467-7652.2007.00273.x

[95] Salazar-Gonzalez JA, Angulo C, Rosales-Mendoza S. Chikungunya virus vaccines: Current strategies and prospects for developing plant-made vaccines. Vaccine. 2015;33: 3650-3658. DOI: 10.1016/j.vaccine.2015.05.104

[96] Ou J, GuoZ, Shi J, Wang X, Liu J, Shi B, Guo F, Zhang C, Yang D. Transgenic rice endosperm as a bioreactor for molecular pharming. Plant Cell Reports. 2014;33:585-594. DOI: 10.1007/s00299-013-1559-2

[97] Lemgo GN, Sabbadini S, Pandolfini T, Mezzetti B. Biosafety considerations of RNAi-mediated virus resistance in fruit-tree cultivars and in rootstock. Transgenic Research. 2013;22:1073-1088. DOI: 10.1007/s11248-013-9728-1

[98] Zhang L, Hou D, Chen X, Li D, Zhu L, Zhang Y, Li J, Bian Z, Liang X, Cai X. Exogenous plant MIR168a specifically targets mammalian LDLRAP1: Evidence of cross-kingdom regulation by microRNA. Cell Research.2012;22:107-126. DOI: 10.1038/cr.2011.158

[99] Ricaño-Rodríguez J, Zavala-González E, Ramírez-Lepe M. Silenciamiento génico en plantas: mecanismos moleculares del ARN de interferencia y aplicaciones biotecnológicas. Revista Fitotecnia Mexicana. 2014;37:339-350.

[100] Auer C, Frederick R. Crop improvement using small RNAs: applications and predictive ecological risk assessments. Trends in Biotechnology. 2009;27:64-651. DOI: 10.1016/j.tibtech.2009.08.005

[101] Casacuberta JM, Devos Y, Jardin Pd, Ramon M, Vaucheret H, Nogué F. Biotechnological uses of RNAi in plants: risk assessment considerations. Trends in Biotechnology. 2015;33:145-147. DOI: 10.1016/j.tibtech.2014.12.003

[102] Scheurer S, Sonnewald S. Genetic engineering of plant food with reduced allergenicity. Frontiers in Bioscience. 2009;1:59-71.DOI: 10.2741/3231.

[103] Saurabh S, Vidyarthi AS, Prasad D. RNA interference: Concept to reality in crop improvement. Planta. 2014;239:543-564. DOI: 10.1007/s00425-013-2019-5

[104] Ramon M, Devos Y, Lanzoni A, Liu Y, Gomes A, Gennaro A, Waigmann E. RNAi-based GM plants: Food for thought for risk assessors. Plant Biotechnology Journal. 2014;12:1271-1273. DOI: 10.1111/pbi.12305

[105] Li JF, Chung HS, Niu Y, Bush J, McCormack M, Sheen J. Comprehensive protein-based artificial microRNA screens for effective gene silencing in plants. Plant Cell. 2013;25: 1507-1522. DOI: 10.1105/tcp.113.112235

Genomics of Plant Stress Response

Genomic Approaches to Developing Molecular Markers Linked to Grey Leaf Spot Resistance Loci in Ryegrasses

Wataru Takahashi

Abstract

Ryegrass grey leaf spot (GLS), which is also called ryegrass blast, is caused by *Magnaporthe oryzae* (anamorph *Pyricularia oryzae*). It is a serious disease in ryegrasses including perennial ryegrass (*Lolium perenne* L.) and Italian ryegrass (*L. multiflorum* Lam.). Heavily infected young seedlings die within days, and grass stands can be seriously damaged by the disease. Thus, the development of GLS-resistant cultivars has become one of the most important objectives in ryegrass breeding. This chapter provides an overview of the current information regarding molecular marker development in the breeding of GLS-resistant ryegrass cultivars. It focuses on the pathology of GLS, heritability and breeding of GLS resistance, and development of molecular markers linked to a major ryegrass GLS resistance gene.

Keywords: Comparative genomics, Forage grasses, *Lolium*, Molecular breeding, Resistance gene

1. Introduction

Perennial ryegrass (*Lolium perenne* L.) and Italian ryegrass (*L. multiflorum* Lam.) are taxonomically related cool-season grasses and are the most cultivated species in the genus *Lolium* in temperate regions. Perennial ryegrass is mainly used as turf and for grazing, whereas Italian ryegrass is primarily grown for hay and silage.

Ryegrass grey leaf spot (GLS), also called ryegrass blast, is a major disease of perennial ryegrass in the United States [1] and Italian ryegrass in Japan [2-4]. Rice blast and ryegrass GLS are caused by a common pathogenic fungal species, *Magnaporthe oryzae* (anamorph *Pyricularia oryzae*) [5]. Severely infected young seedlings die within days, and infected ryegrass stands can cause widespread damage and losses.

Effective GLS management strategies in ryegrass turf include the use of chemical fungicides. However, the high cost of fungicide application is an important limitation for growers managing large turf areas [1]. Additionally, overreliance on fungicides may lead to the development of fungicide-resistant fungal strains [6] and adversely affect nontarget organisms [7], ultimately resulting in adverse ecological consequences. Furthermore, the bioaccumulation of fungicides in domesticated animals (e.g., cattle) and its possible effects on the safety of dairy products are potential problems associated with fungicide use. There are currently no labeled fungicides effective against GLS in the United States [8] and Japan [3]. Therefore, there are a limited number of disease management options.

In this context, cultural management practices such as minimizing drought stress, reducing leaf wetness, avoiding excessive applications of nitrogen, and soil compaction may help to reduce disease severity [9]. However, these practices often do not work efficiently because the disease develops rapidly in susceptible ryegrass cultivars [1]. Thus, integrated management including the use of GLS-resistant cultivars is necessary to establish productive ryegrass cultural systems.

This chapter focuses on ryegrass breeding for the development of GLS-resistant cultivars. The main topics covered herein include pathology of ryegrass GLS, diversity and conventional breeding of GLS-resistant ryegrasses, and development of molecular markers linked to GLS resistance loci.

2. Pathology of ryegrass GLS

2.1. Taxonomy

In 2002, the causal pathogen of GLS of grass species including ryegrasses (*Lolium* species) and rice blast was identified as a new species, *M. oryzae* (anamorph *P. oryzae*). This new species was considered distinct from *Magnaporthe grisea* (anamorph *P. grisea*), which is associated with the grass genus *Digitaria*. The distinction was based on phylogenetic analyses and laboratory mating experiments that showed the two species were not interfertile, although there were no morphological differences between them [5].

In this chapter, the term "*M. oryzae*" is used. However, it is important to note that a formal change from *M. grisea* to *M. oryzae* has not yet occurred. A proposal for changing the name based on the results of [5] is allowed under the International Code of Nomenclature for algae, fungi, and plants (Melbourne Code). A proposal will be submitted to and discussed by the Nomenclature Committee for Fungi of the International Association for Plant Taxonomy [10]. A final decision on a name change will be made during the Nomenclature Session of the International Botanical Congress in 2017 [10].

2.2. Population structure and host specificity

Analysis of genomic DNA using molecular markers is the most powerful method for determining the population structures of the *Magnaporthe* species. Repetitive DNA elements such

as transposons and retrotransposons are often used to generate probes for Southern blotting experiments during DNA fingerprinting [11-15]. This is because of the diversity in copy numbers of elements and the richness of polymorphisms around, within, or among the elements, which might be caused by base substitutions or insertions and deletions. The use of internal transcribed spacer regions between ribosomal DNAs as probes for DNA fingerprinting is also common [12, 13]. Similarly, the internal transcribed spacer regions have been sequenced for population structure analyses [14]. Table 1 lists the repetitive sequences that have been used to analyze the population structure of *Magnaporthe* species associated with grass weeds, turf grasses, and/or forage grasses in addition to major crops such as rice and wheat (*Triticum aestivum*) [11-15].

Target	Feature	Reference	
		Sequence	Result of application
MAGGY	Retrotransposon	[16]	[11, 12]
MGLR-3	Retrotransposon	[17]	[13]
MGR583	Retrotransposon	[18, 19]	[12, 14]
MGR586	Transposon	[18, 20]	[11-14]
Pot2	Transposon	[21]	[11-15]
rDNA	Ribosomal DNA	[22, 23]	[12-14]
RETRO5	Retroelement	[24]	[12]

Table 1. Repetitive DNA sequences for DNA fingerprinting of *Magnaporthe* species associated with grass weeds, turf grasses, and/or forage grasses

In some cases, probes derived from these repetitive DNA sequences cannot clearly distinguish between isolates from different hosts. Restriction fragment length polymorphisms (RFLPs) with single-copy probes derived from long insert-cosmid clones (35–40 kb) are appropriate for the initial comparison of poorly characterized isolates from different hosts [12]. In addition to the repetitive DNA sequences, amplified fragment length polymorphisms (AFLPs) can produce many markers and provide a higher resolution for population structure analyses even within the same *Magnaporthe* lineage [25, 26].

Population structures can be determined in dendrograms constructed by analyzing genetic distances among isolates, which are reflected by differences in the banding patterns obtained during molecular marker analyses. Dendrograms of ryegrass isolates have often revealed genetic similarities between ryegrass isolates and isolates from wheat [12-14, 25] and tall fescue (*Schedonorus arundinaceus*) [12, 25].

In artificial inoculation conditions, isolates from ryegrasses, wheat, and tall fescue can cause serious infections in all hosts. Table 2 summarizes the data from six studies on the pathogenicity of *Magnaporthe* isolates from ryegrasses, tall fescue, wheat, rice, and/or crabgrass [13-15, 25, 27, 28]. The isolates from ryegrasses are generally avirulent, but can be virulent to rice [13,

14]. Conversely, although the rice isolates are thought to be unable to cause serious infections in ryegrasses [13, 14], they are occasionally highly virulent to the plant species [27]. The wheat isolates are avirulent to rice [14, 27], although the rice isolates are virulent to wheat [13, 27]. Some isolates from crabgrass (*Digitariasanguinalis*) are virulent to tall fescue [25] and ryegrasses [25, 28], highly virulent to Italian ryegrass [25] but are avirulent to wheat [14, 25]. Additionally, isolates from perennial ryegrass, wheat, and rice can infect crabgrass, but these are generally not highly virulent to crabgrass [14]. Many isolates from tall fescue are avirulent to crabgrass [25].

Original host[a]	Inoculated host[b]						Reference
	PR	IR	TF	W	R	CG	
Perennial ryegrass (PR)	++		++	++	-		[13]
	++	++		+-	+-	+-	[14]
							[15]
							[25]
							[27]
							[28]
Italian ryegrass (IR)							[13]
							[14]
	++	++	++				[15]
							[25]
							[27]
		++					[28]
Tall fescue (TF)							[13]
							[14]
							[15]
	++	++	++	++-		-	[25]
							[27]
							[28]
Wheat (W)							[13]
	++	++		++	-	+	[14]
							[15]
	++	++	++	++-		-	[25]
	++	++	++	++	-	-	[27]
							[28]
Rice (R)	-		-	+	++		[13]

Original host[a]	Inoculated host[b]						Reference
	PR	IR	TF	W	R	CG	
	-	+		-	++	+-	[14]
							[15]
							[25]
	++-	++	++	+-	++	-	[27]
							[28]
Crabgrass (CG)							[13]
	-	-		-	-	++	[14]
							[15]
	+-	++-	+-	-		++	[25]
							[27]
		+-					[28]

[a]According to [5], the crabgrass isolate might be *M. grisea* and the others might be *M. oryzae*.

[b]+: virulent; ++: highly virulent; -: avirulent; +-: virulent but sometimes fails to infect; ++-: highly virulent but sometimes fails to infect.

Table 2. Pathogenicity and host specificity of *Magnaporthe* species during artificial inoculations

In addition to the isolates listed in Table 2, during artificial inoculations, ryegrasses are highly susceptible to isolates from weeping lovegrass (*Eragrostis curvula*) [25], and susceptible to isolates from finger millet (*Eleusine coracana*) [14], St. Augstinegrass (*Stenotaphrum secunda-tum*) [25, 28], Alexandergrass (*Brachiaria plantaginea*) [27], Pennsylvania smartweed (*Polygonum pensylvanicum*) [28], and soybean (*Glycine max*) [28].

The cross-infections observed during artificial inoculations suggest that "opportunistic" cross-infections may occur in nature [12]. However, population structure analyses based on molecular marker analyses have revealed that although there are genetic differences even in isolates from the same host species, the population structures are generally associated with host differences. This indicates that the host species is a major selective factor for constructing isolate populations, and cross-infections among hosts might not be detectable in nature [25]. Nevertheless, ryegrasses might be infected by tall fescue isolates because these hosts are congeneric [29-31]. Therefore, the isolates from ryegrasses and tall fescue are genetically quite similar [12] or belong to the same lineage in some cases [25]. Additionally, wheat isolates are genetically similar to the ryegrass and tall fescue isolates, and all can cause serious infections in wheat, ryegrass, and tall fescue in artificial inoculation conditions (Table 2). However, the wheat isolates are clearly genetically distinct [12, 25]. This might explain why no epidemics of wheat blast caused by the cross-infection of ryegrass isolates and *vice versa*, have been reported [12]. This may also be the case for weeping lovegrass, in which there are genetic similarities and cross-pathogenicity among hosts [25]. Therefore, isolates from wheat and/or weeping

lovegrass may be progenitors of isolates of ryegrasses and tall fescue rather than being directly responsible for GLS in ryegrasses or tall fescue [12, 25].

3. Diversity and conventional breeding of GLS-resistant ryegrasses

3.1. Heritability and genetic effects of GLS resistance

To breed for GLS-resistant ryegrasses, genetic material conferring resistance to GLS must be identified. For this purpose, researchers have investigated the diversity among resistant phenotypes [32-37]. Although most commercial cultivars and experimental lines are susceptible to GLS, some resistant genotypes have been identified in cultivars and experimental lines of Italian ryegrass [32-34] and perennial ryegrass [32, 35, 36]. Perennial ryegrass might be the more GLS-resistant species as resistant phenotypes are more common than in Italian ryegrass [32]. Additionally, in Italian ryegrass, tetraploid lines were slightly more resistant than diploid lines [33]. This is also the case in perennial ryegrass.

The diversity in GLS resistance has encouraged breeders to continue to attempt to generate GLS-resistant cultivars. In outcrossing plants like ryegrasses, a phenotypic recurrent selection is often used to improve important agronomic traits mainly controlled by genes with an additive effect. The effects of recurrent selection have been observed in Italian ryegrass and GLS-resistant experimental lines have been selected [33, 34], indicating that GLS resistance can be conferred using recurrent selection and is possibly controlled by additive gene effects.

Recurrent selection has also been effective in perennial ryegrass [35, 37]. The broad-sense heritability estimates were very high at 0.92 [35] and 0.95 [37] without any interaction between cultivar and environment. These results suggest that GLS resistance is controlled by strong genetic effects [35, 37]. Further, the phenotypic means of populations composed of selected individuals were dramatically shifted toward the selected GLS resistance. Therefore, GLS resistance was thought to be controlled by a few genes and the frequency of the genes in the selected population rapidly increased during selection cycles [35, 37]. However, much of the additive gene effects cannot be obtained with only one cycle of selection. The genetic gain during the second selection cycle was higher than that of the first cycle in the GLS-resistant phenotype [37].

Narrow-sense heritability and the number of genes having additive effects in GLS resistance are among the most important considerations for breeders because the additive gene effects actually reflect the effect of selection. However, these have not been estimated by the studies mentioned above. Diallel cross analysis is a way to determine narrow-sense heritability, number of genes having additive effects, general combining ability (GCA), and specific combining ability (SCA) of parent plants [38-40]. In perennial ryegrass, diallel crosses involving six and eight parents have been analyzed to investigate the GCA, SCA, narrow-sense heritability, and the number of genes involved in GLS resistance [36]. The GCA and SCA were highly significant and accounted for 80–86% and 7–17% of the total genotypic variance, respectively [36]. The significant SCA values suggest that dominant genes or those that interact

with related genes must have been involved in the parents. The considerably higher GCA values also suggest that GLS resistance is mainly controlled by additive gene effects as previously concluded [35, 37]. The narrow-sense heritability and number of genes having additive effects were estimated to range from 0.57 to 0.76 and 2.1 to 4.4, respectively [36]. Results of the diallel cross analysis were consistent with those of the abovementioned studies [35, 37]. Thus, phenotypic recurrent selection was very effective in improving GLS resistance in ryegrasses. Because of the quantitative additive gene effects, resistant phenotypes in the selected lines would be durable although the possibility that some genes with additive effects might be more important for GLS resistance cannot be ruled out. The gene most responsible for GLS resistance may be inherited by the next generation and act as a quasi-qualitative major partial resistance gene.

3.2. Available GLS-resistant ryegrass cultivars

Although almost all of the commercially available cultivars released before 2004 were very susceptible to GLS [9], many GLS-resistant perennial ryegrass cultivars are currently available in the United States [41]. In contrast, GLS-resistant Italian ryegrass cultivars are very rare, but the diploid cultivar "Sachiaoba" [2] in Japan and the tetraploid cultivar "Jumbo" [42] in the United States have been registered as GLS-resistant in 1998 and 2000, respectively. However, an article published in 2010 reported a lack of annual ryegrass cultivars resistant to *P. grisea* in the United States, which led to the belief that GLS resistance in Italian ryegrass was insufficient [8]. All of these resistant cultivars have partial resistance, and no completely resistant perennial ryegrass or Italian ryegrass cultivars have been released. Therefore, continued breeding for GLS resistance is necessary.

4. Development of molecular markers linked to GLS resistance loci

In addition to conventional breeding, researchers have used molecular breeding techniques involving molecular markers to develop disease-resistant cultivars of major crops. Developing resistance to rice blast is a major focus among plant pathologists, and many molecular markers relevant for the breeding of rice blast-resistant cultivars have been reported [43, 44]. Regarding ryegrasses, research groups in the United States and Japan have found genetic loci for GLS resistance and have identified molecular markers linked to the resistance loci in an Italian × perennial ryegrass hybrid [45-47] and Italian ryegrass [4, 48, 49].

4.1. Molecular marker development for GLS resistance in an Italian × perennial ryegrass hybrid

4.1.1. Mapping population derived from Italian × perennial ryegrass hybrid parents

A research group in the United States developed a mapping population consisting of progeny individuals derived from a cross between Italian × perennial ryegrass hybrid heterozygous parental clones MFA and MFB [45, 46]. The parental clones were obtained in separate crosses

between two different grandparental clones of the perennial ryegrass cultivar "Manhattan" and two different grandparental clones of the Italian ryegrass cultivar "Floregon" (Figure 1). A second-generation mapping population [47] was then developed. The GLS-resistant MF-8 was selected from the first mapping population and crossed with the GLS-susceptible L4B-5 obtained in a cross between a clonal individual of the forage-type perennial ryegrass cultivar "Linn" and a clonal individual of the turf-type perennial ryegrass cultivar "SR4400" (Figure 1).

The grandparental clones and parents of the mapping populations could be asexually maintained and propagated. However, the grandparental clones of the Italian ryegrass cultivar "Floregon" could not be maintained because of the annuality of this species [46]. Similarly, the two mapping populations exhibited perenniality, with each individual capable of being clonally maintained and propagated to produce clonal replicates for multiple experiments [45-47].
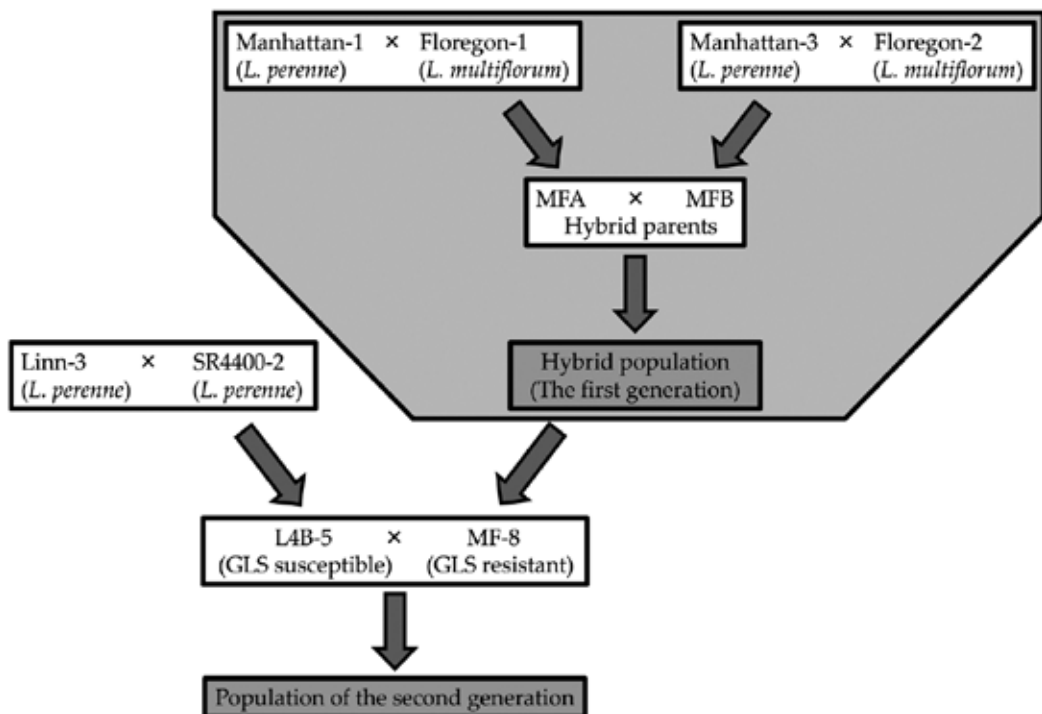
Modified and combined from [45, 47].

Figure 1. Diagram of crosses for the development of mapping populations over two generations.

4.1.2. Phenotyping of GLS resistance/susceptibility in an Italian × perennial ryegrass hybrid

In two previous studies, seven perennial ryegrass isolates obtained from diseased perennial ryegrass fairways and one rice lab strain capable of infecting rice and ryegrass were used in inoculation tests of the parents and grandparents of the first-generation mapping population

[45, 46]. Of these, one of the perennial ryegrass isolates, GG9 [45, 46], and the rice lab strain 6082 [46] were chosen and used for quantitative trait locus (QTL) analyses because of their high sporulation capacity in culture and high virulence [46].

Because the mapping population could be asexually propagated, two inoculation experiments were independently conducted with three or four replicates in one study [45] and four inoculation experiments were completed with four replicates in another [46]. The inoculation experiments were conducted in growth chambers or mist chambers. The GLS resistance/ susceptibility phenotypes of the mapping population were scored based on the rating scale provided in Table 3. In one study, lesion numbers and proportions of resistant lesions were recorded because inoculated individuals often had both resistant and susceptible lesions [45]. In another study, the youngest leaves of each plant were used because symptoms were most severe in these leaves when mixed lesion types occurred on the same plant [46].

Phenotype	Score	Symptoms
Resistant	0	No visible symptoms
	1	Dark brown, non-sporulating 2–3 mm lesions
	2	Dark brown, non-sporulating lesions with a small central necrotic area
Susceptible	3	Circular or small diamond-shaped lesions with prominent dark brown borders and grey or white central sporulating areas
	4	Large, expanding, completely unbordered sporulating lesions, often with chlorotic halos

From [45, 46]

Table 3. Rating scale for grey leaf spot severity in an Italian × perennial ryegrass hybrid

Similar disease reactions and phenotypic segregation patterns were observed in the mapping population inoculated with the perennial ryegrass isolate GG9, but the results were different from those of experiments involving the rice lab strain 6082 [45, 46]. In another study, where the second-generation mapping population was developed, two perennial ryegrass isolates, including GG9, were used. Each isolate was included in two experiments involving four clonal replicates of the mapping population [47]. Similar disease reactions and phenotype segregation patterns were reported for the second-generation mapping population [47]. No symptom-free individuals were observed throughout these studies [45-47]. The results from these three independent studies indicate the existence of different factors regulating the host–pathogen interactions involving perennial ryegrass isolates and a rice lab strain. This is relevant for determining the *Magnaporthe* species population structure based on the host specificities mentioned in Section 2.2.

Similar to the studies mentioned in Section 3.1, the broad-sense heritability for GLS-resistant/ susceptible phenotypes was high in the experiments with the perennial ryegrass isolates with

values of 0.895–0.932 [46] and 0.88 [47]. These results indicate that the GLS resistance of the mapping populations was mainly controlled by genetic effects.

4.1.3. Detection and mapping of GLS resistance loci in an Italian × perennial ryegrass hybrid

Phenotypic data related to GLS resistance/susceptibility have been analyzed to identify GLS resistance loci in mapping populations [45-47]. A genetic linkage map was constructed using RFLP, AFLP, simple sequence repeat (SSR), and random amplified polymorphic DNA markers [45-47]. Isozyme and morphological markers have also been used [47]. The genetic linkage map from [46] was described in detail in another study [50]. Probes for RFLP markers were derived from other well-studied crops such as barley, oat, and rice so that synteny-based comparative studies among different plant species could be conducted with the constructed map [51]. In these studies, two sets of genetic linkage maps composed of seven linkage groups (LGs) derived from both parents were constructed using a two-way pseudo-testcross mapping strategy [52].

In one study, although results were not shown in detail, QTL analysis detected two genomic regions for GLS resistance against the perennial ryegrass isolate GG9 [45]. The identified QTLs were on LG 2 (for proportions of resistant lesions) and LG 4 (for lesion numbers) [45]. The logarithm of odds (LOD) obtained by interval mapping [53] ranged from about 2.0 to 6.0, although the LOD scores were not always significant [45]. In addition to these QTL regions, some regions were noted on LGs 1, 3, and 5, but these were not consistently detected [45].

Isolate GG9 and rice lab strain 6082 were used to inoculate the same population used in [46]. Significant QTLs were detected on LGs 3 and 6 and LGs 2 and 4 for GG9 and 6082, respectively, indicating that GLS resistance against the different isolates was controlled by different genetic effects [46]. Percentages of phenotypic variance explained by the QTLs at the highest LOD scores were 20.1–37.9% for LG 3 and 9.2–10.7% for LG 6 for resistance against GG9, and 8.9–10.0% for LG 2, and 9.9% for LG 4 for resistance against 6082 [46]. The QTL differences between the two isolates were expected because the disease reaction and phenotype segregation of the mapping population were different between the isolates [46] (see Section 4.1.2). Nevertheless, significant QTLs were detected on LGs 2 and 4 for GLS resistance against GG9 and 6082 [45, 46]. However, the QTL relationships between the two studies cannot be confirmed by their location on genetic linkage maps because no marker information linked to the QTLs was provided in [45]. Additionally, the locations of the QTLs for GLS resistance against GG9 differed between the two studies even though the same mapping population was used. This inconsistency was not explained [46], but differences in the phenotype segregation of the mapping population during the GG9 inoculation experiments may have been a factor. That is, in one study, the phenotypic distribution of the mapping population seemed skewed toward resistance in the first experiment, but there was a trend toward susceptibility in the second experiment [45]. In the other study, the patterns of phenotype segregation in the mapping population were consistent and showed a trend toward susceptibility over three experiments [46]. These differences in the same mapping population may have been caused by unknown environmental factors that affected the expression of certain genes in the plant hosts and/or pathogens. Irrespective of the high broad-sense heritability, the values for the phenotypic

variance explained by the QTLs are considered quite low, indicating there might be undetected genetic factors with minor effects on GLS resistance/susceptibility [46].

Although the QTLs for GLS resistance may be unstable and sometimes adversely influenced by environmental factors, the most significant QTL detected on LG 3 [46] might be detectable in the second generation mapping population developed in [47] (Figure 1). The percentage of phenotypic variance explained by the QTL on LG 3 at the highest LOD scores was 9.3–10.8%. Although this is lower than the values reported in [46], it suggests that the QTL is functional in a population with a different genetic background, which is promising for breeding programs focused on developing GLS-resistant ryegrass. However, the nearest RFLP marker (CDO460) closely linked to the major QTL on LG 3 [46] was not mapped in [47]. Therefore, it is necessary to confirm whether the QTL detected in [47] really corresponds to the QTL detected in [46].

4.2. Molecular marker development for GLS resistance in Italian ryegrass

4.2.1. Mapping population derived from a single cross in Italian ryegrass

Marker development studies involving Italian ryegrass have been completed with F_1 mapping populations obtained from a single cross between resistant and susceptible genotypes [4, 49]. Annuality is a more common characteristic among grass species than the perenniality of the previously mentioned Italian × perennial ryegrass hybrid (see Section 4.1). Therefore, it might be difficult to maintain and asexually propagate the Italian ryegrass population to produce clonal replicates like those used in the studies of hybrid populations [45-47]. Regardless, GLS-resistant genotypes, which can involve a resistant parent of the mapping population, are very rare because most Italian ryegrass commercial cultivars are susceptible to GLS, similar to perennial ryegrass. Thus, it would be ideal if the resistant genotypes could at least be maintained. An *in vitro* preservation method [54] can be used to maintain and clonally propagate rare genotypes [55].

4.2.2. Detection of a GLS resistance locus by bulked segregant analysis in Italian ryegrass

A major genetic locus in Italian ryegrass for crown rust resistance has been detected using bulked segregant analysis (BSA) [56], and AFLP markers tightly linked to the locus have been developed [57]. Researchers have attempted to detect a GLS resistance locus in Italian ryegrass [4]. An F_1 mapping population was generated from a single cross between a resistant individual from cultivar "Sachiaoba" [2] as the female parent and a susceptible individual from cultivar "Minamiaoba" as the male parent. The rating scale used for phenotyping the F_1 mapping population is provided in Table 4.

The inoculation test used during phenotyping was completed only once because of the annuality of the plant material. Nevertheless, disease severity in the mapping population segregated in a 1:1 ratio (resistant:susceptible) [4]. This result suggests that resistance is controlled by one genetic locus. Therefore, the resistance locus was considered a suitable target detectable by BSA. As predicted, AFLP markers specific for resistant phenotypes were screened by BSA, and a single genetic linkage map composed of 25 of the screened AFLP

Phenotype	Score	Symptoms
Resistant	0	Plants with no leaf symptoms
	1	Plants with brown spotted or brown spindle-shaped leaf lesions
Susceptible	2	Plants with a few white or grey leaf lesions
	3	Plants with leaves covered in lesions

From [4]

Table 4. Rating scale for grey leaf spot severity in Italian ryegrass

markers was constructed [4]. Additionally, the cleaved amplified polymorphic sequence (CAPS) markers derived from Italian ryegrass expressed sequence tags (ESTs) [58] were mapped. The LG associated with the constructed map could be identified because the CAPS markers had already been assigned to seven Italian ryegrass LGs [59]. As a result, the p56 CAPS marker located on LG 5 was mapped, indicating that the resistance locus was on LG 5. Additionally, a significant QTL was detected by interval mapping. The gene at the identified resistance locus was designated *LmPi1* [4]. Although the results of the QTL analysis, including LOD score and phenotypic variance, were not described in the study, the raw data were analyzed for this chapter. The highest LOD score obtained by interval mapping was 7.36, and the percentage of the phenotypic variance explained by the QTL at the highest LOD score was 19.0%. Although broad-sense heritability of the resistance is unknown, the percentage of the phenotypic variance was unexpectedly low because the strong effect of a major gene was expected based on phenotype segregation data. Similar to the results of the Italian × perennial ryegrass hybrid, the low proportion of the phenotypic variance indicates there might be undetected genetic factors in other genomic regions that have a minor effect on GLS resistance/susceptibility (see Section 4.1.3).

4.2.3. Targeted mapping of rice ESTs to the LmPi1 locus

The sequenced rice genome [60] and expanded EST datasets in various plant species enable comparative genomics studies of model and nonmodel plants, in which collinearity of molecular markers and genes in syntenic regions can be elucidated. Based on syntenic regions, high-resolution mapping of genetic loci associated with agronomic traits is possible. This is true even for nonmodel crops where EST-derived markers can be used to map landmarks and demonstrate synteny among different species [61-63]. Conserved intron-scanning primers (CISPs) can be easily developed and used to study nonmodel species [64]. For CISP development, polymerase chain reaction (PCR) primers are designed within relatively conserved exons nearby boundaries between an exon and a variation-rich intron. Target segments are generated by PCR where the introns are scanned during the extension step. Polymorphisms in the PCR products are detected as variations in the introns including base substitutions or insertions and deletions.

Synteny among ryegrasses, rice, and other grasses such as oat and Triticeae species has been revealed. Ryegrass LG 5, where the previously mentioned *LmPi1* is located, has been shown to be syntenic to rice chromosome (Chr) 9 [51, 65]. Thus, to enhance the single genetic linkage map of *LmPi1*, targeted mapping of rice ESTs to the *LmPi1* locus has been attempted using the F_1 mapping population DNA used to detect the *LmPi1* locus [48]. The CISPs were designed by aligning the rice genome sequence and ESTs on rice Chr 9. Polymorphic PCR products were detected by single-strand conformation polymorphism analysis [48]. Consequently, a single genetic linkage map spanning 66.3 cM composed of 17 CISP markers and the p56 marker tightly linked to *LmPi1* (see Section 4.2.2) was constructed. There was significant collinearity of marker orders between rice Chr 9 and the newly constructed map corresponding to ryegrass LG 5 [48].

Recently, the primer design method involving CISPs has been improved for temperate forage grasses including ryegrasses [66]. Primers were called Conserved Three-prime-End Region (COTER) primers. They were developed from EST sequences of tall fescue and wheat, and eight bases at the 3′ end of each primer were identical to rice orthologues, which provided high transferability in six temperate grasses [66]. The COTER primers have been used for targeted mapping of a locus for brittleness to a single genetic linkage map in a mutant Italian ryegrass line (unpublished data), thereby providing further evidence of the high transferability of these primers.

4.2.4. Detection of a novel major locus for GLS resistance in Italian ryegrass

There has been an attempt to identify a resistance locus using a similar approach to that used to identify *LmPi1* [49]. An F_1 mapping population was generated from a single cross between a resistant individual from the commercial cultivar "Surrey" [67] as the female parent and a susceptible individual from the cultivar "Minamiaoba" as the male parent. As described in Section 3.2, the tetraploid cultivar "Jumbo" [42] has been registered as a GLS-resistant cultivar in the United States. The cultivar was developed by doubling the chromosomes of the diploid "Surrey." Thus, it was reasonable to expect that resistance genotypes existed in "Surrey." However, different genetic factors were expected from the resistant parent because the source material was different from that used in the study of *LmPi1*, which explains why "Surrey" was chosen as the resistant female parent.

4.2.4.1. Artificial inoculation method using detached leaves

A high heritability of target traits enables very precise QTL analyses. However, the severity of GLS symptoms in ryegrasses is influenced by environmental factors such as temperature and humidity [1, 68, 69]. Fluctuations in these factors may prevent accurate phenotyping of GLS resistance/susceptibility of the mapping population, thereby decreasing the heritability of the disease reaction. Accordingly, phenotyping in stable environmental conditions may lead to increased heritability. Additionally, repeated phenotyping in stable environmental conditions can further moderate environmental effects and increase the accuracy of the phenotype evaluation.

Multiple phenotypic evaluations of the Italian ryegrass F_1 mapping population infected with GLS has not been conducted because of the annuality of Italian ryegrass and the fact that GLS is highly lethal to infected plants. Thus, a novel inoculation method, the filter-paper method, has been employed for the phenotypic evaluation of F_1 mapping populations [70]. This method can overcome the difficulties of working with Italian ryegrass because it only requires detached leaves from young seedlings. The rating scale for this method is provided in Figure 2. The scale is similar to those of other studies [45, 46] (Table 3) but differs because the score is based on lesion type and not size. More recently, the filter-paper method has been shown to be applicable to the evaluation of resistance to rice blast [71].

0: No visible symptoms

1: Dark brown, non-sporulating lesions

2: Expanding, dark brown, non-sporulating lesions

3: Small circular or diamond-shaped lesions with sporulating areas

4: Large, expanding lesions with sporulating areas
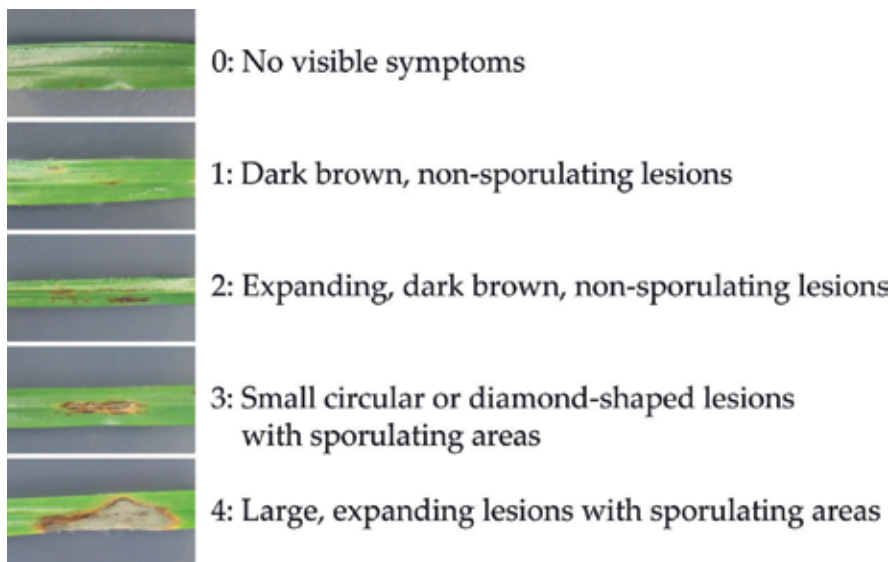
Modified from [70]

Figure 2. Rating scale for grey leaf spot severity used in the filter-paper method.

4.2.4.2. Detection of the LmPi2 locus

Based on the filter-paper method, GLS severity was evaluated twice in young, expanding leaves and fully expanded leaves under controlled inoculation conditions [49]. A significant correlation was observed for all GLS severity scores at different leaf ages, but higher correlation coefficients were found between results from the same leaf stage. Additionally, results of repeated-measures analysis of variance (ANOVA) indicated there were significant differences in GLS severity scores among genotypes for all inoculations, whereas the differences were not significant for inoculated leaves of the same age. This indicated that the results of the filter-paper method were highly reproducible [49]. Because of this method, high broad-sense heritability was determined from the results of the repeated-measures ANOVA, with values of 0.701, 0.779, and 0.665 for young leaves, expanded leaves, and all inoculations, respectively [49].

The ratios for phenotype segregation of the mapping population were 1:1 for young leaves and 3:1 for expanded leaves. Therefore, it was concluded that one or two genes controlled GLS resistance in the mapping population [49]. These results and the high broad-sense heritability mentioned earlier encouraged the use of BSA to identify the most important genes. Preliminary analysis with AFLP markers demonstrated that two markers specific to the resistant parent and resistant bulk were genetically linked. Thus, the two markers along with SSR markers from a reference map of Italian ryegrass [72] were further analyzed. Because the two SSR markers were located on LG 3 in the reference map, the resistance locus was predicted to be located on LG 3. A single genetic linkage map was constructed with the AFLP and SSR markers. Further, ESTs from rice Chr 1 were converted to CISP markers because LG 3 was syntenic to rice Chr 1. Grass anchor RFLP probes located on LG 3 [51, 65] were also converted to CISP markers. The enhanced single genetic linkage map covering 133.6 cM showed significant collinearity with rice Chr 1 in their marker orders [49]. A significant QTL was also detected by interval mapping. The highest LOD scores from interval mapping were 13.8, 15.2, and 17.9 for young leaves, expanded leaves, and total data from four inoculation experiments, respectively [49]. Percentages of phenotypic variance explained by the QTL at the highest LOD scores were 61.0, 68.1, and 69.5% for young leaves, expanded leaves, and total data from four inoculation experiments, respectively [49]. The most important point of this study was that, unlike for *LmPi1*, the broad-sense heritability score (0.665) and percentage of phenotypic variance explained by the QTL at the highest LOD score (69.5%) were very similar. In other words, although only a single genetic linkage map of LG 3 was constructed, most of the genetic factors for the GLS resistance phenotype in the mapping population can be explained by the functions of a single gene.

The detected locus is clearly distinguishable from *LmPi1* because it is located on a different LG. Conversely, the QTL detected in [46] with the highest percentages of phenotypic variance explained was located on the same LG as the detected locus. The two resistance loci could not be distinguished because there was no common marker around the locus that could be used as a landmark. However, there were markers close to both loci on LG 3 of the Italian ryegrass reference genetic linkage map [72]. The genetic distance between the two loci was estimated to be over 25 cM, suggesting the detected locus is probably not the QTL detected in [46]. The detected locus was designated *LmPi2* [49], which is the second identified GLS resistance locus in Italian ryegrass.

5. Conclusion

This chapter summarized the advances that have been made in the molecular breeding of GLS resistance in ryegrasses. Rice blast and GLS are caused by *M. oryzae*, but rice blast has been studied more extensively because of the importance of this staple food crop. Nevertheless, there are still incidences of rice blast leading to considerable yield losses, and numerous issues regarding this disease require further research. The breeding history of rice-blast-resistant cultivars is a major consideration during breeding of GLS-resistant ryegrasses. The breakdown of resistance regulated by a few genes is one of the most important factors related to the

development of rice-blast-resistant cultivars [44]. Similar concerns would apply to the breeding of GLS-resistant ryegrass cultivars if a small number of genes mediated the resistance. Although some genomic regions associated with GLS resistance have been identified, further studies are required in ryegrasses because our knowledge of GLS resistance is more limited than our understanding of rice blast resistance. To establish highly productive cultural system for ryegrasses, synchronized approaches between cultural disease management practices and breeding for GLS resistance, promoted by advances in plant genomics, are necessary.

Acknowledgements

Author details

Wataru Takahashi[*]

Address all correspondence to: twataru@affrc.go.jp

Division of Forage Crop Research, Institute of Livestock and Grassland Science, NARO, Tochigi, Japan

References

[1] Uddin W, Viji G, Vincelli P. Gray leaf spot (blast) of perennial ryegrass turf: An emerging problem for the turfgrass industry. *Plant Dis*. 2003;87:880-889. DOI: 10.1094/PDIS.2003.87.8.880

[2] Mizuno K, Yokohata Y, Oda T, Fujiwara T, Hayashi K, Ozaki R, et al. Breeding of a new variety 'Sachiaoba' in Italian ryegrass (*Lolium multiflorum* L.) and its characteristics. *Bull Yamaguchi Agri Exper Station*. 2003;54:11-24.

[3] Sumida Y, Fukubara H, Inoue T, Kajihara H, Mizuno K, Fujiwara T. A study of the occurrence of dieback on Italian ryegrass (*Lolium multiflorum*) caused by *Pyricularia orizae* – Investigation of facts about occurrence and development of a new resistance test method on breeding. *Bull Yamaguchi Agri Exper Station*. 2003;54:29-36.

[4] Miura Y, Ding C, Ozaki R, Hirata M, Fujimori M, Takahashi W, et al. Development of EST-derived CAPS and AFLP markers linked to a gene for resistance to ryegrass

blast (*Pyricularia* sp.) in Italian ryegrass (*Lolium multiflorum* Lam.). *Theor Appl Genet*. 2005;111:811-818. DOI: 10.1007/s00122-005-0001-8

[5] Couch BC, Kohn LM. A multilocus gene genealogy concordant with host preference indicates segregation of a new species, *Magnaporthe oryzae*, from *M. grisea*. *Mycologia*. 2002;94:683-693. DOI: 10.2307/3761719

[6] Vincelli P, Dixon E. Resistance to QoI (strobilurin-like) fungicides in isolates of *Pyricularia grisea* from perennial ryegrass. *Plant Dis*. 2002;86:235-240. DOI: 10.1094/PDIS. 2002.86.3.235

[7] Pérez-García A, Romero D, de Vicente A. Plant protection and growth stimulation by microorganisms: Biotechnological applications of Bacilli in agriculture. *Curr Opin Biotechnol*. 2011;22:187-193. DOI: 10.1016/j.copbio.2010.12.003

[8] Lemus R, Tomaso-Peterson M. Gray leaf spot (blast) in annual ryegrass. MSU cares.com [Internet]. 2010. Available from: http://msucares.com/pubs/publications/ p2633.html [Accessed: 2015-10-06]

[9] Bonos SA, Murphy JA, Clarke BB. Integrated control of gray leaf spot on perennial ryegrass. Fact Sheet [Internet]. 2006. Available from: https://njaes.rutgers.edu/pubs/ publication.asp?pid=FS1048 [Accessed: 2015-10-06]

[10] Zhang N, Rossman AY, Seifert K, Bennett JW, Cai G, Cai L, et al. Impacts of the international code of nomenclature for algae, fungi, and plants (Melbourne Code) on the scientific names of plant pathogenic fungi. *APS Features*. 2013. DOI: 10.1094/APSFeature-2013-06

[11] Tosa Y, Uddin W, Viji G, Kang S, Mayama S. Comparative genetic analysis of *Magnaporthe oryzae* isolates causing gray leaf spot of perennial ryegrass turf in the United States and Japan. *Plant Dis*. 2007;91:517-524. DOI: 10.1094/PDIS-91-5-0517

[12] Farman ML. *Pyricularia grisea* isolates causing gray leaf spot on perennial ryegrass (*Lolium perenne*) in the United States: Relationship to *P. grisea* isolates from other host plants. *Phytopathology*. 2002;92:245-254. DOI: 10.1094/PHYTO.2002.92.3.245

[13] Viji G, Wu B, Kang S, Uddin W. *Pyricularia grisea* causing gray leaf spot of perennial ryegrass turf: Population structure and host specificity. *Plant Dis*. 2001;85:817-826. DOI: 10.1094/PDIS.2001.85.8.817

[14] Tosa Y, Hirata K, Tamba H, Nakagawa S, Chuma I, Isobe C, et al. Genetic constitution and pathogenicity of *Lolium* isolates of *Magnaporthe oryzae* in comparison with host species-specific pathotypes of the blast fungus. *Phytopathology*. 2004;94:454-462. DOI: 10.1094/PHYTO.2004.94.5.454

[15] Vincelli P, Dixon E, Farman M. Susceptibility of selected cultivars of forage grasses to *Magnaporthe oryzae* isolates from annual ryegrass and relatedness of the pathogen to strains from other grasses. *Forage Grazinglands*. 2008; 6:FG-2008-0226-2001-RS.

[16] Farman ML, Tosa Y, Nitta N, Leong SA. MAGGY, a retrotransposon in the genome of the rice blast fungus *Magnaporthe grisea*. *Mol General Genet*. 1996;251:665-674. DOI: 10.1007/BF02174115

[17] Kang S. Organization and distribution pattern of MGLR-3, a novel retrotransposon in the rice blast fungus *Magnaporthe grisea*. *Fungal Genet Biol*. 2001;32:11-19. DOI: 10.1006/fgbi.2000.1246

[18] Hamer JE, Farrall L, Orbach MJ, Valent B, Chumley FG. Host species-specific conservation of a family of repeated DNA sequences in the genome of a fungal plant pathogen. *Proc Natl Acad Sci*. 1989;86:9981-9985. DOI: 10.1073/pnas.86.24.9981

[19] Kachroo P, Ahuja M, Leong SA, Chattoo BB. Organisation and molecular analysis of repeated DNA sequences in the rice blast fungus *Magnaporthe grisea*. *Curr Genet*. 1997;31:361-369. DOI: 10.1007/s002940050217

[20] Farman ML, Taura S, Leong SA. The *Magnaporthe grisea* DNA fingerprinting probe MGR586 contains the 3′ end of an inverted repeat transposon. *Mol General Genet*. 1996;251:675-681. DOI: 10.1007/BF02174116

[21] Kachroo P, Leong SA, Chattoo BB. Pot2, an inverted repeat transposon from the rice blast fungus *Magnaporthe grisea*. *Mol General Genet*. 1994;245:339-348. DOI: 10.1007/BF00290114

[22] White TJ, Bruns T, Lee S, Taylor J. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ, editors. *PCR Protocols: A Guide to Methods and Applications*. San Diego: Academic Press; 1990. p. 315-322. DOI: 10.1016/B978-0-12-372180-8.50042-1

[23] Free SJ, Rice PW, Metzenberg RL. Arrangement of the genes coding for ribosomal ribonucleic acids in *Neurospora crassa*. *J Bacteriol*. 1979;137:1219-1226.

[24] Farman ML, Eto Y, Nakao T, Tosa Y, Nakayashiki H, Mayama S, et al. Analysis of the structure of the *AVR1-CO39* avirulence locus in virulent rice-infecting isolates of *Magnaporthe grisea*. *Mol Plant-Microbe Interact*. 2002;15:6-16. DOI: 10.1094/MPMI.2002.15.1.6

[25] Tredway LP, Stevenson KL, Burpee LL. Genetic structure of *Magnaporthe grisea* populations associated with St. Augustinegrass and tall fescue in Georgia. *Phytopathology*. 2005;95:463-471. DOI: 10.1094/PHYTO-95-0463

[26] Douhan GW, de la Cerda KA, Huryn KL, Greer CA, Wong FP. Contrasting genetic structure between *Magnaporthe grisea* populations associated with the golf course turfgrasses *Lolium perenne* (perennial ryegrass) and *Pennisetum clandestinum* (kikuyu-grass). *Phytopathology*. 2011;101:85-91. DOI: 10.1094/PHYTO-08-10-0205

[27] Urashima AS, Igarashi S, Kato H. Host range, mating type, and fertility of *Pyricularia grisea* from wheat in Brazil. *Plant Dis*. 1993;77:1211-1216. DOI: 10.1094/PD-77-1211

[28] Trevathan LE. Pathogenicity on ryegrass and cultural variability of Mississippi isolates of *Pyricularia grisea*. *Plant Dis*. 1982;66:592-594. DOI: 10.1094/PD-66-592

[29] Darbyshire SJ. Realignment of *Festuca* subgenus *Schedonorus* with the genus *Lolium* (Poaceae). *Novon*. 1993;3:239-243. DOI: 10.2307/3391460

[30] Soreng RJ, Terrell EE, Wiersema J, Darbyshire SJ. (1488) Proposal to conserve the name *Schedonorus arundinaceus* (Schreb.) Dumort. against *Schedonorus arundinaceus* Roem. & Schult. (*Poaceae: Poeae*). Taxon. 2001;50:915-917. DOI: 10.2307/1223726

[31] Craven KD, Clay K, Schardl CL. Systematics and morphology. In: Fribourg HA, Hannaway DB, West CP, editors. *Tall Fescue for the Twenty-First Century. Agronomy Monograph*. Madison: American Society of Agronomy, Inc.; Crop Science Society of America, Inc.; Soil Science Society of America, Inc.; 2009. p. 11-30. DOI: 10.2134/agronmonogr53.c2

[32] Trevathan LE. Response of ryegrass plant introductions to artificial inoculation with *Pyricularia grisea* under greenhouse conditions. *Plant Dis*. 1982;66:696-697. DOI: 10.1094/PD-66-696

[33] Reith PE, Prine GM, Blount AR. Gray leaf spot: New-old peril to annual ryegrass. *Soil Crop Sci Soc Florida Proc*. 2001;60:143-146.

[34] Reith PE, Prine GM, Blount AR. Selection for gray leaf spot disease resistance in annual ryegrass. *Soil Crop Sci Soc Florida Proc*. 2003;62:69-73.

[35] Bonos SA, Kubik C, Clarke BB, Meyer WA. Breeding perennial ryegrass for resistance to gray leaf spot. *Crop Sci*. 2004;44:575-580. DOI: 10.2135/cropsci2004.5750

[36] Han Y, Bonos SA, Clarke BB, Meyer WA. Inheritance of resistance to gray leaf spot disease in perennial ryegrass. *Crop Sci*. 2006;46:1143-1148. DOI: 10.2135/cropsci2005.07-0217

[37] Bonos SA, Rush D, Hignight K, Langlois S, Meyer WA. The effect of selection on gray leaf spot resistance in perennial ryegrass. *Int Turfgrass Soc Res J*. 2005;10:501-507.

[38] Hayman BI. The theory and analysis of diallel crosses. *Genetics*. 1954;39:789-809.

[39] Yates F. Analysis of data from all possible reciprocal crosses between a set of parental lines. *Heredity*. 1947;1:287-301. DOI: 10.1038/hdy.1947.19

[40] Gilbert NEG. Diallel cross in plant breeding. *Heredity*. 1958;12:477-492. DOI: 10.1038/hdy.1958.48

[41] Buckley R. Touch of gray. Plant and Pest Advisory [Internet]. 2013. Available from: http://plant-pest-advisory.rutgers.edu/touch-of-gray/ [Accessed: 2015-10-06]

[42] Prine GM, Blount AR, Dunavin LS, Mislevy P, Stanley RL, Jr. Registration of 'Jumbo' annual ryegrass. *Crop Sci*. 2002;42:1749. DOI: 10.2135/cropsci2002.1749

[43] Ballini E, Morel J-B, Droc G, Price A, Courtois B, Notteghem J-L, et al. A genome-wide meta-analysis of rice blast resistance genes and quantitative trait loci provides

new insights into partial and complete resistance. *Mol Plant-Microbe Interact*. 2008;21:859-868. DOI: 10.1094/MPMI-21-7-0859

[44] Miah G, Rafii MY, Ismail MR, Puteh AB, Rahim HA, Asfaliza R, et al. Blast resistance in rice: a review of conventional breeding to molecular approaches. *Mol Biol Rep*. 2013;40:2369-2388. DOI: 10.1007/s11033-012-2318-0

[45] Curley J, Sim SC, Jung G, Leong S, Warnke S, Barker RE. QTL mapping of gray leaf spot resistance in ryegrass, and synteny-based comparison with rice blast resistance genes in rice. In: Hopkins A, Wang Z-Y, Mian R, Sledge M, Barker RE, editors. *Molecular Breeding of Forage and Turf. Developments in Plant Breeding*. Dordrecht: Kluwer Academic Publishers; 2004. p. 37-46. DOI: 10.1007/1-4020-2591-2_3

[46] Curley J, Sim SC, Warnke S, Leong S, Barker R, Jung G. QTL mapping of resistance to gray leaf spot in ryegrass. *Theor Appl Genet*. 2005;111:1107-1117. DOI: 10.1007/ s00122-005-0036-x

[47] Curley J, Chakraborty N, Chang S, Jung G. QTL mapping of resistance to gray leaf spot in ryegrass: Consistency of QTL between two mapping populations. *Korean J Turfgrass Sci*. 2008;22:85-100.

[48] Takahashi W, Miura Y, Sasaki T, Takamizo T. Targeted mapping of rice ESTs to the *LmPi1* locus for grey leaf spot resistance in Italian ryegrass. *Eur J Plant Pathol*. 2010;126:333-342. DOI: 10.1007/s10658-009-9549-5

[49] Takahashi W, Miura Y, Sasaki T, Takamizo T. Identification of a novel major locus for gray leaf spot resistance in Italian ryegrass (*Lolium multiflorum* Lam.). *BMC Plant Biol*. 2014;14:303. DOI: 10.1186/s12870-014-0303-6

[50] Warnke SE, Barker RE, Jung G, Sim SC, Rouf Mian MA, Saha MC, et al. Genetic linkage mapping of an annual × perennial ryegrass population. *Theor Appl Genet*. 2004;109:294-304. DOI: 10.1007/s00122-004-1647-3

[51] Sim S, Chang T, Curley J, Warnke SE, Barker RE, Jung G. Chromosomal rearrangements differentiating the ryegrass genome from the Triticeae, oat, and rice genomes using common heterologous RFLP probes. *Theor Appl Genet*. 2005;110:1011-1019. DOI: 10.1007/s00122-004-1916-1

[52] Grattapaglia D, Sederoff R. Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: Mapping strategy and RAPD markers. *Genetics*. 1994;137:1121-1137.

[53] Lander ES, Botstein D. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*. 1989;121:185-199.

[54] Jackson JA, Dale PJ. Callus induction, plant regeneration and an assessment of cytological variation in regenerated plants of *Lolium multiflorum* L. *J Plant Physiol*. 1988;132:351-355. DOI: 10.1016/S0176-1617(88)80119-6

[55] Takahashi W, Komatsu T, Fujimori M, Takamizo T. Screening of regenerable geno-types of Italian ryegrass (*Lolium multiflorum* Lam.). *Plant Prod Sci*. 2004;7:55-61. DOI: 10.1626/pps.7.62

[56] Michelmore RW, Paran I, Kesseli RV. Identification of markers linked to disease-re-sistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations. *Proc Natl Acad Sci U S A*. 1991;88:9828-9832. DOI: 10.1073/pnas.88.21.9828

[57] Fujimori M, Hayashi K, Hirata M, Ikeda S, Takahashi W, Mano Y, et al. Molecular breeding and functional genomics for tolerance to biotic stress. In: Hopkins A, Wang Z-Y, Mian R, Sledge M, Barker RE, editors. *Molecular Breeding of Forage and Turf. Developments in Plant Breeding*. Dordrecht: Kluwer Academic Publishers; 2004. p. 21-35. DOI: 10.1007/1-4020-2591-2_2

[58] Ikeda S, Takahashi W, Oishi H. Generation of expressed sequence tags from cDNA libraries of Italian ryegrass (*Lolium multiflorum* Lam.). *Grassland Sci*. 2004;49:593-598.

[59] Miura Y, Hirata M, Fujimori M. Mapping of EST-derived CAPS markers in Italian ryegrass (*Lolium multiflorum* Lam.). *Plant Breed*. 2007;126:353-360. DOI: 10.1111/j. 1439-0523.2007.01386.x

[60] International Rice Genome Sequencing Project. The map-based sequence of the rice genome. *Nature*. 2005;436:793-800. DOI: 10.1038/nature03895

[61] Perovic D, Stein N, Zhang H, Drescher A, Prasad M, Kota R, et al. An integrated ap-proach for comparative mapping in rice and barley with special reference to the Rph16 resistance locus. *Function Integr Genom*. 2004;4:74-83. DOI: 10.1007/s10142-003-0100-z

[62] Le Cunff L, Garsmeur O, Raboin LM, Pauquet J, Telismart H, Selvi A, et al. Diploid/polyploid syntenic shuttle mapping and haplotype-specific chromosome walking to-ward a rust resistance gene (*Bru1*) in highly polyploid sugarcane (2n ~ 12x ~ 115). *Genetics*. 2008;180:649-660. DOI: 10.1534/genetics.108.091355

[63] Mammadov JA, Steffenson BJ, Saghai Maroof MA. High-resolution mapping of the barley leaf rust resistance gene *Rph5* using barley expressed sequence tags (ESTs) and synteny with rice. *Theor Appl Genet*. 2005;111:1651-1660. DOI: 10.1007/s00122-005-0100-6

[64] Feltus FA, Singh HP, Lohithaswa HC, Schulze SR, Silva TD, Paterson AH. A compa-rative genomics strategy for targeted discovery of single-nucleotide polymorphisms and conserved-noncoding sequences in orphan crops. *Plant Physiol*. 2006;140:1183-1191. DOI: 10.1104/pp.105.074203

[65] Jones ES, Mahoney NL, Hayward MD, Armstead IP, Jones JG, Humphreys MO, et al. An enhanced molecular marker based genetic map of perennial ryegrass (*Lolium per-*

enne) reveals comparative relationships with other Poaceae genomes. *Genome*. 2002;45:282-295. DOI: 10.1139/G01-144

[66] Tamura K, Kiyoshi T, Yonemaru J. The development of highly transferable intron-spanning markers for temperate forage grasses. *Mol Breeding*. 2012;30:1-8. DOI: 10.1007/s11032-011-9593-7

[67] Prine GM. Registration of 'Surrey' annual ryegrass. *Crop Sci*. 1996;36:1713-1714. DOI: 10.2135/cropsci1996.0011183X003600060052x

[68] Moss MA, Trevathan LE. Environmental conditions conducive to infection of ryegrass by *Pyricularia grisea*. *Phytopathology*. 1987;77:863-866. DOI: 10.1094/Phyto-77-863

[69] Li Y, Uddin W, Kaminski JE. Effects of relative humidity on infection, colonization and conidiation of *Magnaporthe orzyae* on perennial ryegrass. *Plant Pathol*. 2014;63:590-597. DOI: 10.1111/ppa.12127

[70] Takahashi W, Miura Y, Sasaki T. A novel inoculation method for evaluation of grey leaf spot resistance in Italian ryegrass. *J Plant Pathol*. 2009;91:171-176. DOI: 10.4454/jpp.v91i1.638

[71] Challagulla V, Bhattarai S, Midmore DJ. *In-vitro* vs *in-vivo* inoculation: screening for resistance of Australian rice genotypes against blast fungus. *Rice Sci*. 2015;22:132-137. DOI: 10.1016/S1672-6308(14)60292-4

[72] Hirata M, Cai H, Inoue M, Yuyama N, Miura Y, Komatsu T, et al. Development of simple sequence repeat (SSR) markers and construction of an SSR-based linkage map in Italian ryegrass (*Lolium multiflorum* Lam.). *Theor Appl Genet*. 2006;113:270-279. DOI: 10.1007/s00122-006-0292-4

Advances in Plant Tolerance to Abiotic Stresses

Geoffrey Onaga and Kerstin Wydra

Additional information is available at the end of the chapter

Abstract

During the last 50 years, it has been shown that abiotic stresses influence plant growth and crop production greatly, and crop yields have evidently stagnated or decreased in economically important crops, where only high inputs assure high yields. The recent manifesting effects of climate change are considered to have aggravated the negative effects of abiotic stresses on plant productivity. On the other hand, the complexity of plant mechanisms controlling important traits and the limited availability of germplasm for tolerance to certain stresses have restricted genetic advances in major crops for increased yields or for improved other traits. However, some level of success has been achieved in understanding crop tolerance to abiotic stresses; for instance, identification of abscisic acid (ABA) receptors (e.g., ABA-responsive element (ABRE) binding protein/ABRE binding factor (AREB/ABF) transcription factors), and other regulons (e.g., *WRKYs*, *MYB/MYCs*, *NACs*, *HSFs*, *bZIPs* and nuclear factor-Y (NF-Y)), has shown potential promise to improve plant tolerance to abiotic stresses. Apart from these major regulons, studies on the post-transcriptional regulation of stress-responsive genes have provided additional opportunities for addressing the molecular basis of cellular stress responses in plants. This chapter focuses on the progress in the study of plant tolerance to abiotic stresses, and describes the major tolerance pathways and implicated signaling factors that have been identified, so far. To link basic and applied research, genes and proteins that play functional roles in mitigating abiotic stress damage are summarized and discussed.

Keywords: abiotic stress, climate change, crop improvement, transcription, regulatory proteins

1. Introduction

Abiotic stress is defined as the negative impact of non-living factors on living organisms in a specific environment. Abiotic stresses, such as drought, salinity, low or high temperatures and other environmental extremes are the major cause of poor plant growth and reduced crop yields in the world [1]. Drought alone affects 45% of the world's agricultural land, whereas

19.5% of irrigated agricultural lands are considered saline [2, 3]. Moreover, 16% of the agricultural rice land of the world suffers from flash flooding [4]. A combination of two or more abiotic stresses, e.g., drought and heat stress also occurs in field situations and causes more severe crop yield reductions than a single stress [5]. With increasing challenges posed by climate change, it is predicted that warming, drought, floods and storm events will become even more frequent and severe, and will further reduce crop yields, especially in the tropics and subtropics.

Abiotic stresses commonly induce overproduction of reactive oxygen species (ROS) causing extensive cellular damage and inhibition of physiological processes in plants. Although anti-oxidative mechanisms would be an immediate endogenic choice of the plants to counter ROS production, this mechanism can be impaired by abiotic stresses causing a rise in ROS intra-cellular concentration and an increase in the damage. To survive under such conditions, plants have evolved intricate mechanisms, allowing optimal responses that enable adaptation or avoidance of the stress. These plant responses are regulated at all levels of organization. At the cellular level, responses include adjustments of the membrane system, modifications of cell wall architecture, changes in cell cycle and cell division, and synthesis of specific endog-enous and low-molecular-weight molecules, such as salicylic acid, jasmonic acid, ethylene and abscisic acid [6]. An overview of changes that may occur under different abiotic stress conditions is presented in Figure 1.

At the genomic level, plant responses include the expression of stress-inducible genes involved in direct plant protection against stresses [3, 7, 8]. A broad range of abiotic stress induced genes are divided into two functional categories: and regulatory proteins. The first group consists of genes encoding for membrane proteins, enzymes for osmolyte biosynthesis, detoxification (glutathione S-transferases, superoxide dismutases, dehydrins, dehydroascorbate reductases, quinine reductases and ascorbate peroxidases) and proteins for macromolecular protection (such as LEA protein, anti-freezing proteins, chaperons and mRNA binding protein) [2]. The second group comprises genes encoding for transcription factors (e.g., *DREBPs*), protein kinases (e.g., *SRK2E*), receptor protein kinases, ribosomal-protein kinases and signal trans-duction proteinases (such as phosphoesterases and phospholipase C). Alterations in the phenylpropanoid pathway in which lignin biosynthesis intermediates are produced also occur under abiotic stress conditions. Moreover, increased accumulation of wall-linked phenolic compounds, for instance, in maize root elongation zone and the polyphenol content in cotton have been linked to stress response [9]. The same authors have shown the role of flavonoids, isoflavonoids, terpenoid and nitrogen-containing secondary metabolites such as glucosino-lates alkaloids in abiotic stress response.

Thus, abiotic stress tolerance in plants is a complex trait, involving many different metabolic pathways and cellular and molecular components.

In the past 100 years, conventional breeding (Figure 2; based on observed variation and controlled mating) approaches have randomly exploited these plant tolerance mechanisms with limited success. Moreover, *in vitro* induced variations have also shown little progress in the improvement of plants against abiotic stresses. These conventional breeding ap-proaches are limited by the complexity of stress tolerance traits coupled with less genetic
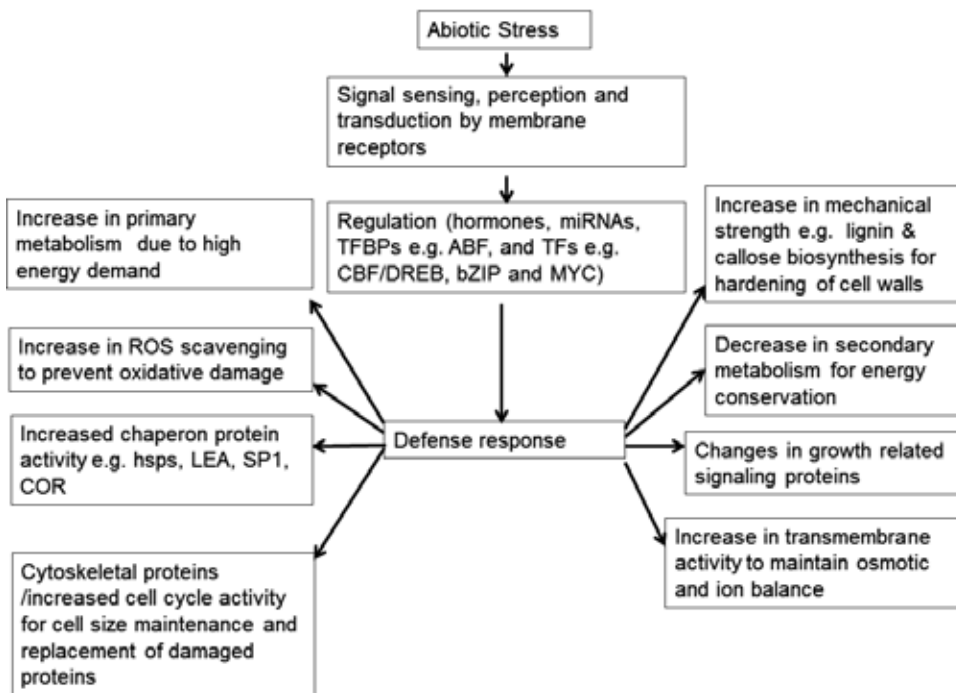
Figure 1. Abiotic stress response in plants. Primary stresses, including drought, salinity, cold, heat, and submergence, are often interconnected and cause cellular damage and secondary stresses, such as osmotic and oxidative stresses. The initial stress signals (e.g., osmotic and ionic effects or membrane fluidity changes) are perceived by membrane receptors that transmit the signals downstream to trigger transcription, which is regulated by hormones, transcription factor binding proteins (TFBPs), miRNAs, and transcription factors (TFs) to precisely activate stress responsive mechanisms to re-establish homeostasis and protect and repair damaged proteins and membranes. Inadequate response at one or several steps in the signaling and gene activation levels may ultimately result in irreversible changes of cellular homeostasis and in the destruction of functional and structural proteins and membranes, leading to cell death.

variation exhibited by most crops due to domestication bottlenecks. The recent reports that the cultivated gene pool of major cereal crops, e.g., rice, maize and wheat, has reduced in genetic variation compared to wild relatives [10–12], raises concern, and could probably undermine the current efforts to identify genetic sources of resistance within the cultivated genepools. It is important, therefore, to consider exploring alternative sources of resistance by incorporating modern techniques into traditional breeding strategies to develop stress-tolerant crops (Figure 2).

Recently, with the support of genomics, targeted genetic studies involving QTL mapping and validation, identification of key regulatory genes, e.g., genes encoding for ABA receptors, developments in transcriptional and post-transcriptional regulation of stress-responsive genes and studies on hormonal interactions during plant response to stress, have provided opportunities for understanding cellular stress responses in plants. Moreover, the emergence of deep sequencing technologies, proteomics, metabolomics and epigenetics, has remarkably provided novel possibilities to understand the biology of plants and consequently to precisely develop stress-tolerant crop varieties. Amongst the techniques that are currently being
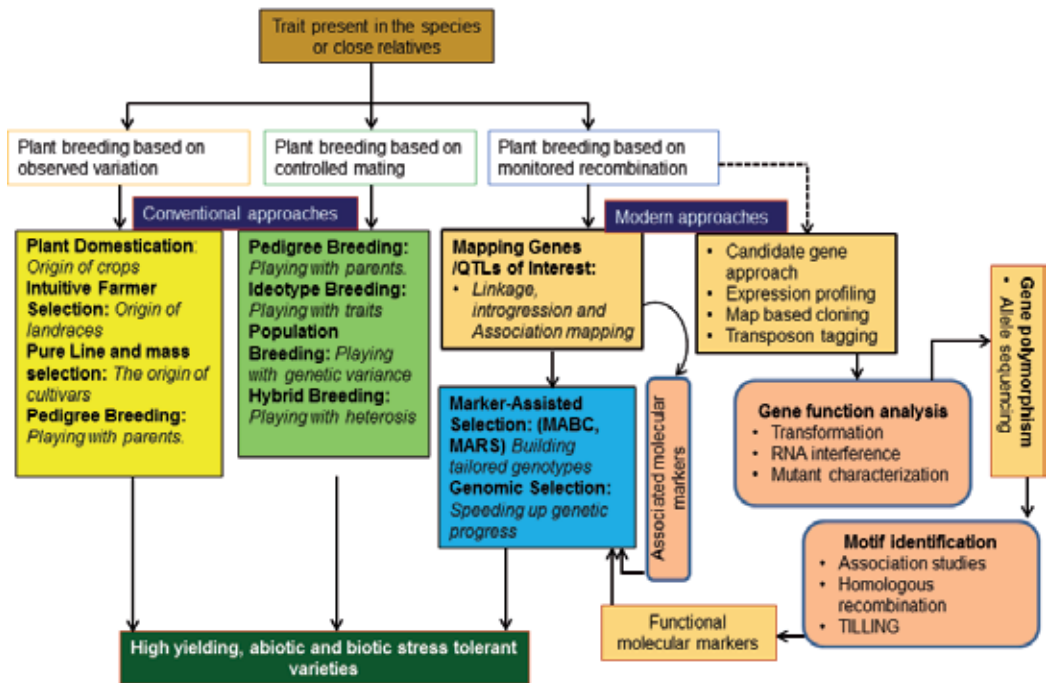
Figure 2. Overview of the traditional and modern approaches in plant breeding. In conjunction with the technological advancements, marker-assisted backcrossing (MABC) and marker-assisted recurrent selection (MARS) schemes, which target an individual marker or set of markers showing significant association with QTLs, are progressively evolving into a modification of MAS, permitting the selection of the desirable genotypes on the basis of genome-wide marker information or genomic selection (GS).

exploited to develop stress-tolerant plants, alongside basic molecular biology, there are molecular breeding methods, including development of functional molecular markers to aid in marker-assisted selection, horizontal gene transfer and genome editing tools such as CRISPR/Cas9, to develop genetically modified plants with new or improved characteristics.

In this chapter, we reviewed the plant responses to various abiotic stresses, and focus on genetic and molecular components that function to confer stress tolerance in plants.

2. Advances in plant tolerance to drought

Drought tolerance in plants is the ability to survive and produce stable yields under water scarcity during various stages of crop growth. Principally, drought stress occurs when the soil water potential falls between −0.5 and −1.5 MPa. This affects plants by decreasing the photosynthetic rate through photo-oxidation and enzyme damage, thereby decreasing the amount of assimilates available for export to the sink organs [13]. Besides this, carbohydrate metabolism in plants is severely altered, ultimately affecting both biological and economical yield [14].

Evidence from several studies has shown that plants respond to drought, like many other abiotic stresses, by inducing cellular damage and secondary stresses, such as osmotic and oxidative stresses. These secondary stresses induce initial stress signals (e.g., osmotic and ionic effects and membrane fluidity changes) that are perceived by membrane receptors (sensors). The perceived signals are transmitted downstream to trigger transcription, which is regulated by phytohormones, transcription factor binding proteins (TFBPs), *cis*-acting elements and miRNAs. Based on the biological functions, the role of these transcriptional regulators and the regulated genes that encode functional proteins or other products to protect plant cells directly from damage is well described [15].

The phytohormone—abscisic acid—acts as a central regulator in the response and adaptation of plants to drought conditions. The various physiological reactions regulated by ABA, including stomatal closure, accumulation of osmoprotectants, changes in gene expression, and other phytohormones have been characterized at the molecular level [16]. The molecular mechanisms of ABA synthesis, transport and signaling in relation to the plant's response to stress are also reasonably well described [17]. ABA signals are perceived by different cellular receptors. The nucleocytoplasmic receptors *PYR/PYL/RCARs* (pyrabactin resistance/pyrabactin resistance-like/regulatory component of ABA receptors) have been suggested to be the primary sensors that bind ABA and inhibit type 2C protein phosphatases (*PP2Cs*) [18]. Inactivation of *PP2Cs* leads to accumulation of active *sucrose non-fermenting-1* (*SNF1*)-related protein kinases (*SnRK2s*), which interacts with ABA-responsive TFs, *ABA-responsive promoter elements* (*ABREs*) and *ABRE*-binding protein/*ABRE*-binding factors (*AREB/ABF*) to regulate transcription of downstream target genes and related physiological processes [19]. Drought also induces changes in calcium ion levels, which activates calcium-dependent protein kinases (*CDPKs*) via calmodulin-like domain. The activated *CDPKs* regulate downstream components of calcium signaling. For instance, *OsCPK4* overexpressing rice plants exhibit increased water-holding capacity under drought or salt stress [20]. Genetic manipulation of *RLK* genes, including *OsSIK1* that acts as a positive regulator of drought stress responses, is also well reported [21]. Other secondary signaling molecules, including phosphatases (serine/threonine phosphatases) and phospholipids such as phosphoinositides, nitric oxide, cAMP and sugars, play an important role in signal transduction [22]. Examples of phosphatases include the wheat phosphatase *TaPP2Ac-1* that exhibited less wilting under water-deficit conditions than non-transformed controls [23].

Numerous TF families such as myeloblastosis oncogene (*MYB*), dehydration-responsive element binding proteins (*DREB*), basic leucine zipper domain (*bZIP*), *WRKYs* and the *NAC* (*NAM*, *ATAF* and *CUC*) are directly or indirectly regulated by endogenous ABA signaling during drought stress [24]. Many *MYB* genes involved in plant response to drought stress are functionally characterized, including *AtMYB15*, which was shown to enhance drought tolerance, and sensitivity to ABA [25]. *WRKY* proteins, including ABA-inducible *OsWRKY45*, *OsWRKY11* and *OsWRKY08,* are upregulated by drought stress [26]. *AP2/ERF* family is another large group of plant-specific TFs that have been demonstrated to be effective in enhancing drought tolerance in plants. For instance, overexpression of *AP2/ERF* genes, e.g., *GmERF3* in soybeans, has been reported to enhance tolerance to drought [27]. In addition, *DREB2s*, e.g.,

ZmDREB2.7, are candidates for drought stress tolerance in maize [28]. The *bZIP* TFs have also been reported to enhance plant tolerance to stress and hormone signal transduction, e.g., *OsbZIP23* in rice [29] and *ZmbZIP72* in maize [30]. Within the *NAC* family, *RD26* (*responsive to dehydration 26*) was the first *NAC* gene identified as a regulator in mediating crosstalk between abscisic acid and jasmonic acid (JA) signaling during drought stress responses in *Arabidopsis* [31]. Overexpression of other *NAC* genes, including *ANAC019*, *ANAC055* and *ANAC072*, has been shown to confer drought tolerance in transgenic *Arabidopsis* [32]. Similarly, overexpression of *SNAC1*, *OsNAC10* and *OsNAC5* driven by a root-specific promoter *RCc3* confers increased drought resistance under field conditions [33, 34]. The nuclear factor Y (*NF-Y*) TFs are emerging as important regulators of drought-stress response, particularly with respect to ABA biosynthesis. For instance, ectopic expression of *Amaranthus hypochondriacus NF-YC* gene (*AhNF-YC*) in *Arabidopsis* and overexpression of Bermuda grass *Cdt-NF-YC1* in rice has shown that these genes confer drought tolerance [35, 36]. *Cdt-NF-YC1* induces expression of both ABA-responsive genes (e.g., *OsRAB16A*, *OsLEA3*, *OsP5CS1* and *OsLIP9*) and signaling genes (e.g., *OsABI2* and *OsNCED3*), as well as, ABA independent genes (e.g., *OsDREB1A*, *OsDREB2B* and *OsDREB1B*). In fact there is an increasing evidence that some NAC genes, e.g., *SNAC3*, contribute to drought resistance and osmotic adjustment independent of ABA [37]. *SNAC3* interacts with *phosphoglycerate mutase*, *cytochrome P450 72A1*, *PP2C*, WD domain-containing protein and *oxidoreductase* to modulate ROS in rice. These findings suggest a complex regulatory mechanisns of drought response and tolerance in plants, involving both ABA and other signaling pathways.

Recent work on inhibitors of phosphoinositide-dependent phospholipases C (*PI-PLCs*) in *Arabidopsis* has also provided considerable insight into the drought-stress-related lipid signaling by identifying links of phosphoinositides to the *DREB2* pathway [38]. Moreover, overexpression of phosphatidylinositol synthase gene *(ZmPIS)* in tobacco plants changed membrane lipids' composition and improved drought stress tolerance [39]. The best characterized lipid derivative, so far, is inositol 1,4,5-trisphosphate (IP_3). IP_3 levels have been shown to increase in response to exogenous ABA in *Vicia faba* guard cell protoplasts and in *Arabidopsis* seedlings, for review see [40]. IP_3 acts as a second messenger involved in releasing Ca^{2+} from internal stores such as vacuoles. This pathway has been reported to induce osmotic-stress-responsive genes, as well as ABA stress-responsive genes [40]. Another lipid derivative, phospholipase D (PLD), has been reported by the same authors to be functionally associated with ABA; and the application of phosphatidic acid (PA), a PLD derivative, has been shown to mimic the effect of ABA in inducing stomatal closure [41]. This could probably suggest that lipid signaling is linked to ABA in drought stress response, and it is worthwhile to study how the different lipid derivatives enter in action, either simultaneously or timely synchronized with ABA.

Downstream of the TFs are numerous responsive genes that function either in a constitutive manner (i.e., also expressed under well-watered conditions) or a drought-responsive manner (i.e., expressed only under pronounced water shortage). Amongst them, genes encoding for receptor-like kinases (RLKs) with Ser/Thr kinase domain could play an important role in optimizing plant responses to drought stress [18]. Other genes that have been shown to be up-

or downregulated by drought stress to enable dehydration avoidance or tolerance in various plant species are documented in several studies [18, 42]. Another process, downstream of transcriptional regulatory networks, is the induction of a large range of genes encoding for enzymes involved in osmotic adjustments, osmoprotection, wax biosynthesis and changes in fatty acid composition (Figure 3). Adjustment of osmotic pressure allows the plant to take up more soil water and maintain turgor and cell function for a longer time under drought.
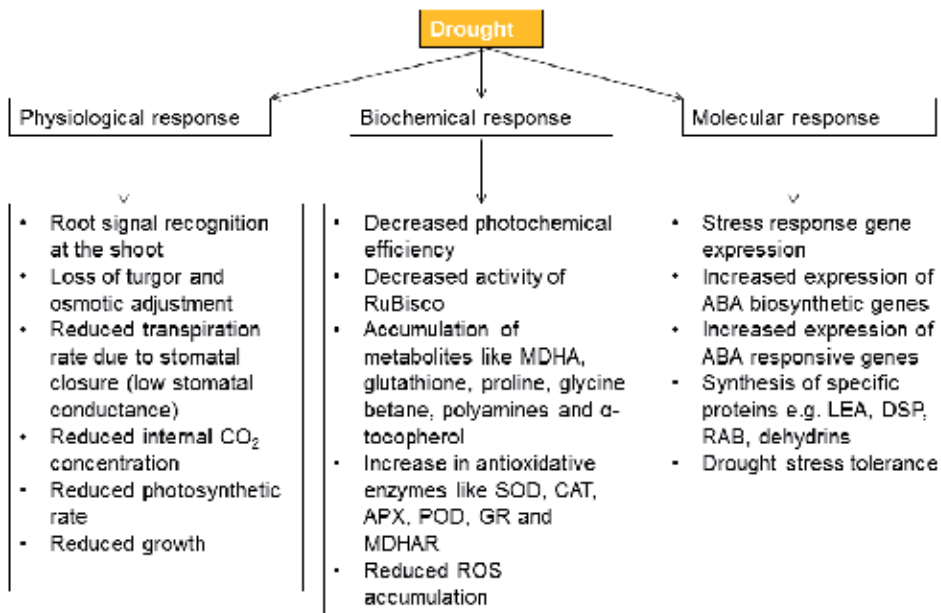
Figure 3. Physiological, biochemical, and molecular basis of drought stress tolerance in plants. Both major and minor changes that occur downstream of the transcriptional regulatory network are shown, although some of them, e.g. pro-line, glycine betaine and other amino acids, were previously shown not to be important in plant resistance to drought stress.

Water-channel proteins, e.g. aquaporins (*AQPs*), and sugar transporters are believed to facilitate the adjustment of osmotic pressure under stress by transporting water and sugars to the cytosol [42]. More recently, *AQPs* encoding genes (e.g., *MaPIP1;1*) were shown to be strongly induced in banana plants exposed to drought [43]. The same authors indicate that overexpression of *MaPIP1;1* in *Arabidopsis* exhibited better growth, reduced water loss and higher survival rates. Li et al. [44] also showed that *AQPs* were elevated under drought stress in Tibetan *Sophora moorcroftiana*, which is consistent with the previous reports [45]. However, the same authors indicate conflicting functions of plasma membrane intrinsic proteins (*PIPs*). For instance, overexpression of *GoPIP1*, cloned from *Galega orientalis*, showed increased sensitivity to drought in transgenic *Arabidopsis* plants. This indicates that AQPs are able to facilitate both tolerance and sensitivity, which warrants further research to delineate *AQPs* that are potentially helpful in improving drought tolerance in plants.

Studies have also shown that the *K* $^+$ *uptake transporter 6* (*KUP6*) subfamily of transporters act as key factors in osmotic adjustment by balancing potassium homeostasis in cell growth and drought stress responses [46]. *KUP6* is apparently under the control of abscisic acid and interacts with ABA-activated *SnRK2*-type protein kinase, *SnRK2E*, resulting in phosphorylation of the *KUP6* C-terminal domain. This indicates that *KUP6* is a downstream target for *SnRK2E* in the control of water stress responses. However, other interacting proteins, and probably hormones, e.g. auxins, could regulate the activity of *KUP6* in the maintenance of water status during drought stress. Indeed, it was reported previously that a variant of *KUP6*, *KUP4/TRH1*, facilitates root-specific auxin distribution [47]. This was substantiated by the findings that triple mutants of the *KUP* genes (i.e., *kup2 kup6 kup8* and *kup6 kup8 gork*) showed enhanced cell expansion and auxin responses in lateral root formation [54]. Moreover, auxin-responsive TFs, *LBD18* and *LBD29*, were highly expressed in the triple mutants in the presence of IAA, indicating that auxin could be modulating K$^+$ and proton fluxes during drought stress.

The biosynthesis of osmoprotectants such as amino acid, amines and carbohydrates is another indispensable strategy for plant resistance to drought stress. The most common osmoprotectants are proline (Pro), γ-aminobutyric acid (GABA), glycine betaine (GB), fructans, starch, mono- and disaccharides, trehalose (Tre) and raffinose family oligosaccharides (RFO). The biosynthesis and transport of trehalose and raffinose is particularly relevant in drought stress response. More recently, genes encoding for trehalose and raffinose biosynthesis were significantly upregulated in the roots and leaves of *Jatropha curcas* under drought [48], suggesting that these compounds may have major impacts on osmotic adjustment and ROS scavenging during drought stress. The same authors indicated that dozens of genes involved in wax biosynthesis, including *KCS* and *WSD*, and their regulators (e.g., *MYB96, CER*) were upregulated more than four-fold in leaves under drought conditions. Overexpression of genes encoding for *MYB96, CER KCS* and *WSD* could probably strengthen the hydrophobic barrier that prevents non-stomatal water loss and increase plant tolerance to drought.

Genes encoding for proteins involved in cellular structure stabilization have also been reported to be induced in plant tolerance to drought. For instance, dehydrins (DHNs) function to protect cells from damage caused by drought stress-induced dehydration [49]. Proteins related to lignin biosynthesis, such as caffeoyl-CoA 3-O-methyl-transferases and class III plant peroxidases, were also found to be induced by drought in wild watermelon [50] and in maize roots [51]. In winter triticale, water-deficit-induced leaf rolling was correlated with a higher level of cell wall-bound phenolics in the leaves [52]. These adaptive mechanism could probably limit water loss by restricting the leaf transpiration surface. In addition, carbon/nitrogen-metabolism-related proteins have been reported to be more abundant in roots of soybean [53], wild watermelon [50] and rapeseed [54] after drought treatment, suggesting an increased energy demand as well as enhanced cellular activities in the root tissues during drought stress. The same authors reported a relative increase in the root growth rate and abundance of root-growth-related small G-protein family members such as *Ran GTPases*, which suggests increased membrane trafficking activity in an effort by the plant roots to absorb water from deep soil layers.

Taken together, the vast amount of data from 'omic' tools provide a basis for identification of more functional genes, which could contribute directly to cellular drought stress tolerance. In addition, understanding expression networks of genes encoding for the aforementioned proteins, especially genes involved in cellular structure stabilization, molecular chaperones, enzymes for detoxification of reactive oxygen species, and those for the biosynthesis of sugars, wax and dehydrins, which are important as protectants [55], may allow for the realization of significant genetic gains in breeding for plant tolerance to drought. Further genomic scale investigations will enable understanding of transcriptional regulators behind co-expressed genes and their association with particular genomic regions (QTLs). Although QTL identification for tracing drought tolerance remains a challenge due to the large number of genes influencing drought tolerance traits, continued investigation into the basis of tolerance in crops like *Jatropha curcas* will probably provide a clearer understanding of drought tolerance. Besides this, the mechanism by which drought tolerance associated protein networks effectively protect PSII and granal stability, as well as maintain photosynthetic competence will need further elucidation.

3. Advances in plant tolerance to heat stress

Temperatures above the normal optimum cause heat stress (HS) at different levels in all living organisms. Heat stress disturbs cellular homeostasis, and causes denaturation and dysfunction in many proteins, leading to severe retardation in growth, development and even death. In plants, the major sites of heat stress injury are the oxygen-evolving complex (OEC) along with associated biochemical reactions in photosystem II (PSII). Ultimately, efficiency of electron transport is reduced or altered affecting electron flow from OEC towards the acceptor side of PSII. These alterations affect the generation of ATP and the regeneration of Rubisco for carbon fixation [56]. Starch synthesis is also negatively affected by heat stress because of the reduced activity of enzymes such as invertase, sucrose phosphate synthase and ADP glucose pyrophosphorylase. Usually, ROS induction and accumulation in the chloroplasts precedes these changes. Accumulated ROS can severely damage DNA and cause autocatalytic peroxidation of membrane lipids and pigments, altering membrane functions and cell semi-permeability. Physiological changes associated with biochemical damage may include a decrease in chlorophyll a:b ratio, inhibitions of stomatal conductance and net photosynthesis, and low plant water potential. These changes ultimately reduce the partitioning of photosynthates, which morphologically manifest as retarded growth, reduced economic yield and harvest index. Scorching and sunburns of leaves and twigs, branches and stems, leaf senescence and abscission, and fruit discoloration and damage are other morphological damages associated with heat stress [57].

Perception of heat stress by plants usually triggers sensors at the plasma membrane and causes a transient opening of Ca^{2+} channels, possibly via modulation of membrane fluidity (Figure 4) [58]. Upon entry of Ca^{2+}, putatively through channels possessing cytosolic C-terminus with a calmodulin-binding domain, multiple kinases are activated.

Figure 4. Hypothetical model for high-temperature signal sensing and induction of molecular pathways leading to plant defence response. Prolonged high-temperature stress causes membrane depolarization leading to Ca^{2+} influx or directly activates apoplastic enzymes including GLPs. Increased levels of cytosolic calcium activate the ROS-producing enzyme, RBOHD, which catalyses ROS production. Effect of temperature on R genes through an unknown pathway is likely to further enhance ROS production. ROS/ Ca^{2+} signaling causes activation of plasma membrane ATPase, which extrude H^+. Alternatively, heat-stress-induced protein damage and protease activity decreases cytosolic pH. Low cytosolic pH and H_2O_2 accumulation reduces CO_2 assimilation, thereby increasing endogenous carbohydrate metabolism. Cytosolic acidification and ATPase activity may also increase accumulation of expansins and methylesterases that eventually affect the cell wall integrity. Activating plasma membrane ATPase is probably reverse phosphorylated by FKBP65 leading to H^+ extrusion and K^+ intrusion. A part from its targeted role in the nucleus, FKBP65 could be targeted to the chloroplast through the tat pathway to activate photosystem II 10 kDa polypeptides or for directing chaperone functions. Activated HSPs probably cause chromatin remodelling and histone displacement. In addition to activating PM ion channels, heat-induced changes in membrane fluidity triggers lipid signaling. Plants deploy phospholipids, including phospholipase D (PLD), PIPK (phosphatidylinositol 4,5-bisphosphate kinase), phosphatidic acid (PA), PIP2 (phosphatidylinositol phosphate kinase) and IP3 (D-myo-inositol-1,4,5-trisphosphate) to specific intracellular locations. The accumulation of lipid signaling molecules also triggers Ca^{2+} influx, which initiates downstream signaling, including activation of CDPKs, hormonal changes, transcription factor activation and secondary metabolism. Question marks indicate the unknown players.

The *MPK6* activity has been particularly shown to increase under heat stress. *MPK6* activates a vacuolar processing enzyme (*VPE*), which has been suggested to play a role in HS-induced programmed cell death [59]. Transcriptional regulators, such as *HSFs*, *WRKY*, *Zat* and *MBF1c*, a transcriptional regulator of *DREB* genes [60], are activated to regulate expression of HSPs and other heat stress response genes.

Heat-induced accumulation of Ca^{2+} in the cytoplasm also activates the ROS-producing enzymes *RBOHD* and *NADPH oxidase*, by direct interaction or through activation of calcium-dependent protein kinases (*CDPK*) that phosphorylate *RBOHD* [61]. When activated, *RBOHD*

catalyzes the production of ROS, causing membrane depolarization and/or initiation of ROS/ redox signaling network, which interacts with the above-mentioned *MBF1c*, *HSFs*, *MAPKs* and *SnRK*s to trigger downstream signaling networks [61].

Calcium/calmodulin-binding protein kinases (*CBK*), which also regulate the expression of *HSPs*, are activated via *CaM3*. A well-known example is the activation of *CBK3*, which enhances thermotolerance in *A. thaliana* seedlings by phosphorylating *HsfA1a* and a *CaM* protein phosphatase (*PP7*) [62]. *PP7* interacts with both *AtCaM3* and *AtHsfA1a*. *AtCaM3* increases thermotolerance by activating *WRKY39* and *HSFs*, indicating that *CBK3* plays a key role in heat stress signaling. The TF *Zat* is necessary for the activation of *WRKYs* and ascorbate peroxidase [63]. *MBF1c* modulates the induction of SA and trehalose, which are regulators of plant stress response [64]. SA has been shown to alleviate heat stress by increasing proline production and restricting the formation of ethylene in heat-stressed plants [65].

Another HS-response-associated signaling pathway was shown in the *Hsp90–ROF1* interaction in the cytoplasm and their subsequent translocation to the nucleus. The *Hsp90–ROF1* complex localizes in the nucleus only in the presence of *HsfA2* [66]. The interaction of these three proteins modulates *HSP* gene expression under HS. Although, *ROF1* has been reported to induce expression of small *HSPs*, which increases plant survival rate under HS, to date the upstream signal that regulates the subcellular localization of *Hsp90–ROF1* remains elusive. Interestingly, just like *MBF1c, ROF1* is involved in calcium-dependent phosphorylation of *HSFs*, which suggests that Ca^{2+}-dependent activation of *RBOHD* or *CDPK*s could be the upstream signal for *ROF1*.

Heat stress also triggers lipid signaling. Activation of *phospholipase D* (*PLD*) and a *phosphatidylinositol 4, 5-bisphosphate kinase* (*PIPK*) increases the accumulation of phosphatidic acid (*PA*), phosphatidylinositol phosphate kinase and D-myo-inositol-1,4,5-trisphosphate (*IP3*); and an active cycling of a G protein appears necessary in this process. The accumulation of lipid signaling molecules could in turn cause the opening of channels and the triggering of Ca^{2+} influx [67].

Downstream effects of heat stress signals have been reported to activate a signaling pathway called unfolded protein response (UPR) in the endoplasmic reticulum, which requires specific calcium signals from the plasma membrane [58]. Within the endoplasmic reticulum, the activity of UPRs involves two signaling pathways: one involving proteolytic processing of membrane-associated *bZIP* TFs and the other involving RNA splicing factor, inositol requiring enzyme-1 (*IRE1*) and its mRNA target [68]. *IRE1* is a dual functional enzyme possessing both serine/threonine protein kinase and endoribonuclease activity. In *Arabidopsis*, heat signals activate *IRE1* to splice *bZIP60* mRNA in the cytosol, causing a frameshift, which triggers the synthesis of a tissue factor without a transmembrane domain, but having a nuclear targeting signal [69]. The *bZIP60* (*bZIP60*(s)) spliced forms activate UPR target genes in the nucleus. A cytosolic form of UPR, which is triggered by the presence of unfolded proteins in the cytosol, was also previously reported [70]. Together, these UPRs are associated with the heat shock promoter elements and the involvement of specific *HSFs*, notably *HSFA2*, regulated by alternative splicing and non-sense-mediated decay. Under severe HS (42–45ºC), a novel post-transcriptional regulatory mechanism governing *HSFA2* expression has also been shown to

occur. Moreover, a new splice variant of *HSFA2-III* is reported to be generated through the use of acryptic 5′ splice site in the intron. *HSFA2-III* can be translated into the small *HSFA2* (*S-HsfA2*), which binds to the TATA box proximal clusters of HS elements (*HSE*) in the *HSFA2* promoter to activate its own gene expression, thus constituting a positive auto-regulatory loop [71]. Although the TFs interacting with *S-HsfA2* are yet to be validated, this finding suggests that severe HS may alter the splicing pattern of *Hsf* genes, generating isoforms that may auto-activate self-expression and consequently rapidly induce the expression of HSPs required for enhanced response to HS.

Apart from *HSFs*, overexpression of *DPB3-1*, which regulates expression of *DREB2A* and *DREB2B*, increases thermotolerance [72]. Other studies have also shown the role of *bZIP28* [73] and *WRKY* proteins in plants thermotolerance [74, 75]. Furthermore, the basic helix-loop-helix (*bHLH*) TF, *phytochrome interacting factor 4* (*PIF4*), was reported to control acclimation to changes in ambient temperature by regulating important hormonal and developmental pathways modulating the acclimation mechanisms [76]. *PIF4* alleles control floral timing by modulating *FLOWERING LOCUS T (FT)*. *PIF4* also controls early inflorescence internode elongation and high-temperature-induced hypocotyl elongation by modulating levels of free indole-3-acetic acid (*IAA*) through the triggering of *YUC8 (YUCCA8)* or *TRYPTOPHAN AMINOTRANSFERASE OF ARABIDOPSIS* (*TAA1*) gene expression [57, 77]. Thus, *PIF4* is a potential regulator of plant responses to high temperature. However, its physical interaction with *cryptochrome 1* (*CRY1*) on nuclear DNA suggests that these two proteins co-regulate temperature responses in plants. Another regulator, *E3 ubiquitin ligase CONSTITUTIVE PHOTOMORPHOGENIC1* (*COP1*), was shown to be essential for plant responses to HS [77]. However, it is not known whether *COP1* signaling is independent of *PIF4*. Orthologs of *PIF4* have been identified in several crop species. Thus, if the interaction with other associated proteins is resolved, *PIF4* has a potential promise to improve plant tolerance to HS in several crops through genetic engineering.

Other components of heat sensing that could be linked to these signaling pathways include the transcriptional modulator, the *nuclear actin-related protein 6* (*ARP6*), which is part of the Snf-2-related CREB-binding activator protein (*SRCAP*) encoding a subunit of the *SWR1* chromatin remodelling complex is necessary for inserting the alternative histone, *H2A.Z*, into nucleosomes, while replacing the core histone H2A [78]. Heat stress induces a decrease in *H2A.Z* occupancy in nucleosomes located at the transcription start site of heat response genes, a process that probably allows nucleosome opening and enhanced transcription of these genes.

Plant adaptation to thermotolerance also involves the activity of superoxide reductase (*SOR*), *S*-nitrosoglutathione reductase (*GSNOR*) and rubisco activase (*RCA*). The functions of these proteins are reasonably well described in a review by [67]. Other commonly reported anti-oxidant enzymes produced by plants under HS include superoxide dismutase (*SOD*), catalase (*CAT*), guaiacol peroxidase (*GPX*), ascorbate peroxidase (*APX*), dehydroascorbate reductase (*DHAR*), glutathione reductase (*GR*), glutathione S-transferase (*GST*) and non-enzymatic anti-oxidants such as flavanoids, anthocyanin, carotenoids and ascorbic acid (*AA*) [60]. The accumulation of other osmolytes such as glycine betaine and trehalose is another well-known adaptive mechanism in plants against HS. Generally, most of these compounds are involved

in ROS removal (anti-oxidants), osmotic adjustment, saturation of membrane-associated lipids, protection of photosynthetic reactions, production of polyamines and protein biosynthesis [94], which enable plants to exhibit basal or acquired thermotolerance. Proline and glycine betaine application considerably reduce the H_2O_2 production, improve the accumulation of soluble sugars and protect the developing tissues from HS [79]. Tocopherol is another important lipid-soluble redox buffer and an important scavenger of singlet oxygen species and other ROS. Moreover α-tocopherol has the highest anti-oxidant activity of all the tocopherol types reported in plants [80].

A number of studies have demonstrated the presence of QTLs associated with most HS-related traits and promise to the use of molecular markers in breeding for heat stress tolerance. More than 50 QTLs have been identified in various crops so far, including maize, wheat, rice, cowpea, lettuce, *Medicago truncatula* and *Brassica napus*. Recent studies in transcriptomics [81, 82], proteomics [83, 84], metabolomics [85, 86] and microRNAs [87] have also provided additional information on the mechanisms controlling plant responses to HS. Understanding the relationship between these mechanisms and the genomic regions mapped and delineated as QTLs would validate the genes controlling plant responses to HS, and subsequently improve genetic gains in plant improvement programmes. Besides, the possibility of developing transgenic plants with enhanced tolerance to HS would also gain significance. This approach has already been demonstrated in cotton [88], *Arabidopsis* [89], tobacco [90] and rice [91], but needs further validation, especially in economically important crops where it has not been applied before. Taken together, heat stress responses discussed here demonstrate that heat stress is a quantitative trait, which requires a combination of several disciplines to improve plant tolerance.

4. Advances in plant tolerance to cold stress

Cold stress occurs at temperatures less than 20ºC and varies with the degree of temperature duration and plant type. Chilling (<20ºC) or freezing (<0ºC) temperatures can trigger the formation of ice in plant tissues, which causes cellular dehydration [92]. Ultimately, cold stress reduces plasma membrane (PM) integrity, causing leakage of intracellular solutes. Cold stress severely affects plant growth and survival, and leads to substantial crop losses in temperate climatic regions and hilly areas of the tropics and subtropics [93]. In rice, for instance, losses due to cold stress can range from 0.5 to 2.5 t/ha and grain yields can drop by up to 26%, especially when low temperatures occur during the reproductive stage [94].

To cope with this adverse condition, plants adapt several strategies such as producing more energy by activation of primary metabolisms, raising the level of anti-oxidants and chaperones, and maintaining osmotic balance by altering cell membrane structure [95]. These mechanisms of plant response to cold stress are closely similar to that of heat stress. However, the difference lies in the fact that membrane rigidification occurs in cold stress as opposed to heat stress. Thus, membrane rigidification is the upstream trigger for the induction cytosolic Ca^{2+} signatures leading to a transient increase in cytosolic Ca^{2+} levels [96]. It is assumed that dimethyl

sulfoxide (DMSO) mediates the perception of membrane rigidification by mechanosensitive Ca^{2+}channels [97]. Other upstream factors such as changes in the metabolic reactions and metabolite concentrations, protein and nucleic acid conformation could contribute to enhance perception of cold stress. These factors as well, either directly or indirectly, induce an increase in cytosolic Ca^{2+}, which is a well-known upstream second messenger, regulating cold regulated (*COR*) gene expression.

Cold-stress-induced cytosolic Ca^{2+} signals can be decoded by different pathways. More recently, Ca^{2+} signal was reported to be transduced directly into the nucleus. The concentration of nuclear Ca^{2+} is monitored by a chimera protein formed by the fusion of aequorin to nucleoaoplasmin, which is also transiently increased after cold shock [95]. Aequorin possesses several EF-hand-type binding sites for Ca^{2+} ions. The binding of Ca^{2+} to these sites causes a conformational change in aequorin which enables the monitoring of Ca^{2+} concentration. It has been reported that nuclear Ca^{2+} concentration peaks at about 5–10s later than the cytosolic Ca^{2+} [95]. The same authors have reported that nuclear Ca^{2+} signal may be initiated from the nuclear envelope and is assumed to be propagated by cytosolic Ca^{2+} transients in plants.

In the cytoplasm, a range of Ca^{2+} sensors have been reported, including calmodulin (*CaM*), CaM-like (*CMLs*), Ca^{2+}-dependent protein kinases (*CDPKs*), Ca^{2+}-and Ca^{2+}/CaM-dependent protein kinase (*CCaMK*), CaM-binding transcription activator (*CAMTA*), calcineurin B-like proteins (*CBLs*) and *CBL*-interacting protein kinases (*CIPKs*) [98]. Some of the sensors work as negative regulators of cold tolerance in plants, e.g., calmodulin3, a SOS3-like or a CBL calcium-binding protein and a protein phosphatase 2C (*AtPP2CA*). The positive regulators, e.g., CDPKs and probably some *CBLs*, relay the Ca^{2+} signal by interacting with and regulating the family of *CIPKs*. For instance, *CBL1* has been shown to regulate cold response by interacting with *CIPK7* [99], whereas *CAMTA3* has been identified as a positive regulator of *CBF2/DREB1C* through binding to a regulatory element (*CG-1, vCGCGb*) in its promoter [100]. Although *CBF2/DREB1C* was previously reported to negatively regulate *CBF1/DREB1B* and *CBF3/DREB1A*, its expression appears to be necessary for integrating cold-inducible calcium signaling with gene expression, but under transient and tight control to avoid repression of freezing tolerance. Both *CBF1/DREB1B* and *CBF3/DREB1A* are required for constitutive expression of cold-inducible genes in *Arabidopsis*, and play an important role in cold acclimation (see discussion below).

Ca^{2+} influx into the cytoplasm also apparently activates phospholipase C (PLC) and D (PLD), which are precursors for IP_3 and PA, respectively. IP_3 activates IP_3-gated Ca^{2+} channels that can amplify Ca^{2+} signatures in the cytoplasm, leading to higher induction of *COR* genes and CBFs, for review see [101].

There are some reports that the chloroplast may also play a role in sensing low temperature [98]. Cold stress is considered to cause excess photosystem II (PSII) excitation pressure, as a result of the imbalance between the capacity for harvesting light energy and the capacity to consume this energy on metabolic activity in the leaves, which probably leads to ROS generation. The damaging effect of ROS on the photosynthetic apparatus presumably leads to photoinhibition, which occurs even under relatively low irradiance [102] and is apparently a mechanism of cold acclimation or freezing tolerance. ROS also acts as the second messenger

and may reprogramme transcriptome changes through induction of Ca^{2+} signatures and activation of MAPKs and redox-responsive TFs. The MAPK cascades in *Arabidopsis* , including AtMEKK1/ANP1 (MAPKKK)–AtMKK2 (MAPKK)–AtMPK4/6 (MAPK), positively regulate cold acclimation in plants [103].

The downstream signals that promote the production of *COR* proteins and cold response to metabolites are reasonably discussed in references [95, 104]. Specific examples include the upregulation of the TFs, *CBF/DREB1s* (CRT (C-repeat)/DRE binding proteins) [103], which initiate the transcription process. The *CBF/DREB1* (mainly *CBF3/DREB1A*) pathway is controlled by a myelocytomatosis oncogene (*MYC*)-type TF, inducer of CBF expression1 (*ICE1*), which binds to the *MYC* recognition cis-elements (*CANNTG*) in the promoter of *CBF3/DREB1A*, and induces the expression of *CBF3/DREB1A* and its regulons during cold acclimation [105]. The function of *ICE1* in cold response is conserved; and overexpression of *Arabidopsis ICE1* improves chilling tolerance and enhances the accumulation of soluble sugars and proline concentration in cucumber [106]. In rice, *OsICE1* and *OsICE2* are induced by cold stress and sequentially upregulate *OsDREB1B*, rice heat shock factor A3 (*OsHsfA3*) and rice trehalose 6-phosphate phosphatase (*OsTPP1*). The *CBF/DREB1s* can bind to *CRT/DRE* cis-elements, A/GCCGAC, in the promoter of *COR* genes to regulate expression of *COR* genes [107]. Moreover, *CBF/DREB1* genes are organized in tandem (*CBF1/DREB1B-CBF3/DREB1A-CBF2/DREB1C*) on *Arabidopsis* chromosome IV and have been reported to be induced at the same time, suggesting that combining these TFs in one genotype could probably improve cold tolerance. However, the inconsistent target specificity amongst the three *CBF* factors in *CBF/DREB1*-overexpressing transgenic plants reveals variability in their roles [108]. Indeed, *CBF2/DREB1C* has been shown to be a negative regulator of both *CBF1/DREB1B* and *CBF3/DREB1A* [109], while *CBF1/DREB1B* and *CBF3/DREB1A* act as positive regulators of cold acclimation by activating the same subset of *CBF/DREB1*-target genes [110]. *CBF1/DREB1B* and *CBF3/DREB1A* are therefore concertedly required to induce the whole *CBF/DREB1*-regulon to complete the development of cold acclimation, while the expression of *CBF2/DREB1C* is tightly controlled to avoid its negative modulation of *CBF1/DREB1B* and *CBF3/DREB1A*. The exact mechanism by which this happens is unknown.

Downstream of these TFs are *COR* genes, which are mainly linked to the onset of tolerance mechanisms and ultimately lead to acclimation. Genes encoding for annexin; hyper-sensitive-induced response (HIR) protein families (e.g., prohibitins and stomatins); dehydrins (e.g., 25 kDa dehydrin-like protein, *ERD14*, and *cold acclimation-specific protein 15* (*CAS15*)); antioxidants (e.g., superoxide dismutase, catalase and ascorbate peroxidase); *HSPs* (e.g., *HSP70* family being the most abundant); defence-related proteins such as protein disulfide isomerase; disease resistance response proteins, peptidylprolyl isomerase *Cyp2* and cysteine proteinase; amino acids, polyamines and polyols; and cellulose synthesis, such as UDP-glucose pyrophosphorylase, are commonly reported in expression studies [111]. Several metabolism-associated proteins, including carbohydrate metabolism enzymes, such as phosphogluconate dehydrogenase, NADP-specific isocitrate dehydrogenase, fructokinase, cytoplasmic malate dehydrogenase, pyruvate orthophosphate dikinase precursors (PPDK), aconitate hydratase, glycine dehydrogenase and enolase, have also been reported to be activated during cold stress

[112]. Thus, several genes and the corresponding proteins are associated with the regulation of the metabolic pathways operating under cold stress.

However, identification of functional polymorphism in these genes remains a daunting task. A similar challenge is observed in the QTLs identified, so far, in various crops, including maize, barley, rice, wheat, sorghum and many other economically important crops. Identification of effective cold sensors also remains elusive, as multiple primary sensors are thought to be involved in sensing low temperatures. Thus, a comprehensive understanding of the defence mechanism from sensors, cold signaling, to the defence response will require further research on both upstream and downstream regulations of *ICE1-CBF/DREB1*-dependent pathway, as well as proteins that may be functioning independent of this pathway.

5. Advances in plant tolerance to salinity

Salinity is increasingly becoming a major threat to crop production, particularly due to inappropriate irrigation regimes and increasing use of brackish water for irrigation. As much as 6% of the total world land is subjected to salinity [113], and more than 20% of irrigated land is affected by salinity [114]. Moreover, major reductions in cultivated land area, crop productivity and quality that have been reported in the recent past are due to salt-induced stress [115]. Climate-change-associated rise in sea levels and coastal floods are expected to further contribute to this phenomenon in the future.

Salt stress in plants occurs when electrical conductivity of saturated soil paste extract (ECe) reaches 4.0 deci-Siemens per meter (dS/m; approximately 40 mM NaCl). The minimum level may, however, vary from crop to crop. For instance, the salinity threshold for rice is 3.0 dS/m [163]. Beyond this threshold, a yield reduction of 12% per dS/m has been reported to occur. When plants gradually accumulate salts, osmotic stress, nutrient imbalance and oxidative stress occur [116]. These salt effects disrupt intracellular ion homeostasis, membrane function and metabolic activity [117]. As secondary effects, salt-induced osmotic stress decreases root epidermal cell division and elongation rates, reducing primary root growth, eventually resulting in inhibition of growth and reduction of crop yields [118].

Alkalinity stress is a heightened version of salinity stress which has been reported to be much harsher than equimolar salinity, especially at neutral pH [119]. Although it is fairly understood that alkalinity causes osmotic challenge and ionic stress, and precipitates nutrients such as metallic micronutrients and phosphates, and disrupts the integrity of root cellular structure, the molecular signals and adaptive mechanisms are not well understood. Because many saline soils are also alkaline due to the presence of sodium (Na) carbonates, in this section we will exclusively focus on salinity, which is wide spread, and has been extensively researched and discussed in several studies.

To cope with saline soils, plants deploy a range of mechanisms that range from exclusion of Na^+ from the cells to tolerance within the cells. When plants are subjected to salinity, a series of responses ranging from genetic molecular expression through biochemical metabolism to physiological processes occur (Figure 5).
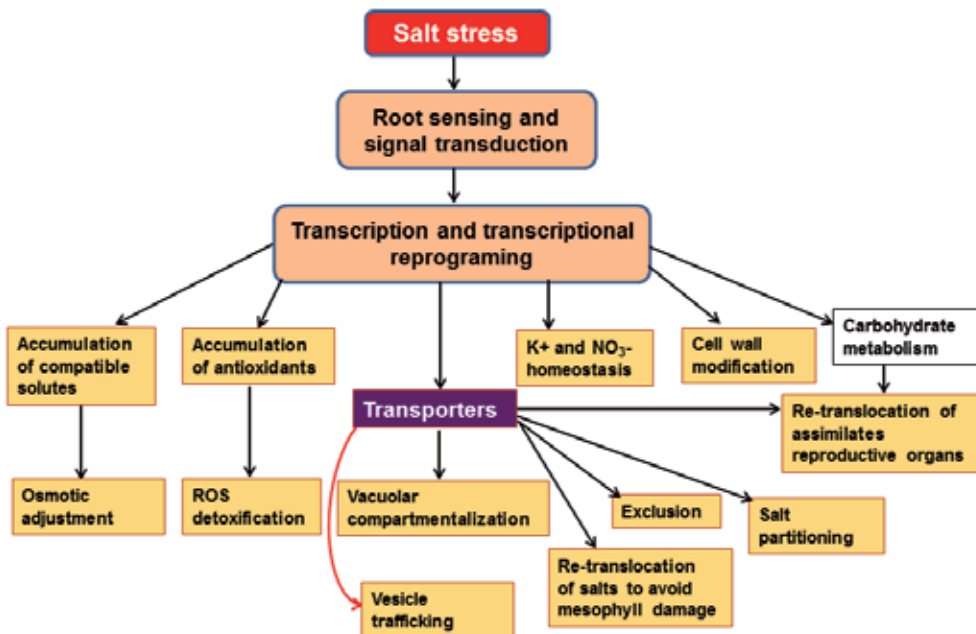
Figure 5. Adaptive mechanisms of salt tolerance. Cellular functions that would apply to all cells within the plant are the first adaptation mechanisms, followed by the functions of specific tissues or organs. Most of these functions are explained in the text (modified from [140].

Amongst the receptor proteins identified as the first detectors of salt stress are G-protein-coupled receptors, ion channel, receptor-like kinase or histidine kinase. These receptors transduce signals that generate secondary signals such as Ca^{2+}, inositol phosphates, ROS, nitric oxide (NO) and ABA. The signaling pathway associated with increased concentration of cytosolic Ca^{2+} is the most reported.

Cytosolic Ca^{2+} activates calcium-dependent protein kinases (*CDPKs*), calcineurin B-like proteins (*CBLs*) and CBL-interacting protein kinases (*CIPKs*) to transduce signals to downstream protein activity and gene transcription [120]. Transcription factors such as calmodulin-binding transcription activators (*CAMTAs*), GT element-binding-like proteins (*GTLs*) and MYBs have been reported to be activated by Ca^{2+}/calmodulin directly [121–123]. Other commonly expressed TFs in response to salt stress include the basic leucine zipper (*bZIP*), e.g., *OsbZIP71* in rice [124], *WRKY* [125], *APETALA2/ETHYLENE RESPONSE FACTOR* (*AP2/ERF*) [126], *MYB* [127], basic helix–loop–helix [128] and *NAC* [42] families. These TFs regulate the expression of genes related to water potential decrease, which results from osmotic stress caused by salinity.

Downstream of these TFs, there are several genes associated with salinity tolerance. The most reported are genes encoding for salt exclusion proteins, e.g., *SOS1*, cation:proton **antiporter family**1 of Na+/H+ anti-porters, salt compartmentalization genes, e.g., *vacuolar H + -pyrophosphatase* [129], and osmotic adjustment, e.g., *pyrroline-5-carboxylate synthetase* [130].

The salt overly sensitive (*SOS*) Ca^{2+} sensor regulatory mechanism is believed to be conserved in higher plants including monocots and dicots [131]. *SOS* consists of three functionally interlinked proteins, *SOS3/SCaBP8–SOS2–SOS1*. *SOS3* mainly functions in the roots, while *CBL10/SCaBP8*, an alternative regulator of *SOS2* that has been described as *SOS3*-like, primarily functions in the shoots. At high Na^+ concentrations, increased influx of Ca^{2+} is perceived by *SOS3* that encodes a myristoylated EF hand (a domain of five serially repeated helix–loop–helix calcium-binding motifs). Upon Ca^{2+} binding, a conformational change occurs and *SOS3* activates the downstream serine/threonine protein kinase, *SOS2*, and recruits it to the plasma membrane. Subsequently, the *SOS3–SOS2* complex stimulates the plasma membrane-localized Na^+/H^+ anti-porter (*SOS1*), leading to the extrusion of the excess Na^+ out of the cells [132]. Different from *SOS3, SOS3*-like proteins (*CBL10/SCaBP8*) are phosphorylated by their interacting protein kinases apparently regulating *CBL/SCaBP–CIPK/PKS* modules [133].

Besides extruding Na^+, the adaptive *SOS* module also links cytosolic Na^+ with Ca^{2+} binding proteins. The PM-localized *NHX7/SOS1* and the intracellular localized cation:proton **antiporter family**1 (*CPA1*) of Na^+/H^+ anti-porters (*NHX1-NHX4*; tonoplast-localized) are a ubiquitous family of transporters that mediate the exchange of K^+ or Na^+ for H^+ while regulating cytoplasmic salt overloads [134]. In the cytosol, increased influx of Ca^{2+} associated with excess Na^+ levels is perceived by Ca^{2+}-binding calmodulins/calmodulin-like proteins, which interact with *NHX1* transporters to sequester excess Na^+ in the vacuole. In *Arabidopsis*, a calmodulin-like protein, *AtCaM15*, regulates the tonoplast localized *AtNHX1* [135]. The interaction of *AtCaM15* with *AtNHX1* occurs in the vacuolar lumen and is dependent on Ca^{2+} and pH. The C-terminus of *AtNHX1* has been shown to localize in the cytosol, which might suggest that this strategic placement is targeted for phosphorylation by protein kinases or for sensing changes in cytosolic pH. However, the protein kinase targeting *AtNHX1* is unknown, and further studies on the interaction of this transporter with other proteins, especially protein kinases, will be necessary.

Interestingly, at moderate salt levels, the role of these transporters is less clear. Indeed, the *nhx1/nhx2* **double mutants** are not sensitive to moderate external Na^+ concentrations, yet they are sensitive to moderate external K^+ concentrations, for review see [134]. Conversely, the trans-Golgi network-localized *NHX* double knockouts, *nhx5/nhx6*, highly respond to moderate salinity and interfere with vesicle trafficking to the vacuole. This suggests that the endosomal *NHXs* are more sensitive to Na^+ accumulation than vacuolar *NHXs*. This difference has implications on Na^+ tolerance in plants. Recently, another *CPA* family member, a cation/H^+ exchanger (*CHX*), *GmSALT3*, was shown to improve shoot Na^+ exclusion and salt tolerance in soybean [136]. Fluorescent protein fusions suggested that *GmSALT3* and other *CHX* proteins are localized to the endoplasmic reticulum, further indicating that endosomal *NHXs* could be more reliable in sensing abnormal Na^+ levels in the cell and has a positive implication on salt tolerance in plants.

Other genes encoding for *Mannose-1-phosphate guanyl transferase* (*OsMPG1*) and **the rice** homologue of Shaker family K+ channel *KAT1* (*OsKAT1*) have also been reported to confer salinity tolerance [137, 138]. *OsMPG1* is an important enzyme for the biosynthesis of ascorbic acid in plants, whereas *OsKAT1* reduces the cellular Na^+ to K^+ ratio by increasing the cellular K^+ content. Another rice potassium transporter (*OsHAK5*) was shown to accumulate more K^+

and less Na^+ when constitutively expressed in *Nicotiana tabacum* cv. Bright Yellow 2 under salinity stress [198]. Several other genes were recently identified by Chen et al. [139] while studying the halophyte seashore Paspalum (*Paspalum vaginatum*).

Another process, downstream of transcriptional regulatory networks, involves accumulation of sufficient solutes (e.g., proline and glycine betaine) to balance extra osmotic pressure in the soil solution to maintain turgor [140]. Moreover, plants can also accumulate sufficient Na^+ and Cl^- to balance those in the soil solution, but this is tightly controlled through strict ionic regulation in various cell compartments ('tissue tolerance'). These tolerance strategies are achieved through a series of ion transporters and their localization in key cell types. Na^+/H^+ anti-porter proteins are the key regulators of these tolerance strategies. Examples include *TaHKT1;5-D* protein, which maintains high cytosolic K^+/Na^+ ratios in bread wheat shoots by restricting Na^+ loads in the root xylem before entering the shoot [141]. Recently, the introgression of the *Triticum monococcum HKT1;5-A* into durum wheat improved shoot Na^+ exclusion and improved grain yield in the field by 25% [142], indicating the significance and functional stability of these transporters even in interspecific hybrids. Additionally, Eswaran et al. [143] used the yeast Full-length cDNA Over-eXpressor (*FOX*) gene hunting to identify several salt-responsive genes in *Jatropha curcas*. The late embryogenesis-abundant protein **(LEA-5),** aquaporins and a cytosolic ascorbate peroxidase-1 (*Apx1*) were amongst the identified genes involved in salinity tolerance. *LEA5* are group 5 LEA genes that have been shown to play roles in the combining of concentrated ions and dehydration [143]. This group of LEA proteins have attracted fewer investigations and will require further studies at salt stress conditions. Aquaporin proteins are members of a large multigenic family that regulates a large proportion of water transport across membranes. Aquaporins are rapidly influenced both transcriptionally and post-translationally, and enhance salt stress tolerance in plants. For instance, a plasma membrane intrinsic protein (*GmPIP1;6*, which belongs to a subfamily of aquaporin specifically located in the PM) in soybean increases shoot Na^+ exclusion and improves the seed yield from a saline field [144]. Orthologous *PIP* proteins are found in *Arabidopsis* , tobacco, barley, rice and wheat. For instance, *GmPIP1;6* is the ortholog of *AtPIP1;2, NtAQP1, HvPIP1;6/1;1* and *TaAQP8.* Overexpression of *NtAQP1* in tobacco increases photosynthetic rate, water use efficiency and yield under salt stress [145]. Overexpression of *TaAQP8, TaNIP* and *TaAQP7* genes in Arabidopsis or tobacco also increases salt tolerance of transgenic plants [146–148]. Root stellar cells also confer control over shoot Cl^- accumulation [149]. The expression of *GmPIP1;6* in roots was recently shown to be correlated with rapid and longer term changes in root hydraulic conductance (L_o) in response to shoot treatments and appeared to be more concentrated in stellar tissue [150]. These results indicated that *GmPIP1;6* could be the protein responsible for the control of root water transport, particularly in response to shoot signals. More recently, overexpression of *GmPIP1;6* was shown to significantly increase salt tolerance of soybean by improving root L_o and Na^+ exclusion, which provided additional evidence that *GmPIP1;6*'s activity is in the stellar tissue. However, as there is no conclusive interactive or independent role of *AQPs* in salt tolerance, *AQPs* could instead be playing an indirect role through their impact on osmotically driven water and solute flow in roots and leaves. Further research will probably provide clear insight as to whether *GmPIP1;6* is responsible for salt regulation in the stellar cells, and whether there are other co-factors involved.

Wheat tonoplast intrinsic protein (*TIP2; 2*) is also reported to enhance salt tolerance [151]. However, the functional role of this protein is regulated by methylation following salt treatment as is *HKT1* in *Arabidopsis* [152]. This suggests that aquaporin methylation could also play a role in regulating salt tolerance in plants and is worth further exploration.

Accumulation of ROS scavenging enzymes has also been reported to lower cellular damage, maintain photosynthetic energy capture, and improve shoot and root growth under saline conditions. For instance, salt-stress-induced accumulation of SOD has been reported to play a protective role in *Canola*, *S. europaea*, *S. chilense* and *K. candel* [153–155]. Furthermore, expression levels of anti-oxidant enzymes *APX* (e.g.,***Apx1)*** , *Trx, Prx, GPX* and *GST* were observed to be enhanced in *Tangut nitraria* [156] under salinity conditions. Moreover, the same authors have reported that a photosynthetic enzyme, Ferredoxin—NADP (+) reductase (*FNR*), activity also increased in *T. nitraria*. Pea plants grown under saline stress also showed an enhancement of both APX activity and S-nitrosylated APX, which suggests that APX plays a significant role in plant tolerance to salt stress. However, apart from ascorbic acid biosynthesis, which has been shown to be modulated by *OsMPG1*, the molecular regulation of most anti-oxidants in response to salinity remains to be explored.

The recent discovery that salt-tolerant plant growth promoting rhizobacteria (PGPR) populations reduce Na^+ concentration in the plant shoots [157] provides further insights into plant tolerance to saline conditions. The PGPRs increase the expression of stress-responsive TFs, induce greater proline synthesis, enhance ROS scavenging and improve plant biomass under salinity stress. Therefore, treatment with rhizospheric organisms, and understanding the mechanisms associated with these PGPRs leading to salt tolerance, is an attractive option to improve crop yields under saline conditions.

Fundamental insights into genetic control of salt tolerance mechanisms have also led to identification of more than 100 QTLs in various crops including *Arabidopsis* , barley, rice and wheat, amongst others. The earlier mentioned **salt overly sensitive** (*SOS*) pathway genes and *AtCIPK16* are amongst the salt tolerance factors spanning several QTLs identified [158]. *CIPK16* is an SNF1-related kinase/CBL-interacting protein kinase, underlying a quantitative trait locus for Na^+ exclusion in the *Arabidopsis* Bay-0×Shahadara mapping population. *CIPK16* was also recently shown to be expressed in barley and improves Na^+ exclusion and biomass in a saline field.

Taken together, several genes and proteins have been shown to enhance salt tolerance in plants. However, the limited number of genes with functional polymorphism for salt tolerance makes it difficult to employ marker-assisted breeding for salt tolerance traits. In addition, the complex molecular mechanisms underlying the difference between seedling and reproductive stage salt tolerance in plants, e.g. rice [159], suggest the need for further research. The importance of the apoplastic bypass flow in delivering Na^+ to the xylem, thus reducing leaf Na^+ concentration and improving tolerance as suggested by [160], is also worth exploring further. Moreover, more insights into the molecular regulation of salt response will provide avenues for combining tolerance mechanisms to develop varieties that are widely adapted to salt stress.

6. Advances in plant tolerance to submergence/flooding

Over the past 25 years, yield losses caused by flooding have been increasing in various parts of the world, including the United States, China, Europe, Pakistan and Australia [161, 162]. Flooding is expected to increase as a result of erratic weather patterns, including frequent and lengthy storms associated with climate change, and could severely affect food production if mitigation measures are not sought.

Generally, submergence/flooding stress results from reduced oxygen levels in the plant root zone due to the low diffusion rate of oxygen in water. Submergence inhibits electron flows that underpin photosynthesis and aerobic respiration from the air causing energy shortfalls that can prove injurious to the plant [162]. Flooding also leads to accumulation of gases such as ethylene and carbondioxide by preventing their diffusive escape and oxidative breakdown [163]. A high concentration of ethylene limits root extension, while carbon dioxide can severely damage plant roots. Trapped carbondioxide may also form bicarbonate ions that can accentuate the effect of high lime content, leading to iron unavailability and chlorosis. The hypoxic environment also leads to restricted production of ATP, forcing cells to rely on glycolysis and fermentation to generate ATP and regenerate NAD^+ to cope with the energy crisis [164]. Moreover, survival through prolonged inundation hypoxia involves the use of inorganic pyrophosphate (PPi) as an alternative energy source and induction of enzymes that reduce reactive oxygen species (ROS) or cytoplasmic acidosis, which are equally energy consuming processes. Because translation is a tremendously energy-intensive process, protein synthesis is affected in such oxygen-deprived conditions. Subsequently, essential metabolic processes slow down affecting the overall growth of the plant. In rice, soybean and wheat, various deleterious effects have been observed, such as suppression or reduction of hypocotyl and root elongation, and suppression of lateral root development [162, 164, 165].

Plant tolerance to submergence/flooding is generally a metabolic adaptation in response to anaerobiosis that enables cells to maintain their integrity so that the plant survives hypoxia without major damages. Several defence-related changes occur in submergence tolerant plants, including anatomical (e.g. formation of higher aerenchyma tissue in the nodal region in rice), physiological (more shoot elongation) and biochemical (inhibition of chlorophyll degradation, less utilization of storage carbohydrates and increased activity of anti-oxidative enzymes). At the molecular level, plants need to adapt these several changes in their gene expression profiles as well as cellular protein profiles. We will focus more on molecular adaptation, with a preference for adaptive QTLs, genes and proteins of significance to crop tolerance to flooding.

One of the early responses to submergence involves the differential regulation of a suite of TFs belonging to the ethylene response factor (*ERFs*) gene family. In rice, a major QTL locus belonging to ERF family, which is responsible for submergence tolerance, was mapped to chromosome 9, designated as Submergence1 (*Sub1*) [166]. This QTL was reported to account for about 70% of the phenotypic variation under submergence [167]. One of the genes adhered to *Sub1* locus is *Sub1A*, which limits shoot elongation during submergence by repressing gibberellic acid (GA) levels and modulating GA signaling. In the process, the consumption of

energy reserves is reduced, and upon de-submergence, genotypes with *SUB1A* are able to resume development when flood water subsides.

Two *ERFs*, *SNORKEL1* (*SK1*) and *SNORKEL2* (*SK2*) from Thai deep water accession C9285, also confer submergence adaptation in deep water rice by inducing rapid internode elongation [168]. SKs have also been found in the genomes of accessions of wild *O. rufipogon* from Asia and *O. glumaepatula* from South America but missing in most cultivated rice varieties, which suggests that an ancient genomic region of *Oryza* was lost during the establishment of rice grown in shallow paddies, but was safeguarded in deep water ecosystems. More recently, two QTLs on chromosome 3 and 12, including *O. sativa-GROWTH-REGULATING FACTOR7* (*OsGRF7*), were reported to be involved in GA-dependent stem elongation and meristem maintenance in deep water rice [169]. *OsGRF7* on chromosome 12 could probably be a regulator of GA responsiveness for internode elongation, whereas a QTL on chromosome 3 and other QTLs may regulate the *DELLA* function or act downstream of GA signaling. The *DELLA* proteins are the key regulators of GA signaling and suppress plant growth in the absence of GA.

In maize, a major QTL, *Subtol6*, was also recently shown to be associated with submergence tolerance [170]. Based on the expression differences between the parent inbreds, *subtol6* is associated with *HEMOGLOBIN2* (*HB2*), a gene which was previously reported to be associated with plant survival in low oxygen or low ATP conditions [171]. The same authors indicate that haemoglobin proteins in maize repress ROS levels and maintain the energy status of maize cells during hypoxia. Other notable candidate genes, including genes related to *ABA-INSEN-SITIVE3* (*ABI3*)/*VIVIPAROUS1* (*RAV1*), genes related to accumulation and metabolism of carbohydrates, e.g., alpha subunit of *PYROPHOSPHATE-DEPENDENT FRUCTOSE-6-PHOSPHATE 1-PHOSPHOTRANSFERASE* (*PFP*) and *ALCOHOL DEHYDROGENASE1* (*ADH1*), have been reported to be highly upregulated in response to submergence [170].

In association with these tolerance genes, a number of other QTLs have also been identified in various crops, including barley, wheat, *Brassica napus*, maize and *Lolium perenne*, amongst others.

In addition to these QTLs studies, several proteins have been reported to enhance submergence tolerance in plants. Enzymes involved in primary metabolism are differentially regulated in response to flooding. For instance, UDP-glucose dehydrogenase, UDP-glucose pyrophos-phorylase, β-glucosidase G4 and rhamnose synthase, aspartate aminotransferase and lipoxy-genase have been reported as early flood-responsive proteins in rice and soybeans [164, 172]. The same authors indicate that phenlypropanoid pathway and cell wall synthesis enzymes decrease in abundance during flooding, which could be an energy-conserving adaptive strategy towards enhanced flooding tolerance.

Together these findings suggest that during flooding several processes are inhibited to reduce energy consumption. It is crucial for the plant to preserve some carbohydrates for release of energy to support further growth when the water level recedes. The regulatory genes in this category may also serve some ABA-mediated water stress recovery and inhibition of GA-induced internodal elongation as quiescence strategies adopted by plants [173]. On the other

hand, avoidance mechanisms employed under deep water conditions involve rapid internode elongation. In *R. palustris*, there are populations that show either the quiescence response or the avoidance response to submergence. This divergence shows that quiescence and avoidance are two strategies that can be employed by plants depending on the duration of flooding. Quiescence can be the optimal strategy for short-duration 'flash' floods, whereas avoidance via growth could be more reliable in prolonged deep flooding. Notwithstanding the above-mentioned tolerance genes and proteins, a deeper insight into the molecular regulation of quiescence and avoidance, and the associated regulatory networks, is still needed to provide sustainable avenues for improving plants specific to either flooding condition or able to grow in both.

7. Advances in plant tolerance to nutrient imbalances

7.1. Tolerance to nutrient deficiency

A total of 21 mineral nutrients are essential for crop growth and development. Most nutrients in the soil are primarily generated from the weathering of the parent material in the Earth's crust. Moreover, nutrient levels can vary widely across locations because of initial influence of the composition of the parent material. In most cases, inadequate replenishment from the parent material and from the adsorbed and complexed fractions causes nutrient deficiencies in the soil. In addition, natural factors, including acidity, alkalinity and human activities such as inadequate fertilization also cause nutrient deficiencies. In countries such as India and China, mineral deficiencies have significantly stagnated or limited crop yields. More than 30% of agricultural soils are boron deficient, not only in China and India, but in the whole world. Moreover, zinc deficiency is even more widespread, affecting approximately 50% of the soils. Significant zinc deficiencies occur in sub-Saharan Africa, Turkey, Iran and Pakistan [174].

Several studies have been conducted on understanding plant nutrition; the most noteworthy being the work of the German scientist Justus von Liebig, who stipulated that plant growth is controlled not only by the total resources (nutrients) available, but also by the scarcest resource (the limiting factor). This submission has stimulated a series of studies on nutrient manage-ment, including plant breeding for tolerance to nutrient deficiencies. Tolerance to nutrient deficiency is associated with the genotype's nutrient use efficiency. Genotypic variation in nutrient use efficiency is closely related to root nutrient acquisition capacity and utilization. In this section, we will focus on nitrogen and phosphorus, the two most limiting nutrients that are essential for several biological processes in plants.

7.1.1. Plant tolerance to nitrogen deficiency

Nitrogen is the most limiting nutrient to plant growth in most ecosystems despite its abun-dance in the atmosphere. This problem occurs because most plants can only take up nitrogen in two solid forms: ammonium ion (NH_4^+) and nitrate ion (NO_3^-). Ammonium is used less by plants because it is extremely toxic if taken up in large concentrations, so inorganic nitrate is the most usable form obtained by plants from the soil solution. Nitrogen-deficiency effect on

crop yields depends on the growth stage at which it occurs, as well as on its duration and extent [175]. However, reduced radiation interception, low radiation use efficiency, poor dry matter partitioning to reproductive organs, reduced leaf area index and decreased protein content of the plant and seed are the common effects of nitrogen deficiency.

Plants react in many different ways to changes in N provision; and physiological and molecular components governing N uptake, assimilation and remobilization during the plant life cycle have been studied extensively in the past decades, for review see [176, 177]. Three types of responses have been recently unraveled: (i) regulation of root N uptake systems, (ii) plasticity of root system architecture and (iii) fast modulation of shoot growth [178]. The first two responses generally improve efficiency of root N uptake under deficient conditions. The upregulation of specific high-affinity membrane transporters and enhanced foraging by the root system are implicated in these responses. When soil conditions for N uptake are seemingly unfavourable, e.g. limited water availability, plants will quickly slow down the overall N demand, as a nutrient conserving adaptive strategy, to prevent N starvation until conditions for N uptake become favourable.

In various plant species, nitrate transporters play a dominant role in N uptake. In *Arabidopsis*, three major families of nitrate transporters have been identified: Chlorate resistant 1 (*CHL1/NRT1*), *NRT2* and chloride channel (*CLC*) [177]. *NRT2* belongs to the high-affinity nitrate transporter group while most of the *NRT1* family members belong to low-affinity nitrate transporters. The only exception, so far, in the latter group is *NRT1.1* that is a dual affinity nitrate transporter. Thus, the high-affinity transporters that have been identified and primarily associated with nitrate uptake from the external environment include *NRT1.2*, *NRT2.1*, *NRT2.2* and the dual affinity transporter, *NRT1.1*.

NRT1.1 is functionally regulated by phosphorylation of a threonine residue, *Thr101*, which facilitates the switching of its activity from a low- to a high-affinity state. *AtNRT1.1*, which is also induced by auxin and is itself an auxin influx facilitator, is a dimer in the asymmetric unit cell despite being monomeric in solution. At low nitrate levels, *AtNRT1.1* is phosphorylated at the dimer interface, dissociates the *NRT1.1* dimer, changes into a high-affinity transporter and represses lateral root (LR) development by promoting basipetal auxin transport out of LR primordia (LRP) [179]. At high nitrate levels, *NRT1 1* is dephosphorylated, adopts a dimeric structure and adapts a low-affinity transporter configuration. In this state, trafficking of auxin out of the LR is blocked, and auxin accumulates in the LR initials promoting LR development. *NRT1.1* is also shown to act upstream of the *MADS box ARABIDOPSIS NITRATE REGULATED1* (*ANR1*) when modulating LR growth [179]. *ANR1* mediates localized N response and modulates the proliferation of LRs in N-dense patches. Moreover, *NRT1.1* has been shown to regulate genes encoding for other nitrate transporters, including *NRT2.1* and *NRT3.1* [180]. However, *NRT1.1* and *NRT2.1* are localized in different cell layers in the roots, and their adaptive/complementary strategy in nitrate uptake is not well elucidated. The *NRT1.1*-auxin modulation and nitrate signaling has also been a topic of interest and requires elucidation [181].

Amongst the *CLC* family members, *CLCa* and *CLCb* function as proton-nitrate exchanges, and have high selectivity for nitrates over chlorides [182]. Both transporters are known to mediate

nitrate accumulation in the plant vacuoles. Besides the above-mentioned transporters, the acquisition of nitrate is also regulated by slow anion channel (*SLAC1*) and *SLAC1* homologue (*SLAH*) and nitrate excretion transporter (*NAXT-1)*. Five *SLAC* genes were previously reported in *Arabidopsis* . Amongst these genes, *SLAC1* and *SLAH3* show nitrate transport activity, but their channel activity is co-regulated by kinases (e.g., *CPK21*) [183]. An efflux component operated by *NAXT-1*, associated with the nitrate transporter 1/peptide transporter (*NRT1/PTR*) family of proteins, mediates nitrate efflux under acid load in the cytosol [184]. Similarly, *NRT1.5*, which loads nitrates into the xylem for root-to-shoot translocation, also mediates nitrate efflux. However, the proton-coupling mechanism of *NAXT1* remains to be elucidated. Two other transporters, *NRT1.8* and *NRT1.9*, have been reported to regulate root-to-shoot nitrate translocation [185, 186]. Both transporters are apparently negative regulators of root-to-shoot nitrate transport. The subsequent nitrate allocation into the vegetative tissues, reproductive tissues and osmotic regulation of guard cells is reasonably described elsewhere [187].

Further studies on signaling, transcriptional and post-translational regulation have revealed evidence that a CBL-interacting protein kinase, *CIPK8*, regulates the activity of nitrate transporters and the expression of nitrate assimilation genes [188]. Like *CIPK8*, *CIPK23* is also suggested to be activated by a CBL protein, *CBL9*, but the exact mechanism is elusive. *CIPK23* directly interacts with *NRT1.1* in the plasma membrane and phosphorylates *NRT1.1* at *Thr101* to adopt a monomeric structure when the nitrate concentration is low. This process helps plants to adapt to low nitrogen levels.

Several TFs have been implicated in regulating *NRT1.1* activity. For instance, the activity of two *bZIP* TFs in *Arabidopsis*, *ELONGATED HYPOCOTYL5* (*HY5*) and *HY5-HOMOLOG* (*HYH*), was suggested to positively modulate *NITRATE REDUCTASE2* (*NIA2*) and negatively modulate *NRT1.1* [189]. The *NODULE INCEPTION* (*NIN*)-like TFs have also been shown to play a central role in the regulation of nitrate-inducible genes [190]. Nitrate signaling activates *NIN*-like transcription factors through their N-terminal regions. The activated factors promote the expression of nitrogen assimilation-related genes and genes encoding regulatory proteins. *NLP7* is the most reported in this family of TFs. *NLP7* is strongly induced in vascular tissues and root hairs, and is required for the induction of several nitrate uptake and assimilatory genes. Thus, *NLP7* is is probably a key regulator of nitrogen utilization mechanisms. More recently, the presence of nitrate in the external solution induced the expression of *NRT* accessory proteins (*NAR*), nitrate reductase, nitrite reductase and genes involved in the GS-GOGAT cycle, in *Arabidopsis*, as well as in maize and other plants [191]. These proteins likely play a role in nitrate sensing.

Strigolactones (SLs), a new class of plant hormones and rhizosphere signaling molecules, also appear to be upregulated in plants under low N conditions [192]; however, the impact of SL levels on root growth is yet to be determined. Changes in root system architecture (RSA) may also be induced depending on the prevailing available organic form of nitrogen, for review see [118]. The most commonly reported organic forms are l-glutamate or carnitine. In *Arabidopsis* seedlings, l-glutamate inhibits cell division in the root apical meristem (PRM) of the primary root (PR) tip and promotes LR formation and outgrowth. However, several

Arabidopsis auxin-signaling mutants display different levels of sensitivity to l-glutamate, suggesting that l-glutamate is rather a signaling molecule as opposed to a nitrogen source [193]. In addition, the rice glutamate receptor mutants display a host of RSA changes, including short PR and LR, reduced cell division and the cell death of root apical meristem [194], further suggesting that l-glutamate is a signaling molecule. l-Glutamate could be a major anchor in the signaling process leading to nitrate uptake and assimilation. This is supported by previous studies that have shown that glutamine synthetase (*GS1*) from alfalfa causes an increase in photosynthesis and growth under low N fertilization regime [195]. Glutamine synthetase also mediates ammonium assimilation into glutamine. Ammonium form of nitrogen is rapidly assimilated into organic nitrogen forms to avoid tissue toxicity, for review see [196]. Several other reviews have documented the genes and proteins regulating nitrogen use efficiency (NUE) in plants. The reader is referred to excellent reviews by [177, 196]. In addition, more than 50 QTLs for nitrogen use efficiency have been reported in plants, though few of them have been validated. Amongst the identified QTLs are nitrogen deficiency response QTLs in rice, nitrogen supply responses and yield in wheat and nitrogen use efficiency in barley.

Collectively, nitrogen use efficiency in plants is controlled by a complex array of physiological, developmental and environmental interactions that are specific to the genotype of a given species. Notwithstanding the aforementioned N uptake and utilization genes and QTLs, an extensive molecular survey of a wide range of genotypes covering the genetic diversity of a crop could provide further evidence on the genetic control of these trait. This can be achieved using the various available 'omics' techniques, combined with agronomic and physiological approaches in order to identify more elements controlling NUE in plants, both universal and specific, for use in crop improvement.

7.1.2. Plant tolerance to phosphorus deficiency

Phosphorus (P) is the second most limiting mineral nutrient in almost all soils, and >30% of the world's arable land has low P [197]. Phosphorus availability is particularly limiting on highly weathered acid soils of the tropics and subtropics due to its fixation by Al and Fe oxides on the surface of clay minerals. Plants take up phosphorus as phosphate (Pi), either directly by the root system or transferred through the fungal symbiont in arbuscular mycorrhizae host plants. Plants have elaborate sensing and signaling mechanisms in response to Pi deficiency, and both local and systemic signaling in response to Pi deficiency have been reported [197]. Key responses in the plant include changes in the root system architecture (RSA), a reduction in photosynthetic rate; increased activity of high-affinity Pi transporter activities; secretion of APases, ribonucleases and organic acids; membrane phospholipid replacement with glycolipids and sulfolipids; and increased availability of anthocyanin and starch [198]. Putative signaling molecules in response to Pi deficiency include sugars, hormones and microRNAs.

Under limiting Pi conditions, plants can monitor Pi deficiency both locally and systemically, and root foraging strategy to explore top soil layers for Pi is employed. The Pi foraging strategy is accomplished through one of the several different RSA and physiological changes [118]. The local external Pi rather than the systemic Pi status of the whole plant regulates the remodelling of RSA [199]. In maize and some species in the *Proteaceae* and *Casuarinaceae* families, the

remodelling of RSA involves production of adventitious roots and cluster roots [200, 201], which increases root surface area for Pi absorption. While a plant Pi receptor is yet to be identified, recent reports have suggested that ethylene biosynthesis and signaling are involved in the Pi-deficiency-triggered remodelling of RSA, for review see [118, 195]. The evidence is supported by previous finding that inhibition of ethylene biosynthesis with 2-*aminoethoxyvinyl glycine* (*AVG*) or ethylene perception with Ag⁺ restricted the low Pi-induced meristem exhaustion of the primary root [202]. Correspondingly, application of Ag⁺ was found to reduce the inhibition of primary root growth triggered by Pi deficiency. Moreover, Pi deficiency induced the formation of aerenchyma in adventitious roots, which is similarly induced by ethylene perception.

At the transcriptional level, Lei et al. [203], using an *Arabidopsis* transgenic line that carries a *LUC* gene fused to the promoter of the high-affinity Pi transporter, *AtPT2*, showed that the transcription of *AtPT2* is induced by Pi starvation. Using this marker line, the authors identified the *Arabidopsis* mutant *etr1/hps2* (*constitutive triple response 1/hyper-sensitive to Pi starvation2*), which showed hyper-induction of the *AtPT2::LUC* gene by Pi deficiency. Interestingly, the expression of *AtPT2* was partially blocked in *ethylene insensitive* 2 (*ein2*) mutants, but was enhanced in *ethylene over producer1* (*eto1*) mutants. A similar expression pattern was observed for several other Pi starvation-induced (*PSI*) genes in the *hps2* (negative regulator of ethylene response) and *ein2* mutants, including high-affinity phosphate transporter, *AtPT1* (*Pht1;1*); a non-coding transcript, *At4*; an APase, *ACP5*; a ribonuclease, *Rxlink*; and *miR399d* [204]. Enhanced transcription of *PSI* genes was also observed in the mutant *hps3* and *hps4*, which are *ETO1* alleles [205, 206]. *ETO1* protein is a member of the broad complex/tramtrack/bric-a-brac (BTB) protein superfamily that participates in substrate recognition during ubiquitin-mediated protein degradation [204, 207]. *ETO1* directly binds to the C-terminal of *ACS5* and mediates its degradation. When *ETO1* is mutated, it causes an overproduction of ethylene in young seedlings [208]. Application of 25 µM ACC to young *Arabidopsis* seedlings under high Pi conditions barely induces the expression of *AtPT2*. Under Pi deficiency, however, 0.5 µM ACC dramatically increases *AtPT2* expression and induces ectopic root-hair development [203]. Thus, these results provide evidence that ethylene production and signaling is involved in the transcriptional responses of plants to Pi deficiency and primarily integrates with other Pi-deficiency-induced signaling pathways.

The other signaling component involving increased transcription of purple acid phosphatase 10 (*AtPAP10*) by Pi starvation in the whole seedlings of *hps3* and *hps4* has been reported [205, 206]. *AtPAP10* is a Pi starvation-induced *APase* (enzymes that scavenge Pi from organophosphate compounds) associated with the root surface. Functional analyses of *atpap10* mutants suggest that *AtPAP10* is important for plant tolerance to Pi starvation. However, the transcription of *AtPAP10* does not significantly increase in *ACC*-treated seedlings or the *constitutive triple response I* (*ctr1*) mutant under Pi deficiency, nor does the accumulation of *AtPAP10* proteins, which could suggest that ethylene has no effect on *AtPAP10* transcription. More recently, Zhang *et al.* [209] have shown that positive regulation of *AtPAP10* depends on sucrose and not ethylene. Moreover, they have also shown that ethylene does not affect *AtPAP10* activity without sucrose, but the opposite is true. This suggests that ethylene could be a local

but indirect signal for *AtPAP10* activity. However, as discussed before, ethylene could be regulating other components of Pi starvation response at the transcriptional level. Song and Liu [204] have demonstrated that accumulation of anthocyanin is lower in *hps2*, *hps3* and *hps4* mutants under low Pi but increases in Pi-starved *ein2* mutants. As mentioned before, accumulation of anthocyanins is an indicator of Pi-deficiency response in plants, thus ethylene could be a negative regulator of Pi-deficiency-induced anthocyanin accumulation probably through the regulation of genes involved in anthocyanin synthesis. Thus, ethylene likely participates at both the transcriptional and post-transcriptional levels, and this has implications on Pi starvation response in plants.

The systemic response to P starvation is also carried out through a complex signaling network that involves other plant hormones [210, 211], sugars [212] and nitric oxide [213], collectively resulting in the alteration of carbohydrate distribution between roots and shoots. Amongst the plant hormones, other than ethylene, auxin likely plays a role in response to Pi starvation. However, ethylene likely exerts its influence through regulating auxin activity, as it has been associated with RSA remodelling [198]. Indeed, ethylene has been reported to interact with auxin and sugars, and changes in auxin transport and localization appear to be at least partially responsible for Pi stress-induced LR development [214]. Decreased sensitivity to CK and GA also appears to be at least partially responsible for Pi-stress-induced LR development [215]. Under low Pi, GA has been shown to repress Pi-induced root architecture changes [216]. Moreover, Pi-deficient plants were shown to accumulate *DELLA* proteins, the negative regulators of GA-induced root growth, which are modulated by auxin.

As discussed before, amongst sugars, sucrose is key to Pi-deficiency response and appears to regulate ethylene activity. Amongst the TFs, phosphate starvation response proteins (e.g., *OsPHR1*, *OsPHR2*, *PvPHR1*, *ZmPHR1* and *TaPHR1*), which bind the promoter sequences of low Pi-induced genes, and their regulator *SMALL UBIQUITIN-LIKE MODIFIER1* [*AtSIZ1*; 217], a small ubiquitin-modified E3 ligase, and the downstream *PHOSPHATE2* (*PHO2*), an E2 conjugase, are involved in Pi-deficiency-related transcriptional changes. Other TFs, including the *bHLH*, *PTF1* (e.g., *OsPTF1* and *ZmPTF1*) and *MYB2P-1* (e.g., *OsMYB2P1*), *MYB62*, *WRKY* (e.g., *WRKY75*, *WRKY6*), *bHLH32* and *ZAT6* are also involved in the signaling network to regulate plant adaptation to P stress, for review see [218].

Based on genetic analysis, two proteins, the P_5 type ATPase encoded by *PHOSPHATE DEFICIENCY RESPONSE2* (*PDR2*), and multicopper oxidase *LOW PHOSPHATE ROOT1* (*LPR1*), were also previously shown to modulate Pi signaling in an endoplasmic-reticulum-localized pathway [219]. *PDR2* is required for maintaining the levels of the root patterning gene, *SCARECROW* (*SCR*), and *SHORT-ROOT* protein (*SHR*) trafficking from stele into endodermis. *PDR2* was proposed to act upstream of *LPR1/LPR2* to adjust meristem activity. A recent study has shown that *LPR1* is a ferroxidase [220]. Mutation of *LPR1* reduces Fe^{3+} levels in the meristemic tissues of Pi-deficient plants. In contrast, increased levels of Fe^{3+} have been reported in *pdr2* mutants leading to high production levels of reactive oxygen species (ROS). ROS signaling increases deposition of callose, which has been suggested to impair the trafficking of *SHR*, thus restricting root tip growth. Thus, PDR2 appears to modulate Pi-deficiency response by limiting Fe^{3+} accumulation in root tips.

More recently, molecular mechanisms defining the phosphate signaling pathway showed that *phosphate uptake 1* (*Pup1*)-specific protein kinase gene, named *phosphorus-starvation tolerance 1* (*PSTOL1*), was confirmed to be involved in regulating root growth and architecture during early stages of rice growth [221]. Allele-specific markers for this gene have been reported recently [222]. Interestingly, *OsPSTOL1* is located within the Kasalath-specific INDEL region and is absent from the rice variety Nipponbare reference genome. Thus, the configuration of the functional mechanism of *PSTOL1* is still elusive. We speculate that *PSTOL1* could be a local sensor of Pi starvation which transduces signals for sucrose or ethylene biosynthesis or both. The interplay of sucrose accumulation and ethylene biosynthesis is apparently the hallmark of Pi starvation response in plants.

The post-transcriptional regulation as well as long-distance signaling is carried out by microRNAs. As mentioned before, *miR399*, which is regulated by *PHR1*, a conserved *MYB* TF, maintains P homeostasis by regulating P transporter *PHO2* [223]. In tomato, overexpression of *Arabidopsis miR399* increases both the Pi accumulation and secretion of acid phosphatase and protons in the roots [223]. Thus, *miR399* is important for Pi acquisition, and could be acting downstream of sucrose and probably ethylene. Overexpression of *miR399* in *Arabidopsis* also increases P uptake and allocation to the shoot. Moreover, P remobilization from older leaves to young leaves is defective in *Arabidopsis miR399* transgenic lines [224]. This suggests that *miR399* is important for allocation and remobilization of P. The targets of *miR399* include a ubiquitin-conjugating E2 enzyme (*UBC24*) encoded by *PHO2*, which is upregulated under P-sufficient conditions and downregulated in P-starved plant roots. Homologues of *PHO2/UBC24* have a conserved structure in many species, and their 5′ UTR regions possess multiple *miR399*-complementary sequences. Thus, the regulatory mechanism of *miR399-PHO2* complex is evolutionarily conserved in angiosperms, making it a potential target for improving P nutrition efficiency in plants.

Strigolactones (SL) have also been shown to be induced by low Pi in many species, including tomato, *Arabidopsis*, pea and rice [225–229]. Strigolactones are terpenoid lactones that function as either endogenous hormones that control plant development or as components of root exudates that promote symbiotic interactions between plants and soil microbes. The production and exudation of SLs may depend on whether the plant is arbuscular mycorrhizal fungi (AMF)-compatible host or an arbuscular mycorrhizal symbiosis (AMS) for Pi and N uptake. A well-known synthetic SL, *GR24*, apparently increases LR formation under low Pi or decreases LR formation under sufficient Pi. In addition, SL biosynthesis (*more axillary growth; max4-1*) and signaling (*max2-1*) mutants have reduced number of root hairs under low Pi condition at the early stages of seedling development. This suggests that SLs mediate plant responses to low Pi; however, the mechanism by which SL exudation affects root growth is not fully understood.

In conclusion, although the molecular components of P stress signaling in plants have been fairly documented, the overall pathway is still less understood and requires further investigation. Nonetheless, the recent developments in whole genome sequencing technologies provide hope for more studies on plants with better P acquisition and utilization. Successes in QTL analysis have also set a stage for subsequent studies. Besides the success story of *PSTOL1*

in rice, QTL analysis in common bean has shown the importance of basal roots and adventitious roots for P acquisition [230–232]. Another study by Yan et al. [233] identified a large number of QTLs for Hþ exudation, root-hair density and length, associated with P efficiency. Additionally, QTLs for root traits related to P efficiency have also been identified in soybean [234, 235]. Moreover, QTLs controlling P deficiency tolerance were mapped by Zhang et al. [344] using 152 RILs derived from a cross between P-stress-tolerant and P-stress-sensitive parents. Thus, future studies will build on these present discoveries to facilitate genetic improvement for Pi-deficiency tolerance.

7.2. Advances in plant tolerance to nutrient toxicities

Metal toxicity is an important factor limiting the growth of plants in many environments. Some metals, such as copper and zinc, are micronutrients at low concentrations and become toxic at higher levels, whereas others (e.g., aluminium, iron, cadmium, chromium and lead) are well known for their toxicity [236]. These elements can be highly phytotoxic and seriously impair plant root growth. However, some crops are able to tolerate toxic environments, without significant display of toxicity symptoms. Three main strategies are employed by such plants to manage toxic soil compounds: (1) Producing root exudates that bind and neutralize the toxin in the rhizosphere, (2) actively transport the compound into the root, but neutralizing and sequestering it in vacuoles for safe accumulation or eliminating it through exudation and (3) excluding the toxic elements by preventing entry into the plant tissues. For the purpose of this chapter, we will focus on aluminium and iron toxicities as these elements have been frequently reported as major constraints in the production of economically important crops.

7.2.1. Plant tolerance to aluminium toxicity

Aluminium (Al) is a light metal that makes up 7% of the Earth's crust and is the third most abundant element after oxygen and silicon. Aluminium toxicity is one of the major constraints to crop productivity worldwide, especially in the acid soils of the tropics and subtropics that comprise almost 50% of all non-irrigated arable land in those regions [118, 237]. The soil pH has a crucial role for Al toxicity to occur, by affecting both solubility and the ability of plant roots to absorb Al. Al solubilizes into its toxic form (Al^{3+}) when the soil pH drops to 5.5 or less, and is most severe in solutions of low ionic strength and low cation concentrations. Al^{3+} is taken up by plants through diffusion [238], and toxic concentrations of >12 μM are detrimental to root growth. Possible exceptions of $Al(OH)_3$ $^{4-}$ toxicity at higher pH values have also been reported [239].

The initial effects of Al^{3+} toxicity on the roots include rapid inhibition of cell division and a reduction in root apical cell expansion and elongation. Consequently, plants develop stubby and brittle roots with swollen malformed root tips. Moreover, lateral root initiation and outgrowth are also inhibited. Root-hair malformation is often reported, and nutrient (mainly P, K, Ca and Mg) and water uptake capacity is impaired [238]. Plant responses in the shoots include reduced stomatal opening, chlorosis, foliar necrosis and reduced photosynthetic activity.

Plant tolerance to aluminium toxicity occurs through (1) external avoidance, which involves root secretion of organic acids to chelate Al^{3+} in the rhizosphere, limiting its diffusion into the roots [240], and (2) true or internal tolerance, which involves regulation of Al^{3+} uptake, and organic acid chelation and sequestration of aluminium bound substrates [241]. In rice, the latter is the main tolerance mechanism, and is apparently associated with the differential expression and transport properties of membrane transporters, e.g., *NRAMP Al $^{3+}$ transporter 1* (*NRAT1*) [242]. Most other plant species also vary significantly in these mechanisms; however, there are some tolerance mechanisms that are largely shared. Cereal crops, such as wheat, barley, sorghum (*Sorghumbicolor* L.) and oat were reported to have simple genetic mechanisms of Al tolerance, whereas rice and maize (*Zea mays* L.) have over time developed complicated inheritance controlled by numerous genes/loci involved [118, 243].

Genetic control of organic acid exudation either rests on the Multidrug and Toxin Efflux (MATE) family encoding a citrate transporter or on the membrane localized Al^{3+}-activated malate transporters (*ALMT*). Several transporters in these families, including *HvAACT1* in barley [244], *TaALMT1* and *TaMATE1* in wheat [245] and *ZmMATE1* and *ZmMATE2* in maize [246] are responsible for organic acid exudation and Al tolerance. Specific markers for *HvAACT1* and the MATE gene, *HvMATE-21*, have been developed and can be used to differentiate tolerant and sensitive barley cultivars. Differences amongst these transporters however exist. For instance, *TaALMT1* encodes a malate transporter on chromosome 4D and is constitutively expressed on root apices, whereas *TaMATE1* reportedly responds to Al stress based on citrate efflux. ZmMATE1 and *ZmMATE2* co-segregate with two major Al-tolerance QTLs [247]. *ZmMATE1* was shown to be induced by Al and enhances Al tolerance, whereas *ZmMATE2* did not respond to Al [246], suggesting variability in their roles. In sorghum, Al tolerance is controlled by *SbMATE*, encoded by a major Al-tolerant locus Alt_{SB} on chromosome 3 [248]. In *Arabidopsis* , two genes were reportedly responsible for Al tolerance: *AtALMT1* that also encodes a malate transporter responsible for malate efflux on chromosome 1 [249] and *AtMATE* that encodes an Al-activated citrate transporter [389]. These two genes function independently, but both are regulated by the C2H2-type zinc finger transcription factor *STOP1* [250], which is also reportedly induced by with low pH tolerance [366]. In rye, *ScALMT1*, which is mainly expressed in the root apex and upregulated by Al, co-segregates with the *Alt4* locus on chromosome 7RS [367]. Another candidate gene *ScAACT1* on chromosome 7RS was mapped to 25 cM from *ScALMT1* [251].

At the transcriptome level, two genes, *SENSITIVE TO ALUMINUM RHIZOTOXICITY1* and *2* (*STAR1* and *2*), which encode the nuclear binding domain and the transmembrane domain, respectively, of an ABC transporter, with specificity for uridine diphosphate (UDP) glucose, are upregulated following root exposure to Al^{3+} [252]. Both *STAR* genes were previously reported to be upregulated by the constitutively expressed rice root *ALUMINUM RESISTANT TRANSCRIPTION FACTOR1* (*ART1*), which also upregulates several other genes implicated in different aluminium tolerance mechanisms [253]. More recently, *ASR5* was reported to act as a key TF that is essential for Al-responsive *STAR1* and other Al response genes [254]. Rice homologues, which encode α-expansin (e.g., *EXPA10*), belong to this family of TFs, and have been implicated in the regulation of root elongation and cell wall elasticity. The members of

EXPA10 decrease cell wall extension potential when exposed to Al^{3+} [255] and are downregulated during Al^{3+} stress. The functions of *STAR1*, *STAR2/ALS3* and *ALS1* in Al tolerance are fairly conserved and ubiquitous in monocot and dicot species. However, these genes are differentially expressed between species. For instance, the expression and induction levels of these genes in response to Al^{3+} stress are higher in the Al-tolerant species of rice than in the Al-sensitive species of *Arabidopsis* , suggesting that Al-tolerant species may require increased expression of these conserved Al-tolerance genes to overcome Al^{3+} stress [256]. The same authors show that Tartary buckwheat shows high expression of the Al-tolerance gene homologues under Al^{3+} stress. Al-tolerance in buckwheat is evolutionarily closer to *Arabidopsis* than rice, suggesting that buckwheat could have rapidly evolved higher expression of Al-tolerance genes to detoxify Al^{3+} than *Arabidopsis* . In addition, the gene duplication of *ART1/STOP1*, *STAR1* and *ALS1* has been suggested to play a significant role in Al tolerance. This is consistent with the previous findings that duplication of key genes responsible for metal translocation and detoxification in *Arabidopsis halleri* facilitates hyper-accumulation of zinc/cadmium [257]. However, further functional analysis by creating knockdown or knockout mutants will be necessary to provide additional insights into the role of each homologous gene in Al detoxification and accumulation in buckwheat.

An *Arabidopsis* cell-wall-associated putative endochitinase, CHITINASE A (*AtCHIA*), likely involved in modulating cell wall extension by regulating chitin levels, has also been suggested to play a role in Al tolerance [258]. Another signal of Al^{3+}-induced cellular response is the induction of *1,3-β-d-glucan synthase*, which leads to the accumulation of callose in root apices, especially in endodermal and cortical cell walls [259, 260]. This callose deposition is suggested to be an inhibitory process that may block symplastic and apoplastic flows. Whether callose deposition represents Al^{3+}-induced injury or a defence response to block further Al^{3+} binding and movement remains to be confirmed.

In *Arabidopsis*, the ethylene receptor gene *ETHYLENE RECEPTOR1* (*ETR1*) and the ethylene signal transducer *ETHYLENE INSENSITIVE2* (*EIN2*) were found to be important for Al^{3+}-induced inhibition of root elongation [261]. These genes apparently regulate Al^{3+}-induced upregulation of the *Arabidopsis* ethylene synthesis genes *1-AMINOCYCLOPROPANE-1-CARBOXYLIC ACID SYNTHASE2*, *6*, and *8* and *1-AMINOCYCLOPROPANE-1-CARBOXYLIC ACID OXIDASE1* and *2*. Moreover, ET was recently shown to negatively regulate aluminium-induced malate efflux from wheat roots and tobacco cells transformed with *TaALMT1* [262], which suggests that ethylene could be a negative regulator of root secretion of organic acids. The upregulation of auxin transporters *PIN FORMED2* (*PIN2)* and auxin influx carrier *AUXIN RESISTANT1* (*AUX1)*, which regulate auxin distribution, is associated with the regulation of root elongation in *Arabidopsis* plants [263]. *AUX1* and *PIN2* are apparently disrupted by ethylene signal that alters auxin distribution and transport in the roots. He et al. [264] suggests that auxin could be involved in aluminium-induced efflux of malic acid acting on anion channels. Thus, auxin/IAA transport could be a target for Al^{3+} toxicity tolerance if the modulation by ET is attenuated. However, considering several phytohormonal changes that occur during Al stress, molecular mechanisms associated with their interplay will require further

elucidation. Recent evidence that microRNAs are involved in Al stress tolerance [265] also provides new insights into understanding the mechanism of Al^{3+} tolerance in plants.

Overall, we expect that major advances in understanding physiological and molecular basis for Al tolerance will happen in the near future, considering that the pace at which new genes are being discovered has improved with new sequencing technologies. The future challenge for studying Al tolerance is the identification of new tolerance mechanisms. The discovery of the key molecular regulators, e.g., *ASR5*, which was recently shown to mediate Al-responsive gene expression to provide Al tolerance in rice, is an indication that several other mechanism of Al tolerance exist in plants. The blocking of Al^{3+} cell wall binding sites in rice may be one of the major mechanisms of aluminium tolerance that will need further investigation. Studies on barley, wheat and maize have shown variation in gene expression associated with variation in gene sequence, which would require further investigation to understand the regulatory networks affected by this sequence polymorphisms.

7.2.2. Advances in plant tolerance to iron toxicity

The problem of iron toxicity occurs in most wetland rice growing areas of the world, primarily in flooded acidic soils, inland and coastal swamps. Some of the irrigated lands in South and Southeast Asia, Africa and South America are affected [266]. In India alone, about 11.7 million hectares of land are affected by iron toxicity. In Burkina Faso, 300 ha of ferrous iron intoxicated soils were abandoned in the Valley du Kou in 1986, most of which remained uncultivated to date [267]. Iron toxicity is also becoming a major rice yield limiting factor in East Africa, including lowland rice cultivation areas of Uganda [268]. Yield losses in the range of 10% to 100% have been reported [266]. Moreover, toxicity at seedling and early vegetative stages can strongly affect plant growth and hinder development, and can result in complete crop failure.

Three major adaptation mechanisms are generally reported for Fe-toxicity tolerance. The details by which rice plants execute these processes and their molecular components are not yet fully understood, but there are some clues from various studies on rice and other plant species. For instance, plant tolerance by root oxidizing power is mediated by diffusion of molecular oxygen from the shoots to the roots through aerenchyma tissue and its subsequent release in the rhizosphere. Oxidation of Fe^{2+} in the rhizosphere results in the precipitation of insoluble iron oxides at the root surface, forming iron plaques. These iron plaques not only reduce Fe^{2+} concentration in the soil solution, but also form a physical barrier against further influx of Fe^{2+} into the roots.

Plant tolerance by retention of iron in the root or shoot involves compartmentalization. Nicotianamine (NA), Fe-NA complex transporters, *VIT* proteins, *FPN2*-like proteins, *MIT*- and *PIC1*-like proteins, organic acids, ferritins, Fe-sulphur and other heme proteins that can sequester Fe are all potential candidates for plant tolerance to excess iron through regulated storage and compartmentalization (Figure 6).

In *Arabidopsis* , apoplasmic Fe is mostly found within the stele [269], suggesting that compartmentalization within the stele could restrict excess Fe from reaching the shoot during transportation towards the aerial parts. Fe^{2+} decreases could also occur in association with an

Figure 6. Iron transport in rice. Fe is taken up into the symplast by transporters in the epidermis (OsIRT1, OsNRAMP, OsZIPL1 and OsYSLs). Proteins encoded by bHLH, IRUNLP1 and IRT2 likely regulate the activities of the above transporters. Radial oxygen loss into the rhizosphere through aerenchyma cells detoxifies part of the excess iron forming insoluble Fe^{3+} at the root surfaces, a process referred to as exclusion. Excess Fe^{2+} travels through the symplastic space to the vasculature, bypassing the waxy Casparian strip on the endodermis. Prior to reaching the xylem, excess iron is retained in the root cell vacuoles, mitochondria and probably detoxified by organic acids within the root cells. Transport into the xylem is mediated by putative chelate effluxers: FRDL1, OsYSLs, TOM1, OsIRT1, PEZ1 and FPN1. In the xylem, iron is carried to the shoot through the transpiration stream either in the form of Fe^{3+} or in both Fe^{3+} and Fe^{2+} forms, and unloaded into the shoot, most likely by YSLs, FRO1 and OsIRT1 proteins. Within the phloem, the rate at which NA, DMA and ITP are synthesized, the kinetic stability of the complexes formed and the oxido-reduction system likely determines the iron speciation. Enzymes involved in NA, DMA and ITP synthesis, including OsIRO2, Os-NAS1, NAAT1 and DMAS1, likely play a significant role in determining iron loading into the phloem. Genes encoding for putative iron effluxers from the phloem to storage organs (VIT, OsNRAMP, HMA3, MTP1, ENA, MIT1, ATM1) are co-regulated with IREG2/FPN2 and YSLs to limit potentially toxic iron in the cytosol, by compartmentalizing in the vacuoles, mitochondria, chloroplast and other non-characterized intracellular vesicles. In the chloroplasts, Fe excess probably promotes NO production. NO is probably involved in activation of the transcription factor (TF) cascades responsible for the regulation of Fe uptake, homeostasis and for the tuning of cellular metabolism, including increased synthesis of ferritins and betalains in chloroplasts and frataxins in the mitochondria. Because NO also triggers the synthesis of ROS, heme biosynthesis likely occurs to compartmentalize excess iron and to limit NO production. Alongside heme biosynthesis, the potent antioxidant system involving SOD and APX probably scavenge and detoxify the excess ROS. Also presented are targets of iron utilization, which could reduce iron overload. This includes synthesis of ferrochelatase (FC) for heme biosynthesis, mitochondrial iron-sulphur cluster (ISC) and plastid-localized sulphur utilization factors (SUF).

alkalization of apoplastic pH, which reduces Fe^{2+} mobility and chemical stability [269]. Alkalization has been reported to be modulated by ethylene [270], suggesting additional role

of ethylene in regulating Fe^{2+} besides its role in aerenchyma formation. Tissue tolerance of Fe toxicity is mediated by detoxification of free radicals. In rice, expression of several genes involved in oxidative stress control, including peroxidases, glutathione transferase (GST) and cytochromes, was upregulated in roots and shoots in response to excess Fe [271]. Similar trends were observed at the protein and enzymatic activity levels of the same genes. Excess iron was reported to induce the activity of superoxide dismutase (SOD) and ascorbate peroxidase (APX) in the leaf sheath and laminae, respectively, in a tolerant variety from *Oryza glaberrima* [272]. The activity of glutathione reductase and peroxidase (POD) was also reported to increase in rice leaf segments exposed to excess iron [273]. Fang et al. [274] also showed that lipid peroxidation resulting from Fe toxicity was inhibited by free radical scavengers such as mannitol and GSH. Moreover, the differential expression of anti-oxidant enzyme activities (SOD, APX, CAT, GR and DHR) was observed between rice varieties contrasting in tolerance of Fe toxicity [275].

Several genetic studies also reflect that iron toxicity tolerance is a complex quantitative trait controlled by a large number of rather small effect quantitative trait loci (QTLs), indicating the involvement of multiple tolerance mechanisms. For instance, Wu et al. [276] identified QTLs for leaf bronzing and shoot dry weight on chromosome 1 and 8, explaining 10–32% of the phenotypic variation. Interestingly, QTLs associated with enzymatic activity of anti-oxidants in rice leaves were detected in the same region [277]. Similarly, Fukuda et al. [278] detected a region on chromosome 3 responsible for high shoot iron content in a susceptible variety, which co-localize with the QTL previously identified by Shimizu et al. [279] for the same trait. Co-localization of most of these QTLs was captured in an integrative genetic map reflecting mapping studies from different conditions of Fe toxicity [277], which substantiates on recurrent chromosomal regions identified in several QTL studies.

A major limitation of iron toxicity tolerance studies, however, is that most of the QTLs associated with iron toxicity tolerance have not been furthered to cloning of tolerance genes. It is thus critical to devote some effort to fine-map the few, but consistent QTLs mentioned herein in order to increase precision and accelerate candidate gene identification. Subsequently, functional validation of several genes identified in microarray studies will need to be explored. Exploring allelic variation of these genes in contrasting genotypes and evaluating the promising alleles in well designed and efficient phenotyping experiments would provide a basis for their use in marker-assisted breeding (MAB) for Fe-toxicity tolerance.

8. Conclusions and perspectives

In this chapter, we have attempted to present the recent advances in crop tolerance to abiotic stresses. Various strategies used by plants to counteract stress, and some success in identifying genomic regions associated with plant tolerance is presented. Interestingly, plants have evolved common regulatory networks in response to abiotic stresses. For instance, drought, salt and cold stress induce calcium influx to activate the downstream second messengers to yield different or similar responses. Calcium influx channels at the membrane (e.g., the recently

reported *hyper-osmolality induced [Ca $^{2+}$] increases 1* (*OSCA1*) from *Arabidopsis thaliana* that is gated by hyper-osmotic stress [280]) act in concert with the membrane-located NADPH-oxidase Respiratory burst oxidase Homolog (RboH), generating apoplastic ROS. Intracellular transduction is conveyed by calcium-binding proteins (e.g., CBLs/CIPKs, CDPKs and calcineurins), a MAP-Kinase cascade and phytohormones (e.g., ABA, ET, JA and SA), which apparently act as integrators of early signals. Depending on the relative temporal patterns of these upstream signals, the activity of TFs and their interacting proteins will decipher specific combinations of genes required to be expressed to boost enzymatic or protein reaction levels necessary to counter the stress perceived. These proteins largely contribute to adaptive response in most plants, e.g., production of compatible osmolytes that helps to reinstall turgidity during drought and synthesis of LEA proteins that prevent protein precipitation. Other examples include chelation/sequestering of ions into cellular compartments in response to toxic elements, induction of anti-oxidative enzymes, induction of molecular chaperones and adaptive regulation of plant hormones. These adaptive strategies and the molecular components involved provide potential molecular genetic targets for enhancing abiotic resistance in crops.

However, many challenges still lie ahead. For example, the regulation of signaling cascades, especially how plants can discriminate the signaling components, and even their specific combinations, to activate specific downstream biological processes for a given stress. A frequent manifestation has been the case of ethylene controversial role in abiotic stress response. Whether the negative regulations associated with ethylene represent a plant strategic mechanism to prime the subsequent useful reaction remains to be confirmed. Also, temporal and specific differences in activation of upstream signaling components will need to be explored to help in identifying molecular components essentially required to counter a given stress. Moreover, the specific downstream components for which much of the studies have been conducted, e.g. transcription factors, transmembrane proteins, transporters, enzymes for osmolyte biosynthesis, hormonal regulators, ROS scavengers and other traits that have been shown to play major roles in plant response to stress, will need classification according to their aptitude and functional significance in response to a given abiotic stress. Morpho-physiological traits associated with stress tolerance would also substantially reinforce the successes in molecular biology if addressed to a greater extent. The use of models for predicting gene effects, particularly when combining multiple traits, will also find greater application in dissecting G × E interactions and will help breeders to improve target varieties. Thus, there is need to integrate molecular tools with precise high-throughput phenotyping and biochemical analysis to confirm the consistency of various molecular findings, and to realize the full benefits of molecular biology in selecting genotypes that are stably tolerant under a given stress, considering the interaction with various environments. Here, we emphasize stresses that have been commonly reported in literature, which would provide a basis for understanding other minor stresses. We also refer to the chapter on biotic stresses and the numerous interactions in signaling pathways and expressions of resistance and tolerance on molecular level towards abiotic and biotic stress in plants. Additional background information can also be found in excellent reviews and references therein.

Acknowledgements

Author details

Geoffrey Onaga[1] and Kerstin Wydra[2*]

*Address all correspondence to: kerstin.wydra@fh-erfurt.de

1 National Crops Resources Research Institute, Kampala, Uganda

2 Erfurt University of Applied Sciences, Faculty of Landscape Architecture, Horticulture and Forestry, Erfurt, Germany

References

[1] Acquaah G. Principles of plant genetics and breeding. 2nd ed. Oxford, UK: Wiley; 2012. 20 pp.

[2] Dos Reis SP, Lima AM, de Souza CRB. Recent molecular advances on downstream plant responses to abiotic stress. International Journal of Molecular Sciences. 2012;13:8628–8647. DOI: 10.3390/ijms13078628

[3] Flowers TJ, Yeo AR. Breeding for salinity resistance in crop plants—where next? Australian Journal of Plant Physiology. 1995;22:875–884. DOI: 10.1071/PP9950875

[4] Colmer TD, Pedersen O. Underwater photosynthesis and respiration in leaves of submerged wetland plants: gas films improve CO_2 and O_2 exchange. New Phytologist. 2008;177:918–926. DOI: 10.1111/j.1469-8137.2007.02318.x

[5] Mittler R. Abiotic stress, the field environment and stress combination. Trends Plant Science. 2006;11:15–19. DOI: 10.1016/j.tplants.2005.11.002

[6] Fujita M, Fujita Y, Noutoshi Y, Takahashi F, Narusaka Y, Yamaguchi-Shinozaki K, Shinozaki K. Crosstalk between abiotic and biotic stress responses: a current view from the points of convergence in the stress signaling networks. Current Opinion in Plant Biology. 2006;9:436–442. DOI: 10.1016/j.pbi.2006.05.014

[7] Chinnusamy V, Schumaker K, Zhu JK. Molecular genetic perspectives on cross-talk and specificity in abiotic stress signaling in plants. Journal of Experimental Botany. 2004;55:225–236. DOI: 10.1093/jxb/erh005

[8]   Krasensky J, Jonak C. Drought, salt, and temperature stress-induced metabolic rearrangements and regulatory networks. Journal of Experimental Botany. 2012;63:1593–1608. DOI: 10.1093/jxb/err460

[9]   Ramakrishna A, Ravishankar GA. Influence of abiotic stress singles on secondary metabolites in plants. Plant Signaling Behavior. 2011;6:1720–1731. DOI: 10.4161/psb.6.11.17613

[10]   Zhu QH, Zheng XM, Luo JC, Gaut BS, Ge S. Multilocus analysis of nucleotide variation of Oryza sativa and its wild relatives: severe bottleneck during domestication of rice. Molecular Biology and Evolution. 2007;24:875–888. DOI: 10.1093/molbev/msm005

[11]   Buckler ES, Thornsberry JM, Kresovich S. Molecular diversity, structure and domestication of grasses. Genetical Research. 2001;77:213–218. DOI: 10.1017/S0016672301005158

[12]   Haudry A, Cenci A, Ravel C, Bataillon T, Brunel D, Poncet C, Hochu I, Poirier S, Santoni S, Glémin S, David J. Grinding up wheat: a massive loss of nucleotide diversity since domestication. Molecular Biology Evolution. 2007;24:1506–1517. DOI: 10.1093/molbev/msm077

[13]   Kim JY, Mahé AB, rangeon J, Prioul JL. A maize vacuolar invertase, IVR2, is induced by water stress. Organ/tissue specificity and diurnal modulation of expression. Plant Physiology. 2000;124:71–84. DOI: 10.1104/pp.124.1.71

[14]   Liu F, Jensen CR, Andersen MN. Drought stress effect on carbohydrate concentration in soybean leaves and pods during early reproductive development: its implication in altering pod set. Field Crops Research. 2004;86(1):1–13. DOI: 10.1016/S0378-4290(03)00165-5

[15]   Hirayama T, Shinozaki K. Research on plant abiotic stress responses in the post genome era: past, present and future. The Plant Journal. 2010;61(6):1041–1052. DOI: 10.1111/j.1365-313X.2010.04124.x

[16]   Umezawa T, Nakashima K, Miyakawa T, Kuromori T, Tanokura M, Shinozaki K, Yamaguchi-Shinozaki K. Molecular basis of the core regulatory network in ABA responses: sensing, signaling and transport. Plant & Cell Physiology. 2010;51:1821–1839. DOI: 10.1093/pcp/pcq156

[17]   Boursiac Y, Leran S, Corratge-Faillie C, Gojon A, Krouk G, Lacombe B. ABA transport and transporters. Trends in Plant Science 2013;18:325–333. DOI: http://dx.doi.org/10.1016/j.tplants.2013.01.007

[18]   Golldack D, Li C, Mohan H, Probst N. Tolerance to drought and salt stress in plants: unraveling the signaling networks. Frontiers in Plant Science. 2014;5:151. DOI: 10.3389/fpls.2014.00151

[19]   Vlad F, Rubio S, Rodrigues A, Sirichandra C, Belin C, Robert N, Leung J, Rodriguez PL, Laurière C, Merlot S. Protein phosphatases 2C regulate the activation of the Snf1-

related kinase OST1 by abscisic acid in *Arabidopsis*. The Plant Cell 2009;21:3170–3184. DOI: 10.1105/tpc.109.069179

[20] Campo S, Baldrich P, Messeguer J, Lalanne E, Coca M, San Segundo B. Overexpression of a calcium-dependent protein kinase confers salt and drought tolerance in rice by preventing membrane lipid peroxidation. Plant Physiology. 2014;165:688–704. DOI: 10.1104/pp.113.230268

[21] Ouyang SQ, Liu YF, Liu P, Lei G, He SJ, Ma B, Zhang WK, Zhang JS, Chen SY. Receptor-like kinase OsSIK1 improves drought and salt stress tolerance in rice (Oryza sativa) plants. The Plant Journal. 2010;62:316–329. DOI: 10.1111/j.1365-313X. 2010.04146.x

[22] Osakabe Y, Yamaguchi-Shinozaki K, Shinozaki K, Lam-Son Phan T. ABA control of plant macro element membrane transport systems in response to water deficit and high salinity. New Phytologist. 2014;202:35–49. DOI: 10.1111/nph.12613

[23] Xu C, Jing R, Mao X, Jia X, Chang X. A wheat (Triticum aestivum) protein phosphatase 2A catalytic subunit gene provides enhanced drought tolerance in tobacco. Annals of Botany. 2007;99:439–450. DOI: 10.1093/aob/mcl285

[24] Banerjee A, Roychoudhury A. WRKY proteins: signaling and regulation of expression during abiotic stress responses. The Scientific World Journal. 2015;2015:807560. DOI: 10.1155/2015/807560

[25] Ding Z, Li S, An X, Liu X, Qin H, Wang D. Transgenic expression of MYB15 confers enhanced sensitivity to abscisic acid and improved drought tolerance in *Arabidopsis thaliana*. Journal of Genetics and Genomics. 2009;36:17–29. DOI: 10.1016/ S1673-8527(09)60003-5

[26] Rushton DL, Tripathi P, Rabara RC, Lin J, Ringler P, Boken AK, Langum TJ, Smidt L, Boomsma DD, Emme NJ, Chen X, Finer JJ, Shen QJ, Rushton PJ. WRKY transcription factors: key components in abscisic acid signaling. Plant Biotechnology Journal. 2012;10:2–11. DOI: 10.1111/j.1467-7652.2011

[27] Zhang G, Chen M, Li L, Xu Z, Chen X, Guo J, Ma Y. Overexpression of the soybean GmERF3 gene, an AP2/ERF type transcription factor for increased tolerances to salt, drought, and diseases in transgenic tobacco. Journal of Experimental Botany. 2009;60:3781–3796. DOI: 10.1093/jxb/erp214

[28] Liu S, Wang X, Wang H, Xin H, Yang X, Yan J Li J, Tran LS, Shinozaki K, Yamaguchi-Shinozaki K, Qin F. Genome-wide analysis of ZmDREB genes and their association with natural variation in drought tolerance at seedling stage of Zea mays L. PLoS Genetics. 2013;9(9):e1003790. DOI: 10.1371/journal.pgen.1003790

[29] Ji X, Liu G, Liu Y, Zheng L, Nie X, Wang Y. The bZIP protein from Tamarix hispida, ThbZIP1, is ACGT elements binding factor that enhances abiotic stress signaling in transgenic *Arabidopsis*. BMC Plant Biology. 2013;13. DOI: 10.1186/1471-2229-13-151

[30] Ying S, Zhang DF, Fu J, Shi YS, Song YC, Wang TY, and Li Y. Cloning and characterization of a maize bZIP transcription factor, ZmbZIP72, confers drought and salt tolerance in transgenic *Arabidopsis*. Planta. 2012;235:253–266. DOI: 10.1007/s00425-011-1496-7

[31] Fujita M, Fujita Y, Maruyama K, Seki M, Hiratsu K, Ohme-Takagi M, Tran LS, Yamaguchi-Shinozaki K, Shinozaki K. A dehydration-induced NAC protein, RD26, is involved in a novel ABA-dependent stress-signaling pathway. The Plant Journal. 2004;39:863–876. DOI: 10.1111/j.1365-313X.2004.02171.x

[32] Tran LS, Nakashima K, Sakuma Y, Simpson SD, Fujita Y, Maruyama K, Fujita M, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. Isolation and functional analysis of *Arabidopsis* stress-inducible NAC transcription factors that bind to a drought responsive cis-element in the early responsive to dehydration stress 1 promoter. The Plant Cell 2004;16:2481–2498. DOI: 10.1105/tpc.104.022699

[33] Jeong JS, Kim YS, Baek KH, Jung H, Ha SH, Do Choi Y, Kim M, Reuzeau C, Kim JK. Root-specific expression of OsNAC10 improves drought tolerance and grain yield in rice under field drought conditions. Plant Physiology. 2010;153:185–197. DOI: 10.1104/pp.110.154773

[34] Jeong JS, Kim YS, Redillas MC, Jang G, Jung H, Bang SW Choi YD, Ha SH, Reuzeau C, Kim JK. OsNAC5 overexpression enlarges root diameter in rice plants leading to enhanced drought tolerance and increased grain yield in the field. Plant Biotechnology Journal. 2013;11:101–114. DOI: 10.1111/pbi.12011

[35] Palmeros-Suárez PA, Massange-Sánchez JA, Martínez-Gallardoa NA, Montero-Vargas JM, Gómez-Leyva JF, Délano-Frier JP. The overexpression of an Amaranthus hypochondriacus NF-YC gene modifies growth and confers water deficit stress resistance in *Arabidopsis*. Plant Science. 2015;240:25–40. DOI: 10.1016/j.plantsci.2015.08.010

[36] Chen M, Zhao Y, Zhuo C, Lu S, Guo Z. Overexpression of a NF-YC transcription factor from Bermuda grass confers tolerance to drought and salinity in transgenic rice. Plant Biotechnology Journal. 2015;13:482–491. DOI: 10.1111/pbi.12270

[37] Fang Y, Liao K, Du H, Xu Y, Song H, Li X, Xiong LA. Stress-responsive NAC transcription factor SNAC3 confers heat and drought tolerance through modulation of reactive oxygen species in rice. Journal of Experimental Botany. 2015;25:3472–3490. DOI: 10.1093/jxb/erv386

[38] Djafi N, Vergnolle C, Cantrel C, Wietrzyñski W, Delage E, Cochet F Puyaubert J, Soubigou-Taconnat L, Gey D, Collin S, Balzergue S, Zachowski A, Ruelland E. The *Arabidopsis* DREB2 genetic pathway is constitutively repressed by basal phosphoinositide-dependent phospholipase C coupled to diacylglycerol kinase. Frontiers in Plant Science. 2013;4:307. DOI: 10.3389/fpls.2013.00307

[39] Zhai SM, Gao Q, Xue HW, Sui ZH, Yue GD, Yang AF, Zhang JR. Overexpression of the phosphatidylinositol synthase gene from Zea mays in tobacco plants alters the

membrane lipids composition and improves drought stress tolerance. Planta. 2012;235:69–84. DOI: 10.1007/s00425-011-1490-0

[40] Zhu J-K. Salt and Drought Stress Signal Transduction in Plants. Annual review of plant biology. 2002;53:247–273. DOI: 10.1146/annurev.arplant.53.091401.143329.

[41] Jacob T, Ritchie S, Assmann SM, Gilroy S. Abscisic acid signal transduction in guard cells is mediated by phospholipase D activity. Proceedings of the National Academy of Sciences of the United States of America. 1999;96:12192–97. DOI: 10.1073/pnas. 96.21.12192

[42] Seki M, Narusaka M, Ishida J, Nanjo T, Fujita M Oono Y, Kamiya A, Nakajima M, Enju A, Sakurai T, Satou M, Akiyama K, Taji T, Yamaguchi-Shinozaki K, Carninci P, Kawai J, Hayashizaki Y, Shinozaki K. Monitoring the expression profiles of 7000 *Arabidopsis* genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. The Plant Journal. 2002;31:279–292. DOI: 10.1046/j.1365-313X. 2002.01359.x

[43] Xu Y, Wei H, Juhua L, Jianbin Z, Caihong J, Hongxia M, Biyu X, Zhiqiang J. A banana aquaporin gene, MaPIP1;1, is involved in tolerance to drought and salt stresses. BMC Plant Biology. 2014;14:59. DOI: 10.1186/1471-2229-14-59

[44] Li H, Yao W, Fu Y, Li S, Guo Q. De Novo Assembly and discovery of genes that are involved in drought tolerance in Tibetan *Sophora moorcroftiana*. PLoS ONE 2015;10(1):e111054. DOI: 10.1371/journal.pone.0111054

[45] Li J, Ban L, Wen H, Wang Z, Dzyubenko N, Chapurin V, Gao H, Wang X. An aquaporin protein is associated with drought stress tolerance. Biochemical Biophysical Research Communication. 2015;459:208–213. DOI: 10.1016/j.bbrc.2015.02.052

[46] Osakabe Y, Arinaga N, Umezawa T, Katsura S, Nagamachi K, Tanaka H, Ohiraki H, Yamada K, Seo SU, Abo M Yoshimura E, Shinozaki K, Yamaguchi-Shinozaki K. Osmotic stress responses and plant growth controlled by potassium transporters in *Arabidopsis*. The Plant Cell 2013;25:609–624. DOI: 10.1105/tpc.112.105700

[47] Vicente-Agullo F, Rigas S, Desbrosses G, Dolan L, Hatzopoulos P, Grabov A. Potassium carrier TRH1 is required for auxin transport in *Arabidopsis* roots. The Plant Journal. 2004;40:523–535. DOI: 10.1111/j.1365-313X.2004.02230.x

[48] Zhang C, Zhang L, Zhang S, Zhu S, Wu P, Chen Y, Li M, Jiang H, Wu G. Global analysis of gene expression profiles in physic nut (Jatropha curcas L.) seedlings exposed to drought stress. BMC Plant Biology. 2015;15:17. DOI: 10.1371/journal.pone.0097878

[49] Wisniewski ME, Bassett CL, Renaut J, Farrell R, Tworkoski T, Artlip TS. Differential regulation of two dehydrin genes from peach (*Prunus persica*) by photoperiod, low temperature and water deficit. Tree Physiology. 2006;26:575–584. DOI: 10.1093/treephys/26.5.575

[50] Yoshimura K, Masuda A, Kuwano M, Yokota A, Akashi K. Programmed proteome response for drought avoidance/tolerance in the root of a C3 xerophyte (wild water-

melon) under water deficits. Plant Cell Physiology. 2008;49:226–241. DOI: 10.1093/pcp/pcm180

[51] Degenhardt B, Gimmler H. Cell wall adaptations to multiple environmental stresses in maize roots. Journal of Experimental Botany. 2000;51:595–603. DOI: 10.1093/jexbot/51.344.595

[52] Hura T, Hura K, Dziurka K, Ostrowska A, Ba⊚czek-Kwinta R, Grzesiak M. An increase in the content of cell wall-bound phenolics correlates with the productivity of triticale under soil drought. Journal of Plant Physiology. 2012;169:1728–1736.61. DOI: 10.1016/j.jplph.2012.07.012

[53] Alam I, Sharmin SA, Kim KH, Yang JK, Choi MS, Lee BH. Proteome analysis of soybean roots subjected to short-term drought stress. Plant and Soil. 2010;333:491–505. DOI: 10.1007/s11104-010-0365-7

[54] Mohammadi PP, Moieni A, Komatsu S. Comparative proteome analysis of drought-sensitive and drought-tolerant rapeseed roots and their hybrid F1 line under drought stress. Amino Acids. 2012;43:2137–2152. DOI: 10.1007/s00726-012-1299-6

[55] Hu H, Xiong L. Genetic engineering and breeding of drought-resistant crops. Annual Reviews of Plant Biology. 2014;65:715–741. DOI: 10.1146/annurev-arplant-050213-040000

[56] Asthir B. Protective mechanisms of heat tolerance in crop plants. Journal of Plant Interactions. 2015;10:202–210. DOI: 10.1080/17429145.2015.1067726

[57] Hasanuzzaman M, Nahar K, Alam MdM, Roychowdhury, Rajib, and Fujita M. Physiological, biochemical, and molecular mechanisms of heat stress tolerance in plants. International Journal of Molecular Science. 2013;14(5):9643–9684. DOI: 10.3390/ijms14059643

[58] Saidi Y, Peter M, Finka A, Cicekli C, Vigh L, Goloubinoff P. Membrane lipid composition affects plant heat sensing and modulates Ca^{2+}-dependent heat shock response. Plant Signaling and Behavior. 2010;5:1530–1533. DOI: 10.4161/psb.5.12.13163

[59] Li Z, Yue HY, Xing D. MAP Kinase 6-mediated activation of vacuolar processing enzyme modulates heat shock-induced programmed cell death in *Arabidopsis*. New Phytologist. 2012;195:85–96. DOI: 10.3389/fpls.2015.00234

[60] Suzuki N, Sejima H, Tam R, Schlauch K, Mittler R. Identification of the MBF1 heat-response regulon of *Arabidopsis thaliana*. The Plant Journal. 2011;66:844–851. DOI: 10.1111/j.1365-313X.2011.04550.x

[61] Miller G, Schlauch K, Tam R, Cortes D, Torres MA, Shulaev V, Dangl JL, Mittler R. The plant NADPH oxidase RBOHD mediates rapid systemic signaling in response to diverse stimuli. Science Signaling. 2009;2. DOI: 10.1126/scisignal.2000448

[62] Liu HT, Gao F, Li GL, Han JL, Liu DL, Sun DY, Zhau RG. The calmodulin-binding protein kinase 3 is part of heat-shock signal transduction in *Arabidopsis thaliana*. The Plant Journal. 2008;55:760–773. DOI: 10.1111/j.1365-313X.2008.03544.x

[63] Baniwal SK, Bharti K, Chan KY, Fauth M, Ganguli A, Kotak S, Mishra SK, Nover L, Port M, Scharf KD, Tripp J, Weber C, Zielinski D, von Koskull-Döring P. Heat stress response in plants: a complex game with chaperones and more than twenty heat stress transcription factors. Journal of Biosciences. 2004;29:471–487.

[64] Suzuki N, Bajad S, Shuman J, Shulaev V, Mittler R. The transcriptional co-activator MBF1c is a key regulator of thermotolerance in *Arabidopsis thaliana*. The Journal of Biological Chemistry. 2008, 283:9269–9275. DOI: 10.1074/jbc.M709187200

[65] Khan MIR, Iqbal N, Masood A, Per TS, Khan NA. Salicylic acid alleviates adverse effects of heat stress on photosynthesis through changes in proline production and ethylene formation. Plant Signaling & Behavior. 2013; 8(11):e26374. DOI: 10.4161/psb.26374.

[66] Meiri D, Breiman A. *Arabidopsis* ROF1 (FKBP62) modulates thermotolerance by interacting with HSP90.1 and affecting the accumulation of HsfA2-regulated sHSPs. The Plant Journal 2009;59:387–399. DOI: 10.1111/j.1365-313X.2009.03878.x

[67] Bokszczanin KL, Fragkostefanakis S. Perspectives on deciphering mechanisms underlying plant heat stress response and thermotolerance. Frontiers in Plant Science. 2013;4:315– 335. DOI: 10.3389/fpls.2013.00315

[68] Sidrauski C, Walter P. The transmembrane kinase Ire1p is a site-specific endonuclease that initiates mRNA splicing in the unfolded protein response. Cell 1997;90:1031–1039. DOI: 10.1016/S0092-8674(00)80369-4

[69] Deng Y, Humbert S, Liu JX, Srivastava R, Rothstein SJ, Howell SH. Heat induces the splicing by IRE1 of a mRNA encoding a transcription factor involved in the unfolded protein response in *Arabidopsis*. Proceedings of the National Academy of Sciences of the United States of America. 2011;108:7247–7252. DOI: 10.1073/pnas.1102117108

[70] Sugio A, Dreos R, Aparicio F, Maule AJ. The cytosolic protein response as a subcomponent of the wider heat shock response in *Arabidopsis*. The Plant Cell 2009;21:642–654. DOI: 10.1105/tpc.108.062596

[71] Liu J, Sun N, Liu M, Liu J, Du B, Wang X, Qi X. An autoregulatory loop controlling *Arabidopsis* HsfA2 expression: role of heat shock-induced alternative splicing. Plant Physiology. 2013;162:512–521. DOI: 10.1104/pp.112.205864

[72] Sato H, Mizoi J, Tanaka H, Maruyama K, Qin F, Osakabe Y, Morimoto K. Ohori T, Kusakabe K, Nagata M, Shinozaki K, Yamaguchi-Shinozakia K. *Arabidopsis* DPB3-1, a DREB2A interactor, specifically enhances heat stress-induced gene expression by forming a heat stress-specific transcriptional complex with NF-Y subunits. The Plant Cell 2014;26:4954–4973. DOI: 10.1105/tpc.114.132928

[73] Gao HB, Brandizzi F, Benning C, Larkin RM. A membrane-tethered transcription factor defines a branch of the heat stress response in *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences of the United States of America. 2008;105:16398–16403. DOI: 10.1073/pnas.0808463105

[74] Li SJ, Fu QT, Huang WD, Yu DQ. Functional analysis of an *Arabidopsis* transcription factor WRKY25 in heat stress. Plant Cell Reports. 2009;28:683–693. DOI: 10.1007/s00299-008-0666-y

[75] Wu XL, Shiroto Y, Kishitani S, Ito Y, Toriyama K. Enhanced heat and drought tolerance in transgenic rice seedlings overexpressing OsWRKY11 under the control of HSP101 promoter. Plant Cell Reports. 2009;28:21–30. DOI: 10.1007/s00299-008-0614-x

[76] Proveniers MCG, van Zanten M. High temperature acclimation through PIF4 signaling. Trends in Plant Science. 2013;18:59–64. DOI: 10.1016/j.tplants.2012.09.002

[77] Ma D, Li X, Guo X, Chu J, Fang S, YanC, Noel JP, Liu H. Cryptochrome 1 interacts with PIF4 to regulate high temperature-mediated hypocotyl elongation in response to blue light. Proceedings of the National Academy of Sciences of the United States of America. 2015;113:224–229. DOI: 10.1073/pnas.1511437113

[78] Erkina TY, Zou Y, Freeling S, Vorobyev VI, Erkine AM. Functional interplay between chromatin remodeling complexes RSC, SWI/SNF and ISWI in regulation of yeast heat shock genes. Nucleic Acids Research. 2010;38:1441–1449. DOI: 10.1093/nar/gkp1130

[79] Rasheed R, Wahid A, Farooq M, Basra SMA. Role of proline and glycinebetaine in improving heat tolerance in sprouting sugarcane (*Saccharum* sp.) buds. Plant Growth Regulation.2011;65:35–45. DOI: 10.1007/s10725-011-9572-3

[80] Kapoor D, Sharma R, Handa N, Kaur H, Rattan A, Yadav P, Gautam V, Kaur R, Bhardwaj R. Redox homeostasis in plants under abiotic stress. Role of electron carriers, energy metabolism mediators and proteinaceous thiols. Frontiers in Environmental Science. 2015;3:13. DOI: 10.3389/fenvs.2015.00013

[81] Priest HD, Fox SE, Rowley ER, Murray JR, Michael TP, Mockler TC. Analysis of global gene expression in *Brachypodium distachyon* reveals extensive network plasticity in response to abiotic stress. PLoS ONE. 2014;9:e87499. DOI: 10.1371/journal.pone.0087499

[82] Zhang Y, Xu L, Zhu X, Gong Y, Xiang F, Sun X, Liu L. Proteomic analysis of heat stress response in leaves of radish (*Raphanus sativus* L.). Plant Molecular Biology Reports. 2013, 31:195–203. DOI: 10.1007/s11105-012-0486-7

[83] Chen X, Zhang W, Zhang B, Zhou J, Wang Y, Yang Q, Ke Y, He H. Phosphoproteins regulated by heat stress in rice leaves. Proteome Science. 2011;9: 37. DOI: 10.1186/1477-5956-9-37

[84] Li W, Wei Z, Qiao Z, Wu Z, Cheng L, Wang Y. Proteomics analysis of alfalfa response to heat stress. PLoS ONE. 2013a;8:e82725. DOI: 10.1371/journal.pone.0082725

[85] Du HM, Wang ZL, Yu WJ, Liu YM, Huang BR. Differential metabolic responses of perennial grass *Cynodon transvaalensis × Cynodon dactylon* (C4) and *Poa pratensis* (C3) to heat stress. Physiologia Plantarum. 2011;141:51–264. DOI: 10.1111/j.1399-3054.2010.01432.x

[86] Yu J, Du H, Xu M, Huang B. Metabolic responses to heat stress under elevated atmospheric CO_2 concentration in a cool-season grass species. Journal of the American Society for Horticultural Science. 2012b;137:221–228. DOI: 10.1111/j.1399-3054.2010.01432.x

[87] Goswami S, Kumar RR, Rai RD. Heat-responsive microRNAs regulate the transcription factors and heat shock proteins in modulating thermostability of starch biosynthesis enzymes in wheat (*Triticum aestivum* L.) under the heat stress. Australian Journal of Crop Science. 2014;8:697—705.

[88] Hozain M, Abdelmageed H, Lee J, Kang M, Fokar M, Allen RD, Holaday AS. Expression of AtSAP5 in cotton up-regulates putative stress-responsive genes and improves the tolerance to rapidly developing water deficit and moderate heat stress. Journal of Plant Physiology. 2012;169:1261—1270. DOI: 10.1016/j.jplph.2012.04.007

[89] Sasaki K, Saito T, Lämsä M, Oksman-Caldentey KM, Suzuki M, Ohyama K, Muranaka T, Ohara K, Yazaki K. Plants utilize isoprene emission as a thermo tolerance mechanism. Plant Cell Physiology. 2007;48:1254—126. DOI: 10.1093/pcp/pcm104

[90] Macková H, Hronková M, Dobrá J, Tureck*á V, Novák O, Lubovská Z, Motyka V, Haisel D, Hájek Prásil ITT, Gaudinová A, Storchová H, Ge E, Werner T, Schmülling T, Vanková R. Enhanced drought and heat stress tolerance of tobacco plants with ectopically enhanced cytokinin oxidase/dehydrogenase gene expression. Journal of Experimental Botany. 2013;64:2805—2815. DOI: 10.1093/jxb/ert131

[91] Feng L, Wang KJ, Li Y, Tan Y, Kong J, H L, Y L, Zhu Y. Over expression of SBPase enhances photosynthesis against high temperature stress in transgenic rice plants. Plant Cell Reports. 2007;26:1635—1646. DOI: 10.1007/s00299-006-0299-y

[92] Chinnusamy V, Zhu J, Zhu JK. Cold stress regulation of gene expression in plants. Trends Plant Science. 2007;12:444–451. DOI: 10.1016/j.tplants.2007.07.002

[93] Sanghera GS, Wani SH, Hussain W, Singh NB. Engineering cold stress tolerance in crop plants. Current Genomics. 2011;12:30–43. DOI: 10.2174/138920211794520178

[94] Ye H, Du H, Tang N, Li X, Xiong L. Identification and expression profiling analysis of TIFY family genes involved in stress and phytohormone responses in rice. Plant Molecular Biology. 2009;71:291–305. DOI: 10.1007/s11103-009-9524-8

[95] Miura K, and Furumoto T. Cold signaling and cold response in plants. International Journal of Molecular Science. 2013;14(3):5312-5337. DOI: 10.3390/ijms14035312

[96] Knight MR, Campbell AK, Smith SM, Trewavas AJ. Transgenic plant aequorin reports the effects of touch and cold-shock and elicitors on cytoplasmic calcium. Nature. 1991;352(6335):524–6. DOI: 10.1038/352524a0

[97] Furuya T, Matsuoka D, Nanmori T. Membrane rigidification functions upstream of the MEKK1-MKK2-MPK4 cascade during cold acclimation in *Arabidopsis thaliana*. FEBS Letters. 2014;588:2025–2030. DOI: 10.1016/j.febslet.2014.04.032

[98] Francia E, Rizza F, Cattivelli L, Stanca AM, Galiba G, Toth B, Hayes PM, Skinner JS, Pecchioni N. Two loci on chromosome 5H determine low-temperature tolerance in a 'Nure' (winter) × 'Tremois' (spring) barley map. Theoretical Applied Genetics. 2004;108:670—680. DOI: 10.1007/s00122-003-1468-9

[99] Huang C, Ding S, Zhang H, Du H, An L. CIPK7 is involved in cold response by interacting with CBL1 in *Arabidopsis thaliana*. Plant Science. 2011;181:57–64. DOI: 10.1016/j.plantsci.2011.03.011

[100] Doherty CJ, Buskirk HAV, Myers SJ, Thomashow MF. Roles for *Arabidopsis* CAMTA transcription factors in cold-regulated gene expression and freezing tolerance. The Plant Cell. 2009;21:972–984. DOI: 10.1105/tpc.108.063958

[101] Chinnusamy V, Zhu J-K, Sunkar R. Gene Regulation during cold stress acclimation in plants. Methods in molecular biology. 2010;639:39–55. DOI: 10.1007/978-1-60761-702-0_3.

[102] Takahashi S, Murata N. How do environmental stresses accelerate photo-inhibition? Trends in Plant Science. 2008;13:178–182. DOI: 10.1016/j.tplants.2008.01.005

[103] Teige M, Scheikl E, Eulgem T, Doczi R, Ichimura K, Shinozaki K, Dangl JL, Hirt H. The MKK2 pathway mediates cold and salt stress signaling in *Arabidopsis*. Molecular Cell. 2004;15:141–152. DOI: 10.1016/j.molcel.2004.06.023

[104] Tarkowski ŁP, Van den Ende W. Cold tolerance triggered by soluble sugars: a multifaceted countermeasure. Frontiers in Plant Science. 2015;6. DOI: 10.3389/fpls.2015.00203

[105] Chinnusamy V, Ohta M, Kanrar S, Lee B H, Hong X, Agarwal M, Zhu JK. ICE1. A regulator of cold-induced transcriptome and freezing tolerance in *Arabidopsis*. Genes and Development. 2003;17:1043–1054. DOI: 10.1101/gad.1077503

[106] Liu L, Duan L, Zhang J, Zhang Z, Mi G, Ren H. Cucumber (Cucumis sativus L.) overexpressing cold-induced transcriptome regulator ICE1 exhibits changed morphological characters and enhances chilling tolerance. Scientia Horticulturae. 2010;124:29–33. DOI: 10.1016/j.scienta.2009.11.018

[107] Medina J, Catala R, Salinas J. The CBFs: three *Arabidopsis* transcription factors to cold acclimate. Plant Science. 2011;180:3–11. DOI: 10.1016/j.plantsci.2010.06.019

[108] Zeller G, Henz SR, Widmer CK, Sachsenberg T, Ratsch G, Weigel D, Laubinger S. Stress-induced changes in the *Arabidopsis thaliana* transcriptome analyzed using

whole-genome tiling arrays. The Plant Journal. 2009;58:1068–1082. DOI: 10.1111/j. 1365-313X.2009.03835.x

[109] Novillo F, Alonso JM, Ecker JR, Salinas J. CBF2/DREB1C is a negative regulator of CBF1/DREB1B and CBF3/DREB1A expression and plays a central role in stress toler-ance in *Arabidopsis*. Proceedings of the National Academy of Sciences of the United States of America. 2004;101:3985–3990. DOI: 10.1073/pnas.0303029101

[110] Novillo F, Medina J, Salinas J. *Arabidopsis* CBF1 and CBF3 have a different function than CBF2 in cold acclimation and define different gene classes in the CBF regulon. Proceedings of the National Academy of Sciences of the United States of America. 2007;104:21002–21007. DOI: 10.1073/pnas.0705639105

[111] Krasensky J, Jonak C. Drought, salt, and temperature stress-induced metabolic rear-rangements and regulatory networks. Journal of Experimental Botany. 2012;63:1593–1608. DOI: 10.1093/jxb/err460

[112] Lee DG, Ahsan N, Lee SH, Lee JJ, Bahk JD, Kang KY, Lee BH. Chilling stress-induced proteomic changes in rice roots. Journal of Plant Physiology. 2009;166:1–11. DOI: 10.1016/j.jplph.2008.02.001

[113] FAO. Land and Plant Nutrition Management Service [Internet]. 2008. Available from: http://wwwfaoorg/ag/agl/agll/spush [Accessed: 2016-05-03]

[114] Munns R, Tester M. Mechanisms of salinity tolerance. Annual Reviews in Plant Biol-ogy. 2008;59:651–681. DOI: 10.1146/annurev.arplant.59.032607.092911

[115] Shahbaz M, Ashraf M. Improving salinity tolerance in cereals. Critical Reviews in Plant Science. 2013;32:237–249. DOI: 10.1080/07352689.2013.758544

[116] Volkov V. Salinity tolerance in plants. Quantitative approach to ion transport starting from halophytes and stepping to genetic and protein engineering for manipulating ion fluxes Frontiers in Plant Science. 2015;6:873. DOI: 10.3389/fpls.2015.00873

[117] Hasegawa PM, Bressan RA, Zhu JK and HJ Bohnert. Plant cellular and molecular re-sponses to high salinity. Annual Reviews in Plant Biology. 2000;51:463–499. DOI: 10.1146/annurev.arplant.51.1.463

[118] Jung JKH, McCouch S. Getting to the roots of it: genetic and hormonal control of root architecture. Frontiers in Plant Science. 2013;4:186. DOI: 10.3389/fpls.2013.00186

[119] Wang H, Wu Z, Chen Y, Yang C, Shi D. Effects of salt and alkali stresses on growth and ion balance in rice (*Oryza sativa* L.). Plant, Soil and Environment. 2011;57:286–294.

[120] Deinlein U, Stephen AB, Horie T, Luo W, Xu G, Schroeder JI. Plant salt tolerance mechanisms. Trends in Plant Science. 2014;19:371–379. DOI: 10.1016/j.tplants. 2014.02.001

[121] Pandey N, Ranjan A, Pant P, Tripathi RK, Ateek F, Pandey HP, Patre UV, Sawant SV. CAMTA 1 regulates drought responses in *Arabidopsis thaliana*. BMC Genomics. 2013;14:216. DOI: 10.1186/1471-2164-14-216

[122] Weng H, Yoo CY, Gosney MJ, Hasegawa PM, Mickelbart MV. Poplar GTL1 is a Ca^{2+}/calmodulin-binding transcription factor that functions in plant water use efficiency and drought tolerance. PLoS ONE. 2012;7:e32925. DOI: 10.1371/journal.pone.0032925

[123] Yoo JH, Park CY, Kim JC, Heo WD, Cheong MS, Park HC, Kim MC, Moon BC, Choi MS, Kang YH, Lee JH, Kim HS, Lee SM, Yoon HW, Lim CO, Yun DJ, Lee SY, Chung WS, Cho MJ. Direct interaction of a divergent CaM isoform and the transcription factor, MYB2, enhances salt tolerance in *Arabidopsis*. Journal of Biological Chemistry. 2005;280:3697–3706. DOI: 10.1074/jbc.M408237200

[124] Liu C, Mao B, Ou S, Wang W, Liu L, Wu Y, Chu C, Wang X. OsbZIP71, a bZIP transcription factor, confers salinity and drought tolerance in rice. Plant Molecular Biology. 2014;84:19–36. DOI: 10.1007/s11103-013-0115-3

[125] Jiang Y, Deyholos MK. Functional characterization of *Arabidopsis* NaCl-inducible WRKY25 and WRKY33 transcription factors in abiotic stresses. Plant Molecular Biology. 2009;69:91–105. DOI: 10.1007/s11103-008-9408-3

[126] Yamaguchi-Shinozaki K, Shinozaki K. Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. In: Novartis Foundation Symposium 236—Rice Biotechnology: Improving Yield, Stress Tolerance and Grain Quality; February 2001; discussion 186-9; 236:176–186; DOI: 10.1002/9780470515778.ch13

[127] Cui MH, Yoo KS, Hyoung S, Nguyen HT, Kim YY, Kim HJ, Ok SH, Yoo SD, Shin JS. An *Arabidopsis* R2R3-MYB transcription factor, AtMYB20, negatively regulates type 2C serine/threonine protein phosphatases to enhance salt tolerance. FEBS Letters. 2013;587:1773–1778. DOI: 10.1016/j.febslet.2013.04.028

[128] Jiang Y, Yang B, Deyholos MK. Functional characterization of the *Arabidopsis* bHLH92 transcription factor in abiotic stress. Molecular Genetics and Genomics 2009;282:503–516. DOI: 10.1007/s00438-009-0481-3

[129] Chen Y, Zong J, Tan Z, Li L, Hu B, Chen C, Chen J, Liu J. Systematic mining of salinity-tolerant genes in Halophyte-*Zoysia matrella* through cDNA expression library screening. Plant Physiology and Biochemistry 2015b;89:44–52. DOI: 10.1016/j.plaphy.2015.02.007

[130] Hur J, Jung KH, Lee CH, An G. Stress-inducible OsP5CS2 gene is essential for salt and cold tolerance in rice. Plant Science 2004;167:417–426. DOI: 10.1016/j.plantsci.2004.04.009

[131] Ismail A, Takeda S, and Nick P. Life and death under salt stress: same players, different timing? Journal of Experimental Botany.2014;65:2963–2979. DOI: 10.1093/jxb/eru159

[132] Quintero FJ, Martinez-Atienza J, Villalta I, Jiang X, Kim WY, Ali Z, Fujii H, Mendoza I, Yun DJ, Zhu JK, and Pardo JM. Activation of the plasma membrane Na/H antiporter Salt-Overly Sensitive 1 (SOS1) by phosphorylation of an auto-inhibitory C-terminal domain. Proceedings of the National Academy of Sciences of the United States of America. 2011;108:2611–2616. DOI: 10.1073/pnas.1018921108

[133] Du W, Lin H, Chen S, Wu Y, Zhang J, Fuglsang AT, Palmgren MG, Wu W, Guo Y. Phosphorylation of SOS3-like calcium-binding proteins by their interacting SOS2-like protein kinases is a common regulatory mechanism in *Arabidopsis*. Plant Physiology. 2011;156:2235–2243. DOI: 10.1104/pp.111.173377

[134] Bassil E, Blumwald E. The ins and outs of intracellular ion homeostasis: NHX-type Cation/H+ transporters. Current Opinion in Plant Biology. 2014, 22:1–6. DOI: 10.1016/j.pbi.2014.08.002

[135] Yamaguchi T, Aharon GS, Sottosanto JB, Blumwald E. Vacuolar Na+/H+ antiporter cation selectivity is regulated by calmodulin from within the vacuole in a Ca2+ and pH-dependent manner. Proceedings of the National Academy of Sciences of the United States of America. 2005;102:16107–16112. DOI: 10.1073/pnas.0504437102

[136] Guan R, Qu Y, Guo Y, Yu L, Liu Y, Jiang J, Chen J, Ren Y, Liu G, Tian L, Jin L, Liu Z, Hong H, Chang R, Gilliham M, Qiu L. Salinity tolerance in soybean is modulated by natural variation in GmSALT3. The Plant Journal. 2014;80:937–950. DOI: 10.1111/tpj.12695

[137] Kumar R, Mustafiz A, Sahoo KK, Sharma V, Samanta S, Sopory SK, Pareek A, Singla-Pareek SL. Functional screening of cDNA library from a salt tolerant rice genotype Pokkali identifies mannose-1-phosphate guanyl transferase gene (OsMPG1) as a key member of salinity stress response. Plant Molecular Biology. 2012;79:555–568. DOI: 10.1007/s11103-012-9928-8

[138] Obata T, Kitamoto HK, Nakamura A, Fukuda A, Tanaka Y. Rice shaker potassium channel OsKAT1 confers tolerance to salinity stress on yeast and rice cells. Plant Physiology. 2007;144:1978–1985. DOI: 10.1104/pp.107.101154

[139] Chen Y, Chen C, Tan Z, Liu J, Zhuang L, Yang Z, Huang B. Functional identification and characterization of genes cloned from halophyte *Seashore paspalum* conferring salinity and cadmium tolerance. Frontiers in Plant Science. 2016;7:102. DOI: 10.3389/fpls.2016.00102

[140] Munns R, Gilliham M. Salinity tolerance of crops – what is the cost? New Phytologist. 2015;208:668–673. DOI: 10.1111/nph.13519

[141] Byrt CS, Xu B, Krishnan M, Lightfoot DJ, Athman A, Jacobs AK, Watson-Haigh NS, Munns R, Tester M, Gilliham M. The Na+ transporter, TaHKT1;5-D, limits shoot Na+

accumulation in bread wheat. The Plant Journal. 2014;80:516–26. DOI: 10.1111/tpj. 12651

[142] Munns R, James RA, Xu B, Athman A, Conn SJ, Jordans C, Byrt CS, Hare RA, Tyerman SD, Tester M, Plett D, Gilliham M. Wheat grain yield on saline soils is improved by an ancestral Na$^+$ transporter gene. Nature Biotechnology. 2012;30:360–4. 10.1038/nbt.2120

[143] Eswaran N, Parameswaran S, Sathram B, Anantharaman B, Kumar GRK, Tangirala SJ. Yeast functional screen to identify genetic determinants capable of conferring abiotic stress tolerance in *Jatropha curcas*. BMC Biotechnology. 2010;10:23. DOI: 10.3389/fpls.2015.00920

[144] Zhou L, Wang C, Liu R, Han Q, Vandeleur RK, Du J, Tyerman S, Shou H. Constitutive overexpression of soybean plasma membrane intrinsic protein GmPIP1;6 confers salt tolerance. BMC Plant Biology. 2014;14:188. DOI: 10.1186/1471-2229-14-181

[145] Sade N, Gebretsadik M, Seligmann R, Schwartz A, Wallach R, Moshelion M. The role of tobacco Aquaporin1 in improving water use efficiency, hydraulic conductivity, and yield production under salt stress. Plant Physiology. 2010;152(1):245–254. DOI: 10.1104/pp.109.145854

[146] Gao Z, He X, Zhao B, Zhou C, Liang Y, Ge R, Shen Y, Huang Z. Overexpressing a putative aquaporin gene from wheat, TaNIP, enhances salt tolerance in transgenic *Arabidopsis*. Plant &Cell Physiology. 2010;51(5):767–775. DOI: 10.1093/pcp/pcq036

[147] Hu W, Yuan Q, Wang Y, Cai R, Deng X, Wang J, Zhou S, Chen M, Chen L, Huang C, Ma Z, Yang G, He G. Overexpression of a wheat aquaporin gene, TaAQP8, enhances salt stress tolerance in transgenic tobacco. Plant & Cell Physiology 2012;53(12):2127–2141. DOI: 10.1093/pcp/pcs154

[148] Zhou QY, Wang L, Cai X, Wang D, Hua XJ, Qu LQ, Lin JX, Chen T. Net sodium fluxes change significantly at anatomically distinct root zones of rice (*Oryza sativa* L.) seedlings. Journal of Plant Physiology. 2011;168(11):1249–1255. DOI: 10.1016/j.jplph.2011.01.017

[149] Henderson SW, Baumann U, Blackmore DH, Walker AR, Walker RR, Gilliham M. Shoot chloride exclusion and salt tolerance in grapevine is associated with differential ion transporter expression in roots. BMC Plant Biology. 2014;14: 273. DOI: 10.1186/s12870-014-0273-8

[150] Vandeleur RK, Sullivan W, Athman A, Jordans C, Gilliham M, Kaiser BN, Tyerman SD. Rapid shoot-to-root signaling regulates root hydraulic conductance via aquaporins. Plant, Cell and Environment. 2014;37(2):520–538. DOI: 10.1111/pce.12175

[151] Xu C, Wang M, Zhou L, Quan T, Xia G. Heterologous expression of the wheat aquaporin gene TaTIP2;2 compromises the abiotic stress tolerance of *Arabidopsis thaliana*. PLoS ONE. 2013;8:e79618. DOI: 10.1371/journal.pone.0079618

[152] Sani E, Herzyk P, Perrella G, Colot V, Amtmann A. Hyperosmotic priming of *Arabidopsis* seedlings establishes a long-term somatic memory accompanied by specific changes of the epigenome. Genome Biology. 2013;14:R59. DOI: 10.1186/gb-2013-14-6-r59

[153] Wang X, Fan P, Song H, Chen X, Li X, Li Y. Comparative proteomic analysis of differentially expressed proteins in shoots of *Salicornia europaea* under different salinity. Journal of Proteome Research. 2009;8:3331–3345. DOI: 10.1021/pr801083a

[154] Zhou S, Sauvé RJ, Liu Z, Reddy S, Bhatti S, Hucko SD, Fish T, Thannhauser TW. Identification of salt-induced changes in leaf and root proteomes of the wild tomato, *Solanum chilense*. Journal of the American Society for Horticultural Science. 2011;136:288–302.

[155] Wang L, Liu X, Liang M, Tan F, Liang W, Chen Y, Lin Y, Huang L, Xing J, Chen W. Proteomic analysis of salt-responsive proteins in the leaves of mangrove *Kandelia candel* during short-term stress. PLoS ONE. 2014;9:e83141. DOI: 10.1371/journal.pone.0083141

[156] Cheng T, Chen J, Zhang J, Shi S, Zhou Y, Lu L, Wang P, Jiang Z, Yang J, Zhang S, Shi J. Physiological and proteomic analyses of leaves from the halophyte *Tangut nitraria* reveals diverse response pathways critical for high salinity tolerance. Front in Plant Science. 2015;6:30. DOI: 10.3389/fpls.2015.00030

[157] Nadeem SM, Ahmad M, Zahir ZA, Javaid A, Ashraf M. The role of mycorrhizae and plant growth promoting rhizobacteria (PGPR) in improving crop productivity under stressful environments. Biotechnology Advances. 2014;32:429–448. DOI: 10.1016/j.biotechadv.2013.12.005

[158] Roy SJ, Huang W, Wang XJ, Evrard A, Schmoeckel SM, Zafar ZU, Tester M. A novel protein kinase involved in Na+ exclusion revealed from positional cloning. Plant, Cell and Environment. 2013;36:553–568. DOI: 10.1111/j.1365-3040.2012.02595.x

[159] Platten JD, Egdane JA, Ismail AM. Salinity tolerance, Na+ exclusion and allele mining of HKT1;5 in *Oryza sativa* and *O. glaberrima*. Many sources, many genes, one mechanism? BMC Plant Biology. 2013;13:32. DOI: 10.1186/1471-2229-13-32

[160] Zhou S, Zhang Z, Tang Q, Lan H, Li Y, Luo P. Enhanced V-ATPase activity contributes to the improved salt tolerance of transgenic tobacco plants overexpressing vacuolar Na+/H+ antiporter AtNHX1. Biotechnology Letters 2011;33:375–380. DOI: 10.1007/s10529-010-0418-2

[161] Bailey-Serres J, Lee SC, Brinton E. Waterproofing crops: effective flooding survival strategies. Plant Physiology. 2012;160:1698 –1709. DOI: 10.1104/pp.112.208173

[162] Jackson MB, Ismail AM. Electrons, water and rice fields: plant response and adaptation to flooding and submergence stress. AoB Plants. 2015;7:plv078. DOI: 10.1093/aobpla/plv078

[163] Arshad M, and Frankenberger WTJ. Production and stability of ethylene in soil. Biology and Fertility of Soils. 1990;10:29-34. DOI: 10.1007/BF00336121

[164] Gibbs J, Greenway H. Mechanisms of anoxia tolerance in plants: growth, survival and anaerobic catabolism. Functional Plant Biology. 2003;30:1–47. DOI: 10.1071/PP98095

[165] Herzog M, Striker GG, Colmer TD, Pedersen O. Mechanisms of waterlogging tolerance in wheat—a review of root and shoot physiology. Plant, Cell and Environment. 2016;39(5):1068–1086. DOI: 10.1111/pce.12676

[166] Jung KH, Seo YS, Walia H, Cao P, Fukao T, Canlas P, Amonpant F, Bailey-Serres J, Ronald PC. The submergence tolerance regulator Sub1A mediates stress-responsive expression of AP2/ERF transcription factors. Plant Physiology. 2010;152:1674–1692. DOI: 10.1104/pp.109.152157

[167] Xu K, Mackill DJ. A major locus for submergence tolerance mapped on rice chromosome 9. Molecular Breeding. 1996;2:219–224. DOI: 10.1007/BF00564199

[168] Hattori Y, Nagai K, Furukawa S, Song XJ, Kawano R, Sakakibara H, Wu J, Matsumoto T, Yoshimura A, Kitano H, Matsuoka M, Mori H, Ashikari M. The ethylene response factors SNORKEL1 and SNORKEL2 allow rice to adapt to deep water. Nature. 2009;460:1026–1030. DOI: 10.1038/nature08258

[169] Nagai K, Kondo Y, Kitaoka T, Noda T, Kuroh T, Angeles-Shim RB, Yasui H, Yoshimura A, Ashikari M. QTL analysis of internode elongation in response to gibberellin in deep water rice. AoB Plants. 2014;6:plu028. DOI: 10.1093/aobpla/plu028

[170] Campbell MT, Proctor CA, Dou Y, Schmitz AJ, Phansak P, Kruger GR, Zhang C, Walia H. Genetic and molecular characterization of submergence response identifies subtol 6 as a major submergence tolerance locus in maize. PLoS ONE. 2015;10(3):e0120385.

[171] Huang S, Hill RD, Wally OSD, Dionisio G, Ayele BT, Jami SK, Stasolla C. Hemoglobin control of cell survival/death decision regulates *in vitro* plant embryogenesis. Plant Physiology. 2014;165:810–825. DOI: 10.11.04/pp.114.239335

[172] Jackson MB, Ram PC. Physiological and molecular basis of susceptibility and tolerance of rice plants to complete submergence. Annals of Botany. 2003;91:227–241. DOI: 10.1093/aob/mcf242

[173] Nishiuchi S, Yamauchi T, Takahashi H, Kotula L, Nakazono M. Mechanisms for coping with submergence and waterlogging in rice. Rice. 2012;5:2. DOI: 10.1186/1939-8433-5-2

[174] Voortman RL. Micronutrients in agriculture and the world food system – future scarcity and implications. Centre for World Food Studies (SOW-VU), VU University, Amsterdam 2012;

[175] Gui QZ, Bao LM, Chang ZR, Bao CL. Timing and level of nitrogen supply affect nitrogen distribution and recovery in two contrasting oat genotypes. Journal of Plant Nutrition and Soil Science. 2012;175:614–621. DOI: 10.1002/jpln.201100279

[176] Hermans C, Hammond JP, White PJ, Verbruggen N. How do plants respond to nutrient shortage by biomass allocation? Trends in Plant Science. 2006;11:610–617. DOI: 10.1016/j.tplants.2006.10.007

[177] Kant S, Bi YM, Rothstein SJ. Understanding plant response to nitrogen limitation for the improvement of crop nitrogen use efficiency. Journal of Experimental Botany. 2011a;62:1499–1509. DOI: 10.1093/jxb/erq297

[178] Nacry P, Bouguyon E, Gojon A. Nitrogen acquisition by roots: physiological and developmental mechanisms ensuring plant adaptation to a fluctuating resource. Plant and Soil.2013;370:1–29. DOI: 10.1007/s11104-013-1645-9

[179] Zhang H, Forde BG. An *Arabidopsis* MADS box gene that controls nutrient-induced changes in root architecture. Science. 1998;279:407–409. DOI:

[180] Kotur Z, Mackenzie N, Ramesh S, Tyerman SD, Kaiser BN, Glass AD. Nitrate transport capacity of the *Arabidopsis thaliana* NRT2 family members and their interactions with AtNAR2.1. New Phytologist. 2012;194:724–731. DOI: 10.1111/j.1469-8137.2012.04094.x

[181] Sun J, Zheng N. Molecular mechanism underlying the plant NRT1.1 dual-affinity nitrate transporter. Frontiers in Physiology. 2015;6:386. DOI: 10.3389/fphys.2015.00386

[182] De Angeli A, Monachello D, Ephritikhine G, Frachisse JM, Thomine S, Gambale F, Barbier-Brygoo H. The nitrate/proton antiporter AtCLCa mediates nitrate accumulation in plant vacuoles. Nature. 2006;442:939–942. DOI: 10.1038/nature05013

[183] Geiger D, Maierhofer T, Al-Rasheid KA, Scherzer S, Mumm P, Liese A, Ache P, Wellmann C, Marten I, Grill E, Romeis T, Hedrich R. Stomatal closure by fast abscisic acid signaling is mediated by the guard cell anion channel SLAH3 and the receptor RCAR1. Science Signaling. 2011;4:ra32. DOI: 10.1126/scisignal.2001346

[184] Segonzac C, Boyer JC, Ipotesi E, Szponarski W, Tillard P, Touraine B, Sommerer N, Rossignol M, Gibrat R. Nitrate efflux at the root plasma membrane: identification of an *Arabidopsis* excretion transporter. The Plant Cell. 2007;19:3760–3777. DOI: 10.1105/tpc.106.048173

[185] Li JY, Fu YL, Pike SM, Bao J, Tian W, Zhang Y, Chen CZ, Zhang Y, Li HM, Huang J, Li LG, Schroeder JI, Gassmann W, Gong JM. The *Arabidopsis* nitrate transporter NRT1.8 functions in nitrate removal from the xylem sap and mediates cadmium tolerance. The Plant Cell 2010;22:1633–1646. DOI: 10.1105/tpc.110.075242

[186] Wang YY, Tsay YF. *Arabidopsis* nitrate transporter NRT1.9 is important in phloem nitrate transport. The Plant Cell 2011;23:1945–1957. DOI: 10.1105/tpc.111.083618

[187] Wang YY, Hsu PK, Tsay YF. Uptake, allocation and signaling of nitrate. Trends in Plant Science. 2012;17:458–467. DOI: 10.1016/j.tplants.2012.04.006

[188] Hu H, Wang Y, Tsay Y. AtCIPK8, a CBL-interacting protein kinase, regulates the low-affinity phase of the primary nitrate response. The Plant Journal. 2009;57:264–278. DOI: 10.1111/j.1365-313X.2008.03685.x

[189] Jonassen EM, Sévin DC, Lillo C. The bZIP transcription factors HY5 and HYH are positive regulators of the main nitrate reductase gene in *Arabidopsis* leaves, NIA2, but negative regulators of the nitrate uptake gene NRT1.1. Journal of Plant Physiology. 2009;166:2071–2076. DOI: 10.1016/j.jplph.2009.05.010

[190] Konishi M, Yanagisawa S. *Arabidopsis* Nin-like transcription factors play a central role in nitrate signaling. Nature Communications. 2013;4:1503. DOI: 10.1038/ncomms2621

[191] Pinton R, Tomasi N, Zanin L. Molecular and physiological interactions of urea and nitrate uptake in plants. Plant Signaling & Behavior. 2015;11:e1076603. DOI: 10.1080/15592324.2015.1076603

[192] Yoneyama K, Xie X, Kusumoto D. Nitrogen deficiency as well as phosphorus deficiency in Sorghum promotes the production and exudation of 5-deoxystrigol, the host recognition signal for arbuscular. Planta. 2007b;227:125–132. DOI: 10.1007/s00425-007-0600-5

[193] Walch-Liu P, Ivanov II, Filleur S, Gan Y, Remans T, Forde BG. Nitrogen regulation of root branching. Annals of Botany. 2006;97:875–881. DOI: 10.1093/aob/mcj601

[194] Li J, Zhu S, Song X, Shen Y, Chen H, Yu J, Yi K, Liu Y, Karplus VJ, Wu P, Deng XW. A rice glutamate receptor-like gene is critical for the division and survival of individual cells in the root apical meristem. The Plant Cell 2006;18 340–349. DOI: 10.1105/tpc.105.037713

[195] Fuentes SI, Allen DJ, Ortiz-Lopez A, Hernandez G. Over-expression of cytosolic glutamine synthetase increases photosynthesis and growth at low nitrogen concentrations. Journal of Experimental Botany. 2001;52:1071–1081. DOI: 10.1093/jexbot/52.358.1071

[196] Reddy MM, Ulaganathan K. Nitrogen nutrition, its regulation and biotechnological approaches to improve crop productivity. American Journal of Plant Sciences. 2015;6:2745–2798. DOI: 10.4236/ajps.2015.618275

[197] Vance CP, Uhde-Stone C, Allan DL. Phosphorus acquisition and use: critical adaptations by plants for securing a nonrenewable resource. New Phytologist. 2003;157:423–447. DOI:

[198] Song L, Liu D. Ethylene and plant responses to phosphate deficiency. Frontiers in Plant Science. 2015;6:796. DOI: 10.3389/fpls.2015.00796.

[199] Thibaud MC, Arrighi JF, Bayle V, Chiarenza S, Creff A, Bustos R, Paz-Ares J, Poirier Y, Nussaume L. Dissection of local and systemic transcriptional responses to phosphate starvation in *Arabidopsis*. The Plant Journal. 2010;64:775–789. DOI: 10.1111/j.1365-313X.2010.04375.x

[200] Wang Z, Rahman AB, Wang G, Ludewig U, Shen J, Neumann G. Hormonal interactions during cluster-root development in phosphate-deficient white lupin (*Lupinus albus* L.). J Plant Physiology. 2015;177:74–82. DOI: 10.1016/j.jplph.2014.10.022

[201] He CJ, Morgan PW, Drew MC. Enhanced sensitivity to ethylene in nitrogen or phosphate-starved roots of *Zea mays* L. during aerenchyma formation. Plant Physiology. 1992;98:137–142. DOI: 10.1104/pp.98.1.137

[202] Chacón-López A, Ibarra-Laclette E, Sánchez-Calderón L, Gutiérrez-Alanís D, Herrera-Estrella L. Global expression pattern comparison between low phosphorus insensitive 4 and WT *Arabidopsis* reveals an important role of reactive oxygen species and jasmonic acid in the root tip response to phosphate starvation. Plant Signaling & Behavior. 2011;6:382–392. DOI: 10.4161/psb.6.3.14160

[203] Lei M, Zhu C, Liu Y, Karthikeyan AS, Bressan RA, Raghothama KG, Liu D. Ethylene signaling is involved in regulation of phosphate starvation-induced gene expression and production of acid phosphatases and anthocyanin in *Arabidopsis*. New Phytologist. 2011b;189:1084–1095. DOI: 10.1111/j.1469-8137.2010.03555.x

[204] Song G, Liu S. Phosphorus speciation and distribution in surface sediments of the Yellow Sea and East China Sea and potential impacts on ecosystem. Acta Oceanologica Sinica. 2015;34:84–91. DOI: 10.1007/s13131-015-0653-4

[205] Yu H, Luo N, Sun L, Liu D. HPS4/SABRE regulates plant responses to phosphate starvation through antagonistic interaction with ethylene signaling. Journal of Experimental Botany 2012;63:4527–4538. DOI: 10.1093/jxb/ers131

[206] Wang L, Dong J, Gao Z, Liu D. The *Arabidopsis* gene HYPERSENSITIVE TO PHOSPHATE STARVATION3 encodes ETHYLENE OVERPRODUCTION1. Plant Cell Physiology 2012;53:1093–1105. DOI: 10.1093/pcp/pcs072

[207] Christians MJ, Gingerich DJ, Hansen M, Binder BM, Kieber JJ, Vierstra RD. The BTB ubiquitin ligases ETO1. EOL1 and EOL2 act collectively to regulate ethylene biosynthesis in *Arabidopsis* by controlling type-2 ACC synthase levels. The Plant Journal. 2009;57:332–345. DOI: 10.1111/j.1365-313X.2008.03693.x

[208] Wang KL, Yoshida H, Lurin C, Ecker JR. Regulation of ethylene gas biosynthesis by the *Arabidopsis*ETO1 protein. Nature. 2004;428:945–950. DOI: 10.1038/nature02516

[209] Zhang Y, Wang X, Lu S, Liu D. A major root-associated acid phosphatase in *Arabidopsis*, AtPAP10, is regulated by both local and systemic signals under phosphate starvation. Journal of Experimental Botany. 2014;65(22):6577–6588. DOI: 10.1093/jxb/eru377.

[210] Li YS, Gao Y, Tian QY, Shi FL, Li LH, Zhang WH. Stimulation of root acid phosphatase by phosphorus deficiency is regulated by ethylene in *Medicago falcata*. Environmental and Experimental Botany. 2011;71:114–120. DOI: 10.1016/j.envexpbot. 2010.11.007

[211] Nacry P, Canivenc G, Muller B, Azmi A, Van Onckelen H, Rossignol M. A role for auxin redistribution in the responses of the root system architecture to phosphate starvation in *Arabidopsis*. Plant Physiology. 2005;138:2061–2074. DOI: 10.1104/pp. 105.060061.

[212] Karthikeyan AS, Varadarajan DK, Jain A, Held MA, Carpita NC, Raghothama KG. Phosphate starvation responses are mediated by sugar signaling in *Arabidopsis*. Planta. 2007;225:907–918. DOI: 10.1007/s00425-006-0408-8

[213] Niu YF, Chai RS, Jin GL, Wang H, Tang CX, Zhang YS. Responses of root architecture development to low phosphorus availability: a review. Annals of Botany. 2013;112(2):391–408. DOI: 10.1093/aob/mcs285.

[214] Karthikeyan AS, Varadarajan DK, Jain A, Held MA, Carpita NC, Raghothama KG. Phosphate starvation responses are mediated by sugar signaling in *Arabidopsis*. Planta. 2007;225(4):907–918. DOI: 10.1007/s00425-006-0408-8

[215] López-Bucio J, Hernandez-Abreu E, Sanchez-Calderon L, Nieto-Jacobo MF, Simpson J, Herrera-Estrella L. Phosphate availability alters architecture and causes changes in hormone sensitivity in the *Arabidopsis* root system. Plant Physiology. 2002;129:244–256. DOI: 10.1104/pp.010934

[216] Jiang C, Gao X, Liao L, Harberd NP, Fu X. Phosphate starvation root architecture and anthocyanin accumulation responses are modulated by the gibberellin-DELLA signaling pathway in *Arabidopsis*. Plant Physiology. 2007;145:1460–1470. DOI: 10.1104/pp.107.103788

[217] iura K, Rus A, Sharkhuu A, Yokoi S, Karthikeyan AS, Raghothama KG, Baek D, Koo YD, Jin JB, Bressan RA, Yun DJ, Hasegawa PM. The *Arabidopsis* SUMO E3 ligase SIZ1 controls phosphate deficiency responses. Proceedings of the National Academy of Sciences of the United States of America. 2005;102:7760–7765. DOI: 10.1073/pnas. 0500778102

[218] Niu YF, Chai RS, Jin GL, Wang H, Tang CX, Zhang YS. Responses of root architecture development to low phosphorus availability: a review. Annals of Botany. 2013;112:391–408. DOI: 10.1093/aob/mcs285

[219] Ticconi CA, Lucero RD, Sakhonwasee S, Adamson AW, Creff A, Nussaume L, Desnos T, Abel S. ER-resident proteins PDR2 and LPR1 mediate the developmental response of root meristems to phosphate availability. Proceedings of the National Academy of Sciences of the United States of America. 2009;106:14174–14179. DOI: 10.1073/pnas.0901778106

[220] Müller J, Toev T, Heisters M, Teller J, Moore KL, Hause G, Dinesh DC, Bürstenbinder K, Abel S. Iron-dependent callose deposition adjusts root meristem maintenance to phosphate availability. Developmental Cell. 2015;33:216–230. DOI: 10.1016/j.devcel. 2015.02.007

[221] Gamuyao R, Chin JH, Pariasca-Tanaka J, Pesaresi P, Catausan S, Dalid C, Slamet-Loedin I, Tecson-Mendoza EM, Wissuwa M, Heuer S. The protein kinase Pstol1 from traditional rice confers tolerance of phosphorus deficiency. Nature. 2012;488:535–539. DOI: 10.1038/nature11346

[222] Pariasca-Tanaka J, Chian J, Drame´ K, Dalid C, Heuer S, Wissuwa m. A novel allele of the P-starvation tolerance gene OsPSTOL1 from African rice (*Oryza glaberrima* Steud) and its distribution in the genus *Oryza*. Theoretical Applied Genetics. 2014;127:1387–1398. DOI: 10.1007/s00122-014-2306-y

[223] Gao N, Yanhua S, Min J, Shen W, Shi W. Transgenic tomato overexpressing ath-miR399d has enhanced phosphorus accumulation through increased acid phospha-tase and proton secretion as well as phosphate transporter. Plant and Soil. 2010;334: 23–136. DOI: 10.1007/s11104-009-0219-3

[224] Chiou T, Aung K, Lin S, Wu C, Chiang S, Su C. Regulation of phosphate homeostasis by microRNA in *Arabidopsis*. The Plant Cell. 2006;18:412–421. DOI: 10.1105/tpc. 105.038943

[225] López-Bucio J, Hernandez-Abreu E, Sanchez-Calderon L, Nieto-Jacobo MF, Simpson J, Herrera-Estrella L. Phosphate availability alters architecture and causes changes in hormone sensitivity in the *Arabidopsis* root system. Plant Physiology. 2002;129:244–256. DOI: 10.1104/pp.010934

[226] Jiang C, Gao X, Liao L, Harberd NP, Fu X. Phosphate starvation root architecture and anthocyanin accumulation responses are modulated by the gibberellin-DELLA sig-naling pathway in *Arabidopsis*. Plant Physiology. 2007;145:1460–1470. DOI: 10.1104/pp.107.103788

[227] Lopez-Raez JA, Charnikhova T, Gomez-Roldan V, Matusova R, Kohlen W, De Vos R, Verstappen F, Puech-Pages V, Bécard G, Mulder P, Bouwmeester H. Tomato strigo-lactones are derived from carotenoids and their biosynthesis is promoted by phos-phate starvation. New Phytologist. 2008;178:863–874. DOI: 10.1111/j. 1469-8137.2008.02406.x

[228] Umehara M, Hanada A, Magome H, Takeda-Kamiya N, Yamaguchi S. Contribution of strigolactones to the inhibition of tiller bud outgrowth under phosphate deficiency in rice. Plant and Cell Physiology. 2010;51:1118–1126. DOI: 10.1093/pcp/pcq084

[229] Kohlen W, Charnikhova T, Liu Q. Strigolactones are transported through the xylem and play a key role in shoot architectural response to phosphate deficiency in nonar-buscular mycorrhizal host. Plant Physiology. 2011;155:974–987. DOI: 10.1104/pp. 110.164640

[230] Foo E, Yoneyama K, Hugill CJ, Quittenden LJ, Reid JB. Strigolactones and the regulation of pea symbioses in response to nitrate and phosphate deficiency. Molecular Plant. 2012;6:76–87. DOI: 10.1093/mp/sss115

[231] Mayzlish-Gati E, De-Cuyper C, Goormachtig S, Beeckman T, Vuylsteke M, Brewer PB, Beveridge CA, Yermiyahu U, Kaplan Y, Enzer Y, Wininger S, Resnick N, Cohen M, Kapulnik Y, Koltai H. Strigolactones are involved in root response to low phosphate conditions in *Arabidopsis*. Plant Physiology. 2012;160:1329–1341. DOI: 10.1104/pp.112.202358

[232] Liao H, Yan XL, Rubio G, Beebe SE, Blair MW, Lynch JP. Genetic mapping of basal root gravitropism and phosphorus acquisition efficiency in common bean. Functional Plant Biology. 2004;31:959–970. DOI: 10.1071/FP03255

[233] Beebe SE, Rojas-Pierce M, Yan XL Matthew W. Blair MM, Pedraza F, Muñoz F, Tohme J, Lynch JP. Quantitative trait loci for root architecture traits correlated with phosphorus acquisition in common bean. Crop Science. 2006;46: 413–423. DOI: 10.2135/cropsci2005.0226

[234] Ochoa IE, Blair MW, Lynch JP. QTL analysis of adventitious root formation in common bean under contrasting phosphorus availability. Crop Science 2006, 46:1609–1621. DOI: 10.2135/cropsci2005.12-0446

[235] Yan X, Liao H, Beebe SE, Blair MW, Lynch JP. QTL mapping of root hair and acid exudation traits and their relationship to phosphorus uptake in common bean. Plant and Soil. 2004;265:17–29. DOI: 10.1007/s11104-005-0693-1

[236] Li YD, Wang YJ, Tong YP, Gao JG, Zhang JS, Chen SY. QTL mapping of phosphorus deficiency tolerance in soybean (*Glycine max* L. Merr.). Euphytica. 2005;142:137–142. DOI: 10.1007/s10681-005-1192-4

[237] Zhang D, Cheng H, Geng L Kan S, Cui S, Meng Q, Gai J, Yu, D. Detection of quantitative trait loci for phosphorus deficiency tolerance at soybean seedling stage. Euphytica. 2009;167:313–322. DOI: 10.1007/s10681-009-9880-0

[238] Hodson MJ. Metal toxicity and tolerance in plants [Internet]. 2012; Available from: http://www.hodsons.org/MartinHodson/Biochemist12.pdf [Accessed: 2016-05-03]

[239] Uexküll HR, Mutert E. Global extent, development and economic impact of acid soils. Plant and Soil. 1995;171:1–15. DOI: 10.1007/BF00009558

[240] Kochian LV. Cellular mechanisms of aluminum toxicity and resistance in plants. Annual Reviews in Plant Biology. 1995;46:237–260. DOI: 10.1146/annurev.pp.46.060195.001321

[241] Kinraide TB. Assessing the rhizotoxicity of the aluminate ion, $Al(OH)_4$. Plant Physiology. 1990;93:1620–1625. DOI: 10.1146/annurev.pp.46.060195.001321

[242]   Delhaize E, Ryan PR, Randall PJ. Aluminum tolerance in wheat (*Triticum aestivum* L.) (II. Aluminum-stimulated excretion of malic acid from root apices). Plant Physiology. 1993;103:695–702. DOI: 10.1104/pp.103.3.695

[243]   Klug B, Horst WJ. Spatial characteristics of aluminum uptake and translocation in roots of buckwheat (*Fagopyrum esculentum*). Physiologia Plantarum. 2010;139:181–191. DOI: 10.1111/j.1399-3054.2010.01355.x

[244]   Li JY, Liu J, Dong D, Jia X, McCouch SR, Kochian LV. Natural variation underlies alterations in Nramp aluminum transporter (NRAT1) expression and function that play a key role in rice aluminum tolerance. Proceedings of the National Academy of Sciences of the United States of America. 2014;111:6503–6508. DOI: 10.1073/pnas.1318975111

[245]   Bian M, Zhou M, Sun D, Li C. Molecular approaches unravel the mechanism of acid soil tolerance in plants. The Crop Journal. 2013;1:91–104. DOI: 10.1016/j.cj.2013.08.002

[246]   Zhou GF, Delhaize E, Zhou MX, Ryan PR. The barley MATE gene, HvAACT1, increases citrate efflux and Al3+ tolerance when expressed in wheat and barley. Annals of Botany. 2013;112:603–612. DOI: 10.1093/aob/mct135

[247]   Sasaki T, Yamamoto Y, Ezaki B, Katsuhara M, Ahn SJ, Ryan PR, Delhaize E, Matsumoto H. A wheat gene encoding an aluminum-activated malate transporter. The Plant Journal. 2004;37:645–653. DOI: 10.1111/j.1365-313X.2003.01991.x

[248]   Maron LG, Pineros MA, Guimarães CT, Magalhaes JV, Pleiman JK, Mao C, Shaff J, Belicuas SN, Kochian LV. Two functionally distinct members of the MATE (multidrug and toxic compound extrusion) family of transporters potentially underlie two major aluminum tolerance QTLs in maize. The Plant Journal. 2010;61:728–740. DOI: 10.1111/j.1365-313X.2009.04103.x

[249]   Ryan PR, Raman H, Gupta S, Horst WJ, Delhaize E. A second mechanism for aluminum resistance in wheat relies on the constitutive efflux of citrate from roots. Plant Physiology. 2009;149:340–351. DOI: 10.1104/pp.108.129155

[250]   Magalhaes JV, Liu J, Guimarães CT, Lana UGP, Alves VM, Wang YH, Schaffert RE, Hoekenga OA, Piñeros MA, Shaff JE, Klein PE, Carneiro NP, Coelho CM, Trick HN, Kochian LV. A gene in the multidrug and toxic compound extrusion (MATE) family confers aluminum tolerance in sorghum. Nature Genetics. 2007;39:1156–1161. DOI: 10.1038/ng2074

[251]   Hoekenga OA, Maron LG, Piñeros MA, Cancado GM, Shaff J, Kobayashi Y, Ryan PR, Dong B, Delhaize E, Sasaki T, Matsumoto H, Yamamoto Y, Koyama H, Kochian LV. AtALMT1, which encodes a malate transporter, is identified as one of several genes critical for aluminum tolerance in *Arabidopsis*. Proceedings of the National Academy of Sciences of the United States of America. 2006;103:9738–9743. DOI: 10.1073/pnas.0602868103

[252] Liu J, Magalhaes JV, Shaff J, Kochian LV. Aluminum-activated citrate and malate transporters from the MATE and ALMT families function independently to confer *Arabidopsis* aluminum tolerance. The Plant Journal. 2009;57:389–399. DOI: 10.1111/j.1365-313X.2008.03696.x

[253] Silva-Navas J, Benito C, Téllez-Robledo B, El-Moneim A, Gallego FJ. The ScAACT1 gene at the Qalt5 locus as a candidate for increased aluminum tolerance in rye (*Secale cereale* L.). Molecular Breeding. 2011;30(2):845–856. DOI: 10.1007/s11032-011-9668-5

[254] Huang J, Xia H, Li Z, Xiong Y, Kong G. Soil aluminum uptake and accumulation by *Paspalum notatum*. Waste Management Research 2009b;27:668–675. DOI:

[255] Fujii M, Yokosho K, Yamaji N, Saisho D, Yamane M, Takahashi H, Sato K, Nakazono M, Ma JF. Acquisition of aluminium tolerance by modification of a single gene in barley. Nature Communications. 2012;3:713. DOI: 10.1038/ncomms1726

[256] Arenhart RA, Lima JC, Pedron M, Carvalho FE, Silveira JA, Rosa SB, Caverzan A, Andrade CM, Schünemann M, Margis R, Margis-Pinheiro M. Involvement of ASR genes in aluminum tolerance mechanisms in rice. Plant Cell & Environment. 2013b; 36:52–67. DOI: 10.1111/j.1365-3040.2012.02553.x

[257] Gao CJ, Xing D, Li LL, Zhang LR. Implication of reactive oxygen species and mitochondrial dysfunction in the early stages of plant programmed cell death induced by ultraviolet-C overexposure. Planta. 2008;227:755–767. DOI: 10.1007/s00425-007-0654-4

[258] Zhu HF, Wang H, Zhu YF, Zou JW, Zhao FJ, Huang CF. Genome-wide transcriptomic and phylogenetic analyses reveal distinct aluminum-tolerance mechanisms in the aluminum-accumulating species buckwheat (*Fagopyrum tataricum*). BMC Plant Biology. 2015;15:16. DOI: 10.1186/s12870-014-0395-z

[259] Hanikenne M, Talke IN, Haydon MJ, Lanz C, Nolte A, Motte P, Kroymann J, Weigel D, Krämer U. Evolution of metal hyperaccumulation required cis-regulatory changes and triplication of HMA4. Nature. 2008;453(7193):391–395. DOI: 10.1038/nature06877

[260] Kwon H-K, Yokoyama R, Nishitani K. A proteomic approach to apoplastic proteins involved in cell wall regeneration in protoplasts of *Arabidopsis* suspension-cultured cells. Plant Cell Physiology. 2005;46:843–857. DOI: 10.1093/pcp/pci089

[261] Bhuja P, McLachlan K, Stephens J, Taylor G. Accumulation of 1,3-beta-D-glucans, in response to aluminum and cytosolic calcium in *Triticum aestivum*. Plant Cell Physiology. 2004;45:543–549. DOI: 10.1093/pcp/pch068

[262] Sivaguru M, Horst WJ, Eticha D, Matsumoto H. Aluminum inhibits apoplastic flow of high-molecular weight solutes in root apices of *Zea mays* L. Journal of Plant Nutrition and Soil Science. 2006;169:679–690. DOI: 10.1002/jpln.200620603

[263] O'Malley RC, Rodriguez FI, Esch JJ, Binder BM, O'Donnell P, Klee HJ, Bleecker AB. Ethylene-binding activity, gene expression levels, and receptor system output for

ethylene receptor family members from *Arabidopsis* and tomato. The Plant Journal. 2005;41:651–659. DOI: 10.1111/j.1365-313X.2004.02331.x

[264] Tian Q, Zhan X, Ramesh S, Gilliham M, Tyerman SD, Zhan W-H. Ethylene negatively regulates aluminium-induced malate efflux from wheat roots and tobacco cells transformed with TaALMT1. Journal of Experimental Botany 2014;65:2415–2426. DOI:

[265] Sun P, Tian Q-Y, Chen J, Zhang W-H. Aluminium-induced inhibition of root elongation in *Arabidopsis* is mediated by ethylene and auxin. Journal of Experimental Botany. 2010;61:347–356. DOI: 10.1093/jxb/erp306

[266] He H, He L, Gu M. Interactions between nitric oxide and plant hormones in aluminum tolerance. Plant Signaling and Behavior. 2012;7:469–471. DOI: 10.4161/psb.19312

[267] He HY, He LF, Gu MH. Role of microRNAs in aluminum stress in plants. Plant Cell Reports. 2014;33:831–836. DOI: 10.1007/s00299-014-1565-z

[268] Sahrawat KL. Iron toxicity in wetland rice and the role of other nutrients. Journal of Plant Nutrition. 2005;27:1471–1504. DOI: 10.1081/PLN-200025869

[269] Keita A, Yacouba H, Hayde LG, Schultz B. A single-season irrigated rice soil presents higher iron toxicity risk in tropical savannah valley bottoms. Open Soil Science. 2013;3:314–322. DOI: 10.4236/ojss.2013.37036

[270] Onaga G, Edema R, Asea G. Tolerance of rice germplasm to iron toxicity stress and the relationship between tolerance, Fe^{+2}, P and K content in the leaves and roots. Archives of Agronomy and Soil Science. 2013a;59:213–229. DOI: 10.1080/03650340.2011.622751

[271] Roschzttardtz H, Conéjéro G, Divol F, Alcon C, Verdeil JL, Curie C, Mari S. New insights into Fe localization in plant tissues. Frontiers in Plant Science. 2013;4:350. DOI: 10.3389/fpls.2013.00350

[272] Staal M, Cnodder TD, Simon D, Vandenbussche F, Straeten DVD, Verbelen JP, Elzenga T, Vissenberg K. Apoplastic alkalinization is instrumental for the inhibition of cell elongation in the *Arabidopsis* root by the ethylene precursor 1-aminocyclopropane-1-carboxylic acid. Plant Physiology. 2011;155:2049–2055. DOI: 10.1104/pp.110.168476

[273] Quinet M, Vromman D, Clippe A, Bertin P, Lequeux H, Dufey I, Lutts S, Lefèvre I. Combined transcriptomic and physiological approaches reveal strong differences between short and long term response of rice (*Oryza sativa*) to iron toxicity. Plant Cell & Environment. 2012;35:1837–1859. DOI: 10.1111/j.1365-3040.2012.02521.x

[274] Majerus V, Bertin P, Swenden V, Fortemps A, Lobréaux S, Lutts S. Organ-dependent responses of the African rice to short-term iron toxicity: ferritin regulation and antioxidative responses. Biologia Plantarum 2007;51(2):303–312. DOI: 10.1007/s10535-007-0060-6

[275] Peng XX, Yu XL, Li MQ, Yamauchi M. Induction of peroxidase by Fe 2+ in detached rice leaves. Plant and Soil. 1996;180:159–163. DOI: 10.1007/BF00015422

[276] Fang WC, Wang JW, Lin CC, Kao CH. Iron induction of lipid peroxidation and effects on antioxidative enzyme activities in rice leaves. Plant Growth Regulation. 2001;35(1):75–80. DOI: 10.1023/A:1013879019368

[277] Stein RJ, Lopes SIG, Fett JP. Iron toxicity in field-cultivated rice: contrasting tolerance mechanisms in distinct cultivars. Theoretical and Experimental Plant Physiology. 2014;26:135–146. DOI: 10.1007/s40626-014-0013-3

[278] Wu LB, Shhadi MY, Gregorio G, Matthus E, Becker M, Frei M. Genetic and physiological analysis of tolerance to acute iron toxicity in rice. Rice 2014;7:8. DOI: 10.1186/s12284-014-0008-3

[279] Dufey I, Mathieu AS, Draye X, Lutts S, Bertin P. Construction of an integrated map through comparative studies allows the identification of candidate regions for resistance to ferrous iron toxicity in rice. Euphytica. 2015a;203:59–69. DOI: 10.1007/s10681-014-1255-5

[280] Fukuda A, Shiratsuchi H, Fukushima A, Yamaguchi H, Mochida H, Terao T, Ogiwara H. Detection of chromosomal regions affecting iron concentration in rice shoots subjected to excess ferrous iron using chromosomal segment substitution lines between Japonica and Indica. Plant Production Science. 2012;15:183–191. DOI: 10.1626/pps.15.183

[281] Shimizu A. QTL analysis of genetic tolerance to iron toxicity in rice (*Oryza sativa* L.) by quantification of bronzing score. Journal of New Seeds. 2009;10(3):171–179. DOI: 10.1080/15228860903064989

[282] Yuan F, Yang H, Xue Y, Kong D, Ye R, Li C, Zhang J, Theprungsirikul L, Shrift T, Krichilsky B, Johnson DM, Swift GB, He Y, Siedow JN, Pei ZM. OSCA1 mediates osmotic-stress evoked Ca^{2+} increases vital for osmosensing in *Arabidopsis*. Nature. 2014;514(7522):367-371. DOI: 10.1038/nature13593

Advances in Plant Tolerance to Biotic Stresses

Geoffrey Onaga and Kerstin Wydra

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/64351

Abstract

Plants being sessile in nature encounter numerous biotic agents, including bacteria, fungi, viruses, insects, nematodes and protists. A great number of publications indicate that biotic agents significantly reduce crop productivity, although there are some biotic agents that symbiotically or synergistically co-exist with plants. Nonetheless, scientists have made significant advances in understanding the plant defence mechanisms expressed against biotic stresses. These mechanisms range from anatomy, physiology, biochemistry, genetics, development and evolution to their associated molecular dynamics. Using model plants, e.g., Arabidopsis and rice, efforts to understand these mechanisms have led to the identification of representative candidate genes, quantitative trait loci (QTLs), proteins and metabolites associated with plant defences against biotic stresses. However, there are drawbacks and insufficiencies in precisely deciphering and deploying these mechanisms, including only modest adaptability of some identified genes or QTLs to changing stress factors. Thus, more systematic efforts are needed to explore and expand the development of biotic stress resistant germplasm. In this chapter, we provided a comprehensive overview and discussed plant defence mechanisms involving molecular and cellular adaptation to biotic stresses. The latest achievements and perspective on plant molecular responses to biotic stresses, including gene expression, and targeted functional analyses of the genes expressed against biotic stresses have been presented and discussed.

Keywords: Biotic stress, climate change, innate immunity, phytohormones

1. Introduction

Biotic stresses are the damage to plants caused by other living organisms such as bacteria, fungi, nematodes, protists, insects, viruses and viroids. Numerous biotic stresses are of historical significance, for instance, the potato blight in Ireland, coffee rust in Brazil, maize leaf blight caused by *Cochliobolus heterostrophus* in the United States and the great Bengal famine

in 1943 [1]. These are some of the major events that devastated food production and led to millions of human deaths and migration to other countries in the past. Presently, the occurrence of new pathogen races and insect biotypes poses further threat to crop production [2]. Pathogens account for about 15% losses in global food production, and are a major challenge in breeding resistant crops. Considering that genetic polymorphism is present in phytopathogenic agents and insect populations, changes in the climatic factors are considered to further influence/modify this polymorphism, causing evolution of aggressive strains or biotypes [3] that will alter the outcome of host-pathogen interaction. Thus, disease or insect pest outbreaks are expected to continue to cause food production losses or even worsen by expanding to the areas they were not prevalent before [4]. This has important implications for the management options available. Using a combination of options provides certainly more reliability. However, in areas where resources are limiting, e.g., the smallholder farming systems in rural Africa and South East Asia, plant breeders are compelled to make the best use of the diverse disease and pest resistance alleles existing in cultivated crop gene pools and their wild relatives. Thus, exploring the mechanisms of resistance regulated by these resistance alleles is required to enable their exploitation for improving the cultivated elite germplasm that support most of the rural poor livelihoods.

Plant mechanisms of resistance to various pathogens and insect pests are known to involve an array of morphological, genetic, biochemical and molecular processes [5]. These mechanisms may be expressed continuously (constitutively) as preformed resistance, or they may be inducible and deployed only after attack. Plant success in deploying these resistance mechanisms is an evolved ability to persist in unfavourable and variable environments [6]. The recent realization that plant mechanisms of disease/insect resistance or susceptibility are related to mechanistic animal immunity [7] has significantly reshaped our view of plant immunity. The identification of plant pattern recognition receptors (PRRs) that sense pathogens' or insect pests'conserved molecules termed pathogen-associated molecular patterns or herbivore-associated molecular patterns (PAMPs/MAMPs/HAMPs)—and the subsequent PAMP-triggered immunity (PTI) [8] is a paradigm for plant-pathogen interaction studies.

On the other hand, the ability of pathogens/insect pests to suppress or evade PTI, as a structural and functional basis of pathogen survival and evolutionary dynamics in their feeding mechanisms has revitalized research on the so-called 'gene-for-gene' effector induced resistance in plants. It is now clear that effectors are important determinants of pathogens' ability to evade the plant's arsenal targeted towards PAMPs/HAMPs. Effector induced resistance or vertical resistance, often interchangeably translated in modern terms as effector triggered immunity (ETI), is the most successful means of controlling pathogens able to evade PTI [6]. ETI engages a compensatory mechanism within the defense network to transcriptionally coordinate and boost the defense output against pathogens. ETI mostly relies on the endogenious NB-LRR protein products encoded by the resistance (R)-genes. Although R gene mediated resistance is generally not durable, ETI is now effectively deployed through pyramiding of several resistance (R)-genes in the same cultivar, which increases resistance durability and spectrum.

Another aspect of resistance that has gained significance in plant defence studies is the systemic acquired resistance (SAR), in which defence proteins accumulate not only at the site of infection

but also systemically in uninfected tissues and/or plants. SAR provides long-term defense against a broad-spectrum of pathogens and insects. Another form of induced resistance, which, in many aspects, is similar to SAR, is induced systemic resistance (ISR). ISR is potentiated by plant growth promoting rhizobacteria (PGPR), many of them belonging to *Pseudomonas* species. Obviously, the sessile nature of plants requires an efficient signalling system capable of detecting, transporting and interpreting signals produced at the plant-pathogen interface, and SAR and ISR provide a practical means to confer a fitness advantage to plants in conditions of high disease pressure, since plants are primed to more quickly and effectively activate their defences ahead of pathogen/ insect attack. Plants also defend themselves through RNA interference to target and inactivate invading nucleic acids from viruses, and more recently fungal pathogens.

These are the aspects that this chapter has addressed to provide background information for a more detailed discussion of the diverse aspects of plant defence patterns, including qualitative and quantitative mechanisms and their associated molecular patterns. Although pathogenic mechanisms would be interesting to the reader, this chapter does not delve extensively into this aspect, except to mention it as a consideration in emphasizing certain aspects of plant resistance. For additional background, the reader is referred to excellent reviews and the references therein that address plant-pathogen interaction.

2. Plant defence mechanisms in response to pathogens

Plants respond to various pathogens through an intricate and dynamic defence system. The mechanism of defence has been classified as innate and systemic plant response. The overview of plant defence response is represented in Figure 1. An innate defence is exhibited by the plant in two ways, viz., specific (cultivar/pathogen race specific) and non-specific (non-host or general resistance) [8]. The molecular basis of non-host resistance is not well studied, but presumably relies on both constitutive barriers and inducible responses that involve a large array of proteins and other organic molecules produced prior to infection or during pathogen attack [9, 10]. Constitutive defences include morphological and structural barriers (cell walls, epidermis layer, trichomes, thorns, etc.), chemical compounds (metabolites, phenolics, nitrogen compounds, saponins, terpenoids, steroids and glucosinolates), and proteins and enzymes [11, 12, 199]. These compounds confer tolerance or resistance to biotic stresses by not only protecting the plant from invasion, but also giving the plant strength and rigidity. The inducible defences, e.g., the production of toxic chemicals, pathogen-degrading enzymes e.g., chitinases and glucanases, and deliberate cell suicide are conservatively used by plants because of the high energy costs and nutrient requirements associated with their production and maintenance. These compounds may be present in their biologically active forms or stored as inactive precursors that are converted to their active forms by host enzymes in response to pathogen attack or tissue damage. Plant defence strategies involving these compounds can fall in either category, innate or SAR. Although innate immunity is of greater efficiency and is the most common form of plant resistance to microbes, both defence strategies depend on the

ability of the plant to distinguish between self and non-self molecules. The molecular bases of these defence mechanisms are discussed below.
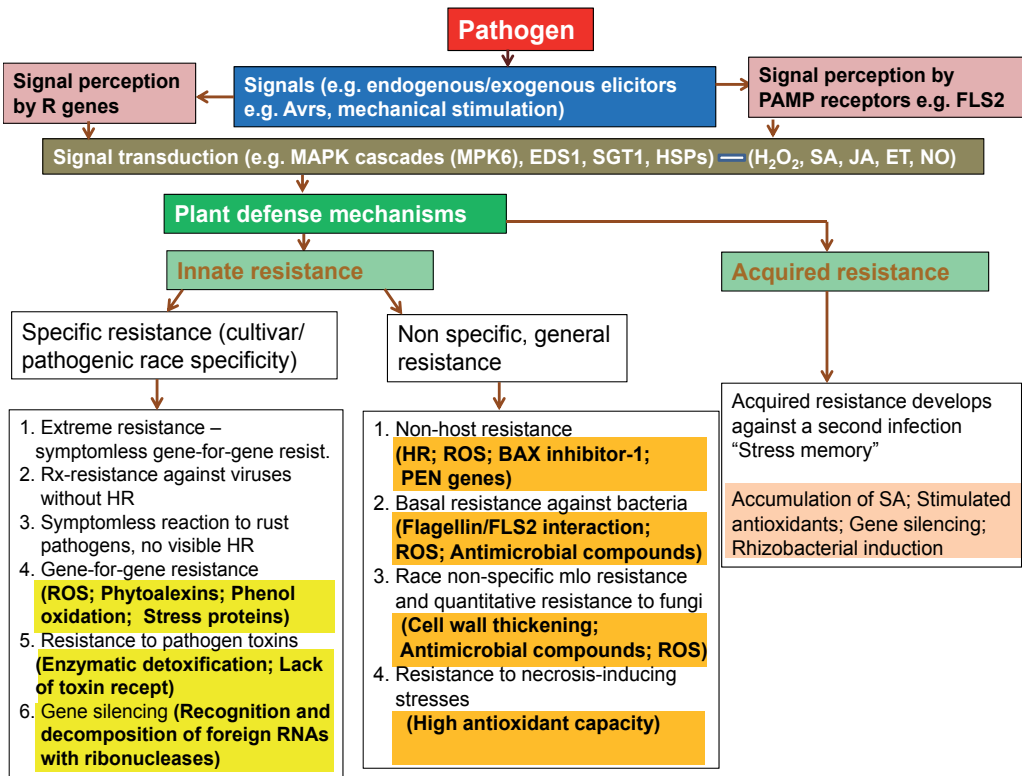

Figure 1. Overview of cellular mechanisms of biotic stress response leading to innate immunity and systemic acquired resistance. Plant PRRs or R-genes perceive PAMPS/DAMPs and effectors, respectively. Inside the cell, an overlapping set of downstream immune responses results from the PTI/ETI continuum. This includes the activation of multiple signaling pathways involving reactive oxygen species (ROS), defense hormones (such as salicylic acid, jasmonic acid and ethylene), mitogen activated protein kinases (MAPK), and transcription factor families, e.g., AP2/ERF, WRKY, MYB, bZIP etc. these signals activate either innate response or acquired immune response or both.

3. Innate immunity

Innate immunity in plants is divided into microbial-associated molecular-pattern-triggered immunity (MTI; also called PTI) and effector-triggered immunity (ETI). In MTI/PTI, innate immunity is defined by receptors for microbe-associated molecules, conserved mitogen-associated protein kinase signalling cascades and the production of antimicrobial peptides/ compounds [13]. Recognition of microbes is divided into two branches, one involving slowly evolving microbial- or pathogen-associated molecular patterns, such as fungal chitin, xylanase or bacterial flagellin, lipopolysaccharides and peptidoglycans [14], and the other that responds

to a compromised 'self', also called damage-associated molecular patterns (DAMPs) [14, 15]. Both PAMPs and DAMPs are recognized by transmembrane pattern recognition receptors (PRRs).

A common strategy employed by adapted pathogens is to secrete effector proteins that avoid or regulate PTI recognition. To counter this stealth afforded by the microbial effectors, plants have evolved an intracellular surveillance involving polymorphic NB-LRR protein products encoded by resistance (R) genes, named after their characteristic feature due to the presence of nucleotide binding (NB) and leucine-rich repeat (LRR) domains [9]. This type of plant defence is referred to as ETI and is synonymous to pathogen race/host plant cultivar-specific plant disease resistance [8].

Generally, PTI and ETI trigger similar defence responses, but ETI is much faster and quantitatively stronger [16]. ETI is often associated with a localized cell death termed the hypersensitive response (HR) that functions to restrict further spread of microbial attack [9, 17]. Hence, the important feature of ETI is the ability to sense microbe-mediated modifications inferred on points of vulnerability in the host, whereas PTI is able to sense infectious-self and non-self. By guarding against weak points or even setting up decoys to confuse invaders, ETI is an efficient defence system for more progressed infections [15, 18], whereas PTI is important for non-host resistance and for basal immunity in susceptible host plant cultivars. In the following section, we will discuss novel insights and overviews on the dynamics of innate immunity in plant defence.

3.1. Pathogen- or microbial-associated molecular-pattern (PAMP/MAMP)-triggered immunity (PTI)

PTI (formerly called basal or horizontal disease resistance) is the first facet of active plant defence and can be considered as the primary driving force of plant-microbe interactions [19]. As discussed before, PTI involves the recognition of conserved, indispensable microbial elicitors known as PAMPs by PRRs of either the receptor-like kinase (RLK) or receptor-like proteins (RLPs) families, which are membranous bound extracellular receptors. RLPs resemble the extracellular domains of RLKs, but lack the cytosolic signalling domain, whereas RLKs have both extracellular and intracellular kinase domains [6]. Instances of hetero-oligomeric complexes between RLKs and RLPs have been reported to occur, and to complement each other in PAMP detection [8], as will be discussed in the following sections. Examples of RLPs include the S locus glycoprotein (*SLG*), *CLAVATA2* and *Xa21D*. RLKs are numerous, and some examples will also be discussed in the following sections. Despite different configurations, both RLKs and RLPs receptors contribute to blocking infection before the microbe gains a hold on the plant.

PAMPs occur throughout the pathogen classes, including bacterial flagellin (*flg22*) and EF-Tu (*elf18*), fungal chitin (*CEBiP*) and mannans of yeast, xylanase (*LeEIX1/2*) and Oomycetes' heptaglucan (*HG*) [17, 19–21]. The early responses induced by PAMPs occur within minutes to hours and are varied, ranging from rapid ion fluxes across the plasma membrane, oxidative burst, activation of mitogen-activated protein kinases (MAPKs) and calcium-dependent protein kinases (CDPKs) to local induction of defence -related genes or pathogen cell wall/cell

membranes lysing enzymes/peptides, e.g., chitinases, glucanases and defensins (Figure 1) [22]. Other responses may include production of antimicrobial phytoalexins, plant cell wall modifications, e.g. deposition of papillae, enriched with (1,3)-β-glucan cell wall polymer, callose, lignin biosynthesis, or changes in cell wall proteins and pectic polysaccharide structures [14, 22, 89, 90, 200]. When the pathogen gains entry and initiates colonization, a concerted effort of both PTI and ETI may be required to restrict further colonization. In the event that ETI is not active, PTI could probably contribute to effective plant resistance as much as ETI, if the capacity to recognize undetected epitopes could be engineered into plants. Some of the examples of PTI that have been shown to contribute to resistance in plants are discussed in the following section.

3.1.1. Specific examples of PTI in plants

3.1.1.1. Flagellin-induced resistance

Flagellin constitutes the main building block of bacterial flagellum, and is so far the best characterized PAMP in plants. A 22 amino acid (*flg22*) peptide-spanning region in the N-terminal part of flagellin of *Pseudomonas syringae* is sufficient to elicit the whole array of typical immune responses in a broad variety of plants [23]. The PRR responsible for flagellin perception in the model plant *Arabidopsis thaliana* is the leucine-rich repeat receptor-like kinase (LRR-RLK) *FLAGELLIN-SENSING 2* (*FLS2*). Functional *FLS2* homologs have been identified in other major groups of higher plants, including tomato, grapevine, *Nicotiana benthamiana* and rice, suggesting that the receptors for the *flg22* epitope of bacterial flagellin are evolutionarily ancient and conserved [14, 24]. Despite evolutionary conservation, *FLS2* proteins from different plant species, such as tomato flagellin receptor (*LeFLS2*), grapevine *(VvFLS2)* and *A. thaliana* (*AtFLS2*), still exhibit different perception specificities to elicitation determinants of flagellins [24–26]. This suggests that the domains found in *FLS* may have undergone some functional innovations that contribute to different perception specificities. Flagellin also seems to be recognized by other means in certain plant species. For instance, in rice, *flg22* epitope does not allow the activation of PRR, but flagellin induces cell death [26]. Moreover, the glycosylation status of flagellin proteins is emerging as a determinant of recognizing adapted and non-adapted bacteria by *Solanaceae* plants, such as tobacco and tomato [27, 28]. More recently, another flagellin, *flgII-28*, was identified in *Solanaceae* [29], though the corresponding PRR is yet to be identified. Both *flg22* and *flgII-28* are physically linked by a stretch of 33 amino acid residues, suggesting that both molecules are detected by the same receptor, *FLS2* [30].

The signalling events triggered in plant cells following *flg22* detection include rapid binding of *FLS2* to *BAK1* (*BRI1-associated kinase 1*) by reciprocal transphosphorylation of their kinase domains [31]. The plasma membrane localized receptor-like cytoplasmic kinase *BOTRYTIS-INDUCED KINASE 1* (*BIK1*) and related *PBS1-LIKE* (*PBL*) kinases associate with *FLS2/BAK1* [32]. The complex formed triggers multiple rapid phosphorylation events resulting in *BIK1* release. *BIK1* plays a central role in conveying signals from not only *FLS2* but also other PRRs, including *EFR*, *CERK1* and the DAMP receptor, *PEPR1/PEPR2*. The signal transduction downstream of *flg22* perception includes a Ca^{2+} burst, activation of *CDPKs* and *RbohD* required

for the ROS burst and induction of *MAPK* cascades. These signalling cascades activate transcriptional reprogrammers such as the *WRKY* TFs, which are required for induction of defence genes [201].

3.1.1.2. Elongation factor (EF-Tu) induced resistance

Elongation factor Tu (EF-Tu) is the most abundant bacterial protein originally isolated from *Escherichia coli*, and acts as PAMP in *Brassicaceae* family members including *A. thaliana* [33]. The conserved *N*-acetylated epitope *elf18* (first 18 amino acids of the protein) is sufficient to trigger defence responses in plants [33, 34]. The shorter peptide, *elf12* (first 12 N-terminal amino acids), comprising the acetyl group, is inactive as an elicitor but acts as a specific antagonist for EF-Tu–related elicitors. EF-Tu is recognized by the *LRR-RLK EF-TU RECEPTOR* (*EFR*) of the same subfamily (*LRRXII*) as *FLS2* [34]. Interestingly, the ability to perceive *elf18* epitope seems restricted to the plant family *Brassicaceae*. However, heterologous expression of *EFR* in the Solanaceae family, e.g., *N. benthamiana* and *Solanum lycopersicum*, makes them more resistant to a range of phytopathogenic bacteria, suggesting that *EFR* can be as well used to engineer broad-spectrum disease resistance in other families [35]. More recently, *EFa50* central region comprising *Lys176* to *Gly225* was found to be fully active as a PAMP in rice and induced H_2O_2 generation and callose deposition [36]. Moreover, *AtEFR*-transformed rice plants were shown to be well responsive to the *Xanthomonas oryzae* derived *elf18* peptide by strongly inducing ROS burst and expression of *OsPBZ1* in transgenic cell cultures [37], further suggesting that *EFR* confers stable resistance across plant families.

The mechanism of *EFR* resistance is mediated by heteromeric complex formation. For instance, in rice, the complex formed between *SOMATIC EMBRYOGENESIS RECEPTOR KINASEs* (*OsSERK2*; an ortholog of *BAK1*) and *XA21* binding protein 24 (*XB24*) is the most important component of *XA21*-mediated defence response. Four *SERK* co-receptor-like kinases interact with *EFR* within seconds to minutes of ligand binding [38], and once the ligand is perceived, *EFR* is rapidly phosphorylated, which triggers downstream signal activation, including the activation and release of *BIK1*. *BIK1* plays a central role in conveying signals, as discussed before (see discussion on flagellin-induced resistance). Interaction between *EFR* and *SERK* also triggers the activation and release of other members of the cytoplasmic receptor-like kinase subfamily VII from the complex. Downstream components of these responses include activation of a RING finger ubiquitin ligase (*XB3*), *MAPKs*, the plant-specific ankyrin-repeat (PANK) containing protein *XB25*, and *WRKY* TFs.

Notwithstanding the *FLS2* and *EFR* PRRs identified so far, relatively fewer PRR genes have been utilized to enhance plant resistance to bacterial pathogens through breeding and transgenic approaches [37], except a few that have been shown to be better adapted to defence signalling. The most famous example is that of *Xa21* gene transferred from *Oryza longistaminata*, which confers high resistance to *X. oryzae* in rice [39]. Heterologous expression of *XA21* in *Citrus sinensis*, *Lycopersicon esculentum* and banana (*Musa* sp.) also conferred moderate resistance to *Xanthomonas axonopodis pv. citri* and resistance to *Ralstonia solanacearum* and *Xanthomonas campestris pv. malvacearum* in experiments under controlled conditions [40–42]. The tomato *RLP Ve1*, which recognizes *Ave1* from *Verticillium dahliae race 1* is another inter-class example that confers stable resistance when transferred and expressed in Arabidopsis for use as a model genetic system [43]. Taken together, *XA21* and *Ve1* are an example of

engineered resistance strategy under controlled conditions, despite their taxonomic restrictions. However, more PRRs recognizing conserved molecular signatures in bacteria will need to be discovered and their complex interaction with the plant's physiology and metabolism and the environment understood, if the ambition of improving crop plants through genetic engineering of broad-spectrum disease resistance by gene transfer is to become more convincing.

3.1.1.3. Plant perception of PAMPs from fungi and oomycetes

Chitin, a homopolymer of β-(1,4)-linked N-acetylglucosamine (GlcNAc) unit, is a major constituent of fungal cell walls and is a classical PAMP [17]. Chitin is an ideal point of attack during plant defence responses since glucosamine polymers are not found in plants. Upon pathogen contact with the host, plant chitinases (hydrolytic enzymes) break down microbial chitin polymers. Interestingly, different plants have evolved mechanisms that employ common factors for chitin perception, and this could be probably the reason for the evolution of pathogen counter measures, e.g., in the biotrophic fungal pathogen *Cladosporium fulvum* [44]. In this context, the reaction of tomato with induction of defense-related, signal transduction and transcription genes to external chitin application supports the role of the described mechanisms [202].

The first chitin-binding PRR was identified in rice as the *lysine motif* (*LysM*)-*RLP*, and was named *chitin*-elicitor *binding* protein (*CEBiP*) [45]. *CEBiP* is a glycoprotein that localizes in the plasma membrane. Upon chitin binding, *CEBiP* homodimerizes and forms a hetero-oligomeric complex with the *Chitin Elicitor Receptor Kinase 1* (*OsCERK1*), the rice ortholog of Arabidopsis *AtCERK1*. The binding thus forms a sandwich-type receptor system for chitin as described in [45, 46]. The mechanism of perception, however, varies between plant species. For example, *AtCERK1* does not seem to employ *CEBiP*-like *LysM-RLPs* to induce typical immune responses such as reactive oxygen species and immune gene expression upon chitin perception [47]. Instead, *AtCERK1* binds directly to octamers of chitin, which in turn induce *AtCERK1* homodimerization and the resultant immune signalling [48]. Arabidopsis *LysM (AtLYM2)*, the closest ortholog of *AtCEBiP*, and the rice *LysM RLPs* (*OsLYP4* and *OsLYP6*) are also able to bind chitin [49]. However, it is not clear whether *AtLYM2/LYK4* also display the putative homodimerization induced by chitin perception. Two other orthologs of *CEBiP*, *AtLYM1* and *AtLYM3*, which specifically bind *PGN*, but not chitin, interact with *AtCERK1*. This indicates that *AtCERK1* is a multifaceted *RLK* that also forms hetero-oligomeric complexes with ligand-binding *RLPs*, probably across different plant families.

Fungal xylanases also function as fungal PAMPs by eliciting defence responses and promoting necrosis [50, 51]. In tomato, ethylene-inducing xylanases (*EIXs*) produced by *Trichoderma* species are perceived by two specific LRR-RLPs receptors, *LeEix1* and *LeEix2* [52]. Both receptors bind *Eixs*, but o*LeEix2* is the primary mediator of defence responses. *LeEix1* heterodimerizes with *LeEix2* upon application of the *Eixs* and attenuates *Eix*-induced internalization and the subsequent signalling of the *LeEix2* receptor [53]. Microbial xyloglucan-specific endoglucanases (XEGs) have also been reported to induce plant defences. Fungal XEGs are inhibited by xyloglucan endoglucanase inhibiting proteins (*XEGIPs*), which so far have been characterized in tomato, carrot and tobacco [54, 55].

Other PRRs that have been identified in plants in response to fungal PAMPs include the *Brassica napus LepR3/Rlm2,* for blackleg resistance, which perceives *AVRLM1* [56]. In *Arabidopsis, Rlm2* interacts with *suppressor of BAK1-interacting receptor-like kinase 1* (*AtSOBIR1*), suggesting that *SOBIR1* is a component of LRR-RLP-mediated resistance against *Leptosphaeria maculans*, which is similar to that formed by rice *OsCERK1 and* Arabidopsis *AtCERK1* [57]. The tomato *Cf* proteins (*Cf2, Cf4* and *Cf9*) that recognize the corresponding effector proteins (*Avr2, Avr4* and *Avr9*) secreted by *C. fulvum* are other PRR-like receptors that were previously identified. *Cf4* interacts with *BAK1* in a manner similar to the rice ligand binding and associated receptor *OsSERK/EFR*.

Wheat and Arabidopsis *RLP1.1* and *RLP30* are also involved in antifungal defence, although the corresponding ligands are unknown so far [58]. Several orphan PAMPs with unknown PRRs, from fungi or oomycetes that can trigger immune signalling have also been identified, including fungal ergosterol [59], oomycete arachidonic acid [60], elicitins (*INF1*) [61], the transglutaminase-derived immunogenic epitope *Pep13* [62], cryptogein [63] and cellulose-binding elicitor lectin (*CBEL*) [64]. Thus, further research is required to understand mechanistically how these orphan PAMPs are involved in PTI.

Taken together, the identification of several potential host plant receptor targets and receptor complexes, and their stability across plant species and in the field will greatly help to improve plant protection. Moreover, identification of several potential microbial molecules that act as PAMPs would increase chances of identifying more potential host plant PRRs for developing crops with higher resistance or inducible resistance.

3.1.1.4. Plant perception of virus PAMPs

Although viral patterns inducing PTI are well known from animal systems, there is no similar pattern reported for plants [48]. Instead, plant resistance to viruses is mediated by post-transcriptional gene silencing of viral RNA or ETI. Nevertheless, infection by compatible viruses can also induce defence responses similar to PTI. Typical PTI cellular responses in plant-virus interactions include ion fluxes, ROS production, ethylene, salicylic acid (SA), MAPK signalling and callose deposition, for review see [65]. Commonly reported genes associated with PRRs in response to viruses include *PEPs* that encode longer peptides (*ProPEP*) from which small peptides (*PEP*) are derived. In Arabidopsis, *AtPEP* interact with two DAMP PRRs, *PEP*-receptor 1 (*PEPR1*) and *PEPR2* [66], both of which interact with *BAK1* upon recognition of *AtPEP*. Thus, *BAK1* is important for antiviral defence in *Arabidopsis.* Indeed, the *bak1* mutants show enhanced susceptibility to three different RNA viruses (*TMV-U1, ORMV* and *TCV*) during compatible interactions [67]. The immune response induced by *PEPR-BAK1* interaction is a classical PTI. Another viral resistance mechanism, which is highly similar to *BAK1* and *BAK1-like Kinase 1* (*BKK1*), is exhibited by the viral nuclear shuttle protein (NSP)-interacting kinases (*NIKs*) from leucine-rich repeats containing receptor-like serine/threonine kinase (LRR-RKs) subfamily [68].

Recent reviews have also suggested that the ribonuclease III-type DICER-like (DCL) enzymes could be acting as PRRs perceiving viral nucleic acids and triggering immune responses equivalent to the zig-zag model first layer [66]. The virus-derived molecules (e.g., dsRNAs)

act as PAMPs, which trigger PTI and RNA interference (RNAi). However, PTI is typically a form of innate immunity, whereas RNAi induces a form of adaptive immunity. Thus, it is clear that a lot remains to be discovered to prove that virus-derived molecules trigger PTI.

3.1.1.5. Plant perception of insect PAMPs

Molecular recognition via ligand-receptor binding phenomena is increasingly becoming important in insect-plant interactions [69]. As reported earlier, the concept of PAMPS has been expanded to include herbivore-associated molecular patterns or damaged-self compounds produced after insect attack [70]. HAMPs isolated and characterized to date include components found in insect oral secretions (proteins, fatty acid-amino acid conjugates (FACs), sulphur-containing fatty acids, as well as plant-derived molecules generated following insect herbivory, including degradation products of ATP synthase and cell walls [71, 72]. The insect oral secretion molecules are released by chewing insects and have been reported to induce ion imbalances, variations in membrane potential, changes in Ca^{2+} fluxes and the generation of reactive oxygen species (ROS), which stimulate downstream signalling events in plants [73]. Ca^{2+} influx is obviously preceded by the opening of calcium channels, and it is likely that these channels are associated with plant receptors tuned to insect elicitors. Recently, a mechanism similar to PTI was reported in Arabidopsis in which LRR-RK *BAK1* was shown to contribute to innate immunity against aphids [69]. Moreover, application of synthetic FACs on wounded *N. attenuate* leaves strongly induced *MAPK* activity, and subsequently wound-induced modifications in the transcriptome, proteome and defensive secondary metabolites [74, 75]. Insect egg ovipositional fluids have also been shown to induce plant defences [76, 77]. Moreover, insect egg deposition on one leaf could induce volatile emission in the other egg-free leaves [77], suggesting that SAR could be involved after detection of insect eggs' associated molecules. An interesting example was reported in the oviposition by *Pieris brassicae*, which triggered SA accumulation and the subsequent induction of PAMP responsive gene expression associated with lectin-domain RK (*LecRK*), *LecRK-I* [78]. Correspondingly, expression of the defence gene *PR-1*, which requires *EDS1*, *SID2* and *NPR1*, was also detected, implicating the SA pathway downstream of the insect egg recognition.

Another mechanism that is closely related to the PAMP receptors in plant resistance to insects is the *Mi-1* gene in tomato. The induction of *Mi-1* confers resistance to *Macrosiphum euphorbiae* [79]. A receptor-like kinase gene *OsLecRK* in rice, which confers basal resistance to *Nilaparvata lugens*, was recently suggested to be a PRR that recognizes molecules secreted by these insects [80]. A similar mechanism was demonstrated in aphid infestation of *Arabidopsis* in which the immune response was apparently triggered by infiltration of aphid saliva [81]. Consistent with this, infiltration of whole aphid extract from *M. persicae* was reported to activate PTI-like responses in *Arabidopsis* [69, 82].

This notwithstanding, the insect HAMP-receptor binding phenomenon that allows plants to detect insects still remains less clear as to whether these responses are exclusively due to the specific perception of herbivores or due to different damage patterns or both.

3.1.1.6. Infection self-perception DAMPs

As discussed before, plants can also sense self-molecules called damage-associated molecular patterns that are available for recognition only after cell/tissue damage. The striking similarities of DAMP perception in animals and plants have been reviewed [83]. A perfect example that was discussed earlier is the *Arabidopsis* plasma membrane LRR receptor kinase (LRR-RK), designated *PEPR1/PEPR2*, which perceives *AtPep* peptides derived from propeptide (Pro-PEPs) encoded by a seven-member multigenic family (*Pep1-Pep7*). Both *PEPR1* and *PEPR2* were reported to be transcriptionally induced by wounding, treatment with methyl jasmonate, *Pep* peptides and pathogen-associated molecular patterns [64, 84]. Moreover, *AtPep* perception is part of a PTI amplification loop and is important for the induction of systemic immunity [85]. In another example, hydroxyproline-containing glycopeptides (*HypSys*) and rapid alkalinization factor (*RALF*) peptides have been shown to induce an *MAPK* cascade in tomato cells [86]. The precursors of *HypSys* and *RALF* are constitutively present in the plant cell walls [14]. Microbial proteases or intracellular proteases release these peptides upon cell injury, making then to act as DAMPs.

Cell wall components derived from the enzymatic activity of highly specific microbial homogalacturonan (HGA) is another good example of DAMPs [87]. The enhanced production of oligogalacturonic acid (OGA) fragments from plant cell walls potentially acts as DAMP, which are perceived by receptors such as *RLK THESEUS1* (*THE1*), *ER* and *WAK1*. Plants may also rely on the recognition of cell wall degrading enzymes (CWDEs) by LRR-RLPs receptors, e.g., *RBPG1* and *LeEIX1-2* [88]. A decisive role of the composition and structure of plant cell wall polysaccharides, specifically of side chains of pectic polysaccharides, in elicitation of plant defence has also been described in tomato interaction with a bacterial pathogen, *R. solanacearum* [89, 90, 203]. Thus, studying the expression of endogenous molecules and microbial cell wall degrading enzymes and their inhibitors, e.g., polygalacturonases (PGs) and polygalacturonase-inhibiting proteins (PGIPs) [204] is a valuable approach to understanding the dynamics of plant-pathogen interactions as well as to develop a strategy to improve plant protection using induced plant endogenous molecules.

3.2. Effector-triggered immunity (ETI)

ETI (formerly called *R*-gene-mediated or vertical resistance) is based on the highly specific, direct or indirect interaction of pathogen effectors and the products of plant *R* genes according to the gene-for-gene theory [14]. As discussed before, R genes encode proteins of the intracellular nucleotide-binding leucine-rich repeat (NB-LRR) class [10]. The NB-LRR consist of N-terminal effector domain, central NB domain and C-terminal LRR domain, which largely vary in plants [91]. Two major subgroups that have distinct N-terminal domains are generally recognized: (1) one group with a Toll–interleukin 1 receptor (*TIR*) domain are called TNLs, and (2) those with a coiled-coil (CC) domain are called CNLs [92].

In Arabidopsis, the CNLs functionally interact with the glycosylphosphatidylinositol (GPI) anchored protein—*NON-RACE SPECIFIC DISEASE RESISTANCE 1* (*NDR1*), a positive regulator of SA accumulation, for signalling [93, 94]. Indeed, an *ndr1* mutation compromises resistance conferred by the CC-NBS-LRR proteins *RPS2*, *RPM1* or *RPS5* to *P. syringae* express-

ing the avirulence effectors *avrRpt2*, *avrB* and *avrRpm1*, or *avrPph3*, respectively [95]. In contrast, multiple TNLs functionally associate with *ENHANCED DISEASE SUCEPTIBILITY 1* (*EDS1*) and *PHYTOALEXIN DEFICIENT 4* (*PAD4*) for signalling. For instance, resistance conferred by the TIR–NBS–LRR protein *RPS4*, which recognizes *avrRps4* in *P. syringae* is compromised in *eds1* mutants [96]. However, resistance mediated by some R genes is independent of *EDS1/PAD4* and *NDR1* or require additional co-activating proteins, suggesting existence of additional components for signal transmission during plant-pathogen interaction. Some of the regulatory components functionally associated with R genes for an effective HR mediated resistance include *RAR1* (required for *Mla12* resistance) and *SGT1* (suppressor of the G$_2$ allele of *skp1*) proteins [97]. *RAR1* interacts with the N-terminal half of *HSP90* that contains the ATPase domain. *HSP90* also specifically interacts with *SGT1* that contains a tetratricopeptide repeat motif and a domain with similarity to the co-chaperone *p23* [98]. These observations suggest that R proteins require several co-activating proteins, although distinct downstream signalling pathways could be involved. There are also some NLRs containing N terminus other than the classical TIR and CC, either because their protein structures are not validated or due to lack of significant homology; they are referred to as non-TIR-type NLRs (nTNLs) or generally referred to as NLRs. Further work on non-sequenced genomes is likely to expand the number of NLRs, and probably refine functional difference associated with NLR repertoires.

Regardless of the NLR class, NB-ARC domain is the core nucleotide-binding fold in NB-LRR proteins. Four distinct subdomains constitute the NB-ARC domain, including nucleotide-binding (NB) fold and *ARC1, -2* and *-3* subdomains. *ARC1* is a four-helix bundle, *ARC2* is a winged-helix fold and *ARC3* is a helical bundle [99]. *ARC1* and *ARC2* are conserved in *Caenorhabditis elegans CED-4*, and plant NB-LRR R proteins, whereas *ARC3* is absent [99]. Throughout the NB-ARC domain in R proteins, numerous conserved motifs (e.g., *hhGRExE*, Walker A or P-loop, Walker B, GxP, *RNBS-A* to *D* and *MHD*) have been reported [100]. A mutation in these conserved motifs has shown their functional importance in the NB-LRR proteins [101], and is apparently a critical factor determining R gene functional effector recognition pattern differences. Generally, pathogen effector recognition by NLR and NLR expression are broadly characterized into (1) direct NLR-Effector interaction or (2) indirect NLR indirect surveillance of effector activities.

3.2.1. Direct NLR-effector interaction

NLRs maintain an ADP-binding inactive state in the absence of effectors. The binding of effectors induces conformational changes in NLRs, which allow ADP/ATP exchange. Consequently, the exchange of nucleotides triggers a second conformational change that activates the NB-LRRs' N-terminus (TIR or CC) to interact with and trigger downstream target processes [102]. However, there is no substantial evidence on direct NLR-effector interaction that underlies resistance specificity in the NLR-effector combinations, apart from the yeast two-hybrid (Y2H) and *in vitro* interaction assays [103, 104]. A few examples that attempt to show the NLR-effector interaction include the Arabidopsis NLR *RPP1* recognition of the oomycete effector *ATR1* leading to *Hyaloperonospora arabidopsidis (Hpa)* resistance [104]. Both the *RPP1*

receptor and *ATR1* alleles from *Hpa* strains can be diverse. This diversity contributes to a spectrum of resistance phenotypes and effectors. For instance, the recognition specificity of *RPP1-WsB* (from the Wassilewskija ecotype) and *RPP1-NdA* (from the Niederzenz ecotype) vary. The *RPP1-NdA* recognizes a small subset of the *ATR1* alleles recognized by *RPP1-WsB*, while the *RPP1-WsB* associates with the cognate *Hpa* effector protein, *Atr1*, through its LRR domain in a recognition-specific manner [105]. Another example is the Arabidopsis NLR *RRS1*, a domain with sequence similarity to *WRKY* TFs, positioned after the LRR. The cognate effectors *AvrRps4* and *PopP2* directly interact with this *WRKY*-like domain to activate the downstream resistance components [106].

Together, the different R proteins have functional domains that can occupy different positions in NLRs. The functional domain positioning differences could be the reason behind several R genes that have been identified in plants. For instance, in rice more than 100 NLRs encoding genes have been described to confer resistance to strains of *Magnaporthe oryzae* [107]. However, only few R proteins encoded by these genes have been characterized, which limits their deployment. A well-known structure for the recognition of *M. oryzae* effectors is that of *AVR-Piz-t*, which adopts a six-stranded β-sandwich structure and contains a single disulphide bond [108]. The *AVR-Pia* and *AVR1-CO39* have also been reported to be recognized by the *R GENE ANALOGs (RGA4/RGA5)* NLR pair [109, 110] through direct binding to a Heavy-Metal Associated domain (HMA; also known as RATX1) integrated into *RGA5* after the LRR position. *RGA4/RGA5* physically interact to prevent cell death mediated by *RGA4* in the absence of *AVR-Pia*; the presence of the effector relieves this suppression, and induces cell death response, a mechanism that could also be described as indirect NLR surveilance. More recently, Maqbool et al. [111] also found that recognition of *AVR-Pik* by *Pik* is by direct binding to the *HMA* domain of *Pik-1*. However, the positioning of the *HMA* domain between the CC and NB-ARC region of *Pik-1* and after the LRR in *RGA5* is a striking difference between *Pik-1* and *RGA5.* These conformational changes underlying direct effector binding could be causing immunity-related signalling differences. However, the intra- and/or inter-molecular complexes mediating output may be conserved [111].

3.2.2. Indirect NLR surveillance of effector activities

During indirect recognition, the NLR guards the host protein by recognizing (monitoring) the modifications caused by the pathogen effector on the guarded protein [10]. The guarded protein can either be the actual effector virulence target or a decoy inviting modification by the pathogen. An example of the indirect recognition of effectors by NLRs was demonstrated in the conserved Arabidopsis protein *RPM1*-interacting protein 4 (*RIN4*). *RIN4* is targeted by multiple bacterial effectors, e.g., *AvrRpt2*, *AvrRpm1* and *AvrB*, and is monitored for effector-induced modification by two plasma membrane CNL receptors, *RPM1* (resistance to *P. syringae* pv. *maculicola 1*) and *RPS2* (*resistance to P. syringae 2*) [112]. *AvrB*-induced phosphorylation and cis/trans isomerization coupled with conformational changes in *RIN4* are sensed by *RPM1* to activate immune signalling [112, 113]. *AvrRpt2*, being a cysteine protease, cleaves *RIN4* and induces *RIN4* degradation. In the absence of *RPM1* and *RPS2*, *RIN4* acts as a negative regulator of basal resistance, and in that capacity appears to be targeted for manipulation by multiple bacterial effectors [114].

The functioning of NLRs as genetically tightly linked pairs to deliver disease resistance was also recently reported [115]. Moreover, Williams et al. [116] demonstrated, by coupling crystal structure and functional analyses, that *RPS4* and *RESISTANT TO RALSTONIA SOLANACEA-RUM 1* (*RRS1*) TIR domains form homo- and hetero-dimers through a common conserved interface that includes a core serine-histidine (SH) motif. Transient expression assays in tobacco revealed that the *RPS4* TIR domain triggers an effector-independent cell death, which is dependent on the SH motif. Co-expression of the *RRS1* TIR domain and *RPS4* TIR impedes the auto-active cell death caused by *RPS4* TIR, and this was found to be dependent on the *RRS1* SH motif. This suggests that an inactive *RRS1*/*RPS4* TIR hetero-dimer and the formation of an active *RPS4* TIR homo-dimer compete to modulate signalling. As discussed before, Cesari et al. [109] investigated the mode of action of *RGA 4* and *5* that associate through their coiled-coil domains. *RGA4* and *RGA5* are tightly linked rice CC-NLRs, which functionally interact to modulate resistance to the rice pathogen *M. oryzae*. *RGA5* modulates an effector independent cell death constitutively induced by *RGA4* signalling. *RGA5* domain on the C-terminus has a heavy-metal-associated domain, which is related to the cytoplasmic copper chaperone *ATX1* from *Saccharomyces cerevisiae* (*RATX1* domain). This domain is an *AVR-Pia* effector interacting domain in *RGA5*. Thus, the formation of the *RGA4*/*RGA5* hetero-complex is crucial to regulate *RGA4* activity in the absence of pathogen in rice. Hence, *RGA4* acts as a signalling component regulated by its interaction with *RGA5* that acts both as a repressor and a receptor that directly binds the *AVR-Pia* proteins. The apparent striking similarity between the *RPS4*/*RRS1* and the *RGA4*/*RGA5* functional models suggests that similarities are likely to be frequent between the different R genes present in dicots and monocots.

3.2.3. Patterns of NLRs signalling in plant defence

Most NLRs respond to the presence of proteins (effectors) delivered by adapted pathogens/parasites. Using suppressor screens, Gabriels et al. [117], identified *NRC1* (*NLR protein required for HR-associated cell death 1*) as a component of fungal resistance modulated by the tomato plasma membrane receptor-like resistance protein *Cf-4* (*C. fulvum 4*). *NRC1* mediates resistance and cell death induced by both membrane receptors and intracellular NLRs. This indicates that *NRC1* is probably a downstream convergence point in ETI initiated at various cell locations. Indeed, silencing of *NRC1* in *N. benthamiana* impairs the HR mediated by several other R proteins including two NLRs, *Rx* and *Mi*. Members of a conserved class of non-canonical CNLs also function in ETI, downstream of NLR effector recognition and have been designated as helper NLRs [118]. Characterization of these non-canonical CNLs is required in order to track their interaction networks.

The downstream components of ETI signalling events partially overlap with PTI response, including activation of *MAPK* cascade and activation of TFs such as *WRKYs* [119]. In Arabidopsis, three CNLs—*activated disease resistance 1* (*ADR1*), *ADR1-L1* and *ADR1-L2*—transduce signals that lead to SA accumulation and induction of downstream *WRKYs* modulated resistance [118]. In rice, the CNL receptor, *panicle blast 1* (*Pb1*), also appears to mediate resistance against rice blast in a mechanism involving interaction with *WRKY45*, a TF involved in induced resistance via SA signalling pathway [120]. Some CNLs directly translocate or

localize in the nucleus to activate defence [121], e.g., *barley mildew A 10* (*MLA10*) and Arabidopsis *RPS4* and *RPS6*. In the nucleus, *MLA10* interacts with *Hordeum vulgare* (Hv) *WRKY1/2*, which are suppressors of basal defence, during incompatible interaction with powdery mildew fungus. A CNL designated as *MLA1*, also from barley, functions in Arabidopsis against *Blumeria graminis* f. sp. *hordei* (*Bgh*) [122]. The *MLA1*-triggered immunity, including host cell death response and disease resistance, is fully retained in Arabidopsis mutant plants that are simultaneously impaired in well-characterized defence-phytohormone pathways (ET, JA and SA). Similar to *MLA1*, co-acting Arabidopsis TNL pair, *RPS4* and *RRS1* (which encodes a WRKY DNA binding domain), confers resistance in cucumber, *N. benthamiana*, and tomato [122].

Another example supporting our understanding of the NLR nuclear activity is the interaction of N immune receptor with the TF *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 6* (*SPL6*) in *N. benthamiana* [123]. The N immune receptor is present in the nucleus, and confers resistance to tobacco mosaic virus (TMV) infection. N receptor associates with *SPL6* at the sub-nuclear bodies only when the cognate effector, *p50*, is present in the cell. A genetic requirement for *SPL6* was not only shown in *N. benthamiana* for N-mediated disease resistance using the yeast two-hybrid system, but also in *A. thaliana* for *RPS4* immune receptor mediated defence against *P. syringae* pv. *tomato* expressing *AvrRps4* effector. Moreover, a number of *RPS4*-mediated defence responsive genes were differentially regulated upon *AtSPL6* silencing, including some of the previously characterized defence responsive genes such as *PAD4*, *PR1*, *ALD1*, *AIG1*, *NUDT6* and *FMO1*. Additional evidence has been shown in Arabidopsis *RPW8* resistance protein, which encodes truncated CNL-like proteins conferring resistance to powdery mildews in *N. tabacum* and *N. benthamiana* as in Arabidopsis. *RPW8* requires SA, *EDS1*, *NPR1* and *PAD4* to be effective. The functional role of *RPW8* is typically similar to a TNL *ADR1*, a close homolog of *N Requirement Gene 1* (*NRG1*), which functions in and beyond innate immunity [124]. These findings present a unique opportunity to further understand how effector-activated immune receptors directly associate with TFs in the nucleus to activate immune responses. Overall, a resistance signalling framework appears to have emerged for plants in which certain specificity-determining (sensor) NLRs initiate the immune response and either auto-activate and contribute to defence or compliment with other signalling NLRs to contribute to defence by conveying or amplifying the signal.

4. Phytohormones in plant defence response to pathogens and insects

Plant defence against pathogen/herbivore attack involves many signal transduction pathways that are mediated by a network of phytohormones. Phytohormones also play a critical role in regulating plant growth and development. Three most reported plant defence response phytohormones against pathogens/insects include salicylic acid (SA), jasmonic acid (JA) and ethylene (ET) [125]. Salicylic acid, a benzoic acid derivative, is an extensively studied important phytohormone in the regulation of plant defence [13]. In Arabidopsis, activation of the SA pathway has been shown to be important in both basal and R gene mediated biotrophic and hemibiotrophic pathogen defence [126, 127]. As discussed before, *NDR1* and *EDS1* act

upstream of SA, while the downstream pathway is modulated by *NONEXPRESSOR OF PR GENES 1* (*NPR1*), and *WRKY45* in rice. *NPR1* is a transcriptional co-activator of a large set of defence-related genes downstream of SA, and it can conditionally regulate *PDF1.2* expression following treatment of plants with SA and MeJA [128]. SA also contributes to the HR-associated resistance via mechanisms that interact with *RBOHD*, a catalyst in ROS generation and cell death [128]. In tobacco, SA significantly increases in resistant plants infected with TMV [129]. A similar response was observed in Ny-1-resistant potatoes after infection with Potato virus Y (PVY) [130].

In response to insect attack, SA regulates plant defence signalling against aphids by modulating the activity of *PAD4*. Indeed, *pad4* mutants, with compromised SA signalling, have increased susceptibility to *Myzus persicae*. Correspondingly, there is a correlation between *pad4* susceptibility and a delay in aphid-induced senescence [131], indicating that SA defence pathways are compromised in *pad4* mutants. Basal SA defences have also been shown to decrease *M. euphorbiae* longevity in tomato. Moreover, SA is necessary for *Mi1.2*-mediated resistance to potato aphids [132]. SA is also a key derivative of SAR in plants. SAR is a 'whole-plant' broad-spectrum resistance response that occurs following an earlier localized exposure to a pathogen [133]. It is well known that ETI can trigger SAR through both local and systemic synthesis of SA, resulting in transcriptional reprogramming of a battery of genes encoding PR proteins [133, 134]. The reports published so far point to different compounds as potential SAR signals [135]. A change in amino acid homeostasis is one of the suggested components in SAR mediated by ETI [136]. Moreover, amino acids have been reported to be precursors of a large array of plant secondary metabolites involved in defence, including signal SA, cell wall components and anthocyanins. Further evidence on the involvement of amino acid homeostasis in plant defence was reported in Arabidopsis *agd2-like defence response protein 1* (*ald1*) mutants. Characterization of the Arabidopsis *ald1* suggested that an amino acid–derived defence signal was generated upstream of SA synthesis [135]. These findings reveal that plants likely employ amino acids and their derivatives to rapidly reprogram SA synthesis and cellular transcription in order to cope with pathogen invasion, even though it appears to be at the expense of growth and development.

SA also interacts with other phytohormones either synergistically or antagonistically [137–138]. There is an obvious cross-talk between JA and SA signalling pathways in pepper to control thionin synthesis as part of the PR response and other defence pathways [139]. Other synergistic examples include the treatment of *N. benthamiana* plants with JA or SA, which was shown to enhance systemic resistance to TMV [140]; Ellis et al. [141] have also shown that SA- and JA-signalling pathways are required to accomplish the defence response necessary to avert pathogen attack. More recently, Arabidopsis mutants with constitutive SA responses were reported to require JA and ethylene signalling for SA mediated resistance [142]. A dominant mutant named *suppressor of SA insensitivity* (*ssi1*), which has constitutive expression of *PR* genes and is resistant to *P. syringae*, was also shown to constitutively express *PDF1.2* and accumulate elevated levels of SA [143]. Although this finding may be intriguing, because SA does not normally induce *PDF1.2* in wild-type plants, it suggests the existence of an intricate signalling network involving SA and JA. Another mutant named *constitutive PR 5 (cpr5)* was shown to

have SA-mediated *NPR1*-independent resistance, which apparently required components of the JA and ET signal pathways [144]. The pre-treatment of plants with JA followed by SA was also shown to remarkably enhance resistance more than otherwise. Moreover, plants impaired in the JA pathway fail to accumulate SA in the leaves or phloem and become highly susceptible to TMV [145]. Conversely, impairing the SA pathway does not affect JA levels, although increased susceptibility is observed [141, 146]. During infection by the pathogen *P. syringae* pv. *tomato* (Pst) DC3000/*AvrRpm1*, JA as a systemic signal for SAR, increases significantly 6 hours after infection and returns to normal 11 hours after infection [147], which suggests that JA may be transiently required for SA accumulation. Further evidence indicates that SAR is compromised in JA-insensitive mutants, *sgt1b/jai4*, *opr3* (JA-biosynthesis mutant) and *jin1* (JA-response mutant). The JA-biosynthesis mutants *dde2* and *opr3* as well as the downstream signalling mutants *coi1*, *jar1* and *jin1*, though intact in SAR, partially require JA biosynthesis for an effective resistance response [148]. Thus, it is possible that JA probably modulates early components of the SA biosynthetic or signalling pathway. However, it seems likely that the synergistic mechanisms may require not only SA and JA, but also ethylene [149, 150], considering that *cpr5* phenotype is suppressed by the *ethylene-insensitive* (*ein2*) mutation.

The negative crosstalk between SA and JA/ET pathways is probably modulated by *TGA1A-RELATED GENE* (*TGA*) factors. *TGA* class of *bZIP* TFs are repressed by plant-specific gluta-redoxins (e.g., *ROXY19*), which are in turn induced by SA. Co-expression of *ROXY19* with *OCTADECANOID-RESPONSIVE ARABIDOPSIS AP2/ERF-domain protein 59* (*ORA59*) and *ETHYLENE INSENSITVE 3* (*EIN3*) complex suppresses *ORA59* promoter activity. Moreover, a study by Van der Does et al. [137] indicated that SA negatively regulates *ORA59* protein accumulation in *35S:ORA59-GFP* overexpressing plants. *ORA59* is a transcriptional regulator of JA/ET-induced defence genes and is activated by either JA or ET and suppressed by SA. More recently, *TGA2*, *TGA5* and *TGA6* were shown to activate the SA-suppression of ET-inducible defence by regulating *ORA59* expression [150]. This suggests that SA-suppresses JA/ET-inducible defence by interfering with *ORA59* activity through regulation of *ROXY-TGA* interaction. Conversely, evidence of SA positive regulation of ET was proposed by Guan et al. [151]. These authors have shown that in Arabidopsis, SA modules ET by potentiating MITO-GEN-ACTIVATED PROTEIN KINASE6 (MPK6) and MPK3, and involves two 1-aminocyclo-propane-1-carboxylic acid synthase (*ACS*; *ACS2* and *ACS6*) isoforms, which are downstream components o MPK signalling pathway. This finding adds another level of complexity to the phytohormones regulatory network and will probably require further elucidation on how this pathway differs from the ORA59 regulated pathway.

On the other hand, most ET dependent defenses are positively modulated by JA. The *JASMONATE ZIM-DOMAIN* (*JAZ*) protein, which directly binds *EIN3/EIL1* and recruits *HISTONE DEACETYLASE 6* (*HDA6*) to repress ET responsive transcription, is repressed in the presence of JA. Thus, accumulation of JA degrades *JAZ* and allows the binding of *EIN3* to the *ERF1* promoter resulting in the transcription of *ERF1* [142, 152]. *EIN3* also directly activates the promoter of *ORA59* that regulates JA/ET-activated defence pathway. Studies on microarray analysis of Arabidopsis plants infected with *Alternaria brassicicola* revealed that nearly half of the genes induced by ET are also induced by JA [153]. This was substantiated by Lorenzo et

al. [154] who reported that JA and ET pathways indeed converge in the transcriptional activation of *ERF1*, which encodes a TF that regulates the expression of pathogen response genes. *ERF* TFs have been reported to exhibit different regulatory roles depending on the species. For instance, in wheat *ERF* gene *TaPIEP1/TaPIE1*, which belongs to the B3 subgroup within the *ERF* subfamily, confers enhanced resistance to the fungal pathogens, *Bipolaris sorokiniana* and *R. cerealis*, when overexpressed in transgenic wheat [155], whereas in cotton GhERF of group IX, which includes *ORA59*, confer resistance to *Xanthomonas campestris pv. malvacearum*. Because *ERF1* integrates signals from the JA and ET defence signalling pathways, the constitutive expression of *ERF* family members activates the expression of several JA/ET-dependent defence genes and induces resistance against necrotrophic pathogens. For instance, expression of several *PR genes* which confer resistance against several necrotrophs (e.g., *PR3* and *PR5d* and *PDF1.2*) is modulated by *ERFs*. These defence genes possess a GCC box in their promoters, which is a direct target for the action of *ERFs* [156].

Although ET has been shown to regulate plant defence responses against fungi and bacteria, ET is probably not essential in plant resistance against viruses. Recently, 1-aminocyclopropane-1- carboxylic acid (ACC) was shown to enhance *TMVcg* accumulation in treated plants [157], which increased susceptibility, suggesting that ET is required for viral infection.

Other phytohormones, such as ABA, gibberellins (GBs), auxins, brassinosteroids and cytokinins (CKs), have recently emerged as defence regulators [158]. ABA, a sesquiterpene compound resulting from the cleavage of γ-carotene, regulates numerous developmental processes and adaptive stress responses in plants. ABA can positively regulate plant defence at the early stages of infection by mediating stomatal closure against invaders, or inducing callose deposition if the pathogen evades the first line of defence [159]. If activated at later stages, ABA can suppress ROS induction and SA or JA signal transduction, thereby negating defences controlled by these two pathways [160].

Cytokinins promote cell division, and are known to play a role in the synthesis and maintenance of chlorophyll and chloroplast development and metabolism. CKs are also involved in the modulation of defence mechanisms, including the induction of resistance against viruses [161, 162], but are known to suppress HR [163]. Cytokinins can however act synergistically with SA signalling [164]. CKs activate the transcriptional regulator *ARABIDOPSIS RESPONSE REGULATOR 2* (*ARR2*), which positively modulates SA signalling by interacting with the SA-responsive factor *TGA3* [165]. *TGA3* induces the binding of *ARR2* to the promoters of *PR-1* and *PR-2* to induce cytokinin-dependent gene transcription. Correspondingly, the *npr1-1* or *NahG* mutants fail to modulate the induction of *ARR2* when treated with CK, indicating that CK modulates signaling components downstream of SA. Moreover, increased transcription of genes involved in SA-biosynthesis and signalling (e.g., *SID1*, *SID2*, *PR-1* and *PR-5*) is observed in *ARR2* over-expressing mutants challenged with *P. syringae* pv. *tomato* (*Pst DC3000*). Thus, CKs synergistically interacts not only with the SA signaling pathway to boost SA dependent induction of plant defence genes but also modulates SA biosynthesis. Cytokinins have also been shown to enhance the production of two antimicrobial phytoalexins, scopoletin and capsidiol in tobacco plants challenged with *P. syringae* pv. *tabaci* (*Pst*) independent of SA signalling [166]. Moreover, cytokinins induce the expression of cell wall invertase, a key

sucrose cleaving enzyme required for carbohydrates supply through an apoplasmic pathway [167]. Invertase is required for plant defence against pathogens, including *Pst*. The glucose target of rapamycin (*TOR*) signalling pathway involved in autophagy apparently modulates the transcriptional dynamics associated with cytokinin-invertase-induced defence pathway by providing the required energy, metabolites and the cell cycle machinery required for cytokinin signal transduction [168]. The link between autophagy and cytokinin signalling was previously suggested [169], but the cytokinin-induced defence system in this interplay is probably a protective mechanism to maintain plant growth and proliferation despite pathogen challenge [170].

Brassinosteroids (BRs) are a class of polyhydroxysteroids that affect many cellular processes including elongation, proliferation, differentiation, membrane polarization and proton pumping [171]. BRs are increasingly becoming important in plant defence against pathogens. The mechanism underlying BR signalling involves the direct binding of BRs such as BL and castasterone to the LRR-RLK (*BRI1*). This interaction is reported to unlock *BRI1* from the negative regulator *BKI1*, followed by heterodimerization of *BRI1* with a co-receptor *BAK1* and phosphorylation of the *BRI1*-interacting signalling kinase (*BSK1*). Other events include the activation of the protein phosphatase *BSU1*. These biochemical changes inhibit the shaggy-like kinase *BIN2*, which culminates into the activation of the homologous TFs, *BZR1* and *BES1/BZR2* [172]. These TFs translocate to the nucleus, interact with BR-responsive promoters, and cause transcriptional changes that eventually lead to defence response. BRs have been demonstrated to enhance plant defence against pathogens. In potato, BRs have been shown to be effective against viral infection from the starting planting materials to the second tuber generation [173]. Furthermore, application of BRs on tobacco plants decreases TMV viral load and restricts infection by other biotrophs [174]. The same authors reported that *BAK1* is essential for plant basal immunity during compatible interactions with RNA viruses. The *BAK1* mutants, *bak1-4* and *bak1-5*, accumulate *turnip crinkle virus* (TCV), *oilseed rape mosaic virus* (ORMV) and TMV to higher levels compared to the WT plants [174]. Thus, *BAK1* could probably be a general regulator of plant defence against biotrophs and hemibiotrophs. BRs have also been reported to interact with other phytohormones, such as GA and auxins, but independent of SA [175]. For details on auxin- and cytokinin-modulated immunity, and GA/BR interaction, the reader is referred to excellent reviews [176, 177]. Furthermore, details on the interaction of BRs and SA, including their effect on SAR marker genes (e.g., *PR-1*, *PR-2* and *PR-5*) can be found in [178].

Taken together, the intricate cross-talk among hormones to cooperate with other signals and to coordinate appropriate induction of defences against pathogens and/or insect pests depends on the pathogen type, physiological stage and environmental and probably circadian regulations.

5. RNAi-mediated plant defence

RNA interference or silencing is one of the emergent crop improvement strategies that involve sequence-specific gene regulation by small non-coding RNAs, which mainly belong to two

categories, i.e., small interfering RNA (siRNA) and microRNA (miRNA). Though these sRNAs differ in biogenesis [179], both regulate the target gene repression through ribonucleoprotein silencing complexes. Plant RNA silencing involves four basic steps, which include introduction of double-stranded RNA (dsRNA) into the cell, processing of dsRNA into 18–25-nt small RNA (sRNA), sRNA 2-O-methylation and sRNA incorporation into effector complexes that interact with target RNA or DNA [180]. The formation of RNA-induced silencing complex (*RISC*) and its incorporation into the antisense strand of siRNAs, which interacts with Argonaute and other effector proteins, precedes the cleavage of the target mRNA. For details about the formation of *RISC* and cleavage of the target mRNA, the reader is referred to comprehensive reviews [179, 181]. For sRNA to meet the target mRNA, it has to move from the point of initiation to the target. Thus, two main movement categories include cell-to-cell (short-range; symplastic movement through the plasmodesmata) and systemic (long-range; through the vascular phloem) movement. These mobile silencing strategies use sRNAs to target mRNA in a nucleotide sequence specific manner. By use of fluorescently labelled 21 and 24-nt siRNAs, Dunoyer et al. [182] demonstrated the movement of siRNAs from cell to cell and over long distances. Such systematic movements enhance systemic silencing of viruses as reported in *N. benthamiana* [183]. Similar systemic movements have been reported in the phloem sap of oilseed rape [184] and pumpkin [185]. Endogenous 21-nt miRNAs (miR399) were also reported to be mobile within the roots [186], and between shoots and roots of rapeseed and pumpkin [187]. Thus, sRNAs can be targeted to most active plant tissues, with transcription activity, to achieve a desirable consequence.

Several RNAi strategies have shown success in plant improvement against biotic stresses. Arabidopsis *miR393* was the first sRNA implicated in bacterial PTI [188], and enhanced *miR393* accumulation was found during sRNA profiling in Arabidopsis challenged with *Pst* [189]. The mechanism of *miR393*-induced resistance involves repression of auxin signalling by negatively regulating the F-box auxin receptors like *transport inhibitor response 1* (*TIR1*). This process restricts *Pst* infection, and, indeed, plants overexpressing *miR393* exhibit effective resistance against *Pst* [188].

RNAi in plant resistance to fungi has also shown promise. For instance, RNAi-mediated suppression of a rice gene *OsSSI2* enhances resistance towards *M. oryzae* and *X. oryzae* [189]. Moreover, RNAi suppression of *OsFAD7* and *OsFAD8*, the two genes encoding for Ω-3 fatty acid desaturase, also enhances resistance against *M. oryzae* [190]. RNAi targeting of lignin production pathway genes aimed at reducing lignin content has also been shown to enhance resistance against *Sclerotinia sclerotiorum* in soybean [191]. Increased resistance to *Blumeria graminis* f. sp. *tritici* in wheat was also demonstrated through RNAi using 24 miRNAs [192]. Nevertheless, the performance of these approaches under environmental conditions has often been unsatisfactory and environmental influences in expression of resistance often remain unpredictable [205].

In response to virus infection, several cases have shown successful crop improvement. For instance, resistance to *African Cassava Mosaic Virus* (CMV) was achieved in transgenic cassava plants producing dsRNA against PSTVd sequences [193]. A similar strategy was successful in transgenic tomato resistance against *Potato Spindle Tuber Viroid* (PSTVd) [194]. RNAi targeting

of the virus coat protein has also been successfully engineered into plants to induce resistance against viruses. For instance, transgenic tobacco plants expressing the CP gene of TMV are resistant to TMV. The resistance of *N. benthamiana* to *Cucumber Green Mottle Mosaic Virus* (CGMMV); and that of *Prunus domestica* to *Plum Pox virus* (PPV) are other examples documented; for review see [179].

In functional biology studies, virus-induced gene silencing (VIGS) has emerged to be one of the most powerful RNA-mediated post-transcriptional gene silencing (PTGS), not only in plant protection against viruses, but also for gene knockouts in functional genomic studies [195, 196].

Although RNAi has the potential to contribute to increased crop productivity, by generating crops with improved resistance against pests and diseases, it would be even better if interaction between sRNAs and their targets is validated in several backgrounds. This would provide valuable insight into mechanisms of post-transcriptional gene regulation and multiple molecular pathways controlling plant stress responses. However, the danger of unintentional silencing of genes with regions of homology to the intended target, and target mutations leading to easier escape from miRNA-directed silencing are still ethical issues. Certain biosafety concerns on the use of RNAi transgenics, especially transcriptional gene silencing by chromatin modification is even a more sensitive and contentious issue, as it is rumoured to lead to hereditary changes associated with adverse effects. Thus, the underlying mechanisms associated with RNAi require further investigations using well-controlled experiments.

6. Modern approaches for improving biotic stress tolerance in plants

Conventional breeding methods still play an important role in the selection of new varieties. However, emerging tools in biotechnology are much needed to maximize the probability of success. One area of biotechnology, molecular marker assisted breeding (MAB), has already made significant impact in improving efficiency of conventional breeding. There are, however, major gaps in the improvement of traits controlled by a large number of small effects, epistatic QTLs displaying significant genotype × environment (G × E) interactions. Thus, accurate indirect selections based on genomic tools that have emerged over the last few decades are continuously being employed to improve the breeding efficiency for such traits. The advantage is that, to date, the genome sequences for more than 55 plant species have been produced and many more are being sequenced [197]. The genome sequence information available enables the identification and development of genomewide markers. Availability of markers covering the whole genomic regions has already shown promise in the development of special populations, such as recombinant inbred lines (RILs), near isogenic lines (NILs), introgression lines (ILs) or chromosome segment substitution lines (CSSLs). Recently, heterogeneous inbred family (HIFs) and multi-parent advanced generation inter-cross (MAGIC) populations, which can serve the dual purpose of permanent mapping populations for precise QTL mapping and for direct or indirect use in variety development, have shown promise in plant breeding. Also, genomewide association (GWA) analysis has been successfully applied to rice, maize, barley, wheat, sesame and other plants. GWA has also been adapted to the "breeding by design"
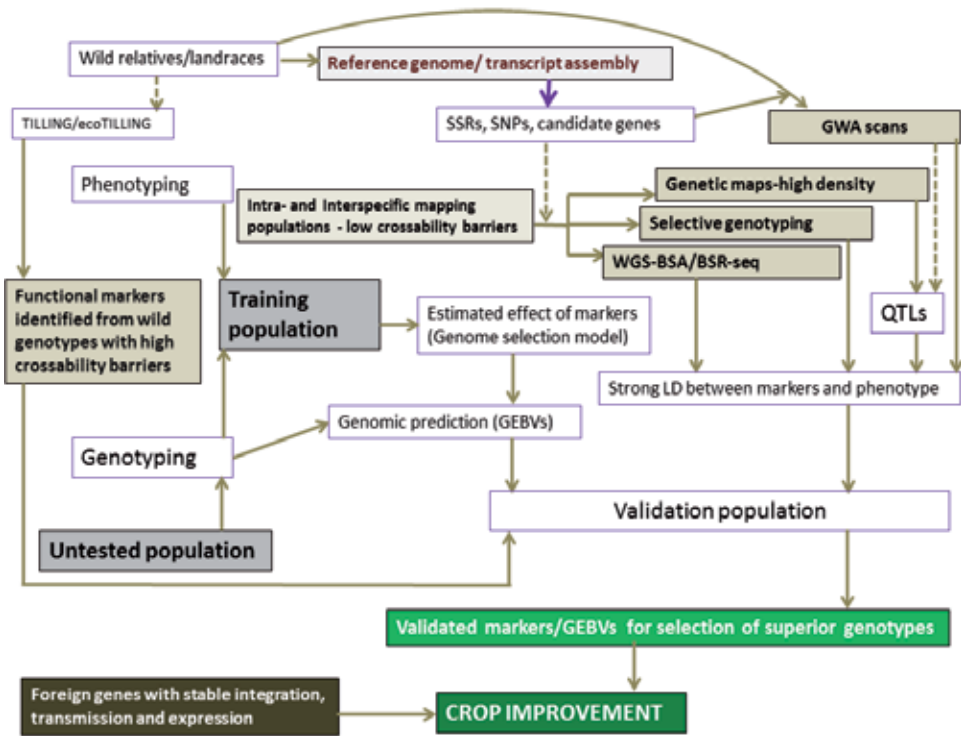
Figure 2. Principle of genomic selection. Two steps are involved; developing a training population to provide phenotypic and genotypic data; effects are estimated for all molecular markers. The second step involves genotyping untested populations and selecting superior genotypes based on their expected phenotypes according to the estimates obtained from the marker effects on the training population (bottom).

approach, often referred to as genome selection (Figure 2), which predicts the outcome of a set of crosses on the basis of molecular markers information.

Recently, a combination of different approaches has been used to develop new rice cultivars referred to as 'Green Super Rice', possessing resistance to multiple insects and diseases, high nutrient efficiency and drought resistance. If fully exploited, the integration of a similar approach with breeding by design or genome selection would help to design new plant types with not only a few selected major loci, but nearly all the functional loci of the genome controlling key desirable traits in commercial cultivars.

Expression studies also present a major area of interest for breeders. Among them, the NGS technologies have become the mainstay of studying complex traits, as direct sequencing of genomes and comparison with reference sequences is increasingly becoming more feasible. Re-sequencing has been performed for model species, e.g., *Arabidopsis*, to understand the whole genome sequence variation, and ultimately discover single nucleotide polymorphisms (SNPs). Similar re-sequencing efforts have been applied in rice, maize, soybean, grape and poplar. Combining re-sequencing with the recent developments in omic biology, including
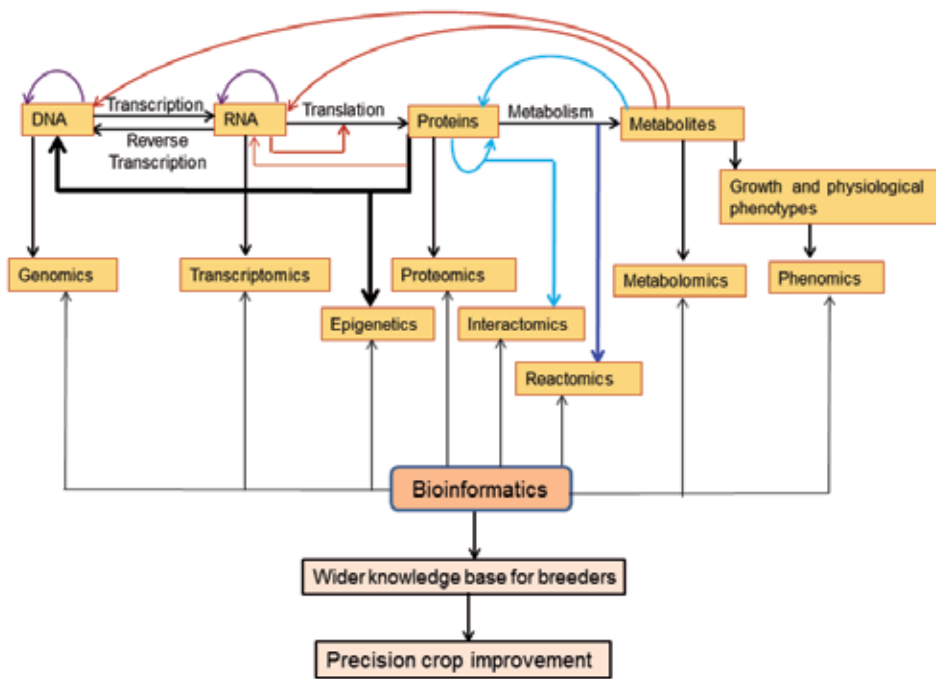
Figure 3. Supportive omic tools for increasing plant breeding efficiency against biotic stresses. Sky blue lines indicate interactions; largest bold black lines indicate epigenetic regulation; red lines indicate regulation; and blue line indicates metabolic reactions.

transcriptomics, proteomics, metabolomics, epigenetics and physiological and biochemical methods (Figure 3) will remarkably provide novel possibilities to understand the biology of plants and consequently to precisely develop stress tolerant crop varieties.

The recent advent of genotyping by sequencing (GBS) approach that minimizes ascertainment biases and the need for prior genome sequence information associated with traditional techniques has also enabled single nucleotide polymorphism marker detection, exposition of QTLs and the discovery of candidate genes controlling stress tolerance. Thus, genome/ transcript profiling when combined with genome variation analysis is a potential area which could prove useful for breeders in the near future [205, 209]. Another newly developed approach, which combines genetical genomics and bulk segregant analysis (BSA) to identify markers linked to genes, shows the possibility of coupling BSA to high throughput sequencing methods. Although there are shortcomings, including errors introduced during NGS proce-dures, this method has proven to be useful in identifying stress tolerance genomic regions in crop plants. A more recent modification that exploits the power of deep sequencing of target-enriched SNP markers to increase the efficiency of BSA analysis is called target-enriched TEX-QTL mapping [197]. The authors propose that by combining a large F2 population size, deeply sequenced markers, and 10–20% bulk size, most QTLs can be identified within two genera-tions. Although it does not currently detect very closely linked QTL, TEX-QTL method is

potentially a useful development in plant breeding. It is envisaged that BSA, by genotyping pooled-segregant sequencing, is likely to increase the reliability and reduce the time required to map all QTL defining the trait of interest and to identify causative superior alleles that can subsequently be used for crop improvement by targeted genetic engineering.

Desirable alleles are also being identified using functional genomic tools, including transformation, insertional mutagenesis, RNAi, the screening of either mutant or natural germplasm collections by means of targeting induced local lesions in genomes (TILLING) or ecotype TILLING (EcoTILLING) methodologies. These strategies enable plant scientists to predict gene functions and allow efficient prediction of the phenotype associated with a given gene, the so-called reverse genetics approach. The availability of a large volume of sequences generated through NGS technologies is significantly increasing the number and quality of candidates for TILLING and EcoTILLING studies. Thus, a number of crops have benefited from these technologies, including Arabidopsis, lotus, barley, maize, pea, melon and rice, for review see [198].

The use of improved recombinant DNA techniques to introduce new traits in early phases of cultivar selection is also currently gaining momentum in plant biology. Techniques such as oligonucleotide-directed mutagenesis (oDM) as well as those based on zinc finger nuclease (ZFN), transcription activator-like effector nuclease (TALeN) and clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated protein 9 (Cas9) system are all capable of specifically modifying a given target sequence leading to genotypes not substantially different from those obtained through traditional mutagenesis. The practical use of these techniques in developing countries and the performance of the germplasm developed through them under environmental conditions [206, 207, 208] is yet to be fully demonstrated.

7. Conclusion and perspective

Plant resistance to biotic stresses is jointly controlled by the plants' anatomy, physiology, biochemistry, genetics, development and evolution. Efforts to understand these mechanisms have generated a lot of data on candidate genes, quantitative trait loci (QTLs), proteins and metabolites associated with plant defences. This chapter has reviewed most of these aspects to provide a reader with background information on the diverse plant defence patterns. Some of the genes and methods that hold promise for improving plant defences are also discussed. Certainly, plant-pathogen/insect interaction is a complex phenomenon that involves various signalling pathways tracking and regulating the pathogens/insect ingress. The interactions leading to effective defence apparently involve activation of both innate and systemic acquired resistance, and require both direct and indirect pathways to rapidly limit the entry or proliferation of biotic agents in the plant. Identifying and harmonizing an efficient defence signalling pathway, which leads to activation of an effective defence strategy, is still a challenge, considering the large number of genes and proteins often expressed in most plant-pathogen/insect interaction studies. However, there are some resistance components that have shown promise, although further studies would be necessary to clarify the signalling patterns in

which such components are involved. Important examples include LRR-RK *BAK1*, which features in several signalling networks leading to plant resistance against a diversity of pathogens and insects, and *NRC1* which mediates resistance and cell death induced by both membrane receptors and intracellular NLRs. *BAK1* forms heteromeric complexes with other receptors, which indicates that BAK1 is a multifaceted receptor capable of PAMP detection, while *NRC1* is probably a downstream convergence point in ETI initiated at various cell locations. Thus, *BAK1* and *NRC1* could probably contribute to effective plant resistance to a diversity of pathogens and insects. However, identification of additional effective receptors will be necessary to counter the stealthy tendencies of most pathogens and insects, and to guarantee the transmission of signals to the downstream components. More studies on adaptability of defence genes or QTLs to changing biotic agents and climatic conditions also need to be conducted in order to limit boom and bust incidences frequently observed in pathosystems.

Acknowledgements

Author details

Geoffrey Onaga[1] and Kerstin Wydra[2*]

*Address all correspondence to: kerstin.wydra@fh-erfurt.de

1 National Crops Resources Research Institute, Kampala, Uganda

2 Erfurt University of Applied Sciences, Faculty of Landscape Architecture, Horticulture and Forestry, Erfurt, Germany

References

[1] Hussain B. Modernization in plant breeding approaches for improving biotic stress resistance in crop plants. Turkish Journal of Agriculture and Forestry. 2015;39:515–530. DOI: 10.3906/tar-1406-176

[2] Sanghera GS, Wani SH, Singh G, Kashyap PL, Singh NB. Designing crop plants for bioticstresses using transgenic approach. International Journal of Plant Research. 2011;24:1–25

[3] Anderson PK, Cunningham AA, Patel NG, Morales FJ, Epstein PR, Daszak P. Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnolo-

gy drivers. Trends in Ecology & Evolution. 2004;19:535–544. DOI: 10.1016/j.tree. 2004.07.021

[4] Ijaz S, Khan AI. Genetic pathways of disease resistance and plants-pathogens interactions. Molecular Pathogens. 2012;3(4):19–26. DOI: 10.5376/mp.2012.03.0004

[5] Howe GA, Jander G. Plant immunity to insect herbivores. Annual Reviews in Plant Biology. 2008;59:41–66. DOI: 10.1146/annurev.arplant.59.032607.092825

[6] Nurnberger T, Kemmerling B. Pathogen-associated molecular patterns (PAMP) and PAMP-triggered immunity. In: Parker J, editor. Molecular Aspects of Plant Disease Resistance. Annual Plant Reviews. Oxford, UK: Wiley. 2009;34:16–47. DOI: 10.1002/9781444301441.ch2

[7] Ausubel FM. Are innate immune signaling pathways in plants and animals conserved? Nature Immunology. 2005;6:973–979. DOI: 10.1038/ni1253

[8] Monaghan J, Zipfel C. Plant pattern recognition receptor complexes at the plasma membrane. Current Opinion in Plant Biology. 2012;15(4):349–357. DOI: 10.1016/j.pbi. 2012.05.006

[9] Kiraly L, Barnaz B, Kiralyz Z. Plant resistance to pathogen infection: forms and mechanisms of innate and acquired resistance. Journal of Phytopathology. 2007;155:385–396. DOI: 10.1111/j.1439-0434.2007.01264.x

[10] Jones JD, Dangl JL. The plant immune system. Nature. 2006;444:323–29. DOI: 10.1038/ nature05286

[11] Ferreira RB, Monteiro S, Freitas R, Santos CN, Chen Z, Batista LM, Duarte J, Borges A, Teixeira AR. The role of plant defence proteins in fungal pathogenesis. Molecular Plant Pathology. 2007;8:677–700. DOI: 10.1111/j.1364-3703.2007.00419.x

[12] Freeman BC, Beattie GA. An Overview of Plant Defenses against Pathogens and Herbivores. The Plant Health Instructor. 2008; DOI: 10.1094/PHI-I-2008-0226-01

[13] War AR, Paulraj MG, War MY, Ignacimuthu S. Herbivore and elicitor-induced resistance in groundnut to Asian armyworm, Spodoptera litura (Fab.) (Lepidoptera: Noctuidae). Plant Signaling & Behavior. 2011;6(11):1769–77. DOI: 10.4161/psb.6.11.17323

[14] Boller T, Felix G. A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. Annual Review in Plant Biology. 2009;60:379–406. DOI: 10.1146/annurev.arplant.57.032905.105346

[15] Malinovsky FG, Fangel JU, Willats WGT. The role of the cell wall in plant immunity. Frontiers in Plant Science. 2014;5:178. DOI: 10.3389/fpls.2014.00178

[16] Wei T, Ou B, Li J, Zhao Y, Guo D, Zhu Y, Guo D, Zhu Y, Chen Z, Gu H, Li C, Qin C, Qu LJ. Transcriptional profiling of rice early response to *Magnaporthe oryzae* identified OsWRKYs as important regulators in rice blast resistance. PLoS ONE. 2013;8(3):e59720. DOI: org/10.1371/journal.pone.0059720

[17] Dodds PN, Rathjen JP. Plant immunity: towards an integrated view of plant–pathogen interactions. Nature Reviews Genetics. 2010;11:539–548. DOI: 10.1038/nrg2812

[18] Van der Hoorn RA, Jones JD. The plant proteolytic machinery and its role in defence. Current Opinion in Plant Biology. 2004;7:400–407. DOI: 10.1016/j.pbi.2004.04.003

[19] Schwessinger B, Zipfel C. News from the frontline: recent insights into PAMP-triggered immunity in plants. Current Opinion in Plant Biology. 2008;11:389–395. DOI: 10.1016/j.pbi.2008.06.001

[20] Belien T, Van Campenhout S, Robben J, Volckaert G. Microbial endoxylanases: effective weapons to breach the plant cell-wall barrier or, rather, triggers of plant defense systems? Molecular Plant Microbe Interaction. 2006;19:1072–1081. DOI: 10.1094/MPMI -19-1072

[21] Mithöfer A, Fliegmann J, Neuhaus-Url G, Schwarz H, Ebel J. The hepta–beta–glucoside elicitor-binding proteins from legumes represent a putative receptor family. Biological Chemistry. 2000;381:705–713. DOI: 10.1515/BC.2000.091

[22] Zipfel C. Early molecular events in PAMP-triggered immunity. Current Opinion in Plant Biology. 2009;12:414–420. DOI: 10.1016/j.pbi.2009.06.003

[23] Felix G, Duran JD, Volko S, Boller T. Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. The Plant Journal. 1999;18:265–276. DOI: 10.1046/j.1365-313X.1999.00265.x

[24] Trdá L, Fernandez O, Boutrot F, Héloir MC, Kelloniemi J, Daire X, Adrian M, Clément C, Zipfel C, Dorey S, Poinssot B. The grapevine flagellin receptor VvFLS2 differentially recognizes flagellin-derived epitopes from the endophytic growth-promoting bacterium *Burkholderia phytofirmans* and plant pathogenic bacteria. New Phytologist. 2014;201:1371–1384. DOI: 10.1111/nph.12592

[25] Robatzek S, Bittel P, Chinchilla D, Kochner P, Felix G, Shiu SH, Boller T. Molecular identification and characterization of the tomato flagellin receptor LeFLS2, an orthologue of Arabidopsis FLS2 exhibiting characteristically different perception specificities. Plant Molecular Biology. 2007;64:539–547. DOI: 10.1007/s11103-007-9173-8

[26] Takai R, Isogai A, Takayama S, Che F. Analysis of flagellin perception mediated by flg22 receptor OsFLS2 in rice. Molecular Plant Microbe Interaction. 2008;21:1635–1642. DOI: 10.1094/MPMI-21-12-1635

[27] Takeuchi K, Taguchi F, Inagaki Y, Toyoda K, Shiraishi T, Ichinose Y. Flagellin glycosylation island in *Pseudomonas syringae* pv. *glycinea* and its role in host specificity. Journal of Bacteriology. 2003;185:6658–6665. DOI: 10.1128/JB.185.22.6658-6665.2003

[28] Taguchi F, Takeuchi K, Katoh E, Murata K, Suzuki T, Marutani M, Kawasaki T, Eguchi M, Katoh S, Kaku H, Yasuda C, Inagaki Y,Toyoda K, Shiraishi T, Ichinose Y. Identification of glycosylation genes and glycosylated amino acids of flagellin in *Pseudomonas syringae* pv. *tabaci*. Cellular Microbiology. 2006;8:923–938. DOI: 10.1111/j. 1462-5822.2005.00674.x

[29] Cai R, Lewis J, Yan S, Liu H, Clarke CR, Campanile F, Almeida NF, Studholme DJ, Lindeberg M, Schneider D, Zaccardelli M, Setubal JC, Morales-Lizcano NP, Bernal A, Coaker G, Baker C, Bender CL, Leman S, Vinatzer BA. The plant pathogen *Pseudomonas syringae* pv. *tomato* is genetically monomorphic and under strong selection to evade tomato immunity. PLoS Pathogens. 2011;7:e1002130. DOI: org/10.1371/journal.ppat.1002130

[30] Clarke CR, Chinchilla D, Hind SR, Taguchi F, Miki R, Ichinose Y, Martin GB, Leman S, Felix G, Vinatzer BA. Allelic variation in two distinct *Pseudomonas syringae* flagellin epitopes modulates the strength of plant immune responses but not bacterial motility. New Phytologist. 2013;200:847–860. DOI: 10.1111/nph.12408

[31] Chinchilla D, Zipfel C, Robatzek S, Kemmerling B, Nürnberger T, Jones JD, Felix G, Boller T. A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defence. Nature. 2007;448:497–500. DOI: 10.1038/nature05999

[32] Lu D, Wu S, Gao X, Zhang Y, Shan L He P. A receptor-like cytoplasmic kinase, BIK1, associates with a flagellin receptor complex to initiate plant innate immunity. Proceedings of the National Academy of Science of the United States of America. 2010;107:496–501. DOI: 10.1073/pnas.0909705107

[33] Kunze G, Zipfel C, Robatzek S, Niehaus K, Boller T, Felix G. The N terminus of bacterial elongation factor Tu elicits innate immunity in Arabidopsis plants. The Plant Cell. 2004;16:3496–3507. DOI: 10.1105/tpc.104.026765

[34] Zipfel C, Kunze G, Chinchilla D, Caniard A, Jones JD, Boller T, Felix G. Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts Agrobacterium-mediated transformation. Cell. 2006;125:749–760. DOI: http://dx.doi.org/10.1016/j.cell.2006.03.037

[35] Lacombe S, Rougon-Cardoso A, Sherwood E, Peeters N, Dahlbeck D, van Esse HP, Smoker M, Rallapalli G, Thomma BP, Staskawicz B, Jones JD, Zipfel C. Interfamily transfer of a plant pattern-recognition receptor confers broad-spectrum bacterial resistance. Nature Biotechnology. 2010;28:365–369. DOI: 10.1038/nbt.1613

[36] Furukawa T, Inagaki H, Takai R, Hirai H, Che FS. Two distinct EF-Tu epitopes induce immune responses in rice and Arabidopsis. Molecular Plant Microbe Interaction. 2014;27(2):113–124. DOI: 10.1094/MPMI-10-13-0304-R

[37] Lu F, Wang H, Wang S, Jiang W, Shan C, Li B, Yang J, Zhang S, Sun W. Enhancement of innate immune system in monocot rice by transferring the dicotyledonous elongation factor Tu receptor EFR. Journal of Integrative Plant Biology. 2015;57(7):641–652. DOI: 10.1111/jipb.12306

[38] Schulze B, Mentzel T, Jehle AK, Mueller K, Beeler S, Boller T, Felix G, Chinchilla D. Rapid heteromerization and phosphorylation of ligand-activated plant transmembrane receptors and their associated kinase BAK1. Journal of Biological Chemistry. 2010;285:9444–9451. DOI: 10.1074/jbc.M109.096842

[39] Song WY, Pi LY, Wang GL, Gardner J, Holsten T, Ronald PC. Evolution of the rice Xa21 disease resistance gene family. The Plant Cell. 1997;9:279–287. DOI: 10.1105/tpc. 9.8.1279

[40] Mendes BMJ, Cardoso SC, Boscariol-Camargo RL, Cruz RB, Mourão Filho FAA, Bergamin Filho A. Reduction in susceptibility to Xanthomonas axonopodis pv. citri in transgenic *Citrus sinensis* expressing the rice Xa21 gene. Plant Pathology. 2010;59:68–75. DOI: 10.1111/j.1365-3059.2009.02148.x

[41] Afroz A, Chaudhry Z, Rashid U, Ali GM, Nazir F, Iqbal J, Khan MR. Enhanced resistance against bacterial wilt in transgenic tomato (*Lycopersicon esculentum*) lines expressing the Xa21 gene. Plant Cell, Tissue and Organ Culture. 2011;104:227–237. DOI: 10.1007/s11240-010-9825-2

[42] Tripathi JN, Lorenzen J, Bahar O, Ronald P, Tripathi L. Transgenic expression of the rice Xa21 pattern-recognition receptor in banana (Musa sp.) confers resistance to *Xanthomonas campestris* pv. *musacearum*. Plant Biotechnology Journal. 2014;12(6):663–673. DOI: 10.1111/pbi.12170

[43] Fradin E, Adb-El-Haliem A, Masini L, van den Berg G, Joosten M, Thomma B. Interfamily transfer of tomato Ve1 mediates Verticillium resistance in Arabidopsis. Plant Physiology. 2011;156:2255–2265. DOI: 10.1104/pp.111.180067

[44] Jashni MK, Mehrabi R, Collemare J, Mesarich CH, de Wit PJGM. The battle in the apoplast: further insights into the roles of proteases and their inhibitors in plant–pathogen interactions. Frontiers in Plant Science. 2015;6:584. DOI: 10.3389/fpls. 2015.00584.

[45] Shimizu T, Nakano T, Takamizawa D, Desaki Y, Ishii-Minami N, Nishizawa Y, Minami E, Okada K, Yamane H, Kaku H, Shibuya N. Two LysM receptor molecules, CE-BiP and OsCERK1, cooperatively regulate chitin elicitor signaling in rice. The Plant Journal. 2010;64:204–214. DOI: 10.1111/j.1365-313X.2010.04324.x

[46] Hayafune M, Berisio R, Marchetti R, Silipo A, Kayama M, Desaki Y, Arima S, Squeglia F, Ruggiero A, Tokuyasu K, Molinaro A, Kaku H, Shibuya N. Chitin-induced activation of immune signaling by the rice receptor CEBiP relies on a unique sandwich-type dimerization. Proceedings of the National Academy of Science of the United States of America. 2014;111:404–413. DOI: 10.1073/pnas.1312099111

[47] Shinya T, Motoyama N, Ikeda A, Wada M, Kamiya K, Hayafune M, Kaku H, Shibuya N. Functional characterization of CEBiP and CERK1 homologs in Arabidopsis and rice reveals the presence of different chitin receptor systems in plants. Plant Cell Physiology. 2012;53:1696–1706. DOI: 10.1093/pcp/pcs113

[48] Zipfel C. Plant pattern-recognition receptors. Trends in Immunology. 2014;35:345–351. DOI: 10.1016/j.it.2014.05.004

[49] Liu T, Liu Z, Song C, Hu Y, Han Z, She J, Fan F, Wang J, Jin C, Chang J, Zhou JM, Chai J. Chitin-induced dimerization activates a plant immune receptor. Science. 2012;336:1160–1164. DOI: 10.1126/science.1218867

[50] Noda J, Brito N, González C. The Botrytis cinerea xylanase Xyn11A contributes to virulence with its necrotizing activity, not with its catalytic activity. BMC Plant Biology. 2010;10:1–15. DOI: 10.1186/1471-2229-10-38

[51] Sella L, Gazzetti K, Faoro F, Odorizzi S, D'Ovidio R, Schafer W, Favaron F. A *Fusarium graminearum* xylanase expressed during wheat infection is a necrotizing factor but is not essential for virulence. Plant Physiology and Biochemistry. 2013;64:1–10. DOI: 10.1016/j.plaphy.2012.12.008.

[52] Ron M, Avni A. The receptor for the fungal elicitor ethylene-inducing xylanase is a member of a resistance-like gene family in tomato. Plant Cell. 2004;16:1604–1615. DOI: 10.1105/tpc.022475

[53] Bar M, Sharfman M, Ron M, Avni A. BAK1 is required for the attenuation of ethylene-inducing xylanase (Eix)-induced defense responses by the decoy receptor LeEix1. The Plant Journal. 2010;63:791–800. DOI: 10.1111/j.1365-313X.2010.04282.x.

[54] Juge N. Plant protein inhibitors of cell wall degrading enzymes. Trends in Plant Science. 2006;11:359–367. DOI: 10.1016/j.tplants.2006.05.006

[55] Bellincampi D, Cervone F, Lionetti V. Plant cell wall dynamics and wall-related susceptibility in plant–pathogen interactions. Frontiers in Plant Science. 2014;5:228. DOI: 10.3389/fpls.2014.00228

[56] Larkan NJ, Lydiate DJ, Parkin IAP, Nelson MN, Epp DJ, Cowling WA, Rimmer SR, Borhan MH. The Brassica napus blackleg resistance gene LepR3 encodes a receptor-like protein triggered by the *Leptosphaeria maculans* effector AVRLM1. New Phytologist. 2013;197:595–605. DOI: 10.1111/nph.12043

[57] Ma L, Borhan MH. The receptor-like kinase SOBIR1 interacts with *Brassica napus* LepR3 and is required for *Leptosphaeria maculans* AvrLm1-triggered immunity. Frontiers in Plant Science. 2015;6:933. DOI: 10.3389/fpls.2015.00933

[58] Zhang W, Fraiture M, Kolb D, Löffelhardt B, Desaki Y, Boutrot FFG, Tör M, Zipfel C, Gust AA, Brunner F. Arabidopsis receptor-like protein30 and receptor-like kinase suppressor of BIR1-1/EVERSHED mediate innate immunity to necrotrophic fungi. The Plant Cell. 2013;25:4227–4241. DOI: 10.1105/tpc.113.117010

[59] Klemptner RL, Sherwood JS, Tugizimana F, Dubery IA, Piater LA. Ergosterol, an orphan fungal MAMP. Molecular Plant Microbe Interaction 2014;15(7):747–761. DOI: 10.1111/mpp.12127

[60] Bostock RM, Laine RA, Kuć JA. Factors affecting the elicitation of sesquiterpenoid phytoalexin accumulation by eicosapentaenoic and arachidonic acids in potato. Plant Physiology. 1982;70:1417–1424. DOI: http://dx.doi.org/10.1104/pp.70.5.1417

[61] Tyler BM. Molecular basis of recognition between Phytophthora pathogens and their hosts. Annual Reviews Phytopathology. 2002;40:137–167. DOI: 10.1146/annurev.phyto. 40.120601.125310

[62] Brunner F, Rosahl S, Lee J, Rudd JJ, Geiler C, Kauppinen S, Rasmussen G, Scheel D, Nurnberger T. Pep-13, a plant defense-inducing pathogen-associated pattern from *Phytophthora transglutaminases*. The EMBO Journal. 2002;21:6681–6688. DOI: 10.1093/emboj/cdf667

[63] Garcia-Brugger A, Lamotte O, Vandelle C, Bourque S, Lecourieux D, Poinsot B, Wendehenne D, Pugin A. Early signaling events induced by elicitors of plant defenses. Molecular Plant Microbe Interaction. 2006;19:711–724. DOI: 10.1094/MPMI-19-0711.

[64] Gaulin E, Drame N, Lafitte C, Torto-Alalibo T, Martinez Y, Ameline-Torregrosa C, Khatib M, Mazarguil H, Villalba-Mateos F, Kamoun S, Mazars C, Dumas B, Bottin A, Esquerre-Tugaye MT, Rickauer M. Cellulose binding domains of a Phytophthora cell wall protein are novel pathogen-associated molecular patterns. The Plant Cell. 2006;18:1766–1777. DOI: 10.1105/tpc.105.038687

[65] Nicaise V. Crop immunity against viruses: outcomes and future challenges. Frontiers in Plant Science. 2014;5:660. DOI: 10.3389/fpls.2014.00660.

[66] Krol E, Mentzel T, Chinchilla D, Boller T, Felix G, Kemmerling B, Postel S, Arents M, Jeworutzki E, Al-Rasheid KA, Becker D, Hedrich R. Perception of the Arabidopsis danger signal peptide 1 involves the pattern recognition receptor AtPEPR1 and its close homologue AtPEPR2. Journal of Biological Chemistry. 2010;285:13471–13479. DOI: 10.1074/jbc.M109.097394

[67] Korner CJ, Klauser D, Niehl A, Domínguez-Ferreras A, Chinchilla D, Boller T, Heinlein M, Hann DR. The immunity regulator BAK1 contributes to resistance against diverse RNA viruses. Molecular Plant Microbe Interaction. 2013;26:1271–1280. DOI: 10.1094/MPMI-06-13-0179-R

[68] Fontes EP, Santos AA, Luz DF, Waclawovsky AJ, Chory J. The geminivirus nuclear shuttle protein is a virulence factor that suppresses transmembrane receptor kinase activity. Genes & Development. 2004;18:2545–2556. DOI: 10.1101/gad.1245904

[69] Prince DC, Drurey C, Zipfel C, Hogenhout SA. The leucine-rich repeat receptor-like kinase BAK1 and the cytochrome P450 PAD3 contribute to innate immunity to aphids in Arabidopsis. Plant Physiology. 2014;164:2207–2219. DOI: 10.1104/pp. 114.235598

[70] Heil M, Ibarra-Laclette E, Adame-Alvarez RM, Martinez O, Ramirez-Chavez E, Molina-Torres J, Herrera-Estrella L. How plants sense wounds: damaged-self recognition is based on plant-derived elicitors and induces octadecanoid signaling. PLoS ONE. 2012;7(2):e30537. DOI: 10.1371/journal.pone.0030537

[71] Wu J, Baldwin IT. Herbivory-induced signaling in plants: perception and action. Plant Cell and Environment. 2009;32:1161–1174. DOI: 10.1111/j. 1365-3040.2009.01943.x

[72] Foyer CH, Verrall SR, Hancock RD. Systematic analysis of phloem-feeding insect-induced transcriptional reprogramming in Arabidopsis highlights common features and reveals distinct responses to specialist and generalist insects. Journal of Experimental Botany. 2015;66:495–512. DOI: 10.1093/jxb/eru491

[73] Maffei ME, Mithöfer A, Boland W. Before gene expression: early events in plant–insect interaction. Trends in Plant Science. 2007;12:310–316. DOI: 10.1016/j.tplants. 2007.06.001

[74] de Oliveira EF, Pallini A, Janssen A. Herbivores with similar feeding modes interact through the induction of different plant responses. Oecologia. 2016;180(1):1–10. DOI: 10.1007/s00442-015-3344-0

[75] Truitt CL, Wei HX, Pare PW. A plasma membrane protein from *Zea mays* binds with the herbivore elicitor volicitin. The Plant Cell. 2004;16:523–532. DOI: 10.1105/tpc. 017723

[76] Hilker M, Meiners T. Early herbivore alert: insect eggs induce plant defense. Journal of Chemical Ecology. 2006;32:1379–1397. DOI: 10.1007/s10886-006-9057-4

[77] Tamiru A, Bruce TJA, Woodcock CM, Caulfield JC, Midega CAO, Ogol CKPO, MayonP, Birkett MA, Pickett JA, Khan ZR. Maize landraces recruit egg and larval parasitoids in response to egg deposition by a herbivore. Ecology Letters. 2011;14:1075–1083. DOI: 10.1111/j.1461-0248.2011.01674.x

[78] Gouhier-Darimont C, Schmiesing A, Bonnet C, Lassueur S, Reymond P. Signalling of *Arabidopsis thaliana* response to *Pieris brassicae* eggs shares similarities with PAMP-triggered immunity. Journal of Experimental Botany. 2013;64:665–674. DOI: 10.1093/jxb/ers362

[79] Rossi M, Goggin FL, Milligan SB, Kaloshian I, Ullman DE, Williamson VM. The nematode resistance gene Mi of tomato confers resistance against the potato aphid. Proceedings of the National Academy of Science of the United States of America. 1998;95:9750–9754. DOI: 10.1073/pnas.95.17.9750

[80] Du B, Zhang W, Liu B, Hu J, Wei Z, Shi Z, He R, Zhu L, Chen R, Han B, He G. Identification and characterization of Bph14, a gene conferring resistance to brown planthopper in rice. Proceedings of the National Academy of Science of the United States of America. 2009;106:22163–22168. DOI: 10.1073/pnas.0912139106

[81] De Vos M, Jander G. Myzus persicae (green peach aphid) salivary components induce defence responses in Arabidopsis thaliana. Plant, Cell and Environment. 2009;32:1548–1560. DOI: 10.1111/j.1365-3040.2009.02019.x

[82] Jaouannet M, Rodriguez PA, Thorpe P, Lenoir CJ, MacLeod R, Escudero-Martinez C, Bos JI. Plant immunity in plant-aphid interactions. Frontiers in Plant Science. 2014;5:663. DOI: 10.3389/fpls.2014.00663

[83] Lotze MT, Zeh HJ, Rubartelli A, Sparvero LJ, Amoscato AA, Washburn NR, Devera ME, Liang X, Tor M, Billiar T. The grateful dead: damage-associated molecular pattern molecules and reduction/oxidation regulate immunity. Immunological Reviews. 2007;220:60–81. DOI: 10.1111/j.1600-065X.2007.00579.x

[84] Yamaguchi Y, Huffaker A, Bryan AC, Tax FE, Ryan CA. PEPR2 is a second receptor for the Pep1 and Pep2 peptides and contributes to defense responses in Arabidopsis. The Plant Cell. 2010;22:508–522. DOI: 10.1105/tpc.109.068874

[85] Ross A, Yamada K, Hiruma K, Yamashita-Yamada M, Lu X, Takano Y, Tsuda K, Saijo Y. The Arabidopsis PEPR pathway couples local and systemic plant immunity. The EMBO Journal. 2014;33:62–75. DOI: 10.1002/embj.201284303

[86] Pearce G, Moura DS, Stratmann J, Ryan CA. RALF, a 5-kDa ubiquitous polypeptide in plants, arrests root growth and development. Proceedings of the National Academy of Science of the United States of America. 2001;98:12843–12847. DOI: 10.1073/pnas.201416998

[87] Liu H, Ma Y, Chen N, Guo S, Liu H, Guo X, Chong K, Xu Y. Overexpression of stress-inducible OsBURP16, the beta-subunit of polygalacturonase 1, decreases pectin contents and cell adhesion, and increases abiotic stress sensitivity in rice. Plant, Cell & Environment. 2014;37:1144–1158. DOI: 10.1111/pce.12223

[88] Bellincampi D, Cervone F, Lionetti V. Plant cell wall dynamics and wall-related susceptibility in plant–pathogen interactions. Frontiers in Plant Science. 2014;5:228. DOI: 10.3389/fpls.2014.00228

[89] Wydra, K. and Beri, H. 2007. Immunohistochemical changes in methyl-ester distribution of homogalacturonan and side chain composition of rhamnogalacturonan I as possible components of basal resistance in tomato inoculated with *Ralstonia solanacearum* Physiological and Molecular Plant Pathology 70: 13-24

[90] Wydra K, Beri H. Structural changes of homogalacturonan, rhamnogalacturonan I and arabinogalactan protein in xylem cell walls of tomato genotypes in reaction to *Ralstonia solanacearum*. Physiological and Molecular Plant Pathology. 2006;68:41–50. DOI: 10.1016/j.pmpp.2006.06.001

[91] Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW: Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis. Plant Cell. 2003;15(4):809–834. DOI: 10.1105/tpc.009308

[92] Griebel T, Maekawa T, Parker JE. NOD-like receptor cooperativity in effector-triggered immunity. Trends in Immunology. 2014;35:562–570. DOI: 10.1016/j.it.2014.09.005

[93] Lu H, Zhang C, Albrecht U, Shimizu R, Wang G, Bowman KD. Overexpression of a citrus NDR1 ortholog increases disease resistance in Arabidopsis. Frontiers in Plant Science. 2014;4:157. DOI: 10.3389/fpls.2013.00157

[94] Century KS, Holub EB, Staskawicz BJ. NDR1, a locus of *Arabidopsis thaliana* that is required for disease resistance to both a bacterial and a fungal pathogen. Proceedings of the National Academy of Science of the United States of America. 1995;92:6597–6601.

[95] Aarts N, Metz M, Holub E, Staskawicz BJ, Daniels MJ, Parker JE. Different requirements for EDS1 and NDR1 by disease resistance genes define at least two R gene-mediated signaling pathways in Arabidopsis. Proceedings of the National Academy of Science of the United States of America. 1998;95:10306–10311. DOI: 10.1073/pnas.95.17.10306

[96] Falk A, Feys BJ, Frost LN, Jones JD, Daniels MJ, Parker JE. EDS1, an essential component of R gene-mediated disease resistance in Arabidopsis has homology to eukaryotic lipases. Proceedings of the National Academy of Science of the United States of America. 1999;96:3292–3297. DOI: 10.1073/pnas.96.6.3292

[97] Rathjen JP, Moffett P. Early signal transduction events in specific plant disease resistance. Current Opinion in Plant Biology. 2003;6:300–306. DOI: 10.1016/S1369-5266(03)00057-8

[98] Takahashi A, Casais C, Ichimura K, Shirasu K. HSP90 interacts with RAR1 and SGT1 and is essential for RPS2-mediated disease resistance in Arabidopsis. Proceedings of the National Academy of Science of the United States of America. 2003;100:11777–11782. DOI: 10.1073/pnas.2033934100

[99] Albrecht M, Takken FLW. Update on the domain architectures of NLRs and R proteins. Biochemical and Biophysical Research Communication. 2006;339:459–462.

[100] Pan Q, Liu YS, Budai Hadrian O, Sela M, Carmel Goren L, Zamir D, Fluhr R. Comparative genetics of nucleotide binding site-leucine rich repeat resistance gene homologues in the genomes of two dicotyledons: tomato and Arabidopsis. Genetics. 2000;155:309–322.

[101] Van Ooijen G, van den Burg HA, Cornelissen BJC, Takken FLW. Structure and function of resistance proteins in solanaceous plants. Annual Reviews Phytopathology. 2007;45:43–72. DOI: 10.1146/annurev.phyto.45.062806.094430

[102] Cui H, Tsuda K, Parker JE. Effector-triggered immunity: from pathogen perception to robust defense. Annual Reviews in Plant Biology. 2015;66:487–511. DOI: 10.1146/annurev-arplant-050213-040012

[103] Deslandes L, Olivier J, Peeters N, Feng DX, Khounlotham M Boucher C, Somssich I, Genin S, Marco Y. Physical interaction between RRS1-R, a protein conferring resistance to bacterial wilt, and PopP2, a type III effector targeted to the plant nucleus.

Proceedings of the National Academy of Science of the United States of America. 2003;100:8024–29. DOI: 10.1073/pnas.1230660100

[104] Fabro G, Steinbrenner J, Coates M, Ishaque N, Baxter L, Studholme DJ, Körner E, Allen RL, Piquerez SJ, Rougon-Cardoso A,Greenshields D, Lei R, Badel JL, Caillaud MC, Sohn KH, Van den Ackerveken G, Parker JE, Beynon J, Jones JD. Multiple candidate effectors from the oomycete pathogen *Hyaloperonospora arabidopsidis* suppress host plant immunity. PLoS Pathogens. 2011;7:e1002348. DOI: 10.1371/journal.ppat. 1002348. pmid:22072967

[105] Chou S, Krasileva K V, Holton JM, Steinbrenner AD, Alber T, Staskawicz BJ. Hyaloperonospora arabidopsidis ATR1 effector is a repeat protein with distributed recognition surfaces. Proceedings of the National Academy of Science of the United States of America. 2011;08:13323–13328. DOI: 10.1073/pnas.1109791108. pmid:21788488

[106] Sarris, Panagiotis F, Duxbury Z, Huh Sung U, Ma Y, Segonzac C, Sklenar J, Derbyshire P, Cevik, V, Rallapalli G, Saucet Simon B, Wirthmueller L, Menke Frank LH, Sohn Kee H, Jones Jonathan DG. A plant immune receptor detects pathogen effectors that target WRKY transcription factors. Cell. 2015;161:1089–1100. DOI: 10.1016/j.cell. 2015.04.024

[107] Liu W, Liu J, Triplett L, Leach JE, Wang GL. Novel insights into rice innate immunity against bacterial and fungal pathogens. Annual Reviews in Phytopathology. 2014;52:213–241. DOI: 10.1146/annurev-phyto-102313-045926

[108] Zhang ZM, Zhang X, Zhou ZR, Hu HY, Liu M, Zhou B, Zhou J. Solution structure of the *Magnaporthe oryzae* avirulence protein AvrPiz-t. Journal of Biomolecular NMR. 2013; 55:219–223. DOI: 10.1007/s10858-012-9695-5

[109] Cesari S, Thilliez G, Ribot C, Chalvon V, Michel C, Jauneau A, Rivas S, Alaux L, Kanzaki H, Okuyama Y, Morel JB, Fournier E, Tharreau D, Terauchi R, Kroj T. The rice resistance protein pair RGA4/RGA5 recognizes the *Magnaporthe oryzae* effectors AVR-Pia and AVR1-CO39 by direct binding. The Plant Cell. 2013;25:1463–1481. DOI: 10.1105/tpc.112.107201

[110] Okuyama Y, Kanzaki H, Abe A, Yoshida K, Tamiru M, Saitoh H, Fujibe T, Matsumura H, Shenton M, Galam DC, Undan J, Ito A, Sone T, Terauchi RA. Multifaceted genomics approach allows the isolation of the rice Pia-blast resistance gene consisting of two adjacent NBS-LRR protein genes. The Plant Journal. 2011;66:467–479. DOI: 10.1111/j.1365-313X.2011.04502.x

[111] Maqbool A, Saitoh H, Franceschetti M, Stevenson C, Uemura A, Kanzaki H, Kamoun S, Terauchi R, Banfield MJ. Structural basis of pathogen recognition by an integrated HMA domain in a plant NLR immune receptor. eLife. 2015;4:e08709. DOI: 10.7554/eLife.08709

[112] Leister RT, Ausubel FM, Katagiri F. Molecular recognition of pathogen attack occurs inside of plant cells in plant disease resistance specified by the Arabidopsis genes

RPS2 and RPM1. Proceedings of the National Academy of Science of the United States of America. 1996;93:15497–15502.

[113] Li M, Ma X, Chiang YHH, Yadeta KA, Ding P, Dong L, Zhao Y, Li X, Yu Y, Zhang L, Shen QH, Xia B, Coaker G, Liu D,Zhou JM.. Proline isomerization of the immune receptor-interacting protein RIN4 by a cyclophilin inhibits effector-triggered immunity in Arabidopsis. Cell Host Microbe. 2014;16:473–483. DOI: 10.1016/j.chom.2014.09.007

[114] Wilton M, Subramaniam R, Elmore J, Felsensteiner C, Coaker G, Desveaux D. The type III effector HopF2 Pto targets Arabidopsis RIN4 protein to promote *Pseudomonas syringae* virulence. Proceedings of the National Academy of Science of the United States of America. 2010;107:2349–2354. DOI: 10.1073/pnas.0904739107

[115] Eitas TK, Dangl JL. NB– LRR proteins: pairs, pieces, perception, partners, and pathways. Current Opinion Plant Biology. 2013;13:472–477. DOI: 10.1016/j.pbi.2010.04.007

[116] Williams SJ, Sohn KH, Wan L, Bernoux M, Sarris PF, Segonzac C, Ve T, Ma Y, Saucet SB, Ericsson DJ, Casey LW, Lonhienne T, Winzor DJ, Zhang X, Coerdt A, Parker JE, Dodds PN, Kobe B, Jones JD. Structural basis for assembly and function of a heterodimeric plant immune receptor. Science. 2014; 344:299–303. DOI: 10.1126/science.1247357

[117] Gabriels SH, Vossen JH, Ekengren SK, van Ooijen G, Abd-El-Haliem AM, van den Berg GC, Rainey DY, Martin GB, Takken FL, de Wit PJ, Joosten MH. An NB-LRR protein required for HR signaling mediated by both extra- and intracellular resistance proteins. The Plant Journal. 2007;50:14–28. DOI: 10.1111/j.1365-313X.2007.03027.x

[118] Bonardi V, Tang S, Stallmann A, Roberts M, Cherkis K, Dangl JL. Expanded functions for a family of plant intracellular immune receptors beyond specific recognition of pathogen effectors. Proceedings of the National Academy of Science of the United States of America. 2011;108:16463–16468. DOI: 10.1073/pnas.1113726108

[119] Muthamilarasan M, Prasad M. Plant innate immunity: an updated insight into defense mechanism. Journal of Biosciences. 2013;38:433–449. DOI: 10.1007/s12038-013-9302-2

[120] Inoue H, Hayashi N, Matsushita A, Xinqiong L, Nakayama A, Sugano S, Jiang CJ, Takatsuji H. Blast resistance of CC-NBLRR protein Pb1 is mediated by WRKY45 through protein-protein interaction. Proceedings of the National Academy of Science of the United States of America. 2013;110:9577–9582. DOI: 10.1073/pnas.1222155110

[121] Shen QH, Saijo Y, Mauch S, Biskup C, Bieri S, Keller B, Seki H, Ulker B, Somssich IE, Schulze-Lefert P.. Nuclear activity of MLA immune receptors links isolate-specific and basal disease-resistance responses. Science. 2007;315(5815):1098–1103. DOI: 10.1126/science.1136372

[122] Narusaka M, Kubo Y, Hatakeyama K, Imamura J, Ezura H, Nanasato Y, Tabei Y, Takano Y, Shirasu K, Narusaka Y. Interfamily transfer of dual NB-LRR genes confers

resistance to multiple pathogens. PLoS One. 2013;8:e55954. DOI: 10.1371/journal.pone.0055954

[123] Padmanabhan MS, Ma S, Burch-Smith TM, Czymmek K, Huijser P, Dinesh-Kumar SP. Novel positive regulatory role for the SPL6 transcription factor in the N TIR-NB-LRR receptor-mediated plant innate immunity. PLoS Pathogens. 2013;9:e1003235. DOI: 10.1371/journal.ppat.1003235

[124] Jacob F, Vernaldi S, Maekawa T. Evolution and conservation of plant NLR functions. Frontiers in Immunology. 2013;4:297. DOI: 10.3389/fimmu.2013.00297

[125] Bari R, Jones JD. Role of plant hormones in plant defence responses. Plant Molecular Biology. 2009;69:473–488. DOI: 10.1007/s11103-008-9435-0

[126] Glazebrook J. Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. Annual Reviews Phytopathology. 2005;43:205–227. DOI: 10.1146/annurev.phyto.43.040204.135923

[127] Spoel SH, Koornneef A, Claessens SM, Korzelius JP, Van Pelt JA, Mueller MJ, Buchala AJ, Métraux JP, Brown R, Kazan K, Van Loon LC, Dong X, Pieterse CM. NPR1 modulates cross-talk between salicylate- and jasmonate-dependent defense pathways through a novel function in the cytosol. The Plant Cell. 2003;15:760–770. DOI: 10.1105/tpc.009159

[128] Mur LA, Kenton P, Atzorn R, Miersch O, Wasternack C. The outcomes of concentration-specific interactions between salicylate and jasmonate signaling include synergy, antagonism, and oxidative stress leading to cell death. Plant Physiology. 2006;140:249–262.

[129] Vlot AC, Dempsey DA, and Klessig DF. Salicylic acid, a multifaceted hormone to combat disease. Annual Reviews Phytopathology. 2009;47:177–206. DOI: 10.1146/annurev.phyto.050908.135202

[130] Baebler Š, Witek K, Petek M, Stare K, Tušek-Žnidarič M, Pompe-Novak M, Renaut J, Szajko K, Strzelczyk-Żyta D, Marczewski W, Morgiewicz K, Gruden K, Hennig J.. Salicylic acid is an indispensable component of the Ny-1 resistance-gene-mediated response against Potato virus Y infection in potato. Journal of Experimental Botany. 2014;65:1095–1109. DOI: 10.1093/jxb/ert447

[131] Pegadaraju V, Knepper C, Reese J, Shah J. Premature leaf senescence modulated by the Arabidopsis PHYTOALEXIN DEFICIENT4 gene is associated with defense against the phloem-feeding green peach aphid. Plant Physiology. 2005;139(4): 1927-1934. DOI: 10.1104/pp.105.070433.

[132] Li Q, Xie QG, Smith-Becker J, Navarre DA, Kaloshian I. Mi-1-Mediated aphid resistance involves salicylic acid and mitogen-activated protein kinase signaling cascades. Molecular Plant Microbe Interaction. 2006;19(6):655-664. DOI: 10.1094/MPMI-19-0655

[133] Fu ZQ, Dong X. Systemic acquired resistance: turning local infection into global defense. Annual Reviews Plant Biology. 2013;64:839–863. DOI: 10.1146/annurev-arplant-042811-105606

[134] Park SW, Kaimoyo E, Kumar D, Mosher S, Klessig DF. Methyl salicylate is a critical mobile signal for plant systemic acquired resistance. Science. 2007;318:113–116. DOI: 10.1126/science.1147113

[135] Song JT, Lu H, McDowell JM, Greenberg JT. A key role for ALD1 in activation of local and systemic defenses in Arabidopsis. The Plant Journal. 2004;40:200–212. DOI: 10.1111/j.1365-313X.2004.02200.x

[136] Beckers GJM, Spoel SH. Fine-tuning plant defence signaling: salicylate versus jasmonate. Plant Biology. 2006;8:1–10. DOI: 10.1055/s-2005-872705

[137] Van der Does D, Leon-Reyes A, Koornneef A, Van Verk MC, Rodenburg N, Pauwels L, Goossens A, Körbes AP, Memelink J, Ritsema T, Van Wees SC, Pieterse CM. Salicylic acid suppresses jasmonic acid signaling downstream of SCFCOI1-JAZ by targeting GCC promoter motifs via transcription factor ORA59. The Plant Cell. 2013;25:744–761. DOI: 10.1105/tpc.112.108548

[138] De Vleesschauwer D, Xu J, Höfte M. Making sense of hormone-mediated defense networking: from rice to Arabidopsis, Frontiers in Plant Science. 2014;5:611. DOI: 10.3389/fpls.2014.00611

[139] Lee SC, Hong JK, Kim YJ and Hwang BK. Pepper gene encoding thionin is differentially induced by pathogens, ethylene and methyl jasmonate. Physiological and Molecular Plant Pathology. 2000;56:207–216. DOI: 10.1006/pmpp.2000.0269

[140] Zhu F, Xi D-H, Yuan S, Xu F, Zhang D-W, Lin H-H. Salicylic acid and jasmonic acid are essential for systemic resistance against tobacco mosaic virus in *Nicotiana benthamiana*. Molecular Plant Microbe Interaction. 2014;27:567–577. DOI: 10.1094/MPMI-11-13-0349-R

[141] Ellis C, Karafyllidis I, Turner JG. Constitutive activation of jasmonate signaling in an Arabidopsis mutant correlates with enhanced resistance to *Erysiphe cichoracearum*, *Pseudomonas syringae*, and *Myzus persicae*. Molecular Plant Microbe Interaction 2002a; 15:1025–1030. DOI: 10.1094/MPMI.2002.15.10.1025

[142] Xu E, Brosche M. Salicylic acid signaling inhibits apoplastic reactive oxygen species signaling. BMC Plant Biology. 2014;14:155. DOI: 10.1186/1471-2229-14-155

[143] Shah J, Kachroo P, Klessig DF. The Arabidopsis ssi1 mutation restores pathogenesis-related gene expression in npr1 plants and renders defensin gene expression salicylic acid dependent. The Plant Cell. 1999;11:191–206

[144] Clarke JD, Volko SM, Ledford H, Ausubel FM, Dong X. Roles of salicylic acid, jasmonic acid, and ethylene in cpr-induced resistance in Arabidopsis. The Plant Cell. 2000;12:2175–2190. DOI: 10.2307/3871113

[145] Alazem M, Lin NS. Roles of plant hormones in the regulation of host–virus interactions. Molecular Plant Pathology. 2014;16:529–540. DOI: 10.1111/mpp.12204

[146] Penninckx IA, Thomma BP, Buchala A, Metraux JP, Broekaert WF. Concomitant activation of jasmonate and ethylene response pathways is required for induction of a plant defensin gene in Arabidopsis. The Plant Cell. 1998;10:2103–2113. DOI: 10.1105/tpc.10.12.2103

[147] Truman W, Bennett MH, Kubigsteltig I, Turnbull C, Grant M. Arabidopsis systemic immunity uses conserved defense signaling pathways and is mediated by jasmonates. Proceedings of the National Academy of Science of the United States of America. 2007;104:1075–1080. DOI: 10.1073/pnas.0605423104

[148] Attaran E, Zeier TE, Griebel T, Zeier J. Methyl salicylate production and jasmonate signaling are not essential for systemic acquired resistance in Arabidopsis. The Plant Cell. 2009;21:954–971. DOI: 10.1105/tpc.108.063164

[149] Tuominen H, Overmyer K, Keinanen M, Kollist H, Kangasjarvi J. Mutual antagonism of ethylene and jasmonic acid regulates ozone-induced spreading cell death in Arabidopsis. The Plant Journal. 2004;39:59–69. DOI: 10.1111/j.1365-313X.2004.02107.x

[150] Zander M, Thurow C, Gatz C. TGA transcription factors activate the salicylic acid-suppressible branch of the ethylene-induced defense program by regulating ORA59 expression. Plant Physiology. 2014; 165:1671–1683. DOI: 10.1104/pp.114.243360

[151] Guan R, Su J, Meng X, et al. Multilayered Regulation of Ethylene Induction Plays a Positive Role in Arabidopsis Resistance against *Pseudomonas syringae*. Plant Physiology. 2015;169(1):299-312. doi:10.1104/pp.15.00659.

[152] Zhu Z, An F, Feng Y, Li P, Xue L, A M, Jiang Z, Kim JM, To TK, Li W, Zhang X, Yu Q, Dong Z, Chen WQ, Seki M, Zhou JM, Guo H. Derepression of ethylene-stabilized transcription factors (EIN3/EIL1) mediates jasmonate and ethylene signaling synergy in Arabidopsis. Proceedings of the National Academy of Science of the United States of America. 2011;108:12539–12544. DOI: 10.1073/pnas.1103959108

[153] Schenk PM, Kazan K, Wilson I, Anderson JP, Richmond T, Somerville SC, Manners JM. Coordinated plant defense responses in Arabidopsis revealed by microarray analysis. Proceedings of the National Academy of Science of the United States of America. 2000;97:11655–11660. DOI: 10.1073/pnas.97.21.11655

[154] Lorenzo O, Piqueras R, Sanchez-Serrano JJ, Solano R. *ETHYLENE RESPONSE FACTOR1* integrates signals from ethylene and jasmonate pathways in plant defense. The Plant Cell 2003;15:165–178. DOI: 10.1105/tpc.007468

[155] Dong N, Liu X, Lu Y, Du LP, Xu HJ, Liu HX, Xin ZY, Zhang ZY. Overexpression of TaPIEP1, a pathogen-induced ERF gene of wheat, confers host-enhanced resistance to fungal pathogen *Bipolaris sorokiniana*. Functional & Integrative Genomics 2010;10:215–226. 10.1007/s10142-009-0157-4

[156] Jia Y, Martin GB. Rapid transcript accumulation of pathogenesis-related genes during an incompatible interaction in bacterial speck disease-resistant tomato plants. Plant Molecular Biology. 1999;40:455–465. DOI: 10.1023/A:1006213324555

[157] Chen L, Zhang L, Li D, Wang F, Yu D. WRKY8 transcription factor functions in the TMV-cg defense response by mediating both abscisic acid and ethylene signaling in Arabidopsis. Proceedings of the National Academy of Science of the United States of America. 2013;110:E1963–E1971. DOI: 10.1073/pnas.1221347110

[158] Robert-Seilaniantz A, Grant M, Jones JDG. Hormone crosstalk in plant disease and defense: more than just JASMONATE-SALICYLATE antagonism. Annual Reviews Phytopathology. 2011; 49:317–343. DOI: 10.1146/annurev-phyto-073009-114447

[159] Asselbergh B, De Vleesschauwer D, Hofte M. Global switches and fine-tuning-ABA modulates plant pathogen defense. Molecular Plant Microbe Interaction. 2008; 21:709–719. DOI: 10.1094/MPMI-21-6-0709

[160] Ton J, Flors V, Mauch-Mani B. The multifaceted role of ABA in disease resistance. Trends in Plant Science. 2009;14:310–317. DOI: 10.1016/j.tplants.2009.03.006

[161] Pogany M, Koehl J, Heiser I, Elstner EF, Barna B. Juvenility of tobacco induced by cytokinin gene introduction decreases susceptibility to Tobacco necrosis virus and confers tolerance to oxidative stress. Physiological Molecular Plant Pathology. 2004;65:39–47. DOI: 10.1016/j.pmpp.2004.10.006

[162] Sano H, Seo S, Koizumi N, Niki T, Iwamura H, Ohashi Y. Regulation by cytokinins of endogenous levels of jasmonic and salicylic acids in mechanically wounded tobacco plants. Plant Cell Physiology. 1996;37:762–769.

[163] Barna B, Smigocki AC, and Baker JC. Transgenic production of cytokinin suppresses bacterially induced hypersensitive response symptoms and increases antioxidative enzyme levels in Nicotiana spp. Phytopathology. 2008;98:1242–1247. DOI: 10.1094/PHYTO-98-11-1242

[164] Pieterse CMJ, Van der Does D, Zamioudis C, Leon-Reyes A, Van Wees SCM. Hormonal modulation of plant immunity. Annual Reviews in Cell Development Biology. 2012;28:489–521. DOI: 10.1146/annurev-cellbio-092910-154055

[165] Choi J, Choi D, Lee S, Ryu CM, Hwang I. Cytokinins and plant immunity: old foes or new friends? Trends in Plant Science. 2011;16:388–394. 10.1016/j.tplants.2011.03.003

[166] Grosskinsky DK, Naseem M, Abdelmohsen UR, Plickert N, Engelke T, Griebel T, Zeier J, Novák O, Strnad M, Pfeifhofer H, van der Graaff E, Simon U, Roitsch T. Cytokinins mediate resistance against *Pseudomonas syringae* in tobacco through increased antimicrobial phytoalexin synthesis independent of salicylic acid signaling. Plant Physiology. 2011;157:815–830. DOI: 10.1104/pp.111.182931

[167] Ehness R, Roitsch T. Co-ordinated induction of mRNAs for extracellular invertase and a glucose transporter in *Chenopodium rubrum* by cytokinins. The Plant Journal. 1997;11:539–548. DOI: 10.1046/j.1365-313X.1997.11030539.x

[168] Xiong Y, Sheen J. The role of target of rapamycin signaling networks in plant growth and metabolism. Plant Physiology. 2014;164(2):499–512. DOI: 10.1104/pp.113.229948

[169] Slavikova S, Ufaz S, Avin-Wittenberg T, Levanony H, Galili G. An autophagy-associated Atg8 protein is involved in the responses of Arabidopsis seedlings to hormonal controls and abiotic stresses. Journal of Experimental Botany. 2008;59:4029–4043. DOI: 10.1093/jxb/ern244

[170] Hwang I, Sheen J, and Müller B. Cytokinin signaling networks. Annual Reviews in Plant Biology. 2012;63:353–380. DOI: 10.1146/annurev-arplant-042811-105503

[171] Xia XJ, Chen Z, and Yu JQ. ROS mediate brassinosteroids-induced plant stress responses. Plant Signal & Behavior. 2010;5:532–534. DOI: 10.4161/psb.10989

[172] De Bruyne L, Höfte M, De Vleesschauwer D. Connecting growth and defense: the emerging roles of brassinosteroids and gibberellins in plant innate immunity. Molecular Plant. 2014;7:943–959. DOI: 10.1093/mp/ssu050

[173] Roux M, Schwessinger B, Albrecht C, Chinchilla D, Jones A, Holton N, Malinovsky FG, Tör M, de Vries S, Zipfel C. The Arabidopsis leucine-rich repeat receptor-like kinases BAK1/SERK3 and BKK1/SERK4 are required for innate immunity to hemibiotrophic and biotrophic pathogens. The Plant Cell. 2011;23:2440–2455. DOI: 10.1105/tpc.111.084301

[174] Korner CJ, Klauser D, Niehl A, Dominguez-Ferreras A, Chinchilla D, Boller T, Heinlein M, Hann DR. The immunity regulator BAK1 contributes to resistance against diverse RNA viruses. Molecular Plant Microbe Interaction. 2013;26:1271–1280. DOI: 10.1094/MPMI-06-13-0179-R

[175] Fridman Y, Savaldi-Goldstein S. Brassinosteroids in growth control: how, when and where? Plant Science. 2013;209:24–31. DOI: 10.1016/j.plantsci.2013.04.002

[176] Naseem M, Dandekar T. The role of auxin-cytokinin antagonism in plant-pathogen interactions. PLoS Pathogens. 2012;8:e1003026. DOI: 10.1371/journal.ppat.1003026

[177] Denancé N, Ranocha P, Oria N, Barlet X, Rivière M-P, Yadeta K, Hoffmann L, Perreau F, Clément G, Maia-Grondard A, van den Berg GC, Savelli B, Fournier S, Aubert Y, Pelletier S, Thomma BP, Molina A, Jouanin L, Marco Y, Goffner D. Arabidopsis wat1 (walls are thin1)-mediated resistance to vascular pathogens is accompanied by cross-regulation of salicylic acid and tryptophan metabolism. The Plant Journal. 2013;73:225–239. DOI: 10.1111/tpj.12027

[178] Nakashita H, Yasuda M, Nitta T, Asami T, Fujioka S, Arai Y, Sekimata K, Takatsuto S, Yamaguchi I, Yoshida S. Brassinosteroid functions in a broad range of disease re-

sistance in tobacco and rice. The Plant Journal. 2003;33:887–898. DOI: 10.1046/j.
1365-313X.2003.01675.x

[179] Kamthan A, Chaudhuri A, Kamthan M, Datta A. Small RNAs in plants: recent devel-
opment and application for crop improvement. Frontiers in Plant Science. 2015;6:208.
DOI: 0.3389/fpls.2015.00208

[180] Ruiz-Ferrer V, Voinnet O. Roles of plant small RNAs in biotic stress responses. An-
nual Reviews in Plant Biology. 2009;60:485–510. DOI: 10.1146/annurev.arplant.
043008.092111

[181] Wilson RC, Doudna JA. Molecular mechanisms of RNA interference. Annual Re-
views in Biophysics. 2013;42:217–239. DOI: 10.1146/annurev-biophys-083012-130404

[182] Dunoyer P, Schott G, Himber C, Meyer D, Takeda A, Carrington JC, Voinnet O.
Small RNA duplexes function as mobile silencing signals between plant cells. Sci-
ence. 2010a;328:912–916. DOI: 10.1126/science.1185880

[183] Hamilton A, Voinnet O, Chappell L, Baulcombe DC. Two classes of short interfering
RNA in RNA silencing. The EMBO Journal. 2002;21:4671–4679. DOI: 10.1093/emboj/
cdf464

[184] Buhtz A, Springer F, Chappell L, Baulcombe DC, Kehr J. Identification and character-
ization of small RNAs from the phloem of *Brassica napus*. The Plant Journal.
2008;53:739–749. DOI: 10.1111/j.1365-313X.2007.03368.x

[185] Yoo B-C, Kragler F, Varkonyi-Gasic E, Haywood V, Archer-Evans S, Lee YM, Lough
TJ, Lucas WJ. A systemic small RNA signaling system in plants. The Plant Cell.
2004;16:1979–2000. DOI: 10.1105/tpc.104.023614

[186] Carlsbecker A, Lee JY, Roberts CJ, Dettmer J, Lehesranta S, Zhou J, Lindgren O, Mor-
eno-Risueno MA, Vatén A, Thitamadee S, Campilho A, Sebastian J, Bowman JL, He-
lariutta Y, Benfey PN. Cell signalling by microRNA165/6 directs gene dose-
dependent root cell fate. Nature. 2010;465:316–321. DOI: 10.1038/nature08977

[187] Pant BD, Buhtz A, Kehr J, Scheible WR. MicroRNA399 is a long distance signal for
the regulation of plant phosphate homeostasis. The Plant Journal. 2008;53:731–738.
DOI: 10.1111/j.1365-313X.2007.03363.x

[188] Navarro L, Dunoyer P, Jay F, Arnold B, Dharmasiri N, Estelle M, Voinnet O, Jones
JD. A plant miRNA contributes to antibacterial resistance by repressing auxin signal-
ling. Science. 2016;312:436–439. DOI: 10.1126/science.1126088

[189] Jiang CJ, Shimono M, Maeda S, Inoue H, Mori M, Hasegawa M, Sugano S, Takatsuji
H. Suppression of the rice fatty-acid desaturase gene *OsSSI2* enhances resistance to
blast and leaf blight diseases in rice. Molecular Plant Microbe Interaction.
2009;22:820–829. DOI: 10.1094/MPMI-22-7-0820

[190] Yara A, Yaeno T, Hasegawa M, Seto H, Montillet JL, Kusumi K. Disease resistance against *Magnaporthe grisea* is enhanced in transgenic rice with suppression of o-3 fatty acid desaturases. Plant Cell Physiol. 2007;48:1263–1274. DOI: 10.1093/pcp/pcm107

[191] Peltier AJ, Hatfield RD, Grau CR. Soybean stem lignin concentration relates to resistance to *Sclerotinia sclerotiorum*. Plant Disease. 2009;93:149–154. DOI: 10.1094/PDIS-93-2-0149

[192] Xin M, Wang Y, Yao Y, Xie C, Peng H, Ni Z, Sun Q. Diverse set of microRNAs are responsive to powdery mildew infection and heat stress in wheat (*Triticum aestivum* L.). BMC Plant Biology. 2010;10:123–134. DOI: 10.1186/1471-2229-10-123

[193] Vanderschuren H, Alder A, Zhang P, Gruissem W. Dose dependent RNAi-mediated geminivirus resistance in the tropical root crop cassava. Plant Molecular Biology 2009;70:265–272. DOI: 10.1007/s11103-009-9472-3

[194] Schwind N, Zwiebel M, Itaya A, Ding B, Wang MB, Krczal G, Wassenegger M. RNAi-mediated resistance to Potato spindle tuber viroid in transgenic tomato expressing a viroid hairpin RNA construct. Molecular Plant Pathology. 2009;10:459–469. DOI: 10.1111/j.1364-3703.2009.00546.x

[195] Ding SW. RNA-based antiviral immunity. Nature Reviews Immunology. 2010;10:632–644. DOI: 10.1038/nri2824

[196] Simon-Mateo C, García JA. Antiviral strategies in plants based on RNA silencing. Biochimica et Biophysica Acta. 2011;1809:722–731. DOI: 10.1016/j.bbagrm.2011.05.011

[197] Guo J, Fan J, Hauser BA, Rhee SY. Target enrichment improves mapping of complex traits by deep sequencing. G3 (Bethesda). 2015;6(1):67–77. DOI: 10.1534/g3.115.023671

[198] Perez-de-Castro AM, Vilanova S, Canizares J, Pascual L, Blanca JM, Diez MJ, Prohens J, Pico B. Application of genomic tools in plant breeding. Current Genomics. 2012;13(3):179–195. DOI: 10.2174/138920212800543084

[199] Dahal, D., Heintz., D., Van Dorsselaer, A., Braun H.-P., Wydra, K. 2009. Pathogenesis and stress related, as well as metabolic proteins are regulated in tomato stems infected with *Ralstonia solanacearum*. Plant Physiology and Biochemistry 47: 838–846

[200] Dahal, D., Pich, A., Braun, H.-P., Wydra, K. 2010. Analysis of cell wall proteins regulated in stems of susceptible and resistant tomato genotypes after inoculation with *Ralstonia solanacearum*: a proteomics approach. Plant Molecular Biology 73, 643-658

[201] Ghareeb, H., Bozsó, Z., Ott, P.G., Repenning, C., Stahl, F., Wydra, K. 2011. Transcriptome of silicon-induced resistance against *Ralstonia solanacearum* in the silicon non accumulator tomato implicates priming effect. Physiol. Mol. Plant Pathol. 75, 83-89

[202] Kiirika, L., Stahl, F. and Wydra, K. 2013. Phenotypic and molecular characterization of resistance induction by single and combined application of chitosan and silicon in

tomato against *Ralstonia solanacearum.* Physiological and Molecular Plant Pathology 81, 1-12

[203] Wydra, K. and Beri, H. 2006. Structural changes of homogalacturonan, rhamnogalacturonan I and arabinogalactan protein in xylem cell walls of tomato genotypes in reaction to *Ralstonia solanacearum.* Physiological and Molecular Plant Pathology Pathology 68: 41-50

[204] Schacht, T., Unger, C., Pich, A., Wydra, K. 2011: Endo- and exopolygalacturonases of *Ralstonia solanacearum* are inhibited by polygalacturonase-inhibiting protein (PGIP) activity in tomato stem extracts. Plant Physiology and Biochemistry 49, 377-387

[205] Dossa, S.G., Oliva, R, Maiss, E., Vera Cruz, C., Wydra, K. 2016. High temperature enhances the cultivated African rice *Oryza glaberrima* resistance to bacterial blight. Plant Disease 100, 380-387, http://dx.doi.org/10.1094/PDIS-05-15-0536-RE

[206] Zinsou, V., Wydra, K., Ahohuendo, B., Hau, B. 2004. Genotype x environment interactions in symptom development and yield of cassava genotypes with artificial and natural cassava bacterial blight infections. European Journal of Plant Pathology 111: 217-233

[207] Banito, A, Kpémoua, K.E. and Wydra, K. 2008. Expression of resistance and tolerance of cassava genotypes to bacterial blight determined by genotype x environment interactions. Journal of Plant Diseases and Plant Protection 115: 152-161

[208] Wydra, K., Banito, A., Kpémoua, K.E. 2007. Characterization of resistance of cassava genotypes to bacterial blight by evaluation of symptom types in different ecozones. Euphytica 155: 337-348

[209] Onaga, G. 2014. Population structure of *Magnaporthe oryzae* from different geographic regions and interaction transcriptomes with different rice genotypes at high temperatures. PhD thesis. Gerog-August Universität Göttingen

Genomics of Salinity Tolerance in Plants

Abdul Qayyum Rao, Salah ud Din, Sidra Akhtar, Muhammad Bilal Sarwar, Mukhtar Ahmed, Bushra Rashid, Muhammad Azmat Ullah Khan, Uzma Qaisar, Ahmad Ali Shahid, Idrees Ahmad Nasir and Tayyab Husnain

Additional information is available at the end of the chapter

Abstract

Plants are frequently exposed to wide range of harsh environmental factors, such as drought, salinity, cold, heat, and insect attack. Being sessile in nature, plants have developed different strategies to adapt and grow under rapidly changing environments. These strategies involve rearrangements at the molecular level starting from transcription, regulation of mRNA processing, translation, and protein modification or its turnover. Plants show stress-specific regulation of transcription that affects their transcriptome under stress conditions. The transcriptionally regulated genes have different roles under stress response. Generally, seedling and reproductive stages are more susceptible to stress. Thus, stress response studies during these growth stages reveal novel differentially regulated genes or proteins with important functions in plant stress adaptation. Exploiting the functional genomics and bioinformatics studies paved the way in understanding the relationship between genotype and phenotype of an organism suffering from environmental stress. Future research programs can be focused on the development of transgenic plants with enhanced stress tolerance in field conditions based upon the outcome of genomic approaches and knowing the mystery of nucleotides sequences hidden in cells.

Keywords: Salt tolerant genes, Salt Tolerance, Transgenic Plants, MicroRNA, Quantative Trait loci

1. Introduction

Nature's rage influences plants in the form of various biotic and abiotic stresses. Extreme abiotic stress conditions, such as salinity, flooding, heat, drought, and cold, as well as heavy metal toxicity and oxidative stress affect plants in many different ways. Human activities exacerbate these stress conditions to a greater extent. All the abiotic and biotic stresses, including

various pathogens, cause havoc to plants eventually limiting their growth and yield potentials. About 50% of crop yields are reduced due to abiotic stresses, making them the major cause of crop failure worldwide [1]. Abiotic stresses are a serious threat to the sustainability of agricultural industry. Naturally, a number of stresses combine with each other and act together; therefore, the negative effects are aggravated to a greater extent when compared to a single stress factor. To combat these stresses, combinations of diverse pathways are triggered [2].

In physical terms, stress is defined as a mechanical force per unit area applied to an object. It is difficult to measure the exact force applied by the stresses because the plants are immotile. This makes it harder to define stress in biological terms. A condition, which may act as a stress for one plant, may be ideal for another plant. Hence, a biological stress can most suitably be defined as a harsh condition or force that impedes the normal functioning of a biological system such as plants [3].

The plasma membrane serves as a barrier that separates a cell from its surrounding environment. Some of the small lipid molecules like steroid hormones are able to pass through this membrane and diffuse into the cytoplasm, whereas the membrane does not allow the water soluble molecules, such as ions, proteins, and other large molecules, to pass through it. Cells start responding when extracellular molecules come in contact with the plasma membrane. This foreign molecule is called an elicitor, and the protein that is present on the cell membrane and interacts with the elicitor is called a receptor. A number of biotic and abiotic stress signals serve as elicitors for the plant cells [4].

2. Salinity stress and its causes

Total amount of dissolved mineral salts in water and soil is termed as salinity [5]. These salts comprise electrolytes of anions (majorly CO_3^{2-}, SO_4^{2-}, Cl^-, NO_3^-, HCO_3^-) and cations (majorly Ca_2^+, K^+, Mg_2^+, Na^+). Salts that are soluble in water get deposited in the upper layer of soil to a greater extent that hinders the agricultural productivity of that land area [6]. Although fewer salts are present in the rainwater, these salts can be accumulated in the soil over a certain period of time. Salts can also be deposited by soil transported by wind from far off places. Impure irrigation water also contributes to the level of deposited salts in the agricultural lands [7].

Salinity stress is one of the main abiotic stresses and is considered as a restraint to crop yield. Increased salinization of cultivable land has disastrous effects worldwide [8]. Hyperosmotic and hyperionic stresses are caused by increased salinity, which can lead to plant death [9]. A number of factors are responsible for causing salinity in a given area such as the extent of precipitation or evaporation and weathering of rocks. Deserts have high salinity due to the fact that the rate of evaporation is greater than the rate of precipitation.

All the key processes within a plant are significantly influenced when the plant is exposed to salt stress [10]. The water stress resulting under salt stress affects the leaf growth and development. Cell division and expansion as well as stomatal opening and closing are negatively influenced by the salinity stress [11]. If the stress condition prevails, then the ionic stress strikes, and eventually, a major decline in photosynthetic rate occurs, and the leaves start to die [12].

Deforestation is a leading cause of salt stress. Heavy salt–rich irrigation is the major cause of salinity in agricultural lands. The process of evapotranspiration is responsible for the retention of excessive amounts of salt in the soil every year. This is due to abundant loss of water as a result of both evaporation and transpiration. Almost all of the main agricultural crops are sensitive to salt stress that results in serious damage to the yields of the crops [13]. Soil contents altered by the deposition of large amounts of salt in the soil, and as a result, soil becomes less porous reducing soil aeration and water transport [14]. Salinity stress and drought stress are quite similar in terms of physiology [15].

3. Stress signaling pathways

The receptors present on the plant cell surface receive the stress signals and transfer them downstream, resulting in the production of secondary messengers, e.g. reactive oxidative stress (ROS), calcium, and inositol phosphates [16]. Calcium level is further controlled by these messengers within the cell. As a result of this disturbance in the intracellular Ca^{2+} level, the Ca^{2+} sensors are triggered, which change their conformation in a calcium-dependent manner [14]. These sensors initiate a phosphorylation cascade by interacting with their respective partners and activate the stress responsive genes or the transcription factors that regulate stress response genes. The products of stress response genes help in plant survival and mitigate the stress conditions. Production of hormones (such as ethylene, salicylic acid, and abscisic acid (ABA) takes place because of changes in gene expression under the stress. Initial signal is amplified by messenger 'Sensor' stress response molecules, and a secondary signaling pathway may be induced. Such molecules which do not take part directly in signaling but play a role in alteration of signaling components are called accessory molecules [17].

The stress responsive genes can be divided into two major categories: early- and late-induced genes. Early-induced genes are prompted immediately after stress signals are received, and most of the times, they express in shorter period. In this category, a number of transcription factors are included because they do not require synthesis of new proteins for their stimulation. In contrast, late-induced genes are expressed slowly under the stress condition, i.e., express after hours of receiving stress signals, and their expression is persistent [18]. In this gene category, major stress responsive genes, such as (COR cold responsive), KIN (cold induced), or RD (responsive to dehydration), and membrane stabilizing proteins, osmolytes, antioxidants, and LEA (late embryogenesis abundant)-like proteins are included [19].

4. Salt tolerance

The percentage of biomass production in saline conditions in comparison with normal growing environment during an extended period is known as salt tolerance. In this regard, vivid variations are found among different plants due to the fact that decline in growth is dependent on the length of time over which plants are growing in the salinity-affected soils. For example,

a salt-tolerant plant like sugar beet may undergo 20% decrease in dry weight when grown in 200 mM NaCl [20]. In contrast, a moderately tolerant plant like cotton may undergo 60% decrease in the dry weight, whereas a sensitive plant like soybean may become dead [21] and a halophyte like *Suaeda maritima* may grow to its full potential in the 200 mM NaCl condition [22]. Evaluation of salt tolerance for perennial plant species can also be done based on survival rate. A marked decline in growth rate is observed in both salt-tolerant as well as nontolerant species during a short time in salt stress. This has been seen in the case of durum wheat and bread wheat, where durum wheat is more salt sensitive [23] and the same was also observed in the case of barley and triticale [15]. This led to realizing the importance of time frame and the mechanisms that different plants use at different growth stages when exposed to salinity.

5. Mechanisms of salt tolerance

The initial discovery of biochemists that enzymes of halophytes and nonhalophytes are equally tolerant to increased levels of NaCl is found to be true [24]. This was explained by the example of enzymes obtained from a halophyte *Atriplex spongeosa* and those obtained from peas or beans that were equally sensitive to NaCl [21, 25]. This is because most enzymes get inhibited at Na^+ concentration more than 100 mM, and some are observed in the case of Cl^-. Even K^+ can also inhibit enzymes when present in 100–200 mM concentrations [26]. Hence, the salt-tolerant mechanisms can be divided into two main categories: (1) preventing or reducing the amount of salt being uptake by plant tissue and (2) reducing the concentration of salt present in the cytoplasm. These both types of mechanisms are found in halophytes, which not only exclude salt very effectively but also quite effectively compartmentalize the excess salt in cell vacuoles. Due to this reason, halophytes are able to grow in saline soils far better and for longer time spans than other plants.

6. Conventional ways to manage salinity

Accumulation of large amounts of salts in the water around the root area is referred to as soil salinity [6]. Plants can tolerate soil salinity by two processes: salt exclusion and salt inclusion [27]. Plants, which are able to eliminate salts from the whole plant or specific plant tissues, are known as salt excluders. Such plants possess low Na^+ and Cl^- content as the membrane permeability prefers K^+ over Na^+ uptake in these plants. On the other hand, salt accumulators can withstand high salt concentrations by two approaches. The first approach is the enduring increased amounts of intercellular salts. The second approach is through the elimination of surplus amounts of salt from the plant because the roots of these plants can absorb salt ions but prevent their harmful effects [28].

To recover the agricultural lands from salt stress and for increased yields, it is necessary to remove excess amounts of salts from the root region. The common strategies used for this purpose are leaching, scraping, and flushing. As these methods were quite costly, new

approaches were introduced for contending salt stress. One of them is the use of halophytes in salinity-affected lands. Halophytes are the plants that can exclude the deposited salts from the soil surface in addition to withstanding high levels of accumulated salts [29]. For this purpose, some halophytes possess salt glands, which are specialized leaf cells having the ability to expel salt. Some others use salt hairs present on stems for this purpose while some have stomatal guard cells, which regulate the rate of transpiration according to the surrounding salt concentration. Another strategy used to protect the plants from the injurious effects of salinity is foliar feeding of nutrients. This enhances plant salt tolerance by relieving plants from Na^+ and Cl^- injury [30].

Soil salinity can also be controlled using better farm management practices. In this regard, improved irrigation methods, such as drip irrigation, can be used to apply controlled amount of water to the land. In rain-fed areas, crop rotation of annual crops with perennial crops (having deep roots) should be practiced to re-establish the equilibrium between rainfall and used water. This will avert the water tables from rising and delivering salts to the surface [31].

7. Genetic responses to salinity

Genetic response in case of salinity stress takes in a complex mechanism that is used by plants to up-regulate or down-regulate (increase or decrease) the production of specific gene products (protein or RNA). These mechanisms have been recognized at different stages of central dogma process like from transcriptional initiation to RNA processing, post-transcriptional modification, and translation, and to the post-translational modification of a protein [32]. Understanding about the transcriptional behavior of plants provides a detailed knowledge about the gene expression at mRNA level. Transcriptional profiling is widely used to screen out candidate genes involved in stress responses. Till now, massive information about the salt responsive genes, transcription factors which either up-regulated or down-regulated, has been identified using transcriptome profiling methodology. Further genomic approaches contribute significant role in encoding, cloning, and characterization of these genes. Gene expression under the certain conditions altered by transcription factors. These factors are considered the most important switches that up-regulate or down-regulate the gene expression. Among them, *bZIP, WRKY, AP2, NAC, C2H2* zinc finger gene, MYB and DREB family proteins comprise a large number of stress-responsive members. These transcription factors have the capacity to alter the gene expression by cis-acting specific binding in the promoter region of broad range of genes.

Up-regulation in the expression of bZIP genes were observed in sensitive wheat cultivar under persistent salinity stress and down-regulation in salt-tolerant variety [33]. It predicts the role of NAC transcription factor in salinity tolerance in both rice and wheat cultivars. In rice, transcriptional regulators, such as DREB1/CBF, DREB2, and AREB/ABF, have been demonstrated to play a significant role in abiotic stress responses [34, 35]. Transcription factors, such as OsNAC5 and ZFP179, show an up-regulation under salinity stress, which may regulate the synthesis and accumulation of proline, sugar, and LEA proteins that in turn play an integral role in stress tolerance [36].

Full-length cDNA is a vital resource for studying the full functional genes in wheat. A group of gene "MYB gene" family analyzed by Zhang et al. [37] that respond to one or more abiotic stress treatments. They isolated 60 full-length cDNA sequences encoding wheat MYB proteins and also construct phylogenetic tree with other wheat, rice, and *Arabidopsis* MYB proteins to understand their evolutionary relationships and the putative functions of wheat MYB proteins based on *Arabidopsis* MYB proteins with known functions. Tissue-specific analysis and abiotic stress response expression profiles were also carried out to find potential genes that participate in the stress signal transduction pathway, including the analysis of transgenic *Arabidopsis* plants expressing the MYB gene, TaMYB32 [38]. In *Arabidopsis*, salt stress results in up-regulation of AtWRKY8 gene expression, which directly binds with the promoter of RD29A [39]. A large number of genes and transcription factors are up-regulated in response to salinity in different plant species, which serve diverse functions [40]. Some of the examples of salt-responsive genes are listed in the Table 1, and these genes are mainly classified into the following functional categories: ion transport or homeostasis (e.g., SOS genes, AtNHX1, and H^+-ATPase), senescence-associated genes (e.g., SAG), molecular chaperones (e.g., HSP genes), and dehydration-related transcription factors (e.g., DREB). Among stress-responsive genes, the SOS transcription gene family is considered to play a very stimulating role in ion homeostasis, thereby conferring salt tolerance [32, 41]. Most of the salinity responsive genes, such as ROS-scavenging and osmotic-regulating genes, are also up-regulated by salinity in salinity tolerant species. Schmidt et al. [42] observed more than 10 extensively up-regulated genes in the halophyte plant species *Spartina alterniflora* under salt stress. Most of these genes were related to osmotic regulation process.

Gene Name	Species	NaCl Concentration	Gene functions	References
SOS1 SOS2 SOS3 AtNHX1	*Brassica juncea* *Brassica campestris*	25 and 50 mM	(1) Plasma membrane Na⁺/K⁺ antiporter (2) Protein kinase (3) Calcium-binding protein (4) Vacuolar Na⁺/K⁺ antiporter	[40]
PRP SAG HSPC025	*Oryza sativa*	50 mM	(1) Proline-rich proteins and cell wall protection (2) Senescence-associated genes, regulatory processes, and cellular signal transduction (3) Heat-shock proteins, protein stabilizing	[43]
OsHSP23.7 OsHSP71.1, OsHSP80.2	*Oryza sativa*	100 mM	Heat-shock proteins, molecular chaperones, folding, assembling and transporting proteins	[44]
AtSKIP	*Arabidopsis thaliana*	150 mM	Transcription factor, transcriptional preinitiation, splicing, and polyadenylation	[45]
OsHsp17.0, OsHsp23.7	*Oryza sativa*	200 mM	Heat-shock proteins, molecular chaperones, folding, assembling and transporting Proteins	[46]

Gene Name	Species	NaCl Concentration	Gene functions	References
DcHsp17.7	*Carrot*	300 mM	Cell viability and membrane stability under heat stress	[47]
JcDREB	*Arabidopsis thaliana*	300 mM	Transcription factor	[48]
katE gene	*Escherichia coli*	150 mM	Membrane stability	[49]
Salt overly sensitive (SOS) genes	*Ipomoea batatas*	120 mmol L^{-1}	Ion homeostasis Improve biochemical indicators	[50]
AtNHX1	*Arabidopsis thaliana*	220 mM	Calcium-binding protein, vacuolar Na$^+$/K$^+$ antiporter	[51]
SNAC1	*Oryza sativa*	200mM	Enhancing root development and reducing transpiration rate Biochemical adjustment	[52]
OsRab7	*Oryza sativa*	250 mM	Vesicle trafficking gene enhanced seedling growth and increased proline content	[53]
PtSOS2	*Populus tremula*	85 mM	Protein kinases enhanced photosynthetic pigments and physiological parameters	[54]
TaSC	*Triticum aestivum*	150mM	Enhanced membrane stability	[55]
PeXTH	*Populus euphratica*	200 mM	Cell viability and membrane stability enhanced water holding capacity	[56]
StP5CS	*Solanum tuberosum*	150 mM	Osmolyte accumulation	[57]
CYP94 (cytochrome P450)	*Oryza sativa*	200 mM	Enhanced CYP94C2b expression	[58]
TaSC	*Triticum aestivum*	120 mM	Regulate the gene expression program	[55]
H3K4me3	*Arabidopsis thaliana*	150 mM	Gene priming, regulate the gene expression program	[56]
WsSGTL1	*Withania somnifera*	100 mM	Stabilized the phenotypic and physiological parameters	[59]
GmPIP1;6	*Glycien max*	100 mM	Multifunctional aquaporin involved in root water transport, photosynthesis, and seed loading	[60]
AtSTO1	*Arabidopsis thaliana*	150 mM	Enhanced the salt tolerance increased concentrations of 9-cis-epoxycarotenoid dioxygenase3	[61]
ONAC045	*Oryza sativa*	200 mM	Functioned as a transcriptional activator	[62]
SOS1	*Nicotina tabacum*	150 mM	(1) Plasma membrane Na$^+$/K$^+$ antiporter (2) Protein kinase	[63]

Gene Name	Species	NaCl Concentration	Gene functions	References
			(3) Calcium-binding protein	
			(4) Vacuolar Na$^+$/K$^+$ antiporter	
mtlD	*Escherichia coli*	200 mM	Enhanced the production of mannitol 1-phosphate dehydrogenase	[64]
glyoxalase II	*Oryza sativa*	200 mM	Detoxification of cytotoxic 2-oxo-aldehydes	[65]
HAL5		100 mM	Regulate Na$^{(+)}$/K$^{(+)}$ homeostasis, lower leaf Na$^{(+)}$ accumulation, reducing Na$^{(+)}$ transport from root to shoot, maintaining Na$^{(+)}$/K$^{(+)}$ homeostasis	[66]
AtSTO1	*Arabidopsis thaliana*	200 mM	Increased concentrations of 9-cis-epoxycarotenoid dioxygenase3, greater overall biomass, greater root biomass, improved photosynthesis, and greater pith size	[61]
TaSTRG	*Triticum aestivum*	200 mM	Higher salt and drought tolerance, lower intracellular Na$^{(+)}$/K$^{(+)}$ ratio, higher survival rate, fresh weight and chlorophyll content, accumulated higher proline and soluble sugar contents and had significantly higher expression levels of putative proline synthetase and transporter genes	[67]

Table 1. Salt responsive genes with their origin and possible functions

Recently, Schmidt et al. [42] characterized root-specific salt-responsive *ERF1* (*SERF1*) transcription factor gene in *Oryza sativa* that showed a root-specific induction upon salt and H$_2$O$_2$ treatment. Plants deficient for *SERF1* are more sensitive to salt stress compared with the wild type, although constitutive overexpression of *SERF1* improves salinity tolerance. Different types of kinases also regulate the activity of transcription factors and have been found to be significant players of plant adaptation to salinity stress. Serra et al. [68] studied the *OsRMC* encodes a receptor-like kinase and described as a negative regulator of salt stress responses in rice. Two transcription factors, *OsEREBP1* and *OsEREBP2*, belonging to the AP2/ERF family were shown to bind to the same GCC-like DNA motif in *OsRMC* promoter and to negatively regulate its gene expression. Basic region/leucine zipper (*bZIP*) TFs (Transcription factors are proteins involved in the process of converting, or transcribing, DNA into RNA. Transcription factors include a wide number of proteins, excluding RNA polymerase, that initiate and regulate the transcription of genes) possesses a basic region that binds DNA and a leucine zipper dimerization motif. A *bZIP* class of ABRE-binding transcription factor, known as *OSBZ8*, has also been identified from rice and has been shown to be highly expressed in salt-tolerant cultivars than

in salt-sensitive one [69]. Moreover, *OSBZ8* has been shown to be activated/phosphorylated by a SNF-1 group of serine/threonine kinase in the presence of Spd during salinity stress [69].

Sairam et al. [70] isolated and analyzed the expression response of wheat *lip19* (encoding *bZIP*-type transcription factors) against cold stress. Further analysis confirmed the upregulation of *Wlip19* gene in a freezing-tolerant wheat cultivar than in a freezing-sensitive cultivar, while under drought and exogenous ABA application, higher activity of *Wlip19* also observed. Heterologous expression of Wlip19 in tobacco has showed a significant increase in abiotic stress tolerance. Alternative splicing of RNA/mRNA played a critical role to cope stress condition especially abiotic stress by switching on/off the transcriptional activities. These splicing factors and spliceosomal proteins mainly involved in plant growth, development process, responses to external environmental factor by affecting the cellular process, cell fate, plant immune/defense system, and tolerance efficiency. All these processes point to critical role of the splicing/alternative splicing under abiotic stress environment [71].

In addition to protein coding genes, recently discovered microRNAs (miRNAs) and endogenous small interfering RNAs (siRNAs) have emerged as important players in plant stress responses. Therefore, post-transcriptional gene regulation plays a crucial role in the plant salt response (Figure 1) [72]. Initial clues suggesting that small RNAs are involved in plant stress responses stem from studies showing stress regulation of miRNAs and endogenous siRNAs, as well as from target predictions for some miRNAs [73]. There has been strong evidence leading to the proposal that miRNAs are hypersensitive to abiotic or biotic stress as well as to diverse physiological processes [74, 75]. Drought, cold, and salinity are major abiotic stresses for plants; all of these conditions strongly induced miR402 overexpression. Numerous studies on plants, such as *Arabidopsis thaliana* and *Oryza sativa*, have been studied with respect to miRNA expression analysis and have revealed an important role for miRNAs in response to abiotic stress.

Various studies with respect to miRNAs profiling under abiotic stress point out the several differentially expressed miRNAs. In response to salt stress, miRNAs, such as miR396, miR394, miR393, miR319, miR171 miR169, miR168, and miR167, were up-regulated, whereas the miR398 was down-regulated in *Arabidopsis*, thus indicate a role for miRNAs in the response to salt stress [76].

Up-regulation of miRS1 and miR159.2 in response to salt stress was observed in *Phaseolous vulgaris* [77]. The expression of miR530a, miR1445, miR1446a-e, miR1447, and miR171l-n was increased, whereas the expression of miR482.2 and miR1450 was decreased during salt stress in *Populus trichocarpa* [76]. Furthermore, two members of miR169 family namely miR169g and miR169n showed enhanced expression during salinity. With the development of genomics tools and computational algorithms to predict and identify the miRNAs in various plant species, the number of miRNAs associated with salt stress response is increasing. A comprehensive understanding of miRNA-based gene regulation under salt stress will definitely help in elucidating the complex network of regulatory factors, proteins, and metabolites.
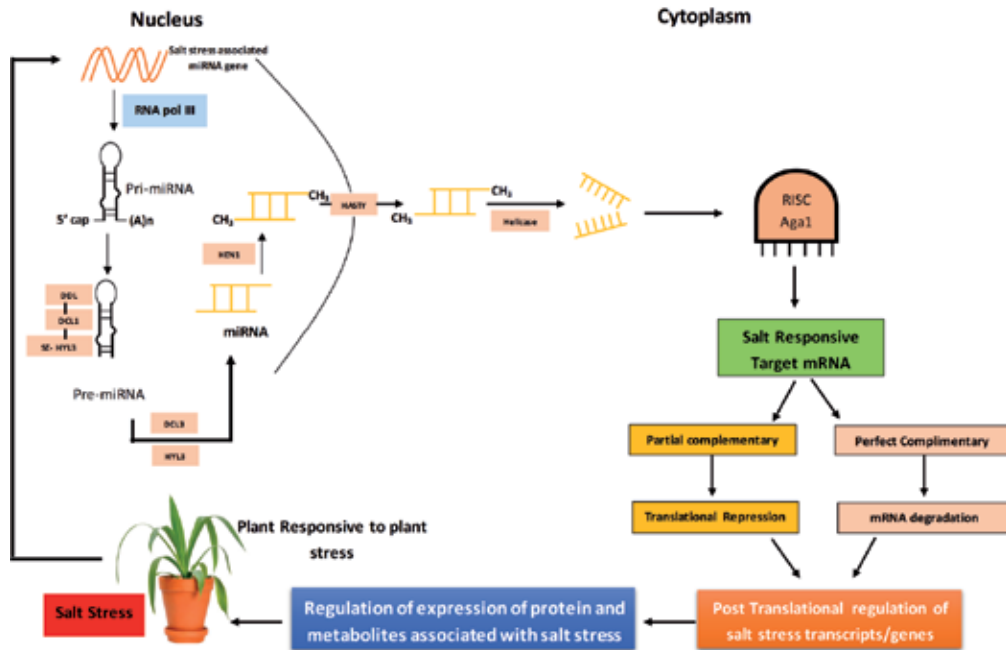
Figure 1. A pathway showing miRNA-mediated post-transcriptional regulation of salt stress–responsive plant genes

8. QTL mapping in relation to plant salinity tolerance

A quantitative trait locus (QTL) is a section of DNA (the locus) that correlates with variation in a phenotype (the quantitative trait). The QTL typically is linked to, or contains, the genes that control that phenotype. Genetic marker is an identifiable fragment of DNA that is linked with a specific point and indicates genetic differences within the genome. Molecular markers act as a 'signs' or 'flags' should not be considered as normal genes. Genetic markers that are tightly linked are referred to as "gene tags" [78]. The DNA markers can be grouped in various categories based on their technical requirements, the number of genetic markers that can be detected throughout the genome, and the amount of genetic variation found at each marker [79]. Restriction fragment length polymorphisms (RFLPs) are one of the earliest types of DNA-based marker system, which detect the variation in restriction fragment length by Southern hybridization, which cause single base changes that led to the creation or removal of a restriction endonuclease recognition site to detect shift in fragment size. Although this technique is an important tool in breeding programs, it has been superseded by microsatellite or simple sequence repeat (SSR) markers and is now rarely used. SSR markers detect variation in the number of short repeat sequences, usually two or three base repeats that allow the detection of multiple alleles. The expressed sequence tag (EST) databases have now opened the opportunity for the identification of single nucleotide polymorphisms (SNPs) that occur at varying frequencies depending on the species and genome region being considered [80].

These DNA marker types could be associated with quantitative traits, which are known as quantitative trait loci (QTLs). Mapping of QTLs for salt tolerance have been a slow process due to the complexity of this trait and poor understanding about it. Ren et al. [81] discovered a gene locus named as QTL SKC1, which codes for a transporter that removes Na^+ from the xylem [82]. Several QTLs have been identified in different crop plants. QTLs for yield and physiological characteristics were identified at a late stage of growth of barley under salinity stress [83]. A total of 10 traits were considered for which 30 QTLs were identified under salt stress and nonstress conditions. Of these 30, 13 QTLs were discovered under salt stress [83]. In white clover, QTLs for salt tolerance were identified at the vegetative stage of plants and the results showed that, in white clover, multiple QTLs are responsible for controlling the salt tolerance [84]. However, QTLs for salt tolerance in tomato were detected at the seedling stage of *Solanum pennellii* and *Solanum lycopersicoides* plants. In *S. pennellii*, four major QTLs were detected, for salt tolerance, on chromosomes 6, 7, and 11, whereas in *Solanum lycopersicoides* six major QTLs were identified under salt stress on chromosomes 4, 6, 9, and 12 [85]. QTLs for salt tolerance in soybean were identified on chromosome 3 [86]. QTLs identified by SSR markers in various plants are given in Table 2.

Crop plants	Locus	Traits	Reference
Wheat (*Triticum aestivum* L.)	*Kna1*	Controls the selectivity of Na+ and K+ transport from root to shoot and maintains high K^+/Na^+ ratio	[88, 89]
	Nax1	Both are involved in decreasing Na+ uptake and enhancing K^+ loading into the xylem	[90, 91]
Rice (*Oryza sativa* L.)	*qRL-7, qDWRO-9a* and *qDWRO-9b qBI-1a*and *qBI-1b*	Play important roles in root length and root dry weight at seedling stage under saline conditions	
	QNa, QNa:K, SKC1/ OsHKT8	Regulate K^+/Na^+ homoeostasis	[92]
	qDM-3 and *qDM-8, qSTR-6*	Improve Na+/K+ ratio under saline conditions	[81]
	qNAK-2 and *qNAK-6*	Improve Na^+/K^+ ratio	[93]
	Saltol	Controls shoot Na^+/K^+ homoeostasis	[94, 95, 96, 97]
	Saltol and *nonSaltol*	Control shoot Na^+/K^+ homoeostasis	
	QKr1.2	Controls K^+ content in root	
Barley (*Hordeum vulgare*)	Five QTL for ST were identified on chromosomes 1H, 2H, 5H, 6H, and 7H, which accounted for more than 50% of the phenotypic variation	Enhance vegetative growth under saline stress	[98]
	A locus *HvNax3* on the short arm of	Reduces shoot Na^+ content by 10–25% in plants grown under salt stress (150 mM NaCl)	[99]

Crop plants	Locus	Traits	Reference
	chromosome 7H in wild barley (*Hordeum vulgare* ssp. *spontaneum*) accession CPI-71284-48		
White clover (*Trifolium repens* L.)	Several QTLs for ST, some at common locations, but each of low scale	Affect ST during vegetative stage	[84]
Soybean (*Glycine max* (L.) Merr.)	A major QTL for ST was identified near the Sat091 SSR marker on linkage group (LG) N Eight QTLs for ST were detected A major QTL for ST was detected	Maintains growth under salt stress Maintains growth under salt stress Maintains healthy growth under salt stress	[100] [101] [102]

Table 2. QTLs for 'Salt Tolerance' (ST) in various plants identified by SSR markers [87]

9. Engineering plants for enhanced salt tolerance: Transgenic approach

Plant breeding strategy for salt tolerance is not much successful due to the reproductive barrier and also as it involves the risk of other undesirable traits transfer. Reproductive barriers and uncontrolled transfer of the traits make the conventional approach of plant breeding and genetics less desirable technique for abiotic stress tolerance in varieties' development. Other advanced techniques like genetic engineering for single gene transfer are considered more powerful to deal with this problem [38]. Transgenic plants are those plants, which have desired gene of interest directly integrated into the plant genome and developed from only a single plant cell. Transgenic plants with improved traits, including resistance to pests, pesticides, diseases or adverse environmental conditions, improved nutritional value, and enhanced product shelf life, have been developed through different genetic engineering techniques. Despite a number of social, political, and legal concerns, many countries are now allowing transgenic crop production in conjunction with their conventional crop production [103]. Transgenic approaches are being successfully pursued by researchers in some crops not only to improve the quality but also to increase the tolerance to abiotic stress, but tolerance trait is a quantitative complex trait and involves a number of genes. Thus, improving crop salt tolerance by genetic engineering is not so easy. Genes that encode ion transport proteins, compatible organic solutes, antioxidants, heat-shock and late embryogenesis abundant proteins, and transcription factors for gene regulation have focused by the biologist for improving the salt tolerance trait in various trait through genetic engineering techniques [104].

Gene	Type of product	Source	Target plant	Reference
coda	Glycine betaine	*Arthrobacter globiformis*	Tomato	[106]
coda	Glycine betaine	*Arthrobacter globiformis*	*Brassica juncea*	[107]
Cox	Glycine betaine	*Arthrobacter pascens*	Rice	[108]
OsTPS1	Trehalose-6-phosphate synthase	Rice	Rice	[109]
TPS1	Trehalose-6-phosphate synthase	Yeast	Tomato	[110]
OstA, ostB	Trehalose	*Escherichia coli*	Rice	[111]
AtTPS1	Trehalose	*Arabidopsis*	Tobacco	[112]
mtlD	Mannitol	Tobacco	Tobacco	[113]
mtlD	Mannitol	Wheat	*Escherichia coli*	[114]
M6PR	Mannitol	Celery	*Arabidopsis*	[115]
S6PDH	Sorbitol	Apple	*Japanese Persimmon*	[116]
P5CS	Proline	Rice	Mouth-bean	[117]
P5CS	Proline	*Vigna acontifolia*	*Nicotiana tabacum*	[118]
SOD2	Na+/H+ antiporter	*Schizosaccharomyces pombe*	*Arabidopsis*	[119]
nhaA	Na+/H+ antiporter	*E. coli*	*Arabidopsis*	[120]
AVP1	Vacuolar H+-pyrophosphates	*Arabidopsis*	Cotton	[121]
AtNHX1	Vacuolar Na+/H+ antiporter	*Arabidopsis*	Tomato	[122]
AgNHX1	Vacuolar Na+/H+ antiporter	*Atriplex gmelini*	Rice	[123]
BnNHX1	Vacuolar Na+/H+ antiporter	Brassica	Tobacco	[124]
GhNHX1	Vacuolar Na+/H+ antiporter	Cotton	Tobacco	[125]
GlyII	GlyoxylaseII	Rice	Tobacco	[126]
OsNAC5	NAC1 Transcription factor	Rice	Rice, *Arabidopsis*	[36]
GmbZIP1	bZIP Transcription factor	Soybean	*Arabidopsis*, tobacco	[127]
TaMYB2A	MYB2A transcription factor	Wheat	*Arabidopsis*	[128]
BrERF4	Ethylene responsive element 4	Brassica	*Arabidopsis*	[129]
MCM6	DNA helicase	Pea	Tobacco	[130]
T30hsp70	Heat-shock protein	*Trichoderma harzianum*	*Arabidopsis*	[130]
HVA1	LEA protein	*Hordeum vulgare L*	Rice	[131]
GhMPK2	MAP kinase	Cotton	Tobacco	[132]

Table 3. List of various genes responsible for salinity tolerance in plants with their role, source, and target plants (transgenic plants).

Plants try to survive with salinity by bringing various metabolic changes, such as a production of osmolytes, antioxidative enzymes, and up-regulating various genes involved in stress response like ion transporters, ion channels, transcriptional factors, and various signaling pathway components. The scientist studied various pathway responses that altered due to the salinity as mentioned above to generate the transgenic plants by transferring the salt-responsive genes into the salt susceptible plants from different genetic background (relatively salt-tolerant plants) or altering the expression of existing genes [105]. There are a number of gene(s) known which are responsible for salinity tolerance when transferred in plants through genetic engineering (Table 3).

Discovery of salt-tolerant genes is essential to induce salt tolerance in crop plants to enable them to grow on saline soils. Successful examples of identification and expression of salt tolerance genes include: over expression of AtNHX1 in *Arabidopsis* [133], tomato [134], *Brassica napus* [122], and cotton [135]. Likewise, overexpression of SOS1 gene in *Arabidopsis* also induces salt tolerance [136]. YCF1 is a yeast protein, which belongs to the ATP-binding cassette transporter family. Expression of this protein in *Arabidopsis* enhanced the salt tolerance in the transgenic plants to a significant level [137]. Since last several years' identification and transformation of salt tolerance genes in crop plants have been done [30]. When *AtNHX1* (vacuolar Na^+/H^+ antiporter from *A. thaliana*) was over expressed in tomato, Brassica [122, 134], and *Arabidopsis* [133], the transformed plants showed enhanced salt tolerance and were able to grow at 200 mM NaCl concentration. On the basis of growth responses of these transgenic plants, it has been concluded that they can grow on saline soils very well. Over expression of another Na^+/H^+ antiporter from *Atriplexgmelini* (*AgNHX1*) in rice enabled the transformed plants to grow at 300 mM concentration of NaCl for 3 days. Similar findings were observed when the same gene was transformed in wheat [138] and maize [139].

Overexpression of Na^+/H^+ antiporter from rice in the same species showed enhanced yield under salt stress [140]. Overexpression of HKT1-1 transporter in root cells surrounding the xylem of *Arabidopsis thaliana*, resulted in more removal of Na^+ ions from the xylem and into specialized compartments in the root tissues preventing the premature injury of shoots and leaves that could occur due to Na^+ accumulation [141]. The reason behind low number of successful transformations of salt-tolerant genes in crops is that these efforts have mostly been restricted to model plants like *Arabidopsis*, rice, and tobacco. Furthermore, there are some problems in applying this technology to other crop plants, such as monocots, due to the difficulty of obtaining series of independent T2 lines because the process is labor intensive and expensive [142].

10. Conclusion

Agriculture is immensely affected by salinity worldwide and is predicted to be a larger problem in near future. The damaging effects of high salinity can be seen in plants at organismic level, leading to immature death or decreased productivity. Some plant species are more tolerant to these detrimental effects than others. Salt stress leads to high yield losses worldwide.

Therefore, the changes aimed at overcoming these issues need to be fully implemented as soon as possible. Information related to the biochemical indicators at the cellular level may act as selection criteria for salt tolerance in different crops. There are many transgenic plants with high stress tolerance towards abiotic stress, yet stress tolerance has complex mechanism that includes multiple physiological and biochemical changes and multiple genes. Transgenic plants, which are commercially valuable, should have relatively high productivity and other traits important for their yield. Genetic modification, moreover, should be combined with marker-assisted breeding along with stress-related genes and QTLs. These strategies must be integrated, and such approaches should be combined to effectively increase plant stress tolerance

Author details

Abdul Qayyum Rao[1*], Salah ud Din[1], Sidra Akhtar[1], Muhammad Bilal Sarwar[1], Mukhtar Ahmed[1], Bushra Rashid[1], Muhammad Azmat Ullah Khan[1], Uzma Qaisar[2], Ahmad Ali Shahid[1], Idrees Ahmad Nasir[1] and Tayyab Husnain[1]

*Address all correspondence to: qayyum.cemb@pu.edu.pk

1 Division of Plant Science, Centre of Excellence in Molecular Biology (CEMB), Pakistan

2 School of Biological Sciences (SBS), University of the Punjab, Pakistan

References

[1] Mahajan S, Tuteja N. Cold, salinity and drought stresses: an overview. Arch Biochem Biophys. 2005;444:139–58. DOI: http://10.1016/j.abb.2005.10.018

[2] Lal R. Soils and sustainable agriculture: a review. Sustainable Agriculture. Springer Netherlands; 2009. pp. 15–23. DOI: http://10.1007/978-90-481-2666-8_3

[3] Hernandez J, Olmos E, Corpas F, Sevilla F, Del Rio L. Salt-induced oxidative stress in chloroplasts of pea plants. Plant Science. 1995;105:151–67. DOI: http://10.1016/0168-9452(94)04047-8

[4] Sondergaard TE, Schulz A, Palmgren MG. Energization of transport processes in plants. Roles of the plasma membrane H+-ATPase. Plant Physiol. 2004;136:2475–82. DOI: http://10.1104/pp.104.048231

[5] Grattan S. Irrigation water salinity and crop production. UCANR Publications, Oakland, CA; 2002. DOI: http://10.1007/s00425-003-1105-5

[6]   Rengasamy P. World salinization with emphasis on Australia. Journal of Experimental Botany. 2006;57:1017-23. DOI: http://10.1093/jxb/erj108

[7]   Eudoxie GD, Wuddivira M.Soil, water, and agricultural adaptations. Impacts of Climate Change on Food Security in Small Island Developing States Ganpat WG, Isaac WA (Editor),. 2014; IGI Global: p. 255. DOI: http://10.4018/978-1-4666-6501-9.ch001

[8]   Wang W, Vinocur B, Altman A. Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. Planta. 2003;218:1–14. DOI: http:// 10.1007/s00425-003-1105-5

[9]   Dai Q-l, Chen C, Feng B, Liu T-t, Tian X, Gong Y-y, et al. Effects of different NaCl concentration on the antioxidant enzymes in oilseed rape (*Brassica napus* L.) seedlings. Plant Growth Regulation. 2009;59:273–8. DOI: http://10.1007/s10725-009-9402-z

[10] Parida AK, Das AB. Salt tolerance and salinity effects on plants: a review. Ecotoxicology and Environmental Safety. 2005;60:324–49. DOI: http://10.1016/j.ecoenv.2004.06.010

[11] Flowers T. Improving crop salt tolerance. Journal of Experimental Botany. 2004;55:307–19. DOI: http://10.1093/jxb/erh003

[12] Cramer GR, Nowak RS. Supplemental manganese improves the relative growth, net assimilation and photosynthetic rates of salt-stressed barley. Physiologia Plantarum. 1992;84:600–5. DOI: http://10.1034/j.1399-3054.1992.840415.x

[13] Flowers T, Yeo A. Breeding for salinity resistance in crop plants: where next? Functional Plant Biology. 1995;22:875–84. DOI: http://10.1071/pp9950875

[14] Kabata-Pendias A. Trace elements in soils and plants. CRC press, oca Raton, USA; 2010. DOI: http://10.1201/9781420039900.ch5`

[15] Munns R, Schachtman D, Condon A. The significance of a two-phase growth response to salinity in wheat and barley. Functional Plant Biology. 1995;22:561–9. DOI: http://10.1071/pp9950561

[16] Kaur N, Gupta AK. Signal transduction pathways under abiotic stresses in plants. Current Science. 2005;88:1771–80. DOI: http://10.1007/978-3-642-25829-9_1

[17] Zhu J-K. Salt and drought stress signal transduction in plants. Annual Review of Plant Biology. 2002;53:247. DOI: http://10.1007/978-1-4020-5578-2_7

[18] Cornic G, Massacci A. Leaf photosynthesis under drought stress. Photosynthesis and the Environment. Springer Netherlands; 1996. pp. 347–66. DOI: http:// 10.1007/0-306-48135-9_14

[19] Kasuga M, Liu Q, Miura S, Yamaguchi-Shinozaki K, Shinozaki K. Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. Nature Biotechnology. 1999;17:287–91. DOI: http:// 10.1002/9780470515778.ch13

[20] Ghoulam C, Foursy A, Fares K. Effects of salt stress on growth, inorganic ions and proline accumulation in relation to osmotic adjustment in five sugar beet cultivars. Environmental and Experimental Botany. 2002;47:39–50. DOI: http://10.1016/s0098-8472(01)00109-5

[21] Greenway H, Munns R. Mechanisms of salt tolerance in nonhalophytes. Annual Review of Plant Physiology. 1980;31:149–90. DOI: http://10.1146/annurev.pp.31.060180.001053

[22] Flowers T, Yeo A. Ion relations of salt tolerance. Solute Transport in Plant Cells and Tissues. Springer Netherlands. 1988:392–416. DOI: http://10.1002/9780470988862.ch15

[23] Francois L, Maas E, Donovan T, Youngs V. Effect of salinity on grain yield and quality, vegetative growth, and germination of semi-dwarf and durum wheat. Agronomy Journal. 1986;78:1053–8. DOI: http://10.2134/agronj1986.00021962007800060023x

[24] Ashraf M, McNeilly T. Salinity tolerance in Brassica oilseeds. Critical Reviews in Plant Sciences. 2004;23:157–74. DOI: http://10.1080/07352680490433286

[25] Flowers T, Troke P, Yeo A. The mechanism of salt tolerance in halophytes. Annual Review of Plant Physiology. 1977;28:89–121. DOI: http://10.1626/jcs.68.supplement2_321

[26] Greenway H, Osmond C. Salt responses of enzymes from species differing in salt tolerance. Plant Physiology. 1972;49:256–9. DOI: http://10.1104/pp.49.2.260

[27] Sykes S. The inheritance of salt exclusion in woody perennial fruit species. Plant and Soil. 1992;146:123–9. DOI: http://10.1007/bf00012004

[28] Badr M, Shafei A. Salt tolerance in two wheat varieties and its relation to potassium nutrition. Journal of Agricultural Research. 2002;35:115–28.

[29] El-Fouly M, Moubarak Z, Salama Z, editors. Micronutrient foliar application increases salt tolerance of tomato seedlings. International Symposium on Techniques to Control Salination for Horticultural Productivity 573; 2000. DOI: http:// http://10.17660/ActaHortic.2002.573.57

[30] Munns R. Comparative physiology of salt and water stress. Plant, Cell & Environment. 2002;25:239–50. DOI: http://10.1046/j.0016-8025.2001.00808.x

[31] Topcu S. Water for agriculture: a major and inefficient consumer. Turkey's Water Policy. Springer Berlin Heidelberg; 2011. p. 93–115. DOI: http://10.1007/978-3-642-19636-2_6

[32] Gupta B, Huang B. Mechanism of salinity tolerance in plants: physiological, biochemical, and molecular characterization. International Journal of Genomics. 2014;2014:701596. DOI: http://10.1155/2014/701596

[33] Johnson RR, Wagner RL, Verhey SD, Walker-Simmons MK. The abscisic acid-responsive kinase PKABA1 interacts with a seed-specific abscisic acid response element-

binding factor, TaABF, and phosphorylates TaABF peptide sequences. Plant Physiology. 2002;130:837–46. DOI: http://10.1104/pp.001354

[34] Mizoi J, Shinozaki K, Yamaguchi-Shinozaki K. AP2/ERF family transcription factors in plant abiotic stress responses. Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms. 2012;1819:86–96. DOI: http://10.1016/j.bbagrm.2011.08.004

[35] Fujita Y, Yoshida T, Yamaguchi-Shinozaki K. Pivotal role of the AREB/ABF-SnRK2 pathway in ABRE-mediated transcription in response to osmotic stress in plants. Physiologia plantarum. 2013;147:15–27. DOI: http://10.1111/j.1399-3054.2012.01635.x

[36] Song S-Y, Chen Y, Chen J, Dai X-Y, Zhang W-H. Physiological mechanisms underlying OsNAC5-dependent tolerance of rice plants to abiotic stress. Planta. 2011;234:331–45. DOI: http://10.1007/s00425-011-1403-2

[37] Zhang L, Zhao G, Jia J, Liu X, Kong X. Molecular characterization of 60 isolated wheat MYB genes and analysis of their expression during abiotic stress. Journal of Experimental Botany. 2012;63:203–14. DOI: http://10.1093/jxb/err264

[38] Roy SJ, Negrão S, Tester M. Salt resistant crop plants. Current Opinion in Biotechnology. 2014;26:115–24. DOI: http://10.1016/j.copbio.2013.12.004

[39] Hu Y, Chen L, Wang H, Zhang L, Wang F, Yu D. *Arabidopsis* transcription factor WRKY8 functions antagonistically with its interacting partner VQ9 to modulate salinity stress tolerance. The Plant Journal. 2013;74:730–45. DOI: http://10.1111/tpj.12159

[40] Chakraborty K, Sairam RK, Bhattacharya R. Differential expression of salt overly sensitive pathway genes determines salinity stress tolerance in Brassica genotypes. Plant Physiology and Biochemistry. 2012;51:90–101. DOI: http://10.1016/j.plaphy.2011.10.001

[41] Hasegawa PM, Bressan RA, Zhu J-K, Bohnert HJ. Plant cellular and molecular responses to high salinity. Annual Review of Plant Biology. 2000;51:463–99. DOI: http://10.1146/annurev.arplant.51.1.463

[42] Schmidt R, Mieulet D, Hubberten H-M, Obata T, Hoefgen R, Fernie AR, et al. SALT-RESPONSIVE ERF1 regulates reactive oxygen species–dependent signaling during the initial response to salt stress in rice. The Plant Cell. 2013;25:2115–31. DOI: http://10.1105/tpc.113.113068

[43] Roshandel P, Flowers T. The ionic effects of NaCl on physiology and gene expression in rice genotypes differing in salt tolerance. Plant and Soil. 2009;315:135–47. DOI: http://10.1007/s11104-008-9738-6

[44] Zou J, Liu A, Chen X, Zhou X, Gao G, Wang W, et al. Expression analysis of nine rice heat shock protein genes under abiotic stresses and ABA treatment. Journal of Plant Physiology. 2009;166:851–61. DOI: http://10.1016/j.jplph.2008.11.007

[45] Gah-Hyun, Zhang X, Chung MS, Lee DJ, Woo YM, Cheong HS, et al. A putative novel transcription factor, AtSKIP, is involved in abscisic acid signalling and confers salt

and osmotic tolerance in *Arabidopsis*. New Phytologist. 2010;185:103–13. DOI: http://10.1111/j.1469-8137.2009.03032.xLim

[46] Zou J, Liu C, Liu A, Zou D, Chen X. Overexpression of OsHsp17. 0 and OsHsp23. 7 enhances drought and salt tolerance in rice. Journal of Plant Physiology. 2012;169:628–35. DOI: http://10.1016/j.jplph.2011.12.014

[47] Song N-H, Ahn Y-J. DcHsp17. 7, a small heat shock protein in carrot, is tissue-specifically expressed under salt stress and confers tolerance to salinity. New Biotechnology. 2011;28:698–704. DOI: http://10.1016/j.nbt.2011.04.002

[48] Tang M, Liu X, Deng H, Shen S. Over-expression of JcDREB, a putative AP2/EREBP domain-containing transcription factor gene in woody biodiesel plant *Jatropha curcas*, enhances salt and freezing tolerance in transgenic *Arabidopsis thaliana*. Plant Science. 2011;181:623–31. DOI: http://10.1016/j.plantsci.2011.06.014

[49] Islam MS, Azam MS, Sharmin S, Sajib AA, Alam MM, Reza MS, et al. Improved salt tolerance of jute plants expressing the katE gene from *Escherichia coli*. Turkish Journal of Biology. 2013;37:206–11. DOI: http://10.3906/biy-1205-52

[50] Shang G, Li Y, Hong Z, LIU C-l, HE S-z, LIU Q-c. Overexpression of SOS genes enhanced salt tolerance in sweetpotato. Journal of Integrative Agriculture. 2012;11:378–86. DOI: http://10.1016/s2095-3119(12)60022-7

[51] Sottosanto JB, Saranga Y, Blumwald E. Impact of AtNHX1, a vacuolar Na^+/H^+ antiporter, upon gene expression during short-and long-term salt stress in *Arabidopsis thaliana*. BMC Plant Biology. 2007;7:18. DOI: http://10.1186/1471-2229-7-18

[52] Liu G, Li X, Jin S, Liu X, Zhu L, Nie Y, et al. Overexpression of rice NAC gene SNAC1 improves drought and salt tolerance by enhancing root development and reducing transpiration rate in transgenic cotton. PLoS One. 2014;9(1): e86895. DOI: http://10.1371/journal.pone.0086895

[53] Peng X, Ding X, Chang T, Wang Z, Liu R, Zeng X, et al. Overexpression of a Vesicle Trafficking Gene, OsRab7, enhances salt tolerance in rice. The Scientific World Journal.;2014: 483526. DOI: http://10.1155/2014/483526

[54] Zhou J, Wang J, Bi Y, Wang L, Tang L, Yu X, et al. Overexpression of PtSOS2 enhances salt tolerance in transgenic poplars. Plant Molecular Biology Reporter. 2014;32:185–97. DOI: http://10.1007/s11105-013-0640-x

[55] Huang X, Zhang Y, Jiao B, Chen G, Huang S, Guo F, et al. Overexpression of the wheat salt tolerance-related gene TaSC enhances salt tolerance in *Arabidopsis*. Journal of Experimental Botany. 2012;63:5463–73. DOI: http://10.1093/jxb/ers198

[56] Han Y, Wang W, Sun J, Ding M, Zhao R, Deng S, et al. *Populus euphratica* XTH over-expression enhances salinity tolerance by the development of leaf succulence in transgenic tobacco plants. Journal of Experimental Botany. 2013;64:4225–38. DOI: http://10.1093/jxb/ert229

[57] Zhang G-C, Zhu W-L, Gai J-Y, Zhu Y-L, Yang L-F. Enhanced salt tolerance of transgenic vegetable soybeans resulting from overexpression of a novel 1-pyrroline-5-carboxylate synthetase gene from *Solanum torvum Swartz*. Horticulture, Environment, and Biotechnology. 2015;56:94–104. DOI: http://10.1007/s13580-015-0084-3

[58] Kurotani K-i, Hayashi K, Hatanaka S, Toda Y, Ogawa D, Ichikawa H, et al. Elevated levels of CYP94 family gene expression alleviate the jasmonate response and enhance salt tolerance in rice. Plant and Cell Physiology. 2015:pcv006. DOI: http://10.1093/pcp/pcv006

[59] Mishra MK, Chaturvedi P, Singh R, Singh G, Sharma LK, Pandey V, et al. Overexpression of WsSGTL1 gene of *Withania somnifera* enhances salt tolerance, heat tolerance and cold acclimation ability in transgenic *Arabidopsis* plants. PLoS One. 2013;8:e63064. DOI: http://10.1371/journal.pone.0063064

[60] Zhou L, Wang C, Liu R, Han Q, Vandeleur RK, Du J, et al. Constitutive overexpression of soybean plasma membrane intrinsic protein GmPIP1; 6 confers salt tolerance. BMC Plant Biology. 2014;14:181. DOI: http://10.1186/1471-2229-14-181

[61] Lawson SS, Michler CH. Overexpression of AtSTO1 leads to improved salt tolerance in *Populus tremula× P. alba*. Transgenic Research. 2014;23:817–26. DOI: http://10.1007/s11248-014-9808-x

[62] Zheng X, Chen B, Lu G, Han B. Overexpression of a NAC transcription factor enhances rice drought and salt tolerance. Biochemical and Biophysical Research Communications. 2009;379:985–9. DOI: http://10.1016/j.bbrc.2008.12.163

[63] Yue Y, Zhang M, Zhang J, Duan L, Li Z. SOS1 gene overexpression increased salt tolerance in transgenic tobacco by maintaining a higher K^+/Na^+ ratio. Journal of Plant Physiology. 2012;169:255–61. DOI: http://10.1016/j.jplph.2011.10.007

[64] auso TD, Thankappan R, Kumar A, Mishra GP, Dobaria JR, Venkat Rajam M. Overexpression of bacterial'mtlD'gene confers enhanced tolerance to salt-stress and water-deficit stress in transgenic peanut ('*Arachis hypogaea*') through accumulation of mannitol. Australian Journal of Crop Science. 2014;3:413. DOI: http://10.1155/2014/125967

[65] Saxena M, Roy SD, Singla-Pareek S-L, Sopory SK, Bhalla-Sarin N. Overexpression of the glyoxalase II gene leads to enhanced salinity tolerance in *Brassica juncea*. The Open Plant Science Journal. 2011;5:23–8. DOI: http://10.2174/1874294701105010023

[66] García-Abellan JO, Egea I, Pineda B, Sanchez-Bel P, Belver A, Garcia-Sogo B, et al. Heterologous expression of the yeast HAL5 gene in tomato enhances salt tolerance by reducing shoot Na^+ accumulation in the long term. Physiologia Plantarum. 2014;152:700–13. DOI: http://10.1111/ppl.12217

[67] Zhou W, Li Y, Zhao BC, Ge RC, Shen YZ, Wang G, Huang ZJ. Overexpression of *TaSTRG* gene improves salt and drought tolerance in rice. Journal of plant physiology. 2009;166:1660-71. DOI: http://10.1016/j.jplph.2009.04.015.

[68] Serra TS, Figueiredo DD, Cordeiro AM, Almeida DM, Lourenço T, Abreu IA, et al. OsRMC, a negative regulator of salt stress response in rice, is regulated by two AP2/ERF transcription factors. Plant Molecular Biology. 2013;82:439–55. DOI: http://10.1007/s11103-013-0073-9

[69] Mukherjee K, Choudhury AR, Gupta B, Gupta S, Sengupta DN. An ABRE-binding factor, OSBZ8, is highly expressed in salt tolerant cultivars than in salt sensitive cultivars of indica rice. BMC Plant Biology. 2006;6:18. DOI: http://:10.1186/1471-2229-6-18

[70] Sairam R, Tyagi A. Physiology and molecular biology of salinity stress tolerance in plants. Current Science. 2004;86:407–21. DOI: http://10.1007/1-4020-4225-6_7

[71] Staiger D, Brown JW. Alternative splicing at the intersection of biological timing, development, and stress responses. The Plant Cell. 2013;25:3640–56. DOI: http://10.1105/tpc.113.113803

[72] Mangrauthia SK, Agarwal S, Sailaja B, Madhav MS, Voleti S. MicroRNAs and their role in salt stress response in plants. Salt Stress in Plants. Springer New York; 2013. pp. 15–46. DOI: http://10.1007/978-1-4614-6108-1_2

[73] Xie F, Stewart CN, Taki FA, He Q, Liu H, Zhang B. High-throughput deep sequencing shows that microRNAs play important roles in switchgrass responses to drought and salinity stress. Plant Biotechnology Journal. 2014;12:354–66. DOI: http://10.1111/pbi.12142

[74] Lu C, Tej SS, Luo S, Haudenschild CD, Meyers BC, Green PJ. Elucidation of the small RNA component of the transcriptome. Science. 2005;309:1567–9. DOI: http://10.1111/pbi.12142

[75] Sunkar R, Zhu J-K. Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. The Plant Cell. 2004;16:2001–19. DOI: http://10.1105/tpc.104.022830

[76] Lu S, Sun YH, Chiang VL. Stress-responsive microRNAs in Populus. The Plant Journal. 2008;55:131–51. DOI: http://10.1111/j.1365-313x.2008.03497.x

[77] Arenas-Huertero C, Pérez B, Rabanal F, Blanco-Melo D, De la Rosa C, Estrada-Navarrete G, et al. Conserved and novel miRNAs in the legume *Phaseolus vulgaris* in response to stress. Plant Molecular Biology. 2009;70:385–401. DOI: http://10.1007/s11103-009-9480-3

[78] Asins M. Present and future of quantitative trait locus analysis in plant breeding. Plant Breeding. 2002;121:281–91. DOI: http://10.1046/j.1439-0523.2002.730285.x

[79] Yeo AR, Koyama ML, Chinta S, Flowers TJ. Salt tolerance at the whole-plant level. Plant Tolerance to Abiotic Stresses in Agriculture: Role of Genetic Engineering. Springer Netherlands; 2000. pp. 107–23. DOI: http://10.1007/978-94-011-4323-3

[80] Horst L, Wenzel G. Molecular marker systems in plant breeding and crop improvement. Springer, Berlin Heidelberg; 2007. DOI: http://10.1007/b137756

[81] Ren Z-H, Gao J-P, Li L-G, Cai X-L, Huang W, Chao D-Y, et al. A rice quantitative trait locus for salt tolerance encodes a sodium transporter. Nature Genetics. 2005;37:1141–6. DOI: http://10.1038/ng1643

[82] Huang S, Spielmeyer W, Lagudah ES, Munns R. Comparative mapping of HKT genes in wheat, barley, and rice, key determinants of Na^+ transport, and salt tolerance. Journal of Experimental Botany. 2008;59:927–37. DOI: http://10.1093/jxb/ern033

[83] Xue D, Huang Y, Zhang X, Wei K, Westcott S, Li C, et al. Identification of QTLs associated with salinity tolerance at late growth stage in barley. Euphytica. 2009;169:187–96. DOI: http://10.1007/s10681-009-9919-2

[84] Wang J, Drayton MC, George J, Cogan NO, Baillie RC, Hand ML, et al. Identification of genetic factors influencing salt stress tolerance in white clover (*Trifolium repens* L.) by QTL analysis. Theoretical and Applied Genetics. 2010;120:607–19. DOI: http://10.1007/s00425-003-1105-5

[85] Li J, Liu L, Bai Y, Finkers R, Wang F, Du Y, et al. Identification and mapping of quantitative resistance to late blight (*Phytophthora infestans*) in *Solanum habrochaites* LA1777. Euphytica. 2011;179:427–38. DOI: http://10.1007/s10681-010-0340-7

[86] Tuyen D, Lal S, Xu D. Identification of a major QTL allele from wild soybean (*Glycine soja Sieb. & Zucc.*) for increasing alkaline salt tolerance in soybean. Theoretical and Applied Genetics. 2010;121:229–36. DOI: http://10.1007/s00122-010-1304-y

[87] Kumar M, Choi J-Y, Kumari N, Pareek A, Kim S-R. Molecular breeding in Brassica for salt tolerance: importance of microsatellite (SSR) markers for molecular breeding in Brassica. Frontiers in Plant Science. 2015;6: 688. DOI: http://10.3389/fpls.2015.00688

[88] Gorham J, Jones RW, Bristol A. Partial characterization of the trait for enhanced K+−Na+ discrimination in the D genome of wheat. Planta. 1990;180:590–7. DOI: http://10.1007/bf02411458

[89] Dubcovsky J, Santa Maria G, Epstein E, Luo M-C, Dvořák J. Mapping of the K^+/Na^+ discrimination locus Kna1 in wheat. Theoretical and Applied Genetics. 1996;92:448–54. DOI: http://10.1007/bf00223692

[90] Lindsay MP, Lagudah ES, Hare RA, Munns R. A locus for sodium exclusion (Nax1), a trait for salt tolerance, mapped in durum wheat. Functional Plant Biology. 2004;31:1105–14. DOI: http://10.1071/fp04111

[91] Huang S, Spielmeyer W, Lagudah ES, James RA, Platten JD, Dennis ES, et al. A sodium transporter (HKT7) is a candidate for Nax1, a gene for salt tolerance in durum wheat. Plant Physiology. 2006;142:1718–27. DOI: http://10.1104/pp.106.088864

[92] Sabouri H, Sabouri A. New evidence of QTLs attributed to salinity tolerance in rice. African Journal of Biotechnology. 2008;7: 4376–4383.

[93] Sabouri H. QTL detection of rice grain quality traits by microsatellite markers using an indica rice (*Oryza sativa* L.) combination. Journal of Genetics. 2009;88:81–5. DOI: http://10.1007/s12041-009-0011-4

[94] Yao M-Z, Wang J-F, Chen H-Y, Zhai H-Q, Zhang H-S. Inheritance and QTL mapping of salt tolerance in rice. Rice Science. 2005;12:25–32.

[95] Thomson MJ, de Ocampo M, Egdane J, Rahman MA, Sajise AG, Adorada DL, et al. Characterizing the Saltol quantitative trait locus for salinity tolerance in rice. Rice. 2010;3:148–60. DOI: http://10.1007/s12284-010-9053-8

[96] Alam R, Sazzadur Rahman M, Seraj ZI, Thomson MJ, Ismail AM, Tumimbang-Raiz E, et al. Investigation of seedling-stage salinity tolerance QTLs using backcross lines derived from *Oryza sativa* L. Pokkali. Plant Breeding. 2011;130:430–7. DOI: http://10.1111/j.1439-0523.2010.01837.x

[97] Ahmadi J, Fotokian M-H. Identification and mapping of quantitative trait loci associated with salinity tolerance in rice (*Oryza Sativa*) using SSR markers. Iranian Journal of Biotechnology. 2011;9:21–30.

[98] Zhou M, Johnson P, Zhou G, Li C, Lance R. Quantitative trait loci for waterlogging tolerance in a barley cross of *Franklin× YuYaoXiangTian Erleng* and the relationship between waterlogging and salinity tolerance. Crop Science. 2012;52:2082–8. DOI: http://10.2135/cropsci2012.01.0008

[99] Shavrukov Y, Gupta NK, Miyazaki J, Baho MN, Chalmers KJ, Tester M, et al. HvNax3—a locus controlling shoot sodium exclusion derived from wild barley (*Hordeum vulgare ssp. spontaneum*). Functional & Integrative Genomics. 2010;10:277–91. DOI: http://10.1007/s10142-009-0153-8

[100] Lee G, Boerma H, Villagarcia M, Zhou X, Carter Jr T, Li Z, et al. A major QTL conditioning salt tolerance in S-100 soybean and descendent cultivars. Theoretical and Applied Genetics. 2004;109:1610–9. http://10.1007/s00122-004-1783-9

[101] Chen H, Cui S, Fu S, Gai J, Yu D. Identification of quantitative trait loci associated with salt tolerance during seedling growth in soybean (*Glycine max* L.). Crop and Pasture Science. 2008;59:1086–91. DOI: http://10.1071/ar08104

[102] Hamwieh A, Tuyen D, Cong H, Benitez E, Takahashi R, Xu D. Identification and validation of a major QTL for salt tolerance in soybean. Euphytica. 2011;179:451–9. DOI: http://10.1007/s10681-011-0347-8

[103] James C. Global status of commercialized biotech/GM crops: 2010. International Serv-ice for the Acquisition of Agri-biotech Applications (ISAAA) Ithaca, NY, USA; 2010. DOI: http://10.1017/s0014479706343797

[104] Ashraf M, Athar H, Harris P, Kwon T. Some prospective strategies for improving crop salt tolerance. Advances in Agronomy. 2008;97:45–110. DOI: http://10.1016/s0065-2113(07)00002-8

[105] Munns R, Tester M. Mechanisms of salinity tolerance. Annual Review of Plant Biolo-gy. 2008;59:651–81. DOI: http://10.1146/annurev.arplant.59.032607.092911

[106] Goel D, Singh AK, Yadav V, Babbar SB, Murata N, Bansal KC. Transformation of to-mato with a bacterial codA gene enhances tolerance to salt and water stresses. Jour-nal of Plant Physiology. 2011;168:1286–94. DOI: http://10.1016/j.jplph.2011.01.010

[107] Dong Y, Ogawa T, Lin D, Koh H-J, Kamiunten H, Matsuo M, et al. Molecular map-ping of quantitative trait loci for zinc toxicity tolerance in rice seedling (*Oryza sativa* L.). Field Crops Research. 2006;95:420–5. DOI: http://10.1016/j.fcr.2005.03.005

[108] Su J, Hirji R, Zhang L, He C, Selvaraj G, Wu R. Evaluation of the stress-inducible pro-duction of choline oxidase in transgenic rice as a strategy for producing the stress-protectant glycine betaine. Journal of Experimental Botany. 2006;57:1129–35. DOI: http://10.1093/jxb/erj133

[109] .Li H-W, Zang B-S, Deng X-W, Wang X-P. Overexpression of the trehalose-6-phos-phate synthase gene OsTPS1 enhances abiotic stress tolerance in rice. Planta. 2011;234:1007–18. DOI: http://10.1007/s00425-011-1458-0

[110] Cortina C, Culiáñez-Macià FA. Tomato abiotic stress enhanced tolerance by trehalose biosynthesis. Plant Science. 2005;169:75–82. DOI: http://10.1016/j.plantsci.2005.02.026

[111] K A, Kim J-K, Owens TG, Ranwala AP, Do Choi Y, Kochian LV, et al. Trehalose accu-mulation in rice plants confers high tolerance levels to different abiotic stresses. Pro-ceedings of the National Academy of Sciences. 2002;99:15898–903. DOI: http://10.1073/pnas.252637799

[112] Almeida AM, Villalobos E, Araújo SS, Leyman B, Van Dijck P, Alfaro-Cardoso L, et al. Transformation of tobacco with an *Arabidopsis thaliana* gene involved in trehalose biosynthesis increases tolerance to several abiotic stresses. Euphytica. 2005;146:165–76. DOI: http://10.1007/s10681-005-7080-0

[113] Tarczynski MC, Jensen RG, Bohnert HJ. Stress protection of transgenic tobacco by production of the osmolyte mannitol. Science. 1993;259:508–10. DOI: http://10.1126/science.259.5094.508

[114] Abebe T, Guenzi AC, Martin B, Cushman JC. Tolerance of mannitol-accumulating transgenic wheat to water stress and salinity. Plant Physiology. 2003;131:1748–55. DOI: http://10.1104/pp.102.003616

[115] Sickler CM, Edwards GE, Kiirats O, Gao Z, Loescher W. Response of mannitol-producing *Arabidopsis thaliana* to abiotic stress. Functional Plant Biology. 2007;34:382–91. DOI: http://10.1071/fp06274

[116] Gao M, Tao R, Miura K, Dandekar AM, Sugiura A. Transformation of Japanese persimmon (Diospyros kaki Thunb.) with apple cDNA encoding NADP-dependent sorbitol-6-phosphate dehydrogenase. Plant Science. 2001;160:837–45. DOI: http://10.1016/s0168-9452(00)00458-1

[117] Su J, Wu R. Stress-inducible synthesis of proline in transgenic rice confers faster growth under stress conditions than that with constitutive synthesis. Plant Science. 2004;166:941–8. DOI: http://10.1016/j.plantsci.2003.12.004

[118] Kishor PK, Hong Z, Miao G-H, Hu C-AA, Verma DPS. Overexpression of [delta]-pyrroline-5-carboxylate synthetase increases proline production and confers osmotolerance in transgenic plants. Plant Physiology. 1995;108:1387–94. DOI: http://10.1016/j.plantsci.2005.05.025

[119] Gao X, Ren Z, Zhao Y, Zhang H. Overexpression of SOD2 increases salt tolerance of *Arabidopsis*. Plant Physiology. 2003;133:1873–81. DOI: http://10.1104/pp.103.026062

[120] Wu L, Fan Z, Guo L, Li Y, Chen Z-L, Qu L-J. Over-expression of the bacterial nhaA gene in rice enhances salt and drought tolerance. Plant Science. 2005;168:297–302. DOI: http://10.1016/j.plantsci.2004.05.033

[121] Pasapula V, Shen G, Kuppu S, Paez-Valencia J, Mendoza M, Hou P, et al. Expression of an *Arabidopsis* vacuolar H+-pyrophosphatase gene (AVP1) in cotton improves drought-and salt tolerance and increases fibre yield in the field conditions. Plant Biotechnology Journal. 2011;9:88–99. DOI: http://10.1111/j.1467-7652.2010.00535.x

[122] Zhang H-X, Hodson JN, Williams JP, Blumwald E. Engineering salt-tolerant Brassica plants: characterization of yield and seed oil quality in transgenic plants with increased vacuolar sodium accumulation. Proceedings of the National Academy of Sciences. 2001;98:12832–6. DOI: http://10.1073/pnas.231476498

[123] Ohta M, Hayashi Y, Nakashima A, Hamada A, Tanaka A, Nakamura T, et al. Introduction of a Na+/H+ antiporter gene from Atriplex gmelini confers salt tolerance to rice. FEBS Letters. 2002;532:279–82. DOI: http://10.1016/s0014-5793(02)03679-7

[124] Wang J, Zuo K, Wu W, Song J, Sun X, Lin J, et al. Expression of a novel antiporter gene from Brassica napus resulted in enhanced salt tolerance in transgenic tobacco plants. Biologia Plantarum. 2004;48:509–15. DOI: http://10.1023/b:biop.0000047145.18014.a3

[125] Wu C-A, Yang G-D, Meng Q-W, Zheng C-C. The cotton GhNHX1 gene encoding a novel putative tonoplast Na+/H+ antiporter plays an important role in salt stress. Plant and Cell Physiology. 2004;45:600–7. DOI: http://10.1093/pcp/pch071

[126] Singla-Pareek S, Reddy M, Sopory S. Genetic engineering of the glyoxalase pathway in tobacco leads to enhanced salinity tolerance. Proceedings of the National Academy of Sciences. 2003;100:14672–7. DOI: http://10.1073/pnas.2034667100

[127] Gao S-Q, Chen M, Xu Z-S, Zhao C-P, Li L, Xu H-j, et al. The soybean GmbZIP1 transcription factor enhances multiple abiotic stress tolerances in transgenic plants. Plant molecular biology. 2011;75:537–53. DOI: http://10.1007/s11103-011-9738-4

[128] Mao X, Jia D, Li A, Zhang H, Tian S, Zhang X, et al. Transgenic expression of TaMYB2A confers enhanced tolerance to multiple abiotic stresses in *Arabidopsis*. Functional & Integrative Genomics. 2011;11:445–65. DOI: http://10.1007/s10142-011-0218-3

[129] Seo YJ, Park J-B, Cho Y-J, Jung C, Seo HS, Park S-K, et al. Overexpression of the ethylene-responsive factor gene BrERF4 from Brassica rapa increases tolerance to salt and drought in *Arabidopsis* plants. Molecules and Cells. 2010;30:271–7. DOI: http://10.1007/s10059-010-0114-z

[130] Montero-Barrientos M, Hermosa R, Cardoza RE, Gutiérrez S, Nicolás C, Monte E. Transgenic expression of the Trichoderma harzianum hsp70 gene increases *Arabidopsis* resistance to heat and other abiotic stresses. Journal of Plant Physiology. 2010;167:659–65. DOI: http://10.1016/j.jplph.2009.11.012

[131] Xu D, Duan X, Wang B, Hong B, Ho T-HD, Wu R. Expression of a late embryogenesis abundant protein gene, HVA1, from barley confers tolerance to water deficit and salt stress in transgenic rice. Plant Physiology. 1996;110:249–57. DOI: http://10.1104/pp.110.1.249

[132] Zhang L, Xi D, Li S, Gao Z, Zhao S, Shi J, et al. A cotton group C MAP kinase gene, GhMPK2, positively regulates salt and drought tolerance in tobacco. Plant Molecular Biology. 2011;77:17–31. DOI: http://10.1007/s11103-011-9788-7

[133] Apse MP, Aharon GS, Snedden WA, Blumwald E. Salt tolerance conferred by overexpression of a vacuolar Na$^+$/H$^+$ antiport in *Arabidopsis*. Science. 1999;285:1256–8. DOI: http://10.1126/science.285.5431.1256

[134] Zhang H-X, Blumwald E. Transgenic salt-tolerant tomato plants accumulate salt in foliage but not in fruit. Nature Biotechnology. 2001;19:765–8. DOI: http://10.1038/90824

[135] He C, Yan J, Shen G, Fu L, Holaday AS, Auld D, et al. Expression of an *Arabidopsis* vacuolar sodium/proton antiporter gene in cotton improves photosynthetic performance under salt conditions and increases fiber yield in the field. Plant and Cell Physiology. 2005;46:1848–54. DOI: http://10.1093/pcp/pci201

[136] 136Shi H, Lee B-h, Wu S-J, Zhu J-K. Overexpression of a plasma membrane Na$^+$/H$^+$ antiporter gene improves salt tolerance in *Arabidopsis thaliana*. Nature Biotechnology. 2003;21:81–5. DOI: http://10.1038/nbt766

[137]   Koh E-J, Song W-Y, Lee Y, Kim KH, Kim K, Chung N, et al. Expression of yeast cadmium factor 1 (YCF1) confers salt tolerance to *Arabidopsis thaliana*. Plant Science. 2006;170:534–41. DOI: http://10.1016/j.plantsci.2005.10.007

[138]   Xue Z-Y, Zhi D-Y, Xue G-P, Zhang H, Zhao Y-X, Xia G-M. Enhanced salt tolerance of transgenic wheat (*Tritivum aestivum* L.) expressing a vacuolar Na^+/H^+ antiporter gene with improved grain yields in saline soils in the field and a reduced level of leaf Na^+. Plant Science. 2004;167:849–59. DOI: http://10.1016/j.plantsci.2004.05.034

[139]   Yin X-Y, Yang A-F, Zhang K-W, Zhang J-R. Production and analysis of transgenic maize with improved salt tolerance by the introduction of AtNHX1 gene. Acta Botanica Sinica—English Edition. 2004;46:854–61. DOI: http://10.1007/s11248-007-9085-z

[140]   Fukuda A, Nakamura A, Tagiri A, Tanaka H, Miyao A, Hirochika H, et al. Function, intracellular localization and the importance in salt tolerance of a vacuolar Na^+/H^+ antiporter from rice. Plant and Cell Physiology. 2004;45:146–59. DOI: http://10.1093/pcp/pch014

[141]   Møller IS, Gilliham M, Jha D, Mayo GM, Roy SJ, Coates JC, et al. Shoot Na^+ exclusion and increased salinity tolerance engineered by cell type–specific alteration of Na^+ transport in *Arabidopsis*. The Plant Cell. 2009;21:2163–78. DOI: http://10.1105/tpc.108.064568

[142]   Yamaguchi T, Blumwald E. Developing salt-tolerant crop plants: challenges and opportunities. Trends in Plant Science. 2005;10:615–20. DOI: http://10.1016/j.tplants.2005.10.002

Edited by Ibrokhim Y. Abdurakhmonov

Plant genomics aims to sequence, characterize, and study the genetic compositions, structures, organizations, functions, and interactions/networks of an entire plant genome. Its development and advances are tightly interconnected with proteomics, metabolomics, metagenomics, transgenomics, genomic selection, bioinformatics, epigenomics, phenomics, system biology, modern instrumentation, and robotics sciences. Plant genomics has significantly advanced over the past three decades in the land of inexpensive, high-throughput sequencing technologies and fully sequenced over 100 plant genomes. These advances have broad implications in every aspect of plant biology and breeding, powered with novel genomic selection and manipulation tools while generating many grand challenges and tasks ahead. This Plant genomics provides some updated discussions on current advances, challenges, and future perspectives of plant genome studies and applications.

IntechOpen