Chapter 3
Bioinformatics and Plant Stress Management

Amrina Shafi and Insha Zahoor

Contents

3.1   Introduction

In nature, plants are simultaneously exposed to a wide range of stresses (abiotic and biotic), which is a major threat towards the living world more precisely to the plants. This stress leads to various physiological and metabolic changes, which in turn negatively hinder growth, development and productivity of plants (Tardieu and Tuberosa 2010). Based on the climate change study, the occurrence and severity of stresses will surge, resulting in a loss (nearly 70%) of agricultural production (Ghosh and Xu 2014). Thus, an important solution for plant protection and yield

A. Shafi (✉)
Department of Biotechnology, School of Biological Sciences, University of Kashmir, Srinagar, Jammu and Kashmir, India

I. Zahoor
Bioinformatics Centre, University of Kashmir, Srinagar, Jammu and Kashmir, India
e-mail: drinsha@uok.edu.in

increase is by designing plants based on a molecular understanding of gene function and on the regulatory networks involved in stress tolerance, growth and development (Shafi et al. 2014, 2015a, b, 2017). Technological advancement has offered a holistic view on systems organization and functionality; however, the ever-growing extensive data poses great challenges for its efficient analysis and interpretation and finally the integration into different crop improvement schemes (Esposito et al. 2016). Latest, ultrahigh-throughput computational studies are crucial to know about the molecular crosstalks of stress conditions on agricultural crop production. Now the challenge is how to integrate multidimensional biological information in a network and model leading to the development of system biology. Most of the plant system biology strategies rely on four main axes, viz. genomics, proteomics, transcriptomics and metabolomics, which provide us with a better platform to identify and understand the molecular systematics and mechanism under stress conditions (Yuan et al. 2008). Genomics deals with the study of genome; transcriptomics includes structural and functional analyses of coding and non-coding RNA or transcriptome; proteomics deals with protein and post-translational protein modification along with their regulatory pathway and metabolomics, a powerful tool to analyse various metabolites and help in identifying the complex network involved in stress tolerance when analysed in an integrated way. Multifaceted molecular regulatory system and biochemical properties which are specifically involved in stress tolerance and adaptation in plants can be easily deciphered with the help of combined 'omics' study (Chawla et al. 2011). Further, bioinformatics has many practical applications in current plant disease management with respect to the study of host-pathogen interactions, understanding the disease genetics, pathogenicity factor of a pathogen and plant-pathogen biological network, which ultimately help in designing best disease management options (Koltai and Volpin 2003).

Plant amends its 'omics' profiles to cope with the changing environment for their survival, tolerance and growth. The main aim of this 'omic' approach is to find out the molecular interaction and their relationship with the signalling cascade and to process the information which in turn connects specific signals with specific molecular responses (Esposito et al. 2016). The era of genomics, proteomics, metabolomics and phenomics of crop stress biology involves transformation, mining and functional ontology annotation, promoter and SNP analysis, gene expression, pathway enrichment analysis, microRNA prediction, subcellular localization, gene structure analysis, comparative analysis, interactome, protein function analysis, tissues-specific and developmental stage expression analysis and simulation and focused on morpho-molecular differences in stress-exposed and stress-affected crop/model plants. These omics approaches can provide new insights and open new horizons for understanding stresses and responses as well as the improvement of plant responses and resistance to stresses (Duque et al. 2013).

Little is known about the 'omics' characterization of abiotic and biotic stress combinations, but recently, several reports have addressed this issue (Suzuki et al. 2014; Kissoudis et al. 2014). The three main domains that must be addressed to take full advantage of plant systems biology are the development of omics technology, integration of data in a usable format and analysis of data within the domain
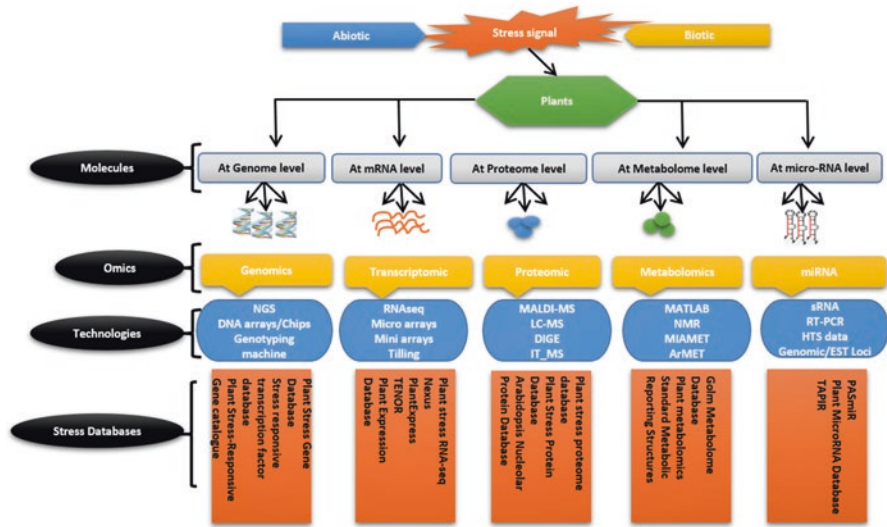
Fig. 3.1 Schematic outline of main 'omics' approaches, their technologies and databases as well as expected outcomes in plant biology and stress research

of bioinformatics. This explicit omics knowledge could subsequently be harnessed by researchers to develop improved crop plants in terms of quality and productivity, showing the enhanced level of abiotic stress tolerance and disease resistance (Singh et al. 2011). In the present chapter, we will introduce the key omics technologies and contemporary innovative technology employed in plant biology and the bioinformatics platforms associated with them (Fig. 3.1). Since the focus of this chapter is the integrated omics approaches in plant stress tolerance, we will describe some of the key concepts, techniques and databases used in bioinformatics, with an emphasis on those relevant to plant stress. It also covers some aspects with regard to the role of application of this endeavour science in today's plant disease management strategies, molecular diagnosis of plant disease in particular to see the application of bioinformatics in detection and diagnosis of plant pathogenic microorganisms.

3.2 Plant Genomics-Related Computational Tools and Databases Under Abiotic Stress

Developments over the past decade, arising predominantly from the human genome project, have led to a new phase of plant genetics known as genomics. 'Genomics' study of all the genes in a given genome includes the identification of gene sequences, intragenic sequences, gene structures and annotations (Duque et al. 2013). This field is the application of the newly available vast amounts of genomic DNA sequence, using a range of novel high-throughput, parallel and other technologies.

The innovation of high-throughput sequencing methods gives scientists the ability to exploit the structure of the genetic material at the molecular level which is known as 'genomics'. Genome sequencing technologies have enabled us to study this part of omics, and it has commenced with the first generation (the 1970s), followed by the next-generation sequencing (NGS) technologies (1990s) as well as the latest third-generation sequencing technologies (El-Metwally et al. 2014b, c). These NGS technologies have got huge impact in plant genome research for the improvement of economically important crops and the understanding of model plant biology. Substantial innovations in platforms for omics-based research and application development provide crucial resources to promote stress-related research in the model and applied plant species (Feuillet et al. 2011). Recent advancement in plant genomics has allowed us to discover and isolate important genes and to analyse functions that regulate yields and tolerance to environmental stress (Govind et al. 2009). Genomics mainly helped in identifying the functional relevance of genes involved in abiotic and biotic stress responses in plants via functional genomic approaches (Ramegowda and Senthil-Kumar 2015). Combinatorial approach using multiple omics platforms and integration of their outcomes is now an effective strategy for clarifying molecular systems integral to improving plant stress tolerance and productivity; this combo approach has helped plant breeders in creating new breeds that can tolerate several biotic and abiotic stresses and, consequently, have increased crop yields as well as pathogen resistance (Shankar et al. 2013; Agarwal et al. 2014). Thus an understanding of plant response towards stresses is enhanced with the application of genomic techniques such as high-throughput analysis of expressed sequence tags (ESTs), large-scale parallel analysis of gene expression, targeted or random mutagenesis and gain-of-function or mutant complementation (Cushman and Bohnert 2000).

Plant genomics study has exploded recently and has become the major boom in plant research due to the rapid increase in plant genomic sequences (Govindaraj et al. 2015). This plant genomic period started from whole-genome sequencing of *Arabidopsis thaliana* (The Arabidopsis Genome Initiative 2000), followed by a draft genome sequence of rice, both *japonica* and *indica* (Yu et al. 2002). Afterwards, the genome sequence of *japonica* rice was completed and published by the International Rice Genome Sequencing Project (International Rice Genome Sequencing Project 2005). Subsequently, the National Science Foundation (NSF) Arabidopsis project (USA) was launched with the stated goal of determining the functions of the 25,000 genes of *Arabidopsis* by 2010. This accumulation of nucleotide sequences of model plants, as well as of applied species such as crops, has provided fundamental information for the design of sequence-based research applications in functional genomics (Somerville and Dangl 2000). Technologies which are included under the canopy of 'genomics' are:

- Automatic DNA sequencing (the machine can read two million base pairs a day)
- Microarrays and DNA chips (tens of thousands of genes can be scanned for activity levels at the same time)
- Automated genotyping machines (assay tens of thousands of DNA diagnostic points a day)

Bioinformatics remains obligatory to projects that seek deciphering of the whole genome of an organism. In fact, soon it will be possible to monitor whole genomes for gene expression on single chips. Once genome sequencing is achieved, one aims to identify and delineate the genomic elements of functional relevance contained within the genome, i.e. 'structural annotation' and assigning biological functions to these elements, referred to as 'functional annotation'.

3.2.1 Genomics Applications in Relation to Abiotic Stress Tolerance

In order to employ applications of genomics field to address the problems of abiotic stress in mandate crops and model plants, approaches like genomic-scale expressed sequence tags (ESTs), genomic sequencing and cDNA microarray analyses have tremendous potentiality in rapidly isolating the candidate genes involved in tolerance mechanisms under stress conditions. Some of the latest techniques used for genomic analysis under stress conditions are as follows:

- Expressed sequence tags (ESTs) are created by partial 'one-pass' sequencing of randomly picked gene transcripts that have been converted into cDNA (Adams et al. 1993). ESTs are often used to be relative collections from stressed and non-stressed plant tissues. A comparison of ESTs of the stressed and non-stressed sample will identify genes that are up-regulated in the stressed tissues and those which are down-regulated or switched off.
- The presence of various key functionalities of full-length cDNA resources in omic space is also essential to establish relevant information resources that provide gateways to these resources as well as to integrate related datasets derived from other omics fields and species (Sakurai et al. 2005).
- cDNA libraries also serve as primary sequence resources for designing microarray probes and as clone resources for genetic engineering to improve crop efficiency (Futamura et al. 2008). Further, candidate genes (induced by stress) which emerge from microarray analyses are ideal for comparative analysis.
- Mini-arrays which are built from collections of ESTs assembled from random cDNA libraries, or from more targeted collections made from cDNAs collected from stressed tissues. Even more targeted will be the special 'stress arrays' made up of all the expressed genes for which there is any evidence of implication.

Omics platforms and their associated databases are also essential for the effective design of approaches making the best use of genomic resources, including resource integration. Various bioinformatics software and tools are being increasingly used to maintain, analyse and retrieve the massive-scale molecular data under stress and non-stressed conditions (Table 3.1). Some of those involved specifically under stress are as follows:

- *Plant Stress Gene Database (PSGD):* It provides information about the genes involved in stress conditions in plants (Prabha et al. 2011). This database includes

Table 3.1 Genomic repositories and stress-related databases

S. No.	Genomics database	Information	References
1.	PlantGDB – resources for Plant Comparative Genomics	Molecular sequence data for all plant species with significant sequencing efforts	Dong et al. (2004)
2.	TAIR (The *Arabidopsis* Information Resource)	Complete genome sequence, gene structure, expression and product information of *Arabidopsis*	Rhee et al. (2003)
3.	Gramene	Online web database resource for plant comparative genomics and pathway analysis based on Ensembl technology	Gupta et al. (2016)
4.	Plant Genome DataBase Japan (PGDBj)	A website that contains information related to genomes of model and crop plants from databases	Nakaya et al. (2017)
5.	PLAZA	Online resource for comparative genomics that integrates plant sequence data and comparative genomic methods	Vandepoele (2017)
6.	Legume Information System (LIS)	Genomic database for the legume family	Dash et al. (2016)
7.	Ensembl Genomes	Analysis and visualization of genomic data and provides access to a variety of data obtained from various sources and analyses, anchored on reference genome sequences	Bolser et al. (2016)
8.	Plant Stress Gene Database	Information about the genes involved in stress conditions in plants	Prabha et al. (2011)
9.	Stress-responsive transcription factor database (STIFDB)	A comprehensive collection of biotic and abiotic stress-responsive genes in *Arabidopsis thaliana* and *Oryza sativa* L. with options to identify probable transcription factor binding sites in their promoters	Shameer et al. (2009)
10.	An Updated Version of Plant Stress-Responsive Transcription Factor database (STIFDB2)	Additional stress signals, stress-responsive transcription factor binding sites and stress-responsive genes in *Arabidopsis* and Rice	Naika et al. (2013)
11.	STIF	Recognition of binding sites of stress-upregulated transcription factors and genes in *Arabidopsis*	Ambika et al. (2008)
12.	Plant Stress-Responsive Gene catalogue (PSRGC)	Database of relationship stress-responsive genes for drought and water with orthologous and paralogous relationships	Wanchana et al. (2018)
13.	PESTD	Database for transcripts with annotated tentative orthologs from crop abiotic stress transcripts	Jayashree et al. (2006)
14.	*Arabidopsis* Stress-Responsive Gene Database	ASRGD database for stress-responsive genes in *Arabidopsis thaliana*	Borkotoky et al. (2013)

(continued)

Table 3.1 (continued)

S. No.	Genomics database	Information	References
15.	Pathogen Receptor Genes Database (PRGDB)	Open and updated space about Pathogen Receptor Genes (PRGs), in which all information available about these genes is stored, curated and discussed	Osuna-Cruz et al. (2018)
16.	Rice SRTFDB	The database which provides information on rice transcription factors drought, salt stress conditions and various developmental stages	Priya and Jain (2013)
17.	QlicRice	A web interface for abiotic stress-responsive QTL and loci interaction channel in rice	Smita et al. (2011)
18.	Drought stress gene database (DroughtDB)	Database for a number of drought stress-associated genes encoding transcription factors that in turn control other various genes involved in diverse physiological and molecular reactions to drought stress	Alter et al. (2015)
19.	MIPS *Oryza sativa* database (MOsDB)	Resource for publicly available sequences of the rice (*Oryza sativa* L.) genome	Wojciech et al. (2003)
20.	MSU rice Genome Annotation Project	Database and Resource is a National Science Foundation project and provides sequence and annotation data for the rice genome	Kawahara et al. (2013)
21.	Plant Genome and Systems Biology (PGSB)	Focuses on the analysis of plant genomes, using bioinformatic techniques, provides a platform for integrative and comparative plant genome research	Spannagl et al. (2016)

259 stress-related genes of 11 species along with all the available information about the individual genes. Stress-related ESTs were also found for *Phaseolus vulgaris*. The database also includes ortholog and paralog of proteins which are coded by stress-related genes.

- *Stress-Responsive Transcription Factor Database (STIFDB V2.0):* It is a comprehensive collection of biotic and abiotic stress-responsive genes in *Arabidopsis thaliana* and *Oryza sativa* L. with options to identify probable transcription factor binding sites in their promoters. In response to biotic stress like bacteria and abiotic stresses like ABA, drought, cold, salinity, dehydration, UV-B, high light, heat, heavy metals, etc., ten specific families of transcription factors in *Arabidopsis thaliana* and six in *Oryza sativa* L. are known to be involved (Shameer et al. 2009).
- *Stress-Responsive Transcription Factor Database (STIFDB2):* Currently it has 38,798 associations of stress signals, stress-responsive genes and transcription factor binding sites predicted using the Stress-responsive Transcription Factor (STIF) algorithm, along with various functional annotation data. As a unique plant stress regulatory genomics data platform, STIFDB2 can be utilized for targeted as well as high-throughput experimental and computational studies to unravel principles of the stress regulome in dicots and gramineae (Naika et al. 2013).

- *STIF (Hidden Markov Model-Based Search Algorithm):* It is used for the recognition of binding sites of stress-upregulated transcription factors and genes in *Arabidopsis* (Ambika et al. 2008).
- *PESTD:* A comparative genomics study on plant responses to abiotic stresses and is a dataset of orthologous sequences. A large amount of sequence information, including those derived from stress cDNA libraries, are used for the identification of stress-related genes and orthologs associated with the stress response. Availability of annotated plant abiotic stress ortholog sets will be a valuable resource for researchers studying the biology of environmental stresses in plant systems, molecular evolution and genomics (Jayashree et al. 2006).
- *Arabidopsis Stress-Responsive Gene Database (ASRGD):* It is a powerful mean for manipulation, comparison, search and retrieval of records describing the nature of various stress-responsive genes in *Arabidopsis thaliana*. About 44 types of different stress factors are related to *Arabidopsis thaliana*, and the database contains 636 gene entries related to stress response with their related information like gene ID, nucleotide and protein sequences and cross-response. The database is based exclusively on published stress-responsive and stress-tolerant genes associated with plants (Borkotoky et al. 2013).
- *The Arabidopsis Information Resource (TAIR):* It contains genetic and molecular biology data for the *Arabidopsis thaliana*, which is more widespread to different aspects apart from the stress response, which makes it difficult to look for only stress-related genes (Swarbreck et al. 2008).
- *Pathogen Receptor Genes Database (PRGDB):* It allows easy access not only to the plant science research community but also to breeders who want to improve plant disease resistance. It offers 153 reference resistance genes and 177,072 annotated candidate pathogen receptor genes (PRGs). Plant diseases display useful information linked to genes and genomes to connect complementary data and better address specific needs. Through a revised and enlarged collection of data, the development of new tools and a renewed portal, PRGdb 3.0 engages the plant science community in developing a consensus plan to improve knowledge and strategies to fight diseases that afflict main crops and other plants (Osuna-Cruz et al. 2018).
- *Rice SRTFDB:* It provides comprehensive expression information on rice transcription factors (TFs) during drought and salinity stress conditions and various stages of development. It will be useful to identify the target TF(s) involved in stress response at a particular stage of development. It also provides curated information for *cis*-regulatory elements present in their promoters, which will be important to study the binding proteins. This database aims to accelerate functional genomics research of rice TFs and understand the regulatory mechanisms underlying abiotic stress responses (Priya and Jain 2013).
- *QlicRice:* This database is designed to host publicly accessible, abiotic stress-responsive quantitative trait loci (QTLs) in rice (*Oryza sativa*) and their corresponding sequenced gene loci. It provides a platform for the data mining of abiotic stress-responsive QTLs, as well as browsing and annotating associated traits, their location on a sequenced genome, mapped expressed sequence tags

(ESTs) and tissue- and growth stage-specific expressions on the whole genome. An appropriate and spontaneous user interface has been designed to retrieve associations to agronomically important QTLs on abiotic stress response in rice (Smita et al. 2011).

- *Drought Stress Gene Database (DroughtDB):* It is a manually curated compilation of molecularly characterized genes that are involved in drought stress response. It includes information about the originally identified gene, its physiological and/or molecular function and mutant phenotypes and provides detailed information about computed orthologous genes in nine model and crop plant species. Thus, DroughtDB is a valuable resource and information tool for researchers working on drought stress and will facilitate the identification, analysis and characterization of genes involved in drought stress tolerance in agriculturally important crop plants (Alter et al. 2015).

3.2.2 Platforms and Resources in the Transcriptome of Plants Under Abiotic Stress/Plant Transcriptomics-Related Computational Tools and Databases

Transcriptome (RNA sequencing or expression profile of an organism) is highly dynamic and involves capturing of the RNA expression profile in spatial and temporal plant organs, tissues and cells within particular conditions (Duque et al. 2013; El-Metwally et al. 2014a). In response to various abiotic stresses, the plant constantly adjusts their transcriptome profile. Thus, transcriptomics study assists in finding genes that are associated with alterations in the plant phenotype under different abiotic or biotic stress conditions (Kawahara et al. 2013). This comprehensive and high-throughput RNAseq analysis finds its applications in plant stress response and tolerance such as searching for abiotic stress candidate genes, predicting tentative gene functions, discovering *cis*-regulatory motifs and providing a better understanding of the plant-pathogen relationship (De Cremer et al. 2013; Agarwal et al. 2014). The recent boom in the availability of online resources, databases and archives of transcriptome data allows for performing novel genome-wide analysis of plant stress responses and tolerance (Duque et al. 2013). Several studies on the transcriptome of different organs and developmental stages of plants under different environmental conditions were observed (Narsai et al. 2010; Zhou et al. 2008). Narsai et al. (2010) identified an exclusively new set of reference genes in rice that are of immense significance, and analysis of their promoter sequence shows the prevalence of some stress regulatory *cis*-element (Zhou et al. 2008).

Different techniques exist to analyse transcriptomic changes in a system under different stress conditions and these are as follows:

- RNA/gene expression profiling is mostly accomplished using microarray, RNA sequencing (RNAseq) through next-generation sequencing (NGS), serial analysis

of gene expression (SAGE) and digital gene expression profiling (Kawahara et al. 2013; Duque et al. 2013; De Cremer et al. 2013).

- Hybridization-based method, such as that used in microarrays and GeneChips, has been well established for acquiring large-scale gene expression profiles for various species (De Cremer et al. 2013).
- Next-generation DNA sequencing application, deep sequencing of short fragments of expressed RNAs, including sRNAs, is quickly becoming an effective tool for use with genome-sequenced species (Harbers and Carninci 2005).
- Quantitative PCR analyses up to a few genes at a time, while microarray analysis allows the simultaneous measurement of transcript abundance for thousands of genes (Joshi et al. 2012).
- Tiling arrays cover the genome at regular intervals to measure transcription without bias towards known or predicted gene structures, the discovery of polymorphisms, analysis of alternative splicing and identification of transcription factor binding sites (Coman et al. 2013). Transcriptome analysis in *Arabidopsis* under abiotic stress conditions using a whole-genome tiling array resulted in the discovery of antisense transcripts induced by abiotic stresses (Matsui et al. 2008).

In the post-genomic era, RNA-Seq provides a global transcriptome profile, which could cover lncRNAs, coding genes and their alternatively spliced isoforms in stress response, and aids plant biologists to expand new insights into molecular mechanisms and responses to biotic and abiotic stress events. Several data portals contain a vast amount of plant RNA-Seq data, such as the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) and the Sequence Read Archive (SRA). However, these data portals mainly serve as raw biological data archives. Large-scale stress-specific RNA-Seq database that can provide comprehensively visualized transcriptome expression profiles and statistical analysis for differential expression has been listed in Table 3.2. Some of these databases are as follows:

- *Plant Stress RNA-Seq Nexus (PSRN):* It is a comprehensive database which includes 12 plant species, 26 plant stress RNA-Seq datasets and 937 samples. PSRN is an open resource for intuitive data exploration, providing expression profiles of coding-transcript/lncRNA and identifying which transcripts are differentially expressed between different stress-specific subsets, in order to support researchers generating new biological insights and hypotheses in molecular breeding or evolution. PSRN was developed with the goal of collecting, processing, analysing and visualizing publicly available plant RNA-Seq data (Li et al. 2018).
- *PlantExpress:* It is a web database as a platform for gene expression network (GEN) analysis with the public microarray data of rice and *Arabidopsis*. PlantExpress has two functional modes: single-species mode is specialized for GEN analysis within one of the species, while the cross-species mode is optimized for comparative GEN analysis between the species. It stores data obtained from three microarrays, namely, the Affymetrix Rice Genome Array, the Agilent

Table 3.2 Transcriptomic repositories and stress-related databases

S. No.	Transcriptomic database	Information	References
1.	Plant Stress RNA-Seq Nexus (PSRN)	Stress-specific transcriptome database in plant cells	Li et al. (2018)
2.	TENOR (Transcriptome Encyclopedia Of Rice)	Database for comprehensive mRNA-Seq experiments in rice	Kawahara et al. (2016)
3.	PlantExpress	A database integrating OryzaExpress and ArthaExpress for single-species and cross-species gene expression network analyses with microarray-based transcriptome data	Kudo et al. (2017)
4.	RiceArrayNet	A database for correlating gene expression from transcriptome profiling, and its application to the analysis of co-expressed genes in rice	Lee et al. (2009)
5.	PLEXdb (Plant Expression Database)	Unified gene expression resource for plants and plant pathogens	Dash et al. (2012)
6.	EGENES	Transcriptome-based plant database of genes with metabolic pathway information and expressed sequence tag indices in KEGG, provides gene indices for each genome	Masoudi-Nejad et al. (2007)
7.	MOROKOSHI	Transcriptome database in *Sorghum bicolor*	Makita et al. (2015)
8.	AgBase	Genome-wide structural and functional annotation and modelling of microarray and other functional genomics data in agricultural species	McCarthy et al. (2006)
9.	Tiling arrays	Measure transcription without bias towards known or predicted gene structures, the discovery of polymorphisms, analysis of alternative splicing and identification of transcription factor binding sites	Coman et al. (2013)
10.	Chickpea Transcriptome Database (CTDB)	Comprehensive information about the chickpea transcriptome. The database contains various information and tools for transcriptome sequence, functional annotation, conserved domain(s), transcription factor families, molecular markers (microsatellites and single nucleotide polymorphisms)	Verma et al. (2015)
11.	TodoFirGene	Omics information of gymnosperms and connect researchers from forest sciences with those in comparative bioinformatics and evolutionary sciences	Ueno et al. (2018)
12.	ROSAcyc	Resource pathway database that allows access to the putative genes and enzymatic pathways, provides useful information on *Rosa*-expressed genes, with thorough annotation and an overview of expression patterns for transcripts with good accuracy	Dubois et al. (2012)

Table 3.2 (continued)

S. No.	Transcriptomic database	Information	References
13.	KEGG PATHWAY Database (Univ. of Kyoto)	Database resource for understanding high-level functions and utilities of the biological system, such as the cell, the organism and the ecosystem, from molecular-level information, especially large-scale molecular datasets generated by genome sequencing and other high-throughput experimental technologies	Kanehisa and Goto (2000)

Rice Gene Expression 4x44K Microarray and the Affymetrix *Arabidopsis* ATH1 Genome Array, with respective totals of 2,678, 1,206 and 10,940 samples. PlantExpress will facilitate understanding of the biological functions of plant genes (Kudo et al. 2017).

- *RiceArrayNet (RAN):* It provides information on co-expression between genes in terms of correlation coefficients (r values). A correlation pattern between Os01g0968800, a drought-responsive element-binding transcription factor; Os02g0790500, a trehalose-6-phosphate synthase; and Os06g0219500, a small heat shock factor, reflecting the fact that genes responding to the same biological stresses is regulated together (Lee et al. 2009).
- *Transcriptome Encyclopedia of Rice (TENOR):* It is a database that encompasses large-scale mRNA sequencing (mRNA-Seq) data obtained from rice under a wide variety of stress conditions. Since the elucidation of the ability of plants to adapt to various growing conditions is a key issue in plant sciences, it is of great interest to understand the regulatory networks of genes responsible for environmental changes. All the resources (novel genes identified from mRNA-Seq data, expression profiles, co-expressed genes and *cis*-regulatory elements) are available in TENOR (Kawahara et al. 2016).

3.2.3   Platforms and Resources in Proteomic of Plants Under Abiotic Stress/Plant Proteomics-Related Computational Tools and Databases

'Proteome' referred to the total expressed protein under certain circumstances in a given organism, organ, cell, tissue or microorganism population, and it comprises all the techniques used in profiling the expressed proteins in a specific context (Tyers and Mann 2003). Similar to the transcriptome, it is an informative approach used to reveal invaluable information when studying plant stress response and tolerance, either in a whole genome or sample scale (Nakagami et al. 2012). It is used for profiling all the expressed proteins under multiple stress conditions and cross-comparing these different sets to identify the proteins which are specifically involved in stress tolerance (Yan et al. 2014). This is an evolving technology for the qualitative large-scale identification and quantification of all protein types in a cell or tissue,

analysis of post-translational modifications and association with other proteins, and characterization of protein activities and structures (Jorrín-Novo et al. 2009).

Proteomics is associated with two types of studies: proteome characterization (identification of all the proteins expressed) and differential proteomics (comparative proteome analysis of control and stressed plants). The proteomic approach has been largely adopted to explore the protein profiles in plants in response to abiotic stress that might lead to the development of new strategies for improving stress tolerance (Helmy et al. 2011). Several types of proteomes can be measured, but whole proteome and the phosphoproteome are the most common proteomes quantified in plant stress tolerance (Helmy et al. 2011, 2012a, b). The main focus of quantitative proteomics is to identify the proteins that are differentially expressed under certain stress response condition (Liu et al. 2015), while phosphoproteomics is closely associated with the identification of proteins activated and functioning in response to particular stress (Zhang et al. 2014). Both whole proteomics and phosphoproteomics can be combined in one comprehensive study to provide a better understanding of the stress (Hopff et al. 2013). The main goal of functional proteomics is the high-throughput identification of all proteins that appeared in cells and/or tissues, but recent rapid technical advances in proteomics have enabled us to progress to the second generation of functional proteomics, including quantitative proteomics, subcellular proteomics and various modifications and protein-protein interactions (Jorrín-Novo et al. 2009).

Two main techniques that are mostly used for quantitative and/or qualitative profiling are protein electrophoresis and protein identification with mass spectrometry. The technology of choice for proteomics is mass spectrometry (MS) including several approaches such as liquid chromatography-mass spectrometry (LC-MS/MS), ion trap-mass spectrometry (IT-MS) and matrix-assisted laser desorption/ionization-mass spectrometry (MALDI-MS) (Helmy et al. 2011, 2012a). These technologies are basically used in measuring the mass and charge of small protein fragments (or 'peptides') that result from protein enzymatic digestion (Helmy et al. 2011; Nakagami et al. 2012). Furthermore, several proteomics labs use protein electrophoresis technologies such as two-dimensional electrophoresis and difference gel electrophoresis (DIGE) in plant proteomics (Duque et al. 2013).

As genome sequencing projects for several organisms have been completed, proteome analysis, which is the detailed investigation of the functions, functional networks and 3D structures of proteins, has gained accumulative consideration. Large-scale proteome datasets available serve as an imperative resource for a better understanding of protein functions in cellular systems, which are controlled by the dynamic properties of proteins (Table 3.3). These properties reflect cell and organ states in terms of growth, development and response to environmental changes. Functional and experimental validation of proteins associated with biotic and abiotic stresses has been employed as the sole criterion for inclusion in the database (Singh et al. 2015). Due to the challenges faced in text/data mining, there is a large gap between the data available to researchers and the hundreds of published plant stress proteomics articles. There are a large number of stress-related databases for proteins (Table 3.3):

Table 3.3 Proteomic databases and resources

S. No.	Proteomic database	Information	References
1.	Protein-Protein Interaction Inhibition Database (2PI2db)	A database containing the structures of protein-protein and protein-modulator complexes that have been characterized by X-ray crystallography or NMR	Basse et al. (2016)
2.	PPDB is a Plant Proteome DataBase	National science foundation-funded project to determine the biological function of each protein in plants. It includes data for two plants that are widely studied in molecular biology, *Arabidopsis thaliana* and maize (*Zea mays*)	Sun et al. (2009)
3.	Plant stress proteome database (PlantPReS)	Enables researchers to perform multiple analyses on the database; the results of each query indicate a series of proteins for which a set of selected criteria are met	Mousavi et al. (2016)
4.	*Arabidopsis* Nucleolar Protein Database (AtNoPDB)	Information on the plant proteins identified to date with a comparison to human and yeast proteins and images of cellular localizations for over a third of the proteins	Brown et al. (2005)
5.	Plant Protein Phosphorylation Database (P³DB)	Integrated database for plant protein phosphorylation can help identify functionally conserved phosphorylation sites in plants using a multi-system approach	Gao et al. (2009)
6.	PhosPhAt (Plant Protein Phosphorylation DataBase)	A database that specifically maintains experimental phosphorylation site data for *Arabidopsis*	Heazlewood et al. (2008) and Durek et al. (2010)
7.	PSPDB (Plant Stress Protein Database)	Current plant stress databases report plant genes without protein annotations specific to these stresses	Singh et al. (2015)
8.	MANET database Molecular Ancestry Network	The database is a bioinformatics database that maps evolutionary relationships of protein architectures directly onto biological networks	Kim et al. (2006)
9.	PhytAMP	Database of plant natural antimicrobial peptides	Hammami et al. (2009)
10.	PRINTS	Compendium of protein fingerprints	Attwood et al. (1994)
11.	PROSITE	Database of protein families and domains	Hulo et al. (2006)
12.	Swiss-Prot	Protein knowledgebase	Bairoch and Apweiler (2000)
13.	Protein DataBank in Europe (PDBe)	European part of the wwPDB for the collection, organization and dissemination of data on biological macromolecular structures	Mir et al. (2018)

(continued)

Table 3.3 (continued)

S. No.	Proteomic database	Information	References
14.	ProteinDatabank in Japan (PDBj)	Maintains a centralized PDB archive of macromolecular structures and provides integrated tools	Kinjo et al. (2017)
15.	Research Collaboratory for Structural Bioinformatics (RCSB)	Protein Data Bank archive information about the 3D shapes of proteins, nucleic acids and complex assemblies that helps students and researchers understand all aspects of biomedicine and agriculture	Rose et al. (2017)
16.	Structural Classification of Proteins (SCOP)	The database is a largely manual classification of protein structural domains based on similarities of their structures and amino acid sequences	Murzin et al. (1995)
17.	InterPro	Classifies proteins into families and predicts the presence of domains and sites	Hunter et al. (2009)
18.	Pfam	Protein families database of alignments and HMMs	Finn et al. (2013)

- *Plant Stress Proteome Database (*PlantPReS*; www.proteome.ir):* It is an open online proteomic database, which currently comprises >35,086 entries from 577 manually curated articles and contains >10,600 unique stress-responsive proteins (Mousavi et al. 2016).
- *Plant Stress Protein Database (PSPDB):* It is one of the largest repositories and a web-accessible resource that covers 2064 manually curated plant stress proteins from a wide array of 134 plant species with 30 different types of biotic and abiotic stresses. Functional and experimental validation of proteins associated with biotic and abiotic stresses has been employed as the sole criterion for inclusion in the database (Singh et al. 2015).

'Proteogenomics' is another comprehensive combo approach of large-scale proteomic data with genomic and/or transcriptomics data to elucidate various innovative regulatory mechanisms (Helmy et al. 2012a). The proteomics data generated by means of MS-based proteomics (high throughput and accuracy) provides a rich source of translation-level information about the expressed proteins that can be used as a source of large-scale experimental evidence for several predictions (Helmy et al. 2012a, b). In a proteogenomics study, the naturally expressed proteins are identified using MS-based proteogenomics followed by mapping them back to the genomic or transcriptomic data (Helmy et al. 2012a). This field has facilitated in elevating our understanding of the biology of plants in general as well as plant stress research in particular. For instance, a large-scale proteogenomics study of *Arabidopsis thaliana* identified 57 new genes and corrected the annotations of hundreds of its genes using intensive sampling from the *Arabidopsis* organs under several conditions (Baerenfaller et al. 2008). Another study reported corrections and new identifications in about 13% of the annotated genes in *Arabidopsis* (Castellana et al. 2008). It also gives information on the

investigation of the host-pathogen relationship (Delmotte et al. 2009), identifying novel effectors in fungal diseases (Cooke et al. 2014), as well as shedding light on the mechanisms of environmental adaptation.

3.2.4   Platforms and Resources in Metabolomics of Plants Under Abiotic Stress/Plant Metabolomics-Related Computational Tools and Databases

The metabolome is the complete pool of metabolites in a cell at any given time and metabolomics refers to techniques and methods used to study the metabolome (Duque et al. 2013). Plants are able to synthesize a diverse group of chemical and biological compounds with different biological activity that is crucial for regulating the response to different types of biotic and abiotic stress (Bino et al. 2004). Therefore, identifying the metabolites produced by the plant under each stress condition by metabolomics plays a significant role to gather information not only about the phenotype but changes in it induced by stress, thereby bridging the gap between phenotype and genotype (Badjakov et al. 2012). Metabolomics may prove to be particularly important in plants due to its ability to elucidate plant cellular systems and permits engineering molecular breeding to improve the growth and productivity of plants in stress tolerance (Fernie and Schauer 2009). Metabolomic approaches allow us to conduct parallel assessments of multiple metabolites, and it is notable that the plant metabolome represents an enormous chemical diversity due to the complex set of metabolites produced in each plant species (Bino et al. 2004). A strong connection between stress metabolites and a particular protein indicates the role of this gene in the stress response process (Urano et al. 2010; Duque et al. 2013; Jogaiah et al. 2013). Metabolic profiling of plants involves a combo of several analytical, separation techniques and with other omics analysis (e.g. transcriptomics or proteomics) to investigate the correlation between metabolite levels and the expression level of genes/proteins (Jogaiah et al. 2013). Thus, metabolomics provides a better understanding of the stress response and tolerance process in model plants such as *Arabidopsis* (Cook et al. 2004) as well as in crops like a common bean (*Phaseolus vulgaris*) (Broughton et al. 2003), and other food crops (Duque et al. 2013).

   This is one of the most rapidly developing technologies, and many notable technological advances have recently been made in instrumentation related to metabolomics; some of them are as follows:

- Major approaches that are used in plant metabolomics research include metabolic fingerprinting which involves separation of metabolites based on the physical and chemical properties using various analytical tools and technologies (Jogaiah et al. 2013).
- Metabolite profiling which includes the study of the alterations in metabolite pool that are induced by stress and finally target analyses.

- Capillary electrophoresis-liquid chromatography-mass spectrometry (CE-MS) is considered the most advanced metabolomics technology (Soga et al. 2002).
- Analytical instruments and separation technologies are employed in metabolomics such as gas chromatography (GC), mass spectrometry (MS) and nuclear magnetic resonance (NMR) (Duque et al. 2013).
- Metabolomics experiment (MIAMET) gives reporting requirements with the aim of standardizing experiment descriptions, particularly within publications (Ernst et al. 2014).
- Standard Metabolic Reporting Structures (SMRS) working group has developed standards for describing the biological sample origin, analytical technologies and methods used in a metabolite profiling experiment (Chen et al. 2015).
- ArMet (architecture for metabolomics) proposal gives a description of plant metabolomics experiments and their results along with a database schema (Castillo-Peinado and de Castro 2016).
- Metabolic flux analysis measures the steady-state flow between metabolites. FluxAnalyzer is a package for MATLAB that integrates pathway and flux analysis for metabolic networks (Rocha et al. 2008).

A number of studies of metabolic profiling in plant species have been performed that have resulted in the publication of related databases (Table 3.4). For instance, metabolic pathways that act in response to environmental stresses in plants were investigated by metabolome analysis using various types of MS coupled with microarray analysis of overexpressors of genes encoding two TFs, DREB1A/CBF3 and DREB2A (Maruyama et al. 2009). Metabolomic profiling was also used to investigate chemical phenotypic changes between wild-type *Arabidopsis* and a knockout mutant of the *NCED3* gene under dehydration stress conditions (Urano et al. 2010). These databases are vast information resources and repositories of large-scale datasets and also serve as tools for further integration of metabolic profiles containing comprehensive data acquired from other omics research (Akiyama et al. 2008). One of the huge databases for metabolites is PlantMetabolomics.org (PM), which is a web portal and database for exploring, visualizing and downloading plant metabolomics data. Widespread public access to well-annotated metabolomics datasets (Table 3.4) is essential for establishing metabolomics as a functional genomics tool. PM can be used as a platform for deriving hypotheses by enabling metabolomic comparisons between genetically unique *Arabidopsis* (*Arabidopsis thaliana*) populations subjected to different environmental conditions (Bais et al. 2015).

3.2.5 Micro RNAs: Attributes in Plant Abiotic Stress Responses and Bioinformatics Approaches on MicroRNA

MicroRNA (miR) represents a major subfamily of endogenously transcribed sequences (21–24 bp) and has been acknowledged as a major regulatory class that inhibits gene expression in a sequence-dependent manner (Eldem et al. 2013). miRs

Table 3.4 Metabolomic databases and resources

S. No.	Metabolomics database	Information	References
1.	Metabolome Tomato Database (MoTo DB)	LC-MS-based metabolome database	Moco et al. (2006)
2.	KOMICS (Kazusa-omics) database	Annotations of metabolite peaks detected by LC-FT-ICR-MS and containing a representative metabolome dataset for the tomato cultivar (Micro-tom)	Iijima et al. (2008)
3.	Golm Metabolome Database (GMD)	Provides public access to custom mass spectra libraries and metabolite profiling experiments as well as to additional information and related tools	Kopka et al. (2005)
4.	MS/MS spectral tag (MS2T) libraries at the Platform for Riken Metabolomics (PRIMe) website	Provides access to libraries of phytochemical LC-MS2 spectra obtained from various plant species by using an automatic MS2 acquisition function of LC-ESI-Q-TOF/MS	Matsuda et al. (2009)
5.	Armec Repository Project	Metabolome data on the potato and serves as a data repository for metabolite peaks detected by ESI-MS	Gomez-Casati et al. (2013)
6.	Plant metabolomics	The NSF-funded multi-institutional project aimed at the development of the *Arabidopsis* metabolomics database	Bais et al. (2015)
7.	Minimum information about a metabolomics experiment (MIAMET)	Reporting requirements with the aim of standardizing experiment descriptions, particularly within publications	Ernst et al. (2014)
8.	Standard Metabolic Reporting Structures (SMRS)	Standards for describing the biological sample origin, analytical technologies and methods used in a metabolite profiling experiment	Chen et al. (2015)
9.	ArMet (architecture for metabolomics)	The proposal gives a description of plant metabolomics experiments and their results along with a database schema	Castillo-Peinado and de Castro (2016)
10.	MATLAB (Metabolic flux analysis) FluxAnalyzer	Measures the steady-state flow between metabolites. It is a package that integrates pathway and flux analysis for metabolic networks	Rocha et al. (2008)
11.	BiGG Models	Knowledge base of genome-scale metabolic network reconstructions, integrating more than 70 published genome-scale metabolic networks into a single database with a set of standardized identifiers called BiGG IDs	King et al. (2016)
12.	BioCyc Database Collection	Microbial genome web portal that combines thousands of genomes with additional information inferred by computer programmes, imported from other databases, and curated from the biomedical literature by biologist curators	Karp et al. (2017)

(continued)

Table 3.4 (continued)

S. No.	Metabolomics database	Information	References
13.	BRENDA (The Comprehensive Enzyme Information System)	Comprehensive enzyme information system, including FRENDA, AMENDA, DRENDA, and KENDA	Scheer et al. (2011)
14.	WikiPathways	Community resource for contributing and maintaining content dedicated to biological pathways	Pico et al. (2008)
15.	Metabolomics at Rothamsted (MeT-RO)	Large- and small-scale metabolomic analyses of any plant or microbial material	Rothamsted Ltd

are small regulators of gene expression in the numerous developmental and signalling pathways and are emerging as important post-transcriptional regulators that may regulate key plant genes responsible for stress tolerance. Plants combat environmental stresses by activating several gene regulatory pathways and studies with different model plants have revealed the role of these miRNAs in response to abiotic stress (Zhou et al. 2010). Plant exposed to abiotic stress causes over- or underexpression of certain miRNA and might even lead to the synthesis of new miRNAs to withstand stress (Khraiwesh et al. 2012). Several studies identified species- and clades-specific miRNA families associated with plant stress-regulated genes (Zhang et al. 2013). The functions of stress-responsive miRNAs can only be studied by understanding the regulatory interaction within the network (Jeong and Green 2013). Identification of a huge number of stress-responsive miRNAs might be helpful in developing new strategies to withstand stress, thereby improving the stress tolerance in plant. With the drastic improvement in genomic tools and methods, novel miRNAs in various plant species involved in abiotic stress response are increasing and are providing us with a better understanding of miRNAs-mediated gene regulation (Wang et al. 2014).

Sequence-based profiling along with computational analysis has played a key role in the identification of stress-responsive miRs. sRNA blot and RT-PCR analysis have played an equally important part in systematically confirming the profiling data (Jagadeeswaran et al. 2010). This has also enabled quantification of their effect on the genetic networks, such that many of the stress-regulated miRs have emerged as potential candidates for improving plant performance under stress. The development and integration of plant computational biology tools and approaches have added new functionalities and perspectives in the miR biology to make them relevant for genetic engineering programmes for enhancing abiotic stress tolerance. So far, three major strategies have been employed for the identification and expression profiling of stress-induced miRs:

- The first approach involves the classical experimental route that included direct cloning, genetic screening or expression profiling.
- The second method involved computational predictions from genomic or EST loci.
- The latest one employed a combo of both as it was based on the prediction of miRs from high-throughput sequencing (HTS) data.

Each of these was followed by experimental validations by northern analysis, PCRs or microarrays. In recent years, high-throughput sequencing and screening protocols have caused an exponential increase in a number of miRs, identified and functionally annotated from various plant species (Jagadeeswaran et al. 2010). The first biological database generated for miR was miRBase, which acts as an archive of miR sequences and annotations (Griffiths-Jones et al. 2008). With the future advancement of genomic tools and methods to identify novel miRNAs in various plant species, the number of miRNAs involved in abiotic stress response is increasing, thus providing us with a better understanding of miRNAs-mediated gene regulation during various abiotic stresses (Table 3.5). Some of the databases comprising miRNAs-related information are:

- *PASmiR:* This database is a complete repository for miRNA regulatory mechanisms involved in plant response to abiotic stresses for the plant stress physiology community. It is a literature-curated and web-accessible database and was developed to provide detailed, searchable descriptions of miRNA molecular regulation in different plant abiotic stresses. It currently includes data from ~200 published studies, representing 1038 regulatory relationships between 682 miRNAs and 35 abiotic stresses in 33 plant species (Zhang et al. 2013).
- *PmiRExAt:* It is a new online database resource that caters plant miRNA expression atlas. The web-based repository comprises of miRNA expression profile and query tool for 1859 wheat, 2330 rice and 283 maize miRNA. The database interface offers open and easy access to miRNA expression profile and helps in identifying tissue preferential, differential and constitutively expressing miRNAs (Table 3.5).
- *Plant MicroRNA Database (PMRD):* miRNA expression profiles are provided in this database, including rice oxidative stress-related microarray data and the published microarray data for poplar, *Arabidopsis*, tomato, maize and rice. The plant miRNA database integrates available plant miRNA data deposited in public databases, gathered from the recent literature, and data generated in-house (Zhang et al. 2010).
- *WMP:* It is a novel resource that provides data related to the expression of abiotic stress-responsive miRNAs in wheat. This database allows the query of small RNA libraries, including in silico predicted wheat miRNA sequences and the expression profiles of small RNAs identified from those libraries (Remita et al. 2016).

3.3   Role of Bioinformatics in Plant Disease Management

Omics studies focused on whole-genome analysis have unlocked a new era for biology in general and for agriculture in particular. Combination of bioinformatics and functional genomics globally has paved way towards a better understanding of plant-pathogen biological interaction which eventually leads to breaking thoughts in the promotion of plant resistance to pests (Koltai and Volpin 2003). Bioinformatics

Table 3.5 Major microRNA repositories and stress databases

S. No.	Databases	Information	References
1.	miRcheck	The tool requires the input of putative hairpin sequences and their secondary structures; the algorithm was first used for identifying conserved miRs in *Arabidopsis* and rice	Jones-Rhoades and Bartel (2004)
2.	PmiRExAt	Plant miRNA expression atlas database and web applications	Gurjar et al. (2016)
3.	UEA sRNA Workbench	It predicts miR from HTS data, trans-acting RNA prediction, the secondary structure of RNA sequences, expression patterns of sRNA loci, alignment of short reads to the genome	Moxon et al. (2008)
4.	TAPIR	Characterizes miR-target duplexes with large loops which are usually not detectable by traditional target prediction tools, fast and canonical FASTA local alignment programme and RNAhybrid for detection of miR-mRNA duplexes	Pearson (2004), Krüger and Rehmsmeier (2006), and Bonnet et al. (2010)
5.	CLCGenomics Workbench	Calculating abundances of sRNA libraries. Workbench provides an interactive visualization to the differential expression and statistical analysis of RNA-Seq and sRNA data	Matvienko et al. (2013)
6.	C-mii	Tool aligns known miRs from different plant species to the EST sequences of the query plant species using blast homology search and has a unique feature of predicting the secondary structures of the miR-target duplexes	Numnark et al. (2012)
7.	miRDeep-P	Specialized tool for identification of plant miR collection of PERL scripts that are used for prediction of novel miRs from deep sequencing data	Friedländer et al. (2008) and Yang and Li (2011)
8.	CleaveLand	The general pipeline, available as a combination of PERL scripts, for detecting miR-cleaved target transcripts from degradome datasets	Addo-Quaye et al. (2008)
9.	ARMOUR	Datasets of rice miRs from various deep sequencing datasets for examining the expression changes with respect to their targets, a valuable tool to biologists for selecting miRs for further functional studies	Tripathi et al. (2015)
10.	PASmiR	Database for the role of miRNA in response to plant abiotic stress	Zhang et al. (2013)
11.	PMRD – plant microRNA database	Plant miRNA database integrates available plant miRNA data deposited in public databases, gleaned from the recent literature, and data generated in-house	Zhang et al. (2010)
12.	WMP	Novel comprehensive wheat miRNA database, including related bioinformatics software	Remita et al. (2016)

has played a great role in plant disease management by understanding the molecular basis of the host-pathogen interaction (Koltai and Volpin 2003). Modern genomics tools, including applications of bioinformatics and functional genomics, allow scientists to interpret DNA sequence data and test hypotheses on a larger scale than previously possible (Anonymous 2005). From past few years, numerous components of the plant signalling system have also been identified that function downstream of the detection molecules such as the pathogen proteins that are used to suppress host defences and drive the infection process (so-called effector proteins) by using molecular biological technologies and genetics approaches (Anonymous 2005). Disease resistance is only one of the several traits under selection in a breeding programme. Thus, bioinformatics has to play an increasing role in integrating phenotypic and pedigree information for agronomic as well as resistance traits (Vassilev et al. 2006). Improved algorithms and increased computing power have made it possible to improve selection strategies as well as to model the epidemiology of pathogens (Michelmore 2003). Some of the key roles of bioinformatics for plant improvements has been enlisted by Vassilev et al. (2005): submitting all sequence data information generated from experimentation into the public domain, through repositories; providing rational annotation of genes, proteins and phenotypes; elaborating relationships both within the plants' data and between plants and other organisms; providing data including information on mutations, markers, maps, functional discoveries; and others.

From past few years, there have been many technological advances in the understanding of plant-pathogen interaction. Omics techniques (genomics, proteomics and transcriptomics) have provided a great opportunity to explore plant-pathogen interactions from a system's perspective and studies on protein-protein interactions (PPIs) between plants and pathogens (Delaunois et al. 2014). Identification of the molecular components as well as the corresponding pathways has provided a relatively clear understanding of the plant immune system. In particular, the study of plant-pathogen interactions has also been stimulated by the emergence of various omics techniques, such as genomics, proteomics and transcriptomics (Schulze et al. 2015). With the availability of massive amounts of data generated from high-throughput omics techniques, network interactions have become a powerful approach to further decipher the molecular mechanisms of plant-pathogen interactions through network biology:

- Genomics is particularly important, and with the rapid development of next-generation sequencing (NGS) technique, numerous plant and pathogen genomes have been fully sequenced.
- Proteomics is a key technique for the analysis of the proteins involved in plant-pathogen interactions (Delaunois et al. 2014).
- DNA microarray and RNA sequencing are two key transcriptomics techniques for acquiring the expression profile of genes on a large scale. Transcriptomics is also important to investigate plant-pathogen interactions and has been employed to learn how plants respond to the pathogen invasion and how pathogens counter the plant defence at the transcript level (Schulze et al. 2015).

Genomic approaches always have a significant impact on efforts to improve plant diseases by increasing the definition of and access to gene pools available for crop improvement (Vassilev et al. 2005). Such an approach to identify key genes and understand their function will result in a quantum leap in plant improvement. Moreover, the ability to examine gene expression will allow us to understand how plants respond to and interact with the physical environment and management practices (Vassilev et al. 2006). This approach will involve the detailed characterization of the many genes that confer resistance, as well as technologies for the precise manipulation and deployment of resistance genes. Plant-pathogen interactions are sophisticated and dynamic in the continually evolving competition between pathogens and plants. Thus genomic studies on pathogens are providing an understanding of the molecular basis of specificity and the opportunity to select targets for more durable resistance (Michelmore 2003). This understanding is fundamental to allow efficient exploitation of plants as biological resources in the development of new cultivars with improved quality and reduced economic, pathogen and abiotic stress resistance and is also vital for the development of new plant diagnostic tools (Vassilev et al. 2006). When plants respond to biotic stress, a series of biological processes rather than a single gene or protein will be changed. Therefore, it is necessary to explore plant-pathogen interactions from a systems perspective (e.g. network level (Mine et al. 2014)). Bioinformatics thus plays several roles in breeding for disease resistance and is important for acquiring and organizing large amounts of information. Some of the databases/repositories (Table 3.6) of plant-pathogen interactions are:

- *PHI-base: A new database for pathogen-host interactions.* It is designed for hosting any type of pathogen-host interaction, and its focus is on genes with functions that have been experimentally verified. These genes are compiled and curated in a way that can be used to bridge the genotype-phenotype gap underlying the interactions between hosts and pathogens (Winnenburg et al. 2006). The mission of PHI-base is to provide expertly curated molecular and biological information on genes proven to affect the outcome of pathogen-host interactions.
- *PathoPlant: A database on plant-pathogen interactions.* This is a relational database to display relevant components and reactions involved in signal transduction related to plant-pathogen interactions. On the organism level, the tables 'plant', 'pathogen' and 'interaction' are used to describe incompatible interactions between plants and pathogens or diseases (Bülow et al. 2004).

3.4   Conclusion and Future Prospects

Bioinformatics is an exclusive approach capable of exploiting and sharing a large amount of omics data. This approach has given a more holistic view of the molecular response in plants when exposed to biotic and abiotic stress, and the integration of various omics studies has revealed a new zone of interactions and regulation (Fig. 3.2). This system biology approach has enabled the identification,

Table 3.6 Abiotic stress databases

S. No	Databases	Information	References
1.	PHI-base (Pathogen-Host Interactions)	First online resource devoted to the identification and presentation of information on fungal and oomycete pathogenicity genes and their host interactions	Winnenburg et al. (2006)
2.	PathoPlant (Plant-Pathogen Interactions)	A new database that combines information of specific plant–pathogen interactions on organism level and data about signal transduction on molecular level related with plant pathogenesis	Bülow et al. (2004)
3.	Fungal Plant Pathogen Database	An internet-based database that crosslinks the digitized genotypic and phenotypic information of individual pathogens at both the species and population levels may allow us to effectively address these problems by coordinating the generation of data and its subsequent archiving	Kang et al. (2002)
4.	Phytopathogenic Fungi and Oomycete EST Database of COGEME	Provides sequences of expressed sequence tags (ESTs) and consequences (cluster assembled ESTs) from 15 plant pathogenic species	Soanes et al. (2002)
5.	fuGIMS database	Integrate functional and sequence information from several plants and animal pathogenic fungi with similar information from *Saccharomyces cerevisiae* available from the GIMS database	Cornell et al. (2003)
6.	DRASTIC Insight	Collates signal transduction information between plants, pathogens, and the environment, including both biotic and abiotic influences on plant disease resistance at the molecular level	Newton et al. (2002)

characterization and functional analysis of plant genes that determine plant's response to various biotic and abiotic factors and understanding of the plant stress interaction. However, so many efforts are still required for detailed analysis of the omic modulation induced by abiotic stress and its interacting partners. This requires the development of reliable and rigorous techniques for firm characterization of the spatiotemporal regulation of omics under stress conditions. The three main domains that must be addressed to take full advantage of systems biology are the development of omics technology, integration of data in a usable format and analysis of data within the domain of bioinformatics. Thus, the perspective of computational/system biology needs to be tapped for performing an extensive analysis among agriculturally important crops for improving crop tolerance to environmental stress. The current surge of affordable omics data encourages researchers to create improved, more integrated and easily accessible plant stress pathway databases. Despite the drawbacks, there is no doubt that bioinformatics is a field that holds great potential for transforming biological research in the coming decades. The expansion and integration of bioinformatics tools and approaches will certainly add new functionalities and perspectives in the stress biology to make them applicable for genetic engineering programmes for enhancing stress tolerance.
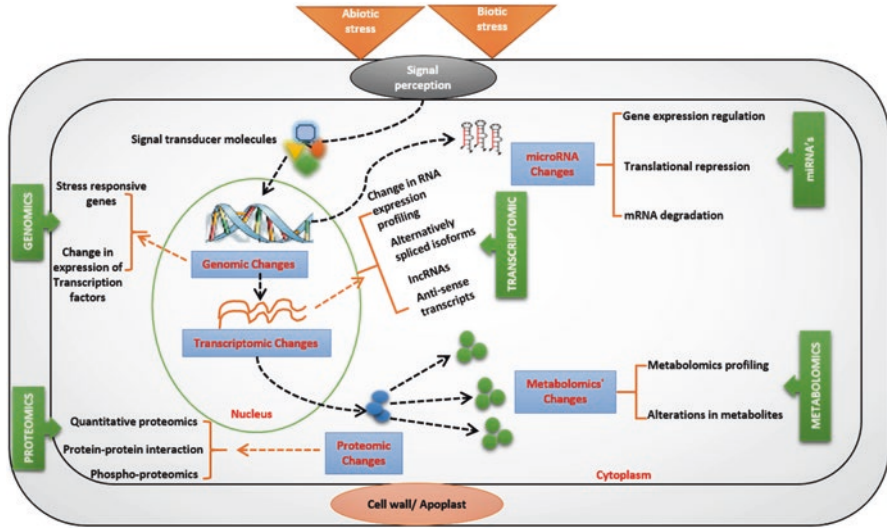
Fig. 3.2 Pathway depicting molecular effects of abiotic and biotic stress at genomic, transcriptomic, metabolomic and microRNA levels inside the plant cell. Both biotic and abiotic stresses have to be first sensed by the plant cell, and then the information is transduced to the appropriate downstream-located pathway(s). Sensors as well as signal transducers might be shared by both types of stressors. After the signal is perceived by the sensors at the cell wall, it transduces the signal towards the nucleus, where modifications occur at genome level, resulting in activation of stress-responsive genes and certain transcription factors (genomics). This activation can occur at the transcript level (transcriptomics), wherein a change in RNA expression levels, alternatively spliced forms of RNA and antisense transcripts occur. Few regulatory small RNAs (microRNA) are also synthesized against a specific type of stress conditions (miRNAs). Alterations, which have been initiated at the cell wall, propagate to the cytosol, in form of proteins which can be identified by quantitative proteomics, protein-protein interaction and phosphor-proteomics (proteomics). After stress signalling, the defence system comes into rescue where large amounts of secondary metabolites are synthesized (metabolomics); these metabolites can be analysed by metabolomic profiling and change in the type of metabolites synthesized

References

Adams MD, Soares MB, Kerlavage AR, Fields C, Venter JC (1993) Rapid cDNA sequencing (expressed sequence tags) from a directionally cloned human infant brain cDNA library. Nat Genet 4:373–380

Addo-Quaye C, Miller W, Axtell MJ (2008) CleaveLand: a pipeline for using degradome data to find cleaved small RNA targets. Bioinformatics 25(1):130–131

Agarwal P, Parida SK, Mahto A et al (2014) Expanding frontiers in plant transcriptomics in aid of functional genomics and molecular breeding. Biotechnol J 9:1480–1492

Akiyama K, Chikayama E, Yuasa H, Shimada Y, Tohge T, Shinozaki K, Hirai MY, Sakurai T, Kikuchi J, Saito K (2008) PRIMe: a web site that assembles tools for metabolomics and transcriptomics. Silicon Biol 8(3):339–345

Alter S, Bader KC, Spannagl M, Wang Y, Bauer E, Schön CC, Mayer KFX (2015) DroughtDB: an expert-curated compilation of plant drought stress genes and their homologs in nine species. J Biol Database Curation 2015:bav046

Ambika S, Susan Mary Varghese SM, Shameer K, Udayakumar M, Sowdhamini R (2008) STIF: Hidden Markov Model-based search algorithm for the recognition of binding sites of Stress-upregulated Transcription Factors and genes in *Arabidopsis thaliana*. Bioinformation 2(10):431–437

Anonymous (2005) Pseudomonas versus Arabidopsis: models for genomic research into plant disease resistance. www.actionbioscience.org

Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408(6814):796

Attwood TK, Beck ME, Bleasby AJ, Parry-Smith DJ (1994) PRINTS–a database of protein motif fingerprints. Nucleic Acids Res 22(17):3590–3596

Badjakov I, Kondakova V, Atanassov A (2012) In: Benkeblia N (ed) Current view on fruit quality in relation to human health in sustainable agriculture and new biotechnologies. CRC Press, Boca Raton, pp 303–319

Baerenfaller K, Grossmann J, Grobei MA et al (2008) Genome-scale proteomics reveals *Arabidopsis thaliana* gene models and proteome dynamics. Science 320:938–941

Bairoch A, Apweiler R (2000) The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. Nucleic Acids Res 28(1):45–48

Bais P, Moon SM, He K, Leitao R, Dreher K, Walk T et al (2015) PlantMetabolomics.org: a web portal for plant metabolomics experiments. Plant Physiol 152(4):1807–1816

Basse M-J, Betzi S, Morelli X, Roche P (2016) 2P2Idb v2: update of a structural database dedicated to orthosteric modulation of protein–protein interactions. Database 2016:baw007

Bino RJ, Hall RD, Fiehn O, Kopka J, Saito K, Draper J, Nikolau BJ, Mendes P, Roessner-Tunali U, Beale MH et al (2004) Potential of metabolomics as a functional genomics tool. Trends Plant Sci 9:418–425

Bolser D, Staines DM, Pritchard E, Kersey P (2016) Ensembl plants: integrating tools for visualizing, mining, and analyzing plant genomics data. In: Edwards D (ed) Plant bioinformatics, Methods in molecular biology, vol 1374. Humana Press, New York, NY

Bonnet E, He Y, Billiau K, Van de Peer Y (2010) TAPIR, a web server for the prediction of plant microRNA targets, including target mimics. Bioinformatics 26:1566–1568

Borkotoky S, Saravanan V, Jaiswal A, et al. (2013) The Arabidopsis stress responsive gene database. Int J Plant Genom 2013:949564

Broughton WJ, Hernández G, Blair M et al (2003) Beans (Phaseolus spp.)—model food legumes. Plant Soil 252:55–128

Brown JW, Shaw PJ, Shaw P, Marshall DF (2005) Arabidopsis nucleolar protein database (AtNoPDB). Nucleic Acids Res 33:D633–D636

Bülow L, Schindler M, Choi C, Hehl R (2004) PathoPlant: a database on plant-pathogen interactions. In Silico Biol 4(4):529–536

Castellana NE, Payne SH, Shen Z et al (2008) Discovery and revision of Arabidopsis genes by proteogenomics. Proc Natl Acad Sci U S A 105:21034–21038

Castillo-Peinado LS, de Castro ML (2016) Present and foreseeable future of metabolomics in forensic analysis. Anal Chim Acta 925:1–5

Chawla K, Barah P, Kuiper M, Bones AM (2011) Systems biology: a promising tool to study abiotic stress responses. Omics Plant Abiotic Stress Tolerance 10:163–172

Chen X, Qi X, Duan LX (2015) Overview. In: Plant metabolomics. Springer, Netherlands, p 1–24

Coman D, Gruissem W, Hennig L (2013) Transcript profiling in Arabidopsis with genome tiling microarrays. In: Tiling arrays: methods and protocols. Humana Press, Totowa, pp 35–49

Cook D, Fowler S, Fiehn O, Thomashow MF (2004) A prominent role for the CBF cold response pathway in configuring the low-temperature metabolome of Arabidopsis. Proc Natl Acad Sci U S A 101:15243–15248

Cooke IR, Jones D, Bowen JK et al (2014) Proteogenomic analysis of the Venturia pirina (Pear Scab Fungus) secretome reveals potential effectors. J Proteome Res 13:3635–3644

Cornell M, Paton NW, Hedeler C, Kirby P, Delneri D, Hayes A, Oliver SG (2003) GIMS: an integrated data storage and analysis environment for genomic and functional data. Yeast 20(15):1291–1306

Cushman JC, Bohnert HJ (2000) Genomic approaches to plant stress tolerance. Curr Opin Plant Biol 3:117–124

Dash S, Van Hemert J, Hong L, Wise RP, Dickerson JA (2012) PLEXdb: gene expression resources for plants and plant pathogens. Nucleic Acids Res 40:D1194–D1201

Dash S, Campbell JD, Cannon EK, Cleary AM, Huang W, Kalberer SR, Karingula V, Rice AG, Singh J, Umale PE, Weeks NT, Wilkey AP, Farmer AD, Cannon SB (2016) Legume information system (LegumeInfo.org): a key component of a set of federated data resources for the legume family. Nucleic Acids Res 44:D1181–D1188

De Cremer K, Mathys J, Vos C et al (2013) RNAseq-based transcriptome analysis of *Lactuca sativa* infected by the fungal necrotroph Botrytis cinerea. Plant Cell Environ 36:1992–2007

Delaunois B, Jeandet P, Clément C, Baillieul F, Dorey S, Cordelier S (2014) Uncovering plant–pathogen crosstalk through apoplastic proteomic studies. Front Plant Sci 5:249

Delmotte N, Knief C, Chaffron S et al (2009) Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. Proc Natl Acad Sci U S A 106:16428–16433

Dong Q, Schlueter SD, Brendel V (2004) PlantGDB, plant genome database and analysis tools. Nucleic Acids Res 32:D354–D359

Dubois A, Carrere S, Raymond O, Pouvreau B, Cottret L, Roccia A, Onesto JP, Sakr S, Atanassova R, Baudino S, Foucher F, Le Bris M, Gouzy J, Bendahmane M (2012) Transcriptome database resource and gene expression atlas for the rose. BMC Genomics 13:638

Duque AS, de Almeida AM, da Silva AB, da Silva JM, et al (2013) Abiotic stress—plant responses and applications in agriculture. InTech, Chapter 3, p 40–101

Durek P, Schmidt R, Heazlewood JL, Jones A, Maclean D, Nagel A, Kersten B, Schulze WX (2010) PhosPhAt: the *Arabidopsis thaliana* phosphorylation site database. An update. Nucleic Acids Res 38:D828–D834

Eldem V, Okay S, Ünver T (2013) Plant microRNAs: new players in functional genomics. Turk J Agric For 37:1–21

El-Metwally S, Ouda OM, Helmy M (2014a) Next generation sequencing technologies and challenges in sequence assembly, 1st edn. Springer, ISBN: 978-1-4939-0714-4

El-Metwally S, Ouda OM, Helmy M (2014b) First- and next-generations sequencing methods. Next gener seq technol Challenges seq assem. Springer, New York, pp 29–36

El-Metwally S, Ouda OM, Helmy M (2014c) New horizons in next-generation sequencing. Next gener seq technol Challenges seq assem. Springer, New York, pp 51–59

Ernst M, Silva DB, Silva RR, Vêncio RZ, Lopes NP (2014) Mass spectrometry in plant metabolomics strategies: from analytical platforms to data acquisition and processing. Nat Prod Rep 31(6):784–806

Esposito A, Colantuono C, Ruggieri V, Chiusano ML (2016) Bioinformatics for agriculture in the next-generation sequencing era. Chem Biol Technol Agric 3(1):9

Fernie AR, Schauer N (2009) Metabolomics-assisted breeding: a viable option for crop improvement? Trends Genet 25(1):39–48

Feuillet C, Leach JE, Rogers J, Schnable PS, Eversole K (2011) Crop genome sequencing: lessons and rationales. Trends Plant Sci 16(2):77–88

Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer ELL, Tate J, Punta M (2013) Pfam: the protein families database. Nucleic Acids Res 42(D1):D222–D230

Friedländer MR, Chen W, Adamidi C, Maaskola J, Einspanier R, Knespel S et al (2008) Discovering microRNAs from deep sequencing data using miRDeep. Nat Biotechnol 26:407–415

Futamura N, Totoki Y, Toyoda A, Igasaki T, Nanjo T, Seki M et al (2008) Characterization of expressed sequence tags from a full-length enriched cDNA library of *Cryptomeria japonica* male strobili. BMC Genomics 9:383

Gao J, Agrawal GK, Thelen JJ, Xu D (2009) Helmy Nucleic Acids Res 37:D960–D962

Ghosh D, Xu J (2014) Abiotic stress responses in plant roots: a proteomics perspective. Front Plant Sci 5:6

Gomez-Casati DF, Zanor MI, Busi MV (2013) Metabolomics in plants and humans: applications in the prevention and diagnosis of diseases. Biomed Res Int 2013:1–11

Govind G, Harshavardhan VT, Patricia JK et al (2009) Identification and functional validation of a unique set of drought induced genes preferentially expressed in response to gradual water stress in peanut. Mol Gen Genomics 281:607

Govindaraj M, Vetriventhan M, Srinivasan M (2015) Importance of genetic diversity assessment in crop plants and its recent advances: an overview of its analytical perspectives. Genet Res Int 2015:431487

Griffiths-Jones S, Saini HK, vanDongen S, Enright AJ (2008) miRBase: tools for microRNA genomics. Nucleic Acids Res 36:D154–D158

Gupta P, Naithani S, Tello-Ruiz MK, Chougule K, D'Eustachio P, Fabregat A et al (2016) Gramene database: navigating plant comparative genomics resources. Curr Plant Biol 8:10–15

Gurjar AKS, Singh Panwar A, Gupta R, Mantri SS (2016) PmiRExAt: plant miRNA expression atlas database and web applications. Database 2016:baw060

Hammami R, Ben Hamida J, Vergoten G, Fliss I (2009) PhytAMP: a database dedicated to antimicrobial plant peptides. Nucleic Acids Res 37:D963–D968

Harbers M, Carninci P (2005) Tag-based approaches for transcriptome research and genome annotation. Nat Methods 2:495–502

Heazlewood JL, Durek P, Hummel J, Selbig J, Weckwerth W, Walther D, Schulze WX (2008) PhosPhAt: a database of phosphorylation sites in Arabidopsis thaliana and a plant-specific phosphorylation site predictor. Nucleic Acids Res 36:D1015–D1021

Helmy M, Tomita M, Ishihama Y (2011) OryzaPG-DB: rice proteome database based on shotgun proteogenomics. BMC Plant Biol 11:63

Helmy M, Sugiyama N, Tomita M, Ishihama Y (2012a) Mass spectrum sequential subtraction speeds up searching large peptide MS/MS spectra datasets against large nucleotide databases for proteogenomics. Cell Mech 17:633–644

Helmy M, Sugiyama N, Tomita M, Ishihama Y (2012b) The rice proteogenomics database oryza PG-DB: development, expansion, and new features. Front Plant Sci 3:65

Hopff D, Wienkoop S, Lüthje S (2013) The plasma membrane proteome of maize roots grown under low and high iron conditions. J Proteome 91:605–618

Hunter S, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Bork P, Das U, Daugherty L (2009) InterPro: the integrative protein signature database. Nucleic Acids Res 37:D211–D215

Hulo N (2006) The PROSITE database. Nucleic Acids Res 34(90001):D227–D230

International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. Nature 436:793–800

Iijima Y, Nakamura Y, Ogata Y, Tanaka K'i, Sakurai N, Suda K, Suzuki T, Suzuki H, Okazaki K, Kitayama M, Kanaya S, Aoki K, Shibata D (2008) Metabolite annotations based on the integration of mass spectral information. Plant J 54(5):949–962

Jagadeeswaran G, Zheng Y, Sumathipala N, Jiang H, Arrese EL, Soulages JL et al (2010) Deep sequencing of small RNA libraries reveals dynamic regulation of conserved and novel microRNAs and microRNA-stars during silkworm development. BMC Genomics 11:52

Jayashree B, Crouch JH, Prasad PVNS, Hoising-ton D (2006) A database of annotated tentative orthologs from crop abiotic stress transcripts. Bioinformation 1:225–227

Jeong DH, Green PJ (2013) The role of rice microRNAs in abiotic stress responses. J Plant Biol 56:187–197

Jones-Rhoades MW, Bartel DP (2004) Computational identification of plant MicroRNAs and their targets, including a stress-induced miRNA. Mol Cell 14(6):787–799

Jogaiah S, Govind SR, Tran L-SP (2013) Systems biology-based approaches toward understanding drought tolerance in food crops. Crit Rev Biotechnol 33:23–39

Jorrín-Novo JV, Maldonado AM, Echevarría-Zomeño S, Valledor L, Castillejo MA, Curto M, Valero J, Sghaier B, Donoso G, Redondo I (2009) Plant proteomics update (2007–2008): second-generation proteomic techniques, an appropriate experimental design, and data analysis to fulfill MIAPE standards, increase plant proteome coverage and expand biological knowledge. J Proteome 72(3):285–314

Joshi R, Karan R, Singla-Pareek SL, Pareek A (2012) Microarray technology. In: Gupta AK, Pareek A, Gupta SM (eds) Biotechnology in medicine and agriculture: principles and practices. IK International Publishing House Pvt. Ltd., New Delhi, pp 273–296

Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28(1):27–30

Kang S, Ayers JE, Dewolf ED, Geiser DM, Kuldau G, Moorman GW, Mullins E, Uddin W, Correll JC, Deckert G, Lee YH, Lee YW, Martin FN, Subbarao K (2002) The internet-based fungal pathogen database: a proposed model. Phytopathology 92(3):232–236

Karp PD, Billington R, Caspi R, Fulcher CA, et al. (2017) The BioCyc collection of microbial genomes and metabolic pathways. Brief Bioinform. https://doi.org/10.1093/bib/bbx085

Kawahara Y, de la Bastide M, Hamilton JP, Kanamori H, McCombie et al (2013) Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. Rice 6:4

Kawahara Y, Oono Y, Wakimoto H, Ogata J (2016) TENOR: database for comprehensive mRNA-Seq experiments in rice. Plant Cell Physiol 57(1):e7

Khraiwesh B, Zhu JK, Zhu J (2012) Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. Biochim Biophys Acta 1819:137–148

Kim HS, Mittenthal JE, Caetano-Anolles G (2006) MANET:tracing evolution of protein architecture in metabolic networks. BMC Bioinforma 7:351

King ZA, Lu JS, Dräger A, Miller PC, Federowicz S, Lerman JA, Ebrahim A, Palsson BO, Lewis NE (2016) BiGG Models: a platform for integrating, standardizing, and sharing genome-scale models. Nucleic Acids Res 44(D1):D515–D522

Kinjo AR, Bekker GJ, Suzuki H, Tsuchiya Y, Kawabata T, Ikegawa Y, Nakamura H (2017) Protein Data Bank Japan (PDBj): updated user interfaces, resource description framework, analysis tools for large structures. Nucleic Acids Res 45:D282–D288

Kissoudis C, van de Wiel C, Visser RGF, van der Linden G (2014) Enhancing crop resilience to combined abiotic and biotic stress through the dissection of physiological and molecular cross-talk. Front Plant Sci 5:e207

Koltai H, Volpin H (2003) Agricultural genomics: an approach to plant protection. Eur J Plant Pathol 109:101–108

Kopka J, Schauer N, Krueger S, Birkemeyer C, Usadel B, Bergmuller E, Dormann P, Weckwerth W, Gibon Y, Stitt M, Willmitzer L, Fernie AR, Steinhauser D (2005) GMD@CSB.DB: the Golm metabolome database. Bioinformatics 21(8):1635–1638

Krüger J, Rehmsmeier M (2006) RNAhybrid: microRNA target prediction easy, fast and flexible. Nucleic Acids Res 34:W451–W454

Kudo T, Terashima S, Takaki Y, Tomita K et al (2017) PlantExpress: a database integrating OryzaExpress and ArthaExpress for single-species and cross-species gene expression network analyses with microarray-based transcriptome data. Plant Cell Physiol 58(1):e1

Lee TH, Kim YK, Pham THM, Song SI et al (2009) Correlating gene expression from transcriptome profiling, and its application to the analysis of coexpressed genes in rice. Plant Physiol 151(1):16–33

Li JR, Liu CC, Sun CH, Chen YT (2018) Plant stress RNA-seq nexus: a stress-specific transcriptome database in plant cells. BMC Genomics 19:966

Liu B, Zhang N, Zhao S et al (2015) Proteomic changes during tuber dormancy release process revealed by iTRAQ quantitative proteomics in potato. Plant Physiol Biochem 86:181–190

Makita Y, Shimada S, Kawashima M, Kondou-Kuriyama T, Toyoda T, Matsui M (2015) MOROKOSHI: transcriptome database in *Sorghum bicolor*. Plant Cell Physiol 56:1

Maruyama K, Takeda M, Kidokoro S, Yamada K, Sakuma Y, Urano K, Fujita M, Yoshiwara K, Matsukura S, Morishita Y, Sasaki R (2009) Metabolic pathways involved in cold acclimation identified by integrated analysis of metabolites and transcripts regulated by DREB1A and DREB2A. Plant Physiol 150(4):1972–1980

Masoudi-Nejad A, Goto S, Jauregui R et al (2007) EGENES: transcriptome-based plant database of genes with metabolic pathway information and expressed sequence tag indices in KEGG. Plant Physiol 144(2):857–866

Matsuda F, Yonekura-Sakakibara K, Niida R, Kuromori T, Shinozaki K, Saito K (2009) MS/MS spectral tag-based annotation of non-targeted profile of plant secondary metabolites. Plant J 57(3):555–577

Matsui A, Ishida J, Morosawa T, Mochizuki Y, Kaminuma E, Endo TA, Okamoto M, Nambara E, Nakajima M, Kawashima M, Satou M (2008) Arabidopsis transcriptome analysis under drought, cold, high-salinity and ABA treatment conditions using a tiling array. Plant Cell Physiol 49(8):1135–1149

Matvienko M, Kozik A, Froenicke L, Lavelle D, Martineau B, Perroud B et al (2013) Consequences of normalizing transcriptomic and genomic libraries of plant genomes using a duplex-specific nuclease and tetramethylammonium chloride. PLoS One 8(2):e55913

McCarthy FM, Wang N, Bryce Magee G, Nanduri B, Lawrence ML, Camon EB, Barrell DG, Hill DP, Dolan ME, Paul Williams W, Luthe DS, Bridges SM, Burgess SC (2006) AgBase: a functional genomics resource for agriculture. BMC Genomics 7(1):229

Michelmore RW (2003) The impact zone: genomics and breeding for durable disease resistance. Curr Opin Plant Biol 6:397–404

Mine A, Sato M, Tsuda K (2014) Toward a systems understanding of plant–microbe interactions. Front Plant Sci 5:423

Mir S, Alhroub Y, Anyango S, Armstrong DR, Berrisford JM, Clark AR et al (2018) PDBe: towards reusable data delivery infrastructure at protein data bank in Europe. Nucleic Acids Res 46:D486–D492

Moco S, Bino RJ, Vorst O, Verhoeven HA, de Groot J, van Beek TA, Vervoort J, de Vos CHR (2006) A liquid chromatography-mass spectrometry-based metabolome database for tomato. Plant Physiol 141(4):1205–1218

Mousavi SA, Pouya FM, Ghaffari MR, Mirzaei M, Ghaffari A, Alikhani M, Ghareyazie M, Komatsu S, Haynes PA, Salekdeh GH (2016) PlantPReS: a database for plant proteome response to stress. J Proteome 143:69–72

Moxon S, Schwach F, Dalmay T, Maclean D, Studholme DJ, Moulton V (2008) A toolkit for analysing large-scale plant small RNA datasets. Bioinformatics 24:2252–2253

Murzin AG, Brenner SE, Hubbard T, Chothia C (1995) SCOP: a structural classification of proteins database for the investigation of sequences and structures. J Mol Biol 247:536–540

Naika M, Shameer K, Mathew OK, Gowda R, Sowdhamini R (2013) STIFDB2: an updated version of plant stress-responsive transcription factor database with additional stress signals, stress-responsive transcription factor binding sites and stress-responsive genes in Arabidopsis and rice. Plant Cell Physiol 54:1–15

Nakagami H, Sugiyama N, Ishihama Y, Shirasu K (2012) Shotguns in the front line: phosphoproteomics in plants. Plant Cell Physiol 53:118–124

Nakaya A, Ichihara H, Asamizu E, Shirasawa S, Nakamura Y, Tabata S, Hirakawa H (2017) Plant genomics databases. Methods in molecular biology, vol 1533. Humana Press, New York, pp 45–77

Narsai R, Ivanova A, Ng S, Whelan J (2010) Defining reference genes in Oryza sativa using organ, development, biotic and abiotic transcriptome datasets. BMC Plant Biol 10:56

Newton AC, Lyon GD, Marshall B (2002) DRASTIC: a database resource for analysis of signal transduction in cells. BSPP Newsl 42:36–37

Numnark S, Mhuantong W, Ingsriswang S, Wichadakul D (2012) C-mii: a tool for plant miRNA and target identification. BMC Genomics 13:S16

Osuna-Cruz CM, Paytuvi-Gallart A, Di Donato A, Sundesha V, Andolfo G, Aiese Cigliano R, Sanseverino W, Ercolano MR (2018) PRGdb 3.0: a comprehensive platform for prediction and analysis of plant disease resistance genes. Nucleic Acids Res 46:D1197–D1201

Pearson W (2004) Finding protein and nucleotide similarities with FASTA. Curr Protoc Bioinforma Chapter 3:Unit3.9

Pico AR, Kelder T, van Iersel MP, Hanspers K, Conklin BR, Evelo C (2008) WikiPathways: pathway editing for the people. PLoS Biol 6(7):e184

Prabha R, Ghosh I, Singh DP (2011) Plant stress gene database: a collection of plant genes responding to stress condition. ARPN J Sci Tech 1(1):28–31

Priya P, Jain M (2013) RiceSRTFDB: a database of rice transcription factors containing comprehensive expression, cis-regulatory element and mutant information to facilitate gene function analysis. J Database Curation Bat027:1–7

Ramegowda V, Senthil-Kumar M (2015) The interactive effects of simultaneous biotic and abiotic stresses on plants: mechanistic understanding from drought and pathogen combination. J Plant Physiol 176:47–54

Remita AM, Lord E, Agharbaoui Z, Leclercq M et al (2016) WMP: a novel comprehensive wheat miRNA database, including related bioinformatics software. Curr Plant Biol 7(8):31–33

Rhee SY, Beavis W, Berardini TZ, Chen G, Dixon D, Doyle A et al (2003) The Arabidopsis Information Resource (TAIR): a model organism database providing a centralized, curated gateway to Arabidopsis biology, research materials and community. Nucleic Acids Res 31(1):224

Rocha I, Förster J, Nielsen J (2008) Design and application of genome-scale reconstructed metabolic models. Microbi Gene Essentiality: Protoc Bioinforma 416:409–431

Rose PW, Prlić A, Altunkaya A, Bi C, Bradley AR, Christie CH et al (2017) The RCSB protein data Bank: integrative view of protein, gene and 3D structural information. Nucleic Acids Res 45:D271–D281

Sakurai T, Satou M, Akiyama K, Iida K, Seki M, Kuromori T et al (2005) RARGE: a large-scale database of RIKEN Arabidopsis resources ranging from transcriptome to phenome. Nucl Acids Res 33:D647–D650

Scheer M, Grote A, Chang A, Schomburg I, Munaretto C, Rother M, Söhngen C, Stelzer M, Thiele J, Schomburg D (2011) BRENDA, the enzyme information system in 2011. NucleicAcids Res 39:D670–D676

Schulze S, Henkel SG, Driesch D, Guthke R, Linde J (2015) Computational prediction of molecular pathogen–host interactions based on dual transcriptome data. Front Microbiol 6:65

Shafi A, Dogra V, Gill T, Ahuja PS, Sreenivasulu Y (2014) Simultaneous over-expression of PaSOD and RaAPX in transgenic *Arabidopsis thaliana* confers cold stress tolerance through increase in vascular lignifications. PLoS One 9:e110302

Shafi A, Gill T, Sreenivasulu Y, Kumar S, Ahuja PS, Singh AK (2015a) Improved callus induction, shoot regeneration, and salt stress tolerance in Arabidopsis overexpressing superoxide dismutase from *Potentilla atrosanguinea*. Protoplasma 252:41–51

Shafi A, Chauhan R, Gill T, Swarnkar MK, Sreenivasulu Y, Kumar S, Kumar N, Shankar R, Ahuja PS, Singh AK (2015b) Expression of SOD and APX genes positively regulates secondary cell wall biosynthesis and promotes plant growth and yield in Arabidopsis under salt stress. Plant Mol Biol 87:615–631

Shafi A, Pal AK, Sharma V, Kalia S, Kumar S, Ahuja PS, Singh AK (2017) Transgenic potato plants overexpressing SOD and APX exhibit enhanced lignification and starch biosynthesis with improved salt stress tolerance. Plant Mol Biol Rep 35:504–518

Shameer K, Ambika S, Varghese SM, Karaba N, Udayakumar M, Sowdhamini R (2009) STIFDB-Arabidopsis stress responsive transcription factor dataBase. Int J Plant Genomics 2009:583429

Shankar A, Singh A, Kanwar P et al (2013) Gene expression analysis of rice seedling under potassium deprivation reveals major changes in metabolism and signaling components. PLoS One 8:e70321

Singh VK, Singh AK, Chand R, Kushwaha C (2011) Role of bioinformatics in agriculture and sustainable development. Int J Bioinform Res 3(2):221–226

Singh B, Bohra A, Mishra S, Joshi R, Pandey S (2015) Embracing new-generation 'omics' tools to improve drought tolerance in cereal and food-legume crops. Biol Plant 59(3):413–428

Smita S, Lenka SK, Katiyar A, Jaiswal P, Preece J, Bansal KC (2011) QlicRice: a web interface for abiotic stress responsive QTL and loci interaction channels in rice. Database 1:1–9

Soanes DM, Skinner W, Keon J, Hargreaves J, Talbot NJ (2002) Genomics of phytopathogenic fungi and the development of bioinformatic resources. Mol Plant-Microbe Interact 15(5):421–427

Soga T, Ueno Y, Naraoka H et al (2002) Simultaneous determination of anionic intermediates for Bacillus subtilis metabolic pathways by capillary electrophoresis electrospray ionization mass spectrometry. Anal Chem 74:2233–2239

Somerville C, Dangl J (2000) Genomics. Plant biology in 2010. Science 290:2077–2078

Spannagl M, Nussbaumer T, Bader KC, Martis MM, Seidel M, Kugler KG, Gundlach H, Mayer KFX (2016) PGSB PlantsDB: updates to the database framework for comparative plant genome research. Nucleic Acids Res 44:D1141–D1147

Sun Q, Zybailov B, Majeran W, Friso G, Olinares PD, van Wijk KJ (2009) PPDB, the Plant Proteomics Database at Cornell. Nucleic Acids Res 37:D969–D974

Suzuki N, Rivero RM, Shulaev V, Blumwald E, Mittler R (2014) Abiotic and biotic stress combinations. New Phytol 203:32–43

Swarbreck D, Wilks C, Lamesch P et al (2008) The Arabidopsis information resource (TAIR): gene structure and function annotation. Nucleic Acids Res 36:D1009–D1014

Tardieu F, Tuberosa R (2010) Dissection and modelling of abiotic stress tolerance in plants. Curr Opin Plant Biol 13:206–212

Tripathi A, Goswami K, Mishra NS (2015) Role of bioinformatics in establishing microRNAs as modulators of abiotic stress responses: the new revolution. Front Physiol 26:286

Tyers M, Mann M (2003) From genomics to proteomics. Nature 422:193–197

Ueno S, Nakamura Y, Kobayashi M, Terashima S (2018) TodoFirGene: developing transcriptome resources for genetic analysis of Abies sachalinensis. Plant Cell Physiol 59(6):1276–1284

Urano K, Kurihara Y, Seki M, Shinozaki K (2010) 'Omics' analyses of regulatory networks in plant abiotic stress responses. Curr Opin Plant Biol 13:132–138

Vandepoele K (2017) A guide to the PLAZA 3.0 plant comparative genomic database. Plant genomics databases. Methods Mol Biol 1533. Humana Press, New York:183–200

Vassilev D, Leunissen JA, Atanassov A, Nenov A, Dimov G (2005) Application of bioinformatics in plant breeding. Wageningen University, Netherland

Vassilev D, Nenov A, Atanassov A, Dimov G, Getov L (2006) Application of bioinformatics infruit plant breeding. J Fruit Ornamental Plant Res 14:145–162

Verma M, Kumar V, Patel RK, Garg R, Jain M (2015) CTDB: an integrated chickpea transcriptome database for functional and applied genomics. PLoS One 10:0136880

Wanchana SS, Thongjuea VJ, Ulat M, Anacleto R, Mauleon M, Conte M, Rouard M, Wang B, Sun YF, Song N, Wei JP, Wang XJ et al (2018) MicroRNAs involving in cold, wounding and salt stresses in Triticumaestivum L., Plant Physiology and Biochemistry In press. Nucleic Acids Res 36:D943–D946

Wang B, Sun YF, Song N, Wei JP, Wang XJ et al (2014) MicroRNAs involving in cold, wounding and salt stresses in *Triticum aestivum* L. Plant Physiol Biochem 80:90–96

Winnenburg R, Baldwin TK, Urban M, Rawlings C, Köhler J, Hammond-Kosack KE (2006) Nucleic Acids Res 34:D459–D464

Wojciech M, Karlowski, Schoof H, Janakiraman V, Stuempflen V, Mayer KFX (2003) MOsDB: an integrated information resource for rice genomics. Nucleic Acids Res 31(1):190–192

Yan S, Du X, Wu F et al (2014) Proteomics insights into the basis of interspecific facilitation formaize (*Zea mays*) in faba bean (Vicia faba)/maize intercropping. J Proteome 109:111–124

Yang X, Li L (2011) miRDeep-P: a computational tool for analyzing the microRNA transcriptome in plants. Bioinformatics 27:2614–2615

Yu J, Hu S et al (2002) A draft sequence of rice genome. Science 296:79–92

Yuan JS, Galbraith DW, Dai SY, Griffin P, Stewart CN Jr (2008) Plant systems biology comes of age. Trends Plant Sci 13:165–171

Zhang Z, Yu J, Li D, Zhang Z, Liu F, Zhou X, Wang T, Ling Y, Su Z (2010) PMRD: plant microRNA database. Nucleic Acids Res 38:D806–D813

Zhang S, Yue Y, Sheng L, Wu Y, Fan G, Li A, Hu X, Shang Guan M, Wei C (2013) PASmiR: a literature-curated database for miRNA molecular regulation in plant response to abiotic stress. BMC Plant Biol 13:1–8

Zhang M, Lv D, Ge P et al (2014) Phosphoproteome analysis reveals new drought response and defense mechanisms of seedling leaves in bread wheat (*Triticum aestivum* L.). J Proteome 109:290–308

Zhou X, Wang G, Sutoh K, Zhu JK, Zhang W (2008) Identification of cold-inducible microRNAs in plants by transcriptome analysis. Biochim Biophys Acta 1779:780–788

Zhou L, Liu Y, Liu Z, Kong D, Duan M et al (2010) Genome-wide identification and analysis of drought-responsive microRNAs in *Oryza sativa*. J Exp Bot 61:4157–4168