12
Omics Techniques in Crop Research: An Overview

Bashasab Fakrudin, Roberto Tuberosa, and Rajeev K. Varshney

12.1
Introduction

Omics is a collective, broad discipline largely referring to analysis of the interactions of biological information obtained from the profiling of the genome, transcriptome, proteome, metabolome, and several other relevant -omes. While phase one of omics technologies aims at nontargeted identification of transcripts, proteins, and metabolites (essentially gene products) in a given biological sample, phase two deals with a very challenging analysis of data eventually leading to the dissection of the qualitative and quantitative dynamics of biological systems. Essentially, the omics science is enabled by a host of diverse, high-throughput technologies and platforms [1]. The full range of omics technologies can now be applied to understand the same fundamental biological processes [2]. Mapping and defining the relationships among genes, proteins, and metabolites require relative comparison of the networks that eventually help in understanding the regulatory mechanisms. A diverse but converging approaches such as forward and reverse genetics and transgenics (overexpression and knockdowns) can define the function of a gene to the specific phenotype, the omic technologies aim at revealing the function of each and every gene in the genome, which collectively contribute toward elucidating the networks and better understanding the whole plant phenotype [3–5]. Access to omics tools at an affordable price is becoming a reality, which together with a large inventory of candidate genes, proteins, and metabolites and their databases deduced from profiling efforts in model systems and crop plants have speeded up the analysis of biological functions operating in various plant stress responses [2, 6]. These new strategies have begun to piece together the physiological and phenotypic observations with information on transcription and transcript regulation, the behavior of proteins, protein complexes and pathways, and the metabolites and metabolite fluxes, finally shedding light on evolutionary adaptive diversifications of organisms.

12.2
Transcriptomics

Transcriptomics, a global mRNA expression profiling of a particular tissue, is essentially genome wide, yielding information about the transcriptional differences between defined states of tissues. Elucidated global differences in gene expression are expected to help in the understanding of genes and pathways involved in biological processes: gene statements showing similarity in quantitative and qualitative expression are functionally related and would be the result of possible common genetic regulation [2, 7]. Rapid sequencing of many eukaryotic genomes has provided unprecedented opportunities to understand gene function, genome structure, and genome evolution [8]. However, an accurate annotation of all expressed genes in the sequenced genomes remains one of the most challenging tasks. Therefore, genomic resources and platforms provide new opportunities for crop research and breeding programs [8–10].

Transcriptomics can be used to understand taxonomic position to gain a deeper understanding of molecular and physiological bases of complex phenotypes such as crop response to abiotic stresses. Common platform technologies used for genome-wide or high-throughput analysis of gene expression are microarrays, serial analysis of gene expression (SAGE), massively parallel signature sequencing (MPSS), and next-generation sequencing platforms (NGSPs) [11–13]. High-throughput quantitative real-time PCR has proven to be a platform of choice for validation of a large number of genes elucidated through omics approaches in the tissue and treatment of choice [14].

12.2.1
Closed Omics Technologies

Microarray technology represents a "closed" profiling strategy limited by the target genes/gene statements imprinted on gene chips. Microarray technology has become a useful tool for the analysis of genome-scale gene expression. This technology was first demonstrated by analyzing 48 *Arabidopsis* genes for differential expression in roots and shoots [15]. Microarrays are artificially constructed grids of known DNA samples such that each element of the grid probes a specific RNA sequence wherein the RNA transcripts from the target sample are captured and quantified. To date, many different protocols and types of microarrays such as oligonucleotide and cDNA arrays, commercially available whole-genome arrays and custom-made, tissue-specific arrays have been developed. All these invariably require (a) isolation of RNA from target sample, (b) conversion of RNA to either cDNA or cRNA, (c) a simultaneous incorporation of either fluorescent nucleotides or a tag that is later used for fluorescent labeling, (d) hybridization to a chosen microchip, (e) washing and labeling (depending on the protocol adopted), (f) scanning under laser light and image processing to extract data, and (g) data analysis [16, 17]. Bioinformatic clustering tools are required for delineation of closely related expression patterns of genes [17]. The sources for cDNA arrays are generally PCR products from expressed sequence tags (ESTs) or from cell (tissue)-specific sources, which are "spotted" on glass slides. In the case of oligonucleotide

arrays, relatively short, 25-mer, oligonucleotides specific for transcripts of interest are generated by photolithography and solid-phase DNA synthesis [18]. Various modifications to these approaches and the use of fluorophore dyes have allowed a more accurate and reliable expression profiling.

Microarrays have been extensively used in most of the experimental systems including major crop plants [19–22]. For instance, in *Arabidopsis*, full-length cDNA libraries from plants under different conditions, such as drought-treated and cold-treated plants, have been developed [23, 24] and a set of 1300 full-length cDNAs were monitored for the expression patterns under drought and cold stresses [25]. This platform has been widely used in crop research for elucidation of differentially expressed genes in crop plants as a result of biotic and abiotic stress interactions [7, 21, 26].

There are two microarray-based methods for genotyping in crop research: (a) one involves arraying thousands of short oligonucleotides on glass slides for detection of many single-nucleotide polymorphism (SNP) loci in target DNA, which is particularly well suited for genotyping thousands of markers, and the other (b) involves arraying amplified PCR products on glass slides to detect a few SNPs. The latter is useful to detect limited number of SNPs in large number of samples [27, 28]. Biotin-terminated, allele-specific PCR products are spotted unpurified on glass slides coated with streptavidin and visualized through fluorescent oligonucleotides attached to the allele-specific PCR primers. These approaches of genotyping hold great promise in high-throughput genotyping the candidate genes and their trait association in crop plants. A maize expression array containing 57 452 genes has been developed and used in the context of maize nitrogen utilization, root growth under drought, water, and phosphate stress, seed development, photosynthesis, pathogen response, aluminum stress in roots, tassel development, and hybrid vigor (www.maizearray.org). Furthermore, microarray-based gene expression technology is a powerful tool to also monitor changes in the expression of a large number of genes simultaneously and provide new insights into physiological and biochemical pathways of abiotic and biotic stress tolerance and identify novel candidate genes that can be used in plant breeding programs [6, 17, 29].

12.2.2
Open Omics Technologies

In the past decade, various sequencing-based strategies, such as ESTs [30], full-length cDNA [24], SAGE [31, 32], and MPSS have been developed for transcriptome studies [33, 34]. These approaches have contributed valuable resources for gene discovery and genome annotation, but their application in most molecular studies has been limited. In contrast to microarray technology, these technologies are of "open" architectural systems that can be used to identify novel genes and to quantify differentially expressed mRNAs.

12.2.2.1 ESTs, SAGE, and MPSS
Single-pass sequencing of cDNAs to generate ESTs has been a much-used method of elucidation of genes and has contributed a lot of entries to public DNA databases. The error-prone ESTs have remained not only a powerful means of gene discovery but also

a source of biologically informative probes in genome mapping and cloning studies. ESTs have become an invaluable resource for gene discovery, genome annotation, alternative splicing, SNP discovery, molecular markers for population analysis, and expression analysis in animal, plant, and microbial species [16, 35]. Generally, EST and full-length cDNA sequencing techniques are not deep enough to isolate rare transcripts responsible for complex traits or address transcript variability that persists within and between closely related pathways and phenotypic traits. Sequencing millions of cDNA clones from various tissues can sample only about 60% of the expressed genes [36]. Although various computer-based gene prediction methods play a role in genome annotation, experimental data are an essential evidence for determination of gene structure and function [12]. This limitation has been addressed through high-throughput and short tag-based approaches such as SAGE and MPSS, and much recently employing NGSP. Notably, these technologies are most useful for gene expression analysis in plant species whose genome has been sequenced [11, 37].

The SAGE library construction involves several tedious steps before tags can be cloned into a plasmid vector. The process includes isolation of short tags (14–26 bp) from the 3′ or 5′ ends of transcripts, ditag formation, and concatenation and sequencing of SAGE clones. Taking advantage of the high-throughput made possible by the classical SAGE technology [32], new approaches for cloning of 5′-end-specific sequencing tags from mRNA-cap analysis gene expression (CAGE), *trans*-spliced exon-coupled RNA-end determination (TEC-RED), and 5′ serial analysis of gene expression (5′SAGE), gene identification signatures (GIS), the tags comprising information from both terminal ends, were developed [12, 31]. However, the time-consuming procedure of colony picking and storage and the high cost of sequencing individual clones in SAGE library construction have discouraged the use of this approach in many biological studies [31, 38].

The MPSS strategy involves *in vitro* cloning of cDNA molecules on the surface of microbeads and nongel-based sequencing of millions of tags (17–20 bp). It is considered to be more sensitive over SAGE technology [36, 38]. The multiple location-matching of 17–21 bp tags from SAGE or MPSS libraries in a sequenced genome is problematic when mapping tags to the EST or genomic sequence. To obtain accurate matches for positive tags in the genome, longer transcripts are required. This is usually accomplished using techniques such as rapid amplification of cDNA ends (RACE) or generation of longer cDNA fragments using the GLGI method. Integration of pyrosequencing in sequencing technology with SAGE tags has resulted in an increased sensitivity for deep transcript profiling: robust analysis of 5′ gene expression (5′RATE), which involves the use of pyrosequencing of ditag libraries, achieves higher sensitivity of transcript profiling. It consists of three major steps including 5′-oligocapping of mRNA, *Nla*III tag and ditag generation, and pyrosequencing of *Nla*III tags. Complicated steps such as purification and cloning of concatemers, colony picking, and plasmid DNA purification are eliminated and the conventional Sanger sequencing method is replaced with the newly developed pyrosequencing method [39]. Taken together, these techniques provide a panoramic profile of the entire pool of mRNA transcripts that make up the transcriptome.

12.2.2.2 Next-Generation Sequencing

At present, numerous strategies and platforms are under development including sequencing by synthesis (SBS), sequencing by hybridization, and nanopore sequencing. In 2005, two new sequencing technologies were introduced. Both are based on sequencing by synthesis: the 454 system (http://www.454.com) using pyrosequencing technology and the Solexa/Illumina system (http://www.illumina.com) that depends on detection of fluorescence signals [40]. These methods employ parallel sequencing in millions of reactions that generate a very large number of data points. The read lengths are averaging 100–230 bp and 300–400 bp for 454FLX and 454Titanium, respectively, and 35–105 bp for Illumina Solexa platforms. The Applied Biosystems SOLiD (http://www3.appliedbiosystems.com) is another addition with a greater potential in transcriptome sequencing and gene discovery approaches. These platforms offer a variety of experimental approaches for characterizing a transcriptome, discovering genes, small RNAs, and variations in homologues [8, 11, 41]. These sequencing technologies collectively are referred to as next-generation sequencing (NGS) technologies. Potential applications of NGS technologies in gene expression analysis and crop breeding research have been accounted and compared in detail [8, 9, 13, 42, 43]. In addition, several organizations are working on third-generation sequencing technology mainly based on single-molecule synthesis [44]. Furthermore, both the nanopore sequencing and the transmission electron microscopy-based sequencing hold greater promise as third-generation sequencing technologies [41, 45, 46]. The generation of millions of tags at low cost makes these technologies the system of choice for gene expression analysis. A reduction in the cost of sequencing services, which is expected to happen in the near future, will have a major positive impact not only on gene expression studies but also on molecular breeding in agri-hort crops and tree species.

12.3
Metabolomics

Metabolomics is considered the ultimate level of postgenomic analysis as it can reveal changes in metabolite fluxes that are controlled by only minor changes within gene expression measured using transcriptomics and/or by analyzing the proteome that elucidates posttranslational control over enzyme activity [47, 48]. Metabolome refers to a set of metabolites that are formed within a biological system and their types and levels can be regarded as the ultimate response of biological systems to genetic or environmental changes [47]. Central to metabolomics is a range of metabolite fingerprinting and profiling technologies and extraction methods, which profile an entire extract without bias; the richest metabolite profiles will most easily be obtained by employing a range of extraction methods and analytical instruments due to the fact that none is without bias toward certain groups of compounds. Hence, metabolomics is the study of final downstream product of a genomic response as the total quantitative collection of small molecular weight compounds (metabolites) present in a cell/tissue type to whole organism [49]. Metabolomics as of now is considered a

technically demanding interdisciplinary research field that requires expertise in the fields of biology, analytical chemistry, organic chemistry, chemometrics, and informatics sciences. Metabolomic analysis consists of three distinct experimental parts: (a) preparation of the sample, (b) acquisition of data using analytical chemical methods, and (c) data mining using appropriate chemometric methods [47]. Essentially, all these steps are strongly interrelated and interdependent.

Two main metabolite profiling strategies are (i) mass spectrometry (MS) and (ii) nuclear magnetic resonance (NMR). The gas chromatography–mass spectrometry (GC-MS), gas chromatography–time-of-flight mass spectrometry (GC-TOF-MS), and liquid chromatography–mass spectrometry (LC-MS) are extensively used MS-based techniques in metabolite analyses. The GC-MS technology enables the identification and quantification of over a few hundred primary metabolites within a single extract [47, 50]. The GC-TOF-MS offers fast scan times, resulting in an improved peak deconvolution and higher sample throughput. On the other hand, LC-MS measures a far broader range of metabolites including primary and secondary metabolites [51]. In addition to this, capillary electrophoresis–mass spectrometry (CE-MS) and fourier-transformation cyclotron resonance–mass spectrometry (FT-ICR-MS) are also used. CE-MS is considered a highly sensitive methodology to detect low-abundance metabolites in plant samples [47]. The FT-ICR-MS relies solely on very high-resolution mass analysis, which potentially enables the measurement of the empirical formula for thousands of metabolites, although it is somewhat limited by the lack of chromatographic separation. NMR approaches rely on the detection of magnetic nuclei of atoms after application of a constant magnetic field for metabolite profiling [52]. NMR can provide subcellular information and it is easier to derive atomic information for flux modeling from NMR than from MS-based approaches [47]. In plant systems, metabolomics approach has already been used to study metabolomic changes during a variety of stresses, for example, temperature [53], water and salinity [54], sulfur [55], phosphorus [56], and oxidative [57] and heavy metals [58]. These tools have recently been turned to evaluation of the natural variance apparent in metabolite composition. Metabolomics approaches have great value in both phenotyping and diagnostic analyses in plants that might eventually enable metabolomics-assisted breeding in crop plants [59].

12.4
Proteomics

Proteomics is the qualitative and quantitative comparison of proteomes under different conditions to further unravel biological processes. The term proteome refers to all proteins expressed by a genome in the targeted tissues at a defined time point. It encompasses a broad range of tools and techniques in determining the identity and quantity of the expressed proteins in cells/tissues, their 3D structure, and other interacting partners that help to disclose gene function. Changes that occur at the protein level can be traced to genetic sequences, thus forming a unique cross reference to the complex biological phenomenon being investigated. This involves

separation, identification, determination of function, and its interactions with other proteins and biological molecules [60]. Protein profiling techniques allow a rapid comparison of complex samples and direct investigation of tissue specimens. In addition, proteomics has been complemented by the analysis of posttranslational modifications and techniques for the quantitative comparison of different proteomes [60, 61].

Techniques such as matrix-assisted laser desorption/ionization (MALDI) has been employed for rapid determination of proteins in particular mixtures besides electrospray ionization (ESI). Most proteins function in collaboration with other proteins, and one goal of proteomics is to identify which proteins interact. This is especially useful in determining potential partners in cell signaling cascades [60]. There has been extensive research over the last few years to study the technical aspects of proteomics in plants [62] and studies have been conducted in *Arabidopsis*, rice [63]), maize [64], barley [65], and chickpea [66]. Proteomics not only enables the study of protein–protein interaction but also helps in identification of multisubunit complexes [67]. Furthermore, proteomics can act as a powerful approach to organize and identify the proteome through development of 2DE gel protein reference maps of subproteomes in different plant species.

12.5
Interactomics

The complexity of an organism or even a complex trait of an organism cannot be completely explained by mere total number of genes alone. The organism would utilize combinatorial complexity to manifest required form of growth and/or resources in time and space [5, 68]. Therefore, to elucidate the complete functioning of an organism, one not only need to learn the biochemical function(s) of every protein and every domain but also need to discover all protein–protein interactions: developing an interactome is a prerequisite to understand the complex web of interactions that link biological molecules in a cell [3, 4, 69]. The network of all interactions is called the interactome that thus aims to compare such networks of interactions. The interactome maps are important in defining gene function and understanding the function of macromolecular complexes. Various experimental approaches for creating large-scale protein–protein interaction maps in plants have been reviewed by Morsy *et al.* [70].

The study of the interactome requires collection of a large amount of data from a single organism under a small number of perturbations. The two-hybrid screening (Y2H), tandem affinity purification, X-ray tomography, and optical fluorescence microscopy are generally employed for this purpose [71]. High-throughput versions of some of these methods have already been developed, although there is room for further improvement. It is envisaged to combine microarrays and mass spectrometry to enhance throughput of the technique. Tandem affinity purification (TAP) is a technique for studying protein–protein interactions. It involves creating a fusion protein with a designed piece, the TAP tag, on the end. The protein of interest with the

TAP tag first binds to beads coated with IgG, the TAP tag is then broken apart by an enzyme, and finally a different part of the TAP tag binds reversibly to beads of a different type [72]. After the protein of interest has been washed through two affinity columns, it can be examined for binding partners. Interactomic approaches are increasingly becoming relevant to gain a comprehensive understanding of both basic and applied aspects of complex plant–environment interactions [35, 73].

12.6
Genomics (or High-Throughput Genotyping) and Phenomics

The omic strategies, as mentioned above, can highlight candidate genes, metabolites, proteins, and so on that are responsible for a particular phenotype [74, 75]. Importantly, this information can be used in applied breeding programs through molecular marker technologies [8, 76]. Although many kinds of marker technologies have been developed, at present, simple sequence repeats (SSRs) or microsatellite, diversity array technologies (DArT), and SNPs are the marker systems of choice [37, 77–80]. Among these systems, SNP and DArT marker systems can be used for very high-throughput genotyping.

Several high-throughput SNP genotyping platforms are available with varying levels of suitability in practical use. Two platforms, namely, Illumina's GoldenGate assay and whole-genome genotyping Infinium assay, hold practical significance. In the former platform, the genomic DNA is activated using paramagnetic particles and PCR-amplified using three oligos and a universal PCR primer pair for each SNP. Two of the oligos used are allele-specific oligos that, upon ligation to the target allele, extends and ligates to the third locus-specific oligo (LSO) that contains SNP-specific tag and sequence complementary to the universal primer. The universal primer carries allele-specific fluorescent label and contains an address sequence that helps in binding the amplified product to the beads of fiber optic array. Genotyping is done in multiple of 96. GoldenGate assays have been developed for several crop species such as barley [81], wheat [82], maize [83], and common bean [84]. SNP genotyping based on GoldenGate assay has been very successful in constructing a genetic map and trait mapping [81, 84]. Crop-specific efforts are at different levels of success in developing first-generation GoldenGate assays for SNP genotyping worldwide. It is expected that in the next couple of years, the SNP GoldenGate assays would be available and large-scale use of SNPs will become integral to genetics and breeding efforts in most of the crop plants [9, 43].

The whole-genome profiling Infinium assay is done through comparative genomic hybridization. The change in the allele composition is measured through varying signal intensities. This assay includes whole-genome amplification to increase the amount of DNA followed by fragmentation and capturing onto bead array through SNP-specific primer. The primer anneals adjacent to a SNP and extension takes place that involves incorporation of hapten-labeled nucleotide corresponding to the SNP allele. Incorporated hapten-labeled nucleotides are detected by adding fluorescent-labeled antibodies during various steps to amplify the signal. Development of Infinium assays is in progress for some crop species such as soybean, maize, and so on.

The advent of a number of omics technologies and especially high-throughput genotyping has now made phenotyping the priority in crop research. Phenotypes that can be studied across species are more attractive, particularly given the rapid development in transgenic modeling. With advances in high-throughput genotyping technologies, the rate-limiting step of large-scale genetic investigations has been the accurate high-throughput phenotyping in a large number of samples. Phenomics is an emerging transdiscipline dedicated to the systematic study of phenotypes on a genome-wide scale. It is the systematic measurement and analysis of qualitative and quantitative traits, including clinical, biochemical, and imaging methods, for the refinement and characterization of a phenotype. Phenomics require deep phenotyping, the collection of a wide breadth of phenotypes with fine resolution, and phenomic analysis composed of constructing heat maps, cluster analysis, text mining, and pathway analysis (Figure 12.1). Many technologies have been developed to help explain the phenotypic consequences of genetic and/or environmental modifications in areas such as functional genomics, pharmaceutical research, and metabolic engineering [85–87]. The advances in metabolomics
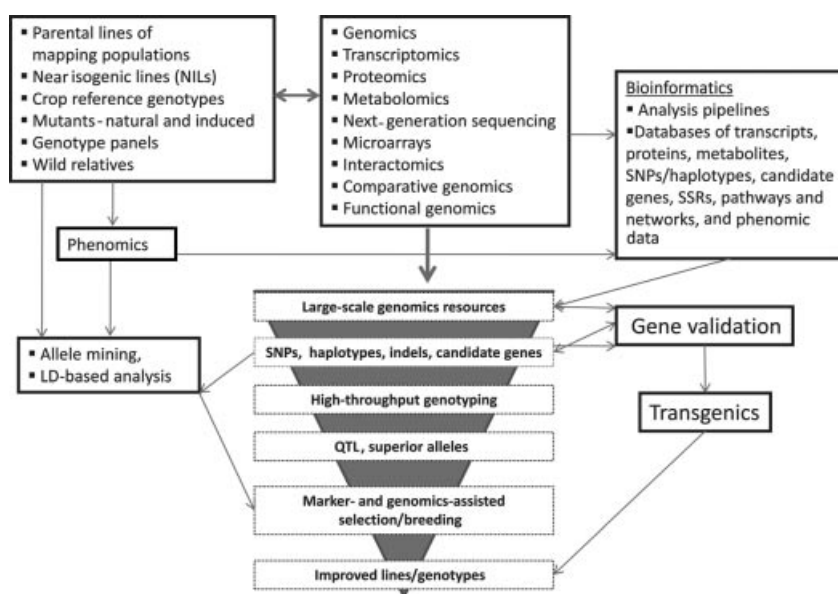

Figure 12.1 An overview of omics technologies and their applications in crop research and breeding. The omics platforms have potential to generate large-scale genomic resources, protein, and metabolite profiles from specific tissues that on their own and in combination with parallel but converging platforms can aid in pinning down the superior alleles, QTL, and candidate genes, besides contributing towards a deeper understanding of complex traits for crop improvement. Recent advances in these platform technologies and the bioinformatic pipelines have led to an accelerated development of robust and high-throughput marker systems. These advancements have ushered in new opportunities and strategies in crop research and breeding of even orphan crops.

and proteomics have a direct bearing on the large-scale phenotyping with a greater sensitivity in a high-throughput manner.

12.7
Integrated Omics Technology Approach

Availability of a number of omics technologies coupled with the vast quantity of genotyping data and volumes of precise phenotyping data opens an opportunity for displaying these techniques in an integrated approach in crop research and breeding [8, 10, 37, 42, 88]. This approach is expected to aid in gaining a better understanding of complex traits and environmental interactions [6]. This would be true both at cellular and whole-plant/crop level. Single-cell analysis was once considered beyond the capacity of omics technologies, but the recent examples of single-cell genomics, transcriptomics, proteomics, and metabolomics indicate an accelerated change largely owing to the rapidly emerging technologies that range from micro/nanofluidics to microfabricated interfaces for mass spectrometry to second- and third-generation automated, high-precision DNA sequencers. Such integration will enable the identification of genes and gene products, and can elucidate the functional relationships between genotype and observed phenotype, thereby permitting a system-wide analysis from genome to phenome, enabling accurate trait mapping, introgression of superior alleles, and in some cases the cloning of major QTL [88] for hitherto considered complex characters such as abiotic and biotic stress tolerance (Figure 12.1; see also Refs [89, 90]). Results of such an integration of omics technologies in model systems and selected crops is highly encouraging. Within next half a decade or so, omics technologies should be available for crop research and breeding in most of the crop plants.

References

1 Holmes, C., McDonald, F., Jones, M. *et al.* (2010) *OMICS.*, **14**, 327–332.

2 North, H., Baud, S., Debeaujon, I. *et al.* (2010) *Plant J.*, **61**, 971–981.

3 Hecker, M., Lambeck, S., Toepfer, S. *et al.* (2009) *Biosystems.*, **96**, 86–103.

4 Long, T.A., Brady, S.M., and Benfey, P.N. (2008) *Annu. Rev. Cell Dev. Biol.*, **24**, 81–103.

5 Sauer, U., Heinemann, M., and Zamboni, N. (2007) *Science*, **316**, 550–551.

6 Bohnert, H.J., Gong, Q., Li, P. *et al.* (2006) *Curr. Opin. Plant Biol.*, **9**, 180–188.

7 Abuqamar, S., Luo, H., Laluk, K. *et al.* (2009) *Plant J.*, **58**, 347–360.

8 Edwards, D. and Batley, J. (2010) *Plant Biotechnol. J.*, **8**, 2–9.

9 Varshney, R.K. and Dubey, A. (2009) *J. Plant Biochem. Biotechnol.*, **18**, 127–138.

10 Yano, M. and Tuberosa, R. (2009) *Curr. Opin. Plant Biol.*, **12**, 103–106.

11 Costa, V., Gallo, M.A., Letizia, F. *et al.* (2010) *PPAR Res.* DOI: 10.1155/2010/409168.

12 Harbers, M. and Carninci, P. (2005) *Nat. Methods.*, **2**, 495–502.

13 Mardis, E.R. (2008) *Trends Genet.*, **24**, 133–141.

14 Czechowski, T., Bari, R.P., Stitt, M. *et al.* (2004) *Plant J.*, **38**, 366–379.

15 Schena, M., Shalon, D., Davis, R.W. *et al.* (1995) *Science*, **270**, 467–470.

16 Sreenivasulu, N., Kishor, P.B.K., Varshney, R.K. *et al.* (2002) *Curr. Sci.*, **83**, 965–973.

17 Sreenivasulu, N., Sunkar, R., Wobus, U. *et al.* (2010) *Methods Mol. Biol.*, **639**, 71–93.

18 Gupta, P.K., Roy, J.K., and Prasad, M. (1999) *Curr. Sci.*, **77**, 875–884.

19 Buchanan, C.D., Lim, S., Salzman, R.A. *et al.* (2005) *Plant Mol. Biol.*, **58**, 699–720.

20 Oztur, Z.N., Talame, V., Deyholos, M. *et al.* (2002) *Plant Mol. Biol.*, **48**, 551–573.

21 Rabbani, M.A., Maruyama, K., Abe, H. *et al.* (2003) *Plant Physiol.*, **133**, 1755–1767.

22 Rensink, W.A., Iobst, S., Hart, A. *et al.* (2005) *Funct. Integr. Genomics.*, **5**, 201–207.

23 Seki, M., Carninci, P., Nishiyama, Y. *et al.* (1998) *Plant J.*, **15**, 707–720.

24 Seki, M., Narusaka, M., Ishida, J. *et al.* (2002) *Plant J.*, **31**, 279–292.

25 Seki, M., Narusaka, M., Abe, H. *et al.* (2001) *Plant Cell.*, **13**, 61–72.

26 Legay, S., Lamoureux, D., Hausman, J.F. *et al.* (2009) *Plant Cell Rep.*, **28**, 1799–1816.

27 Flavell, A.J., Bolshakov, V.N., Booth, A. *et al.* (2003) *Nucleic Acids Res.*, **31**, e115.

28 Hou, P., Ji, M., Li, S. *et al.* (2004) *Clin. Chem.*, **50**, 1955–1957.

29 Conesa, A., Forment, J., Gadea, J. *et al.* (2007) Microarray technology in agricultural research, in *Microarray Technology Through Application*, Taylor & Francis, Abingdon, pp. 173–209.

30 Adams, M.D., Kelley, J.M., Gocayne, J.D. *et al.* (1991) *Science*, **252**, 1651–1656.

31 Gowda, M., Jantasuriyarat, C., Dean, R.A. *et al.* (2004) *Plant Physiol.*, **134**, 890–897.

32 Velculescu, V.E., Zhang, L., Vogelstein, B. *et al.* (1995) *Science*, **270**, 484–487.

33 Brenner, S., Johnson, M., Bridgham, J. *et al.* (2000) *Nat. Biotechnol.*, **18**, 630–634.

34 Meyers, B.C., Vu, T.H., Tej, S.S. *et al.* (2004) *Nat. Biotechnol.*, **22**, 1006–1011.

35 Torto, T.A., Li, S., Styer, A. *et al.* (2003) *Genome Res.*, **13**, 1675–1685.

36 Sun, M., Zhou, G., Lee, S., Chen, J., Shi, R.Z., and Wang, S.M. (2004) *BMC Genomics*, **5**, 1.

37 Joosen, R.V., Ligterink, W., Hilhorst, H.W. *et al.* (2009) *Curr. Genomics*, **10**, 540–549.

38 Gowda, M., Venu, R.C., Raghupathy, M.B. *et al.* (2006) *BMC Genomics*, **7**, 310.

39 Gowda, M., Li, H., and Wang, G.L. (2007) *Nat. Protoc.*, **2**, 1622–1632.

40 Porreca, G.J., Zhang, K., Li, J.B. *et al.* (2007) *Nat. Methods*, **4**, 931–936.

41 Gupta, P.K. (2008) *Trends Niotechnol.*, **26**, 1135.

42 Gupta, P.K., Rustgi, S., and Mir, R.R. (2008) *Heredity*, **101**, 5–18.

43 Varshney, R.K., Nayak, S.N., May, G.D. *et al.* (2009) *Trends Biotechnol.*, **27**, 522–530.

44 Hudson, M. (2008) *Mol. Ecol. Resour.*, **8**, 3–17.

45 McCarthy, A. (2010) *Chem. Biol.*, **17**, 675–676.

46 Schadt, E.E., Turner, S., and Kasarskis, A. (2010) *Hum. Mol. Genet.*, **19** (R2), R227–R240.

47 Keurentjes, J.J. (2009) *Curr. Opin. Plant Biol.*, **12**, 223–230.

48 Saito, N., Robert, M., Kitamura, S. *et al.* (2006) *J. Proteome Res.*, **5**, 1979–1987.

49 Kim, H.K. and Verpoorte, R. (2010) *Phytochem. Anal.*, **21**, 4–13.

50 Roessner, U., Luedemann, A., Brust, D. *et al.* (2001) *Plant Cell.*, **13**, 11–29.

51 Tohge, T., Nishiyama, Y., Hirai, M.Y. *et al.* (2005) *Plant J.*, **42**, 218–235.

52 Fernie, A.R., Trethewey, R.N., Krotzky, A.J. *et al.* (2004) *Nat. Rev. Mol. Cell Biol.*, **5**, 763–769.

53 Kaplan, F., Kopka, J., Sung, D.Y. *et al.* (2007) *Plant J.*, **50**, 967–981.

54 Cramer, G.R., Ergul, A., Grimplet, J. *et al.* (2007) *Funct. Integr. Genomics*, **7**, 111–134.

55 Nikiforova, V.J., Daub, C.O., Hesse, H. *et al.* (2005) *J. Exp. Bot.*, **56**, 1887–1896.

56 Hernandez, G., Ramirez, M., Valdes-Lopez, O. *et al.* (2007) *Plant Physiol.*, **144**, 752–767.

57 Baxter, C.J., Redestig, H., Schauer, N. *et al.* (2007) *Plant Physiol.*, **143**, 312–325.

58 Le Lay, P., Isaure, M.P., Sarry, J.E. *et al.* (2006) *Biochimie*, **88**, 1533–1547.

59 Fernie, A.R. and Schauer, N. (2009) *Trends Genet.*, **25**, 39–48.

60 Schulze, W.X. and Usadel, B. (2010) *Annu. Rev. Plant Biol.*, **61**, 491–516.

61 Hirsch, J., Hansen, K.C., Burlingame, A.L. *et al.* (2004) *Am. J. Physiol. Lung Cell Mol. Physiol.*, **287**, L1–L23.

62 van Wijk, K.J. (2001) *Plant Physiol.*, **126**, 501–508.

63 Tsugita, A., Kamo, M., Kawakami, T. *et al.* (1996) *Electrophoresis.*, **17**, 855–865.

64 Chang, W.W., Huang, L., Shen, M. *et al.* (2000) *Plant Physiol.*, **122**, 295–318.

65 Witzel, K., Weidner, A., Surabhi, G.K. *et al.* (2010) *Plant Cell Environ.*, **33**, 211–222.

66 Pandey, A., Choudhary, M.K., Bhushan, D. *et al.* (2006) *J. Proteome Res.*, **5**, 3301–3311.

67 Washburn, M.P., Wolters, D., and Yates, J.R., 3rd (2001) *Nat. Biotechnol.*, **19**, 242–247.

68 Cesareni, G., Ceol, A., Gavrila, C. *et al.* (2005) *FEBS Lett.*, **579**, 1828–1833.

69 D'Alessandro, A., Zolla, L., and Scaloni, A. (2011) *Mol. Biosyst.*, **7** (3), 889–898.

70 Morsy, M., Gouthu, S., Orchard, S. *et al.* (2008) *Trends Plant Sci.*, **13**, 183–191.

71 Chamousset, D., Mamane, S., Boisvert, F.M. *et al.* (2010) *Proteomics.*, **10**, 3045–3050.

72 Xu, X., Song, Y., Li, Y. *et al.* (2010) *Protein Expr. Purif.*, **72**, 149–156.

73 Gonzalez-Fernandez, R., Prats, E., and Jorrin-Novo, J.V. (2010) *J. Biomed. Biotechnol.*, 932527.

74 Shulaev, V., Cortes, D., Miller, G. *et al.* (2008) *Physiol. Plant.*, **132**, 199–208.

75 Tuberosa, R., Gill, B.S., and Quarrie, S.A. (2002) *Plant Mol. Biol.*, **48**, 445–449.

76 Varshney, R.K. and Tuberosa, R. (2007) Genomics-Assisted Crop Improvement: An Overview in *Genomics Applications in Crops*, vol. **2**, Springer, Dordrecht, pp. 1–12.

77 Comadran, J., Thomas, W.T., van Eeuwijk, F.A. *et al.* (2009) *Theor. Appl. Genet.*, **119**, 175–187.

78 Grewal, T.S., Rossnagel, B.G., Pozniak, C.J. *et al.* (2008) *Theor. Appl. Genet.*, **116**, 529–539.

79 Jimenez-Gomez, J.M. and Maloof, J.N. (2009) *BMC Plant Biol.*, **9**, 85.

80 Wenzl, P., Suchankova, P., Carling, J. *et al.* (2010) *Theor. Appl. Genet.*, **121**, 465–474.

81 Rostoks, N., Mudie, S., Cardle, L. *et al.* (2005) *Mol. Genet. Genomics*, **274**, 515–527.

82 Akhunov, E., Nicolet, C., and Dvorak, J. (2009) *Theor. Appl. Genet.*, **119**, 507–517.

83 Barbazuk, W.B., Emrich, S.J., Chen, H.D. *et al.* (2007) *Plant J.*, **51**, 910–918.

84 Hyten, D.L., Song, Q., Fickus, E.W. *et al.* (2010) *BMC Genomics.*, **11**, 475.

85 Bilder, R.M., Sabb, F.W., Cannon, T.D. *et al.* (2009) *Neuroscience.*, **164**, 30–42.

86 Ihlow, A., Schweizer, P., and Seiffert, U. (2008) *BMC Plant Biol.*, **8**, 6.

87 Jewett, M.C., Hofmann, G., and Nielsen, J. (2006) *Curr. Opin. Biotechnol.*, **17**, 191–197.

88 Salvi, S. and Tuberosa, R. (2007) Cloning QTLs in Plants, in *Genomics-Assisted Crop Improvement: Genomics Approaches and Platforms*, vol **1**, Springer, Dordrecht, pp. 207–226.

89 Collins, N.C., Tardieu, F., and Tuberosa, R. (2008) *Plant Physiol.*, **147**, 469–486.

90 Xue, S., Li, G., Jia, H. *et al.* (2010) *Theor. Appl. Genet.*, **121**, 147–156.