Available online at www.ijpab.com

DOI: http://dx.doi.org/10.18782/2320-7051.3072

ISSN: 2320 – 7051

Int. J. Pure App. Biosci. **5 (6):** 1075-1085 (2017)

International Journal of Pure & Applied Bioscience

Review Article

OPEN ACCESS

OMICS Technologies towards Seed Quality Improvement

Komala, N. T.*, Gurumurthy, R. and Surendra, P.

Department of Seed Science and Technology, College of Agriculture, Dharwad,

University of Agricultural Sciences, Dharwad, India

*Corresponding Author E-mail: ntkomala@gmail.com

Received: 2.06.2017 | Revised: 11.07.2017 | Accepted: 18.07.2017

ABSTRACT

Seeds are the repository of the genetic potential of crop species and their varieties resulting from continuous improvement through plant breeding techniques. It plays an integral role in the global food supply and accounts more than 70 per cent of the calories that we consume on a daily basis. In order to meet increasing global food demand for the growing population, the research as to be done at molecular level to improve seed yield and quality. To developing an established modern seed industry, many characteristics of economically important crops could potentially improve through state of art is OMICS technology; OMICS technologies play an essential role for improving seed quality parametes. Genomics, transcriptomics, proteomics, interactomics and metabolomics have provided biologists with comprehensive data concerning global gene function and identification, dynamic changes between gene expression, protein identification, function and interaction of entire protein networks of seed developmental physiology and morphology are vital for transcending the gap between fundamental biology knowledge and application towards improved seeds which are important for establishing modern seed industry and plant breeding programs for increased food production.

Key words: *Seed Quality, Genomics, Trascriptomics, Proteomics, Metabolomics.*

INTRODUCTION

A seed is an embryonic plant enclosed in a protective outer covering. The formation of the seed is part of the process of reproduction in seed plants, the spermatophytes, including the gymnosperm and angiosperm plants. Seeds are the product of the ripened ovule, after fertilization by pollen and some growth within the mother plant. The embryo is developed from the zygote and the seed coat from the integuments of the ovule. The importance of quality seeds has been recognized from the time immemorial. The old scripture, Manu Smriti says "Subeejam Sukshetre Jayate Sampadyathe" i.e., Good seed in good soil yields abundantly. Seed quality has been treated as sacred, being an important factor in the improvement of agriculture and agrarian societies. In order to improve the quality of seed traditional techniques used to identify favourable crop charecteristics for use in plant breeding are often inadequate in determining specific gene trait association. Therefore, there is a need to take up research at molecular level for that OMICS technology plays an important role.

Cite this article: Komala, N.T., Gurumurthy, R. and Surendra, P., OMICS Technologies towards Seed Quality Improvement, *Int. J. Pure App. Biosci.* **5(6):** xxx-xxx (2017). doi: http://dx.doi.org/10.18782/2320-7051.3072

Technologies that measure some characteristic of a large family of cellular molecules, such as genes, proteins, or small metabolites, have been named by appending the suffix "OMICS". OMICS refers to the collective technologies used to explore the roles, relationships, and actions of the various types of molecules that make up the cells of an organism (Figure 1). They are aimed primarily at the universal detection of genes (Genomics: characterizes genome wide expression of DNA), mRNA (Transcriptomics: study of RNA regulation), proteins (Proteomics: study global protein function and expression) and metabolites (Metabolomics: identification and quantification of all metabolites) in a specific biological sample in a non-targeted and non-biased manner. This can also be referred to as high-dimensional biology; the integration of these techniques is called systems biology[45]. The basic aspect of these approaches is that a complex system can be understood more thoroughly if considered as a whole. Systems biology and OMICS experiments differ from traditional studies, which are largely hypothesis driven or reductionist. By contrast, systems biology experiments are hypothesis-generating, using holistic approaches where no hypothesis is known or prescribed but all data are acquired and analysed to define a hypothesis that can be further tested[24].

Genomics is the new science that deals with the discovery and noting of all the sequences in the entire genome of a particular organism. The genome can be defined as the complete set of genes inside a cell. Seed genomics is the study of genomes and expression of genes that are required to make a seed. Determining the genomic sequence, however, is only the beginning of genomics. Once this is done, the genomic sequence is used to study the function of the numerous genes (functional genomics), to compare the genes in one organism with those of another (comparative genomics), or to generate the 3-D structure of one or more proteins from each protein family, thus offering clues to their function (structural genomics). Sequencing technologies have played a crucial role in unraveling the genome. Therefore, the DNA sequencing methods include 1. Basic DNA sequencing (a. Sanger method: is based on the DNA polymerase-dependent synthesis of a complementary DNA strand in the presence of natural 2′-deoxynucleotides (dNTPs) and 2′, 3′-dideoxynucleotides (ddNTPs) that serve as nonreversible synthesis terminators[46] and it sequence only small stretch of nucleotide and b. Maxam gilbert method or chemical cleavage). 2. Advanced DNA sequencing (a. Short gun sequencing: used to sequence larger sections of DNA in attempt to obtain the entire genome, the method developed during Human Genome Project). 3. Next generation sequencing technology, reads the DNA fragments to adaptors by first fragmenting the DNA and the ligating the DNA fragments to adopters that are randomly read during DNA synthesis (a. Solid sequencing: massively parallel sequencing by ligation. b. Ilumina sequencing: sequencing by synthesis of single-molecule arrays with reversible terminators and c. Pyrosequencing: "sequencing by synthesis" principle). So for there are more than 25 plant model organisms were applied for whole genome sequencing, such as beans, rice, grasses, maize, grape, sorghum, banana, wheat, etc.(http://www.arabidopsis.org/). Gene expression studies have led to the discovery of genes that contribute to cell wall weakening during germination, such as expansion, and other genes such as those involved in energy metabolism[6,33].

One of the greatest achievements of plant biology is the completion of the whole genome sequences of model plants such as Arabidopsis thaliana and rice. In Arabidopsis ~27000 genes were predicted based on nucleotide sequence information; however, only half of these genes have been functionally annotated based on sequence similarity to known genes, and among these the function of only ~11% has been confirmed with direct experimental evidence[34]. Golden rice is another example of the successful engineering of a very important cereal grain, rice, to accumulate β-carotene, a precursor of vitamin A in the endosperm to alleviate a global health

problem associated with vitamin A deficiency[5]. Another nutritional quality sought after in wheat grain is dietary fiber primarily composed of cell wall polysaccharides, β-glucan (20%) and arabinoxylans (70%)[48]. Microarray and bioinformatic analyses have been used to identify a putative wheat β-glucan synthase gene which was down regulated using RNAi to change the levels of β-glucan in the endosperm thus enabling the manipulation of quality and content of dietary fiber in the grain[39]. The wheat genome has recently been sequenced and insights gained from its characterization will provide the necessary genetic toolkit to further improve the grain quality[7]. Lipid peroxidation plays a major role in seed longevity and viability. In rice grains, lipid peroxidation is catalyzed by the enzyme lipoxygenase 3 (*LOX3*). The suppression of *LOX3* expression in rice endosperm increased grain storability. The germination rate of TS-91 (antisense *LOX3* transgenic line) was much higher than the wild type[52]. Furukawa *et al*[13]., reported that introduction of the DFR gene into an *Rcrd* mutant resulted in red-colored rice, which was brown in the original mutant, demonstrating that the *Rd* locus encodes the DFR protein. Accumulation of proanthocyanidins was observed in the transformants by the introduction of the *Rd* gene into the rice *Rcrd* line. Revilla *et al*[43]., compared sweet corn inbred line P39 and EP44 living and dead seeds after 20 and 22 years of seeds storage using markers and found polymorphic SSRs in six known genes, including pathogenesis-related protein 2, superoxide dismutase 4, catalase 3, opaque endosperm 2, and metallothionein1 that were related to germination. Therefore, genetic variability among aged seeds of inbreds was useful for preliminary association analysis to identify candidate genes.

TRANSCRIPTOMICS

Is the study of the transcriptom- the complete set of RNA transcripts that are produced by the genome, under specific circumstances or in specific cell-using high-throughput methods, such as microarray analysis. Comparison of transcriptomes allows the identification of genes that are differentially expressed in distinct cell populations, or in response to different treatments. As a result, transcriptomics is the dynamic link between genomics and proteomics and thus can elaborate on the complex cellular processes responsible for adapting to environmental conditions. Transcriptomics data has provided a valuable resource for molecular studies of many important agricultural crops, such as cereals and legumes.

Large scale transcriptomics studies are rapidly gaining power owing to the development of new technologies. Microarrays brought the first revolutionary step in transcriptomics by allowing large scale quantitative assessments of genes expression[4] at a low cost. In the past few years, the use of microarray for evolutionary transcriptomics has been largely abandoned in favour of more powerful technologies based on high through put sequencing. RNA sequencing has numerous advantages over microarrays even for interaspecies transcriptomic studies, as it enables a wider dynamic range for transcript detection[51], provides a better quantification of expression levels and more comparability with proteomics studies[12]. Combined microarrays, Affymetrix fluorescent fluorophores Gene Chip's, MPSS (Massively Parallel Signature Sequencing) and next-generation RNA-Seq have contributed greatly to our current understanding of the seed transcriptome. Le *et al*[27]., reported the use of a soybean Affymetrix GeneChip to profile RNAs from laser-captured suspensor and embryo-proper tissues of scarlet runner bean. Le *et al*[26]., published the *Arabidopsis* seed transcriptome at seven stages of development from ovule to seedling and identified putative regulators of seed development. At each stage of development, between 8779 and 13,722 distinct mRNAs were detected at the level of the GeneChip with 15,563 unique transcripts detected over all stages of seed development. Of these, only 2% (289) of the transcripts were considered seed-specific with the vast majority being specific to a given stage of development (e.g., globular-cotyldeon).

Recently, ionomics and metabolomics were coupled for a comprehensive assessment of GM and non-GM soybean lines[25]. Li *et al*[28]., found the highest metabolic flux during early seed fill by integrating metabolomics and transcriptomics analysis. RNA sequencing of brown and yellow-coated *B. juncea* revealed three dihydroflavonol reductase genes and three anthocyanin reductase genes that were highly expressed in the brown-seeded variety with almost no detectable expression in the yellow-seeded variety[31]. Figueroa *et al*[10]., generated deep sequencing small RNA reads from two cowpea genotypes (CB46 and IT93K503-1) that grew under well-watered and drought stress conditions. They mapped small RNA reads to cowpea genomic sequences and identified 157 miRNA genes that belong to 89 families. Among 44 drought-associated miRNAs, 30 were upregulated in drought condition and 14 were downregulated. RNA-seq technology also used to explore the transcriptome of a single plant cell type, the *Arabidopsis* mal meiocyte, detecting the expression of approximately 20,000 genes[56].

PROTEOMICS

Is the large scale study of the total complement of proteins in a given sample, has been applied to all aspects of seed biology mainly using model species such as Arabidopsis or important agricultural crops such as corn and rice. Since proteins are responsible for most metabolic processes in the seed, in addition to being important structural components in the cytoskeleton, membranes, the cell wall, *etc.*, it makes excellent sense to describe the proteome of a seed, a seed tissue, a specific cell type or a subcellular compartment. However, proteomics is also a powerful tool for detecting changes in the protein composition in response to developmental or environmental stimuli, so-called differential proteomics. Currently, proteomics is playing an important role in: (i) understanding plant biology, (ii) developing plant biomarkers for human health and food security and (iii) food analysis and bio-safety issues[38]. Most of the proteomic studies of seed development have used whole seeds as the experimental material and only in a few cases has the proteome of embryo, endosperm and/or seed coat been studied separately (Table 1).

Table 1: Proteomic studies on seed development

Species	Development stage			Tissue	Reference
	Histo differentiation	Reserve deposition	Maturation drying		
Arabidopsis thaliana	5,7 DAF	9,11,13 DAF		Whole seed	Hajduch et al[18]., Meyer et al[36].,
Brassica campestri	10, 16, 20, 25 DAF	35 DAF		Whole seed	Li et al[29].,
Brassica napus (rapeseed)	2 WAF	3, 4, 5 WAF	6 WAF	Whole seed	Hajduch et al[17]., Meyer et al[36].,
Glycine max (soybean)	2, 3 WAF	4, 5, 6 WAF		Whole seed	Hajduch et al., Agrawal et al[1]., Meyer et al[36].,
Hordeum vulgare (barley)		Stage 80, 82	Stage 85, 86, 87	Whole seed	Finnie et al[11].,
Jatropha curcas	5, 10, 15 DAF	20, 25, 30 DAF		Endosperm	Liu et al[32].,
Medicago truncatula		12, 14, 16, 18, 20 DAF 12, 14, 16, 20, 24, 36 DAF		Whole seed Embryo, endosperm, seedcoat	Gallardo et al[15].,Gallardo et al[14].,
Oryza sativa (rice)	5, 7 DAF 6, 8 DAF	13 DAF 10, 12, 14, 16, 18, 20 DAF	21, 30 DAF	Embryo Whole seed	Xu et al[53]., Xu et al[54].,
Ricinus communis(castor)	2, 3 WAF	4, 5, 6 WAF		Whole seed	Houston et al[21].,
Triticum aestivum (wheat)	Stage I, II	Stage III	Stage IV,V	Whole seed	Guo et al.,
Zea mays (maize)		17, 22, 25, 28 DAF	40, 65 AF	Embryo and endosperm Embryo and endosperm	Jin et al[23]., Wang et al[50].,

DAF: days after flowering.
WAF: weeks after flowering.

The frequently used separation and identification technologies are gel based approaches, most combining 2-dimensional polyacrylamide gel electrophoresis (separated according to their net charge in the first dimension and according to their molecular mass in the second dimension) followed by liquid chromatography and the acquisition of protein structural information for the purposes of protein identification by mass spectrometry (it measure the masses and relative concentrations of atoms and molecules. It makes use of the basic magnetic force on a moving charged particle). In this way, qualitative and quantitative changes in the proteome during seed development, desiccation tolerance, germination, dormancy release, vigor alteration and responses to environmental factors have all been studied[49].

In Brassica napus, seed yield and quality are related to sulfate availability, but the seed metabolic changes in response to sulfate limitation remain largely unknown. To address this question, proteomics and biochemical studies were carried out on mature seeds obtained from plants grown under low sulfate applied at the bolting (LS32), early flowering (LS53), or start of pod filling (LS70) stage. The protein quality of all low-sulfate seeds was reduced and associated with a reduction of S-rich seed storage protein accumulation (as Cruciferin Cru4) and an increase of S-poor seed storage protein (as Cruciferin BnC1). Proteins involved in plant stress response, such as dehydroascorbate reductase and Cu/Zn-superoxide dismutase, were also accumulated in LS53 and LS32 seeds, and this might be a consequence of reduced glutathione content under low S availability. If S reserves are sufficient, B. napus seeds are able to employ specific metabolic compensations to maintain seed yield and/or quality in the case of S limitation[20]. To address the problem of iron-deficiency anemia, one of the most prevalent human micronutrient deficiencies globally, iron-biofortified rice was produced using three transgenic approaches: 1. Enhanced Fe storage in grains via expression of the Fe storage protein ferritin using endosperm-specific promoters. 2. Enhanced Fe translocation through overproduction of the natural metal chelator nicotianamine. 3. Enhanced Fe flux into the endosperm through expression of the Fe (II)-NA transporter OsYSL2 under the control of an endosperm-specific promoter and sucrose transporter promoter in rice seeds[35]. Barros et al[3]., applied three profiling technologies to compare the transcriptome, proteome, and metabolome of two transgenic maize lines with the respective control line, and revealed that the environment was shown to play an important effect in the protein, gene expression and metabolite levels of the maize samples. Elevated ozone reduces the grain yield in Kirara 397 and Takanari, but not in Koshihikari. Ozone exposure decreased proteins associated with photosynthesis and glycolysis in Kirara 397, but not in Koshihikari. Especially, Kiran 397 showed a remarkable decrease in proteins related to photosynthetic electron transport, suggesting that the elevated ozone suppressed the photosynthetic apparatus in Kirara 397, thereby suppressing yield[47]. Proteomics also used to identify pathogenesis-related (PR) proteins in rice[57], soybean[55]. Stress responsive proteins were accumulated during the seed maturation in order to survive from dessication. Among them, HSPs were the largest group. Proteins including universal stress proteins (USP), dehydrins, DNA J family, and late embryogenesis abundant (LEA) proteins can also be regarded as desiccation stress responsive proteins[19,58]. Proteome analyses on pod and seed development were performed in Lotus japonicus[9,41].

METABOLOMICS

Identification and quantification of all metabolites in a biological system. Profiling the metabolome may actually provide the most 'functional' information of all of the OMICS technologies by giving a broad view of the biochemical status of an organism that can be used to monitor significant metabolite variations. Metabolomics is an advanced, specialised form of analytical biochemistry.

Metabolomics technology focuses on the detection of small molecules which often plays key role in seed quality, disease resistance, antioxidant activity etc. Important small metabolites include organic acids, amino acids, sugars, volatile metabolites and secondary metabolites such as alkaloids, phenolic components and also pigments such as carotenoids and anthocyanins[45]. Seed development can be divided into three stages:morphogenesis, maturation, and desiccation. During these processes, various metabolites such as amino acids, sugar alcohols (e.g. erythritol, arabitol, sorbitol and mannitol), betaines (e.g. trigonelline) and oligosaccharides (such as trehalose) accumulate in seeds[40]. Seed metabolites are directly related to seed quality and nutritional values[37], so information on the chemical compositions of major crop seeds is essential for modern agricultural breeding programs that utilize new technologies, such as metabolic engineering, to improve seed yield and nutritional values.

Chromatography (Liquid or Gas) are usually used to separate complex plant extracts into the individual components; Mass Spectrometry (MS) is then used to detect and if possible, quantify the metabolites present. Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS), liquid chromatography mass spectrometry (LC-MS), capillary electrophoresis mass spectrometry (CE-MS), Alternatively, Nuclear Magnetic Resonance (NMR) are routinely used in plant sciences[42]. Metabolomics is primarily distinguished from more established technologies by the high throughput nature of the approach which also generates a complex dataset per analysis. As a consequence, the technology relies heavily on recent advances made in bioinformatics and information technology in order to process, store and mine the complex data matrices for biological information. Metabolomics is usually used either for 'fingerprinting' samples to perform comparative analyses to detect differences of for 'profiling' where individual differential metabolites, perhaps linked to key quality traits, are identified for further analysis. In Arabidopsis, genome-wide association study (GWAS) analysis associated several lignin precursors with cinnamoyl- CoA, a target for improving the quality of lignocellulosic biomass by genetic engineering[44], identified two major polymorphic loci controlling glucosinolate variation in natural populations related to plant defense[8], and verified that allelic variation at BCAT2 is responsible for the natural variation of seed branched-chain amino acid levels[2].

Lin et al[30]., investigated the seed metabolomes of 29 common soybean cultivars through combined gas chromatography-mass spectrometry and ultra-performance liquid chromatography-tandem mass spectrometry. 169 named metabolites were identified and subsequently used to construct a metabolic network of mature soybean seed. Among the 169 detected metabolites, 104 were found to be significantly variable in their levels across tested cultivars. Metabolite markers that could be used to distinguish genetically related soybean cultivars were also identified, and metabolite- metabolite correlation analysis revealed some significant associations within the same or among different metabolite groups. Findings from this work may potentially provide the basis for further studies on both soybean seed metabolism and metabolic engineering to improve soybean seed quality and yield. Hu et al[22]., profiled 121 metabolites in mature seeds of a wide panel Oryza sativa japonica and indica cultivars, revealing correlations between the metabolic phenotype and geographic origin of the rice seeds. Moreover, japonica and indica subspecies differed significantly not only in the relative abundances of metabolites but also in their corresponding metabolic association networks. These findings provide important insights into metabolic adaptation in rice subgroups, bridging the gap between genome and phenome, and facilitating the identification of genetic control of metabolic properties that can serve as a basis for the future improvement of rice quality via metabolic engineering.

SUMMARY

For many plants, seed is essential to reproduce and disperse their progenies. It is also an adaptive strategy for plant survival under stresses. Seed germination, which is an early and crucial stage in plant life cycle, refers to the physiological process starting from the uptake of water by the dry seed and ending with the radicle protrusion. Seed is vital for propagation of spermatophytes in biome and as food source for inhabitants of the earth. Studies on OMICS provide platform for new avenues to explore molecular networks and pathways governing seed filling, maturation, germination, and seedling formation. Genomics provides an overview of the complete set of genetic instructions provided by the DNA, while transcriptomics looks in to gene expression pattern. Proteomics studies dynamic protein products and their interaction, while metabolomics is also an intermediate step in understanding organism's entire metabolism.

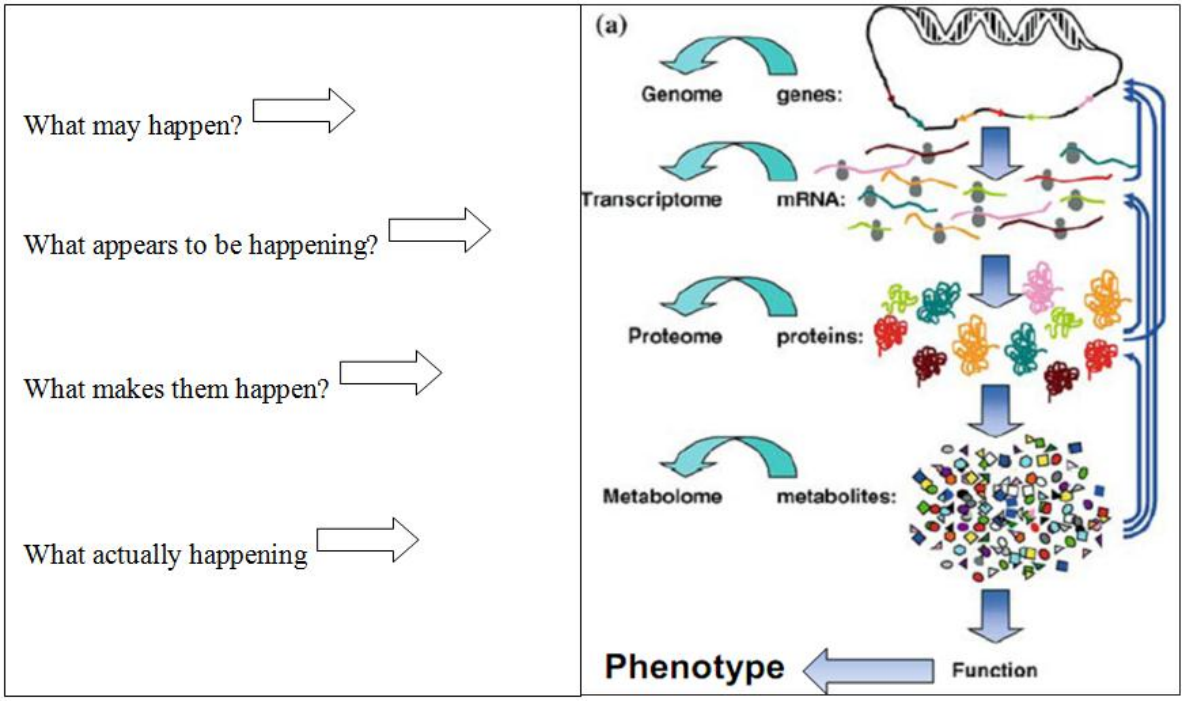

Fig. 1: Schematic of the 'omic hierarchy: genomics, transcriptomics, proteomics, and metabolomics[16]

CONCLUSION

The twenty-first century OMICS technologies play an essential role to improve seed development, seed quality and crop yield. The OMICS branches are equally important to get clear picture of the biological system. Investigations on gene functions at certain seedlings developmental stage such as seed filling or embryogenesis could reveal critical component regulating important metabolic process that could be exploited for improving seed quality. Advances in genetic engineering and transgenic technologies are contributing to applications and translation of several of the key recent discoveries in crop plants. These include development of oil seed crops with improved nutritional and higher oil content. It is expected that future discoveries coming from integrated systems biology based approaches will further enrich our knowledge base and accelerate the identification of key genes and genetic pathways to meet the challenges of improving the seed traits in crop species.

REFERENCES

1. Agrawal, G.K., Hajduch, M., Graham, K., Thelen, J.J., In-depth investigation of the soybean seed filling proteome and comparison with a parallel study of rapeseed. Plant Physiol., **148:** 504-518 (2008).

2. Angelovici, R., Lipka, A.E., Deason, N., Jorge, S.G., Lin, H., Cepela, J., Buell, R., Gore, M.A. and Penna, D.D., Genome-wide analysis of branched chain amino acid levels in Arabidopsis seeds. Plant Cell., **25:** 4827-4843 (2013).
3. Barros, E., Lezar, S., Anttonen, M.J., Dijk, J.P., Rohlig, R.M., Kok, E.J., Engel, K.H., Comparison of two GM maize varieties with a near isogenic non GM variety using transcriptomics, proteomics and metabolomics. Plant Biotechnol J., **8:** 436-451 (2010).
4. Bergmann, S., Ihmels, J. and Barkai, N., Iterative signature algorithm for the analysis of large-scale gene expression data. Phys. Rev. E Stat. Nonlin Soft Matter Phys., **67:** 031902 (2003).
5. Beyer, P., Al-Babili, S., Ye, X., Lucca, P., Schaub, P., Welsch, R., Potrykus, I., Golden rice: introducing the β-carotene biosynthesis pathway into rice endosperm by genetic engineering to defeat vitamin a deficiency. J. Nutr., **132:** 506-510 (2002).
6. Bradford, K.J, Chen, F., Cooley, M.B., Dahal, P., Downie, B., Fukunaga, K.K., Gee, O. H., Gurusinghe, S., Mella, R.A., Nonogaki, H., Wu, C.T., Yang, H., Yim, K.O., Gene expression prior to radicle emergence in imbibed tomato seeds. In: seed biology: advances and applications. Proceedings of the sixth international workshop on seeds, Merida, Mexico, 1999, Black M, Bradford KJ,Vazquez-Ramos J (eds), CABI, CABI International, Wallingford, pp 231-251 (2000).
7. Brenchley, R., Spannagl, M., Pfeifer, M., Barker, G.L.A., Amore, R., Allen, A.M., McKenzie, N., Kramer, M., Kerhornou, A. and Bolser, D., Analysis of the bread wheat genome using whole-genome shotgun sequencing. Nature, **491:** 705-710 (2012).
8. Chan, E.K., Rowe, H.C. and Kliebenstein, D.J., Understanding the evolution of defense metabolites in Arabidopsis thaliana using genome wide association mapping. Genetics. **185:** 991-1007 (2010).
9. Dam, S., Laursen, B.S., Ornfelt, J.H., Jochimsen, B., Staerfeldt, H.H., Friis, C., The proteome of seed development in the model legume Lotus japonicus. Plant Physiol., **149:** 1325-1340 (2009).
10. Figueroa, B.E.B., Gao, L., Diop, N.N., Wu, Z., Ehlers, J.D., Roberts, P.A., Close, T.J., Zhu, J.K., and Liu, R., Identification and comparative analysis of drought-associated microRNAs in two cowpea Genotypes. BMC Plant Biology. **11:** 127-130 (2011).
11. Finnie, C., Melchior, S., Roepstorff, P., Svensson, B., Proteome analysis of grain filling and seed maturation in barley. Plant Physiol., **129:** 1308-1319 (2002).
12. Fu, X., Guo. S., Yan, Z., Xu, Y., Hu, H., Menzel, C., Chen, W., Li., Y., Zeng, R. and Khaitovich, P., Estimating accuracy of RNA-seq and microarrays with proteomics. BMC Genomics. **10:** 161 (2009).
13. Furukawa, T., Maekawa, M., Oki, T., Suda, I., Iida, S., Shimada, H., Takamure, I. and Kadowaki, K., The Rc and Rd genes are involved in proanthocyanidin synthesis in rice pericarp. The Plant J., **49:** 91-102 (2006).
14. Gallardo, K., Firnhaber, C., Zuber, H., Hericher, D., Belghazi, M., Henry, C., Kuster, H., Thompson, R., A combined proteome and transcriptome analysis of developing Medicago truncatula seeds: evidence for metabolic specialization of maternal and filial tissues. Mol. Cell. Proteomics, **6:** 2165-2179 (2007).
15. Gallardo, K., Signor, L.C., Vandekerckhove, J., Thompson, R.D., Burstin, J., Proteomics of Medicago truncatula seed development establishes the time frame of diverse metabolic processes related to reserve accumulation. Plant Physiol., **133:** 664-682 (2003).
16. Goodacre, R., School of Chemistry, The University of Manchester, PO Box 88, Sackville Street, Manchester, M60 1QD, UK (2005).
17. Hajduch, M., Casteel, J.E., Hurrelmeyer, K.E., Song, Z., Agrawal, G.K., Thelen,

J.J., Proteomic analysis of seed filling in Brassica napus. Developmental characterization of metabolic isozymes using high-resolution two-dimensional gelelectrophoresis. Plant Physiol., **141:** 32-46 (2006).

18. Hajduch, M., Hearne, L.B., Miernyk, J.A., Casteel, J.E., Joshi, T., Agrawal, G.K., Song, Z., Zhou, M.Y., Xu, D., Thelen, J.J., Systems analysis of feed filling in Arabidopsis: using general linear modeling to assess concordance of transcript and protein expression. Plant Physiol., **152:** 2078-2087 (2010).
19. He, D., Han, C., Yao, J., Shen, S. and Yang, P., Constructing the metabolic and regulatory pathways in germinating rice seeds through proteomic approach. Proteomics, **11:** 2693-2713 (2011).
20. Hooghe, P.D., Dubousset, L., Gallardo, K., Kopriva, S., Avice, J.C. and Trouverie, J., Evidence for proteomic and metabolic adaptations associated with alterations of seed yield and quality in Sulfur-limited Brassica napus L. Mol. Cell. Proteomics. **13:** 1165-1183 (2014).
21. Houston, N.L., Hajduch, M., Thelen, J.J., Quantitative proteomics of seed filling in castor: comparison with soybean and rapeseed reveals differences between photosynthetic and nonphotosynthetic seed metabolism. Plant Physiol., **151:** 857-868 (2009).
22. Hu, C., Shi, J., Quan, S., Cui, B., Kleessen, S., Nikoloski, Z., Tohge, T., Alexander, D., Guo, L., Lin, H., Wang, J., Cui, X., Rao, J., Luo, Q., Zhao, X., Fernie, A.R. and Zhang, D., Metabolic variation between japonica and indica rice cultivars as revealed by non-targeted metabolomics. Sci. rep., **4:** 5067 (2014).
23. Jin, X.N., Fu, Z.Y., Ding, D., Li, W.H., Liu, Z.H., Tang, J.H., Proteomic identification of genes associated with maize grain filling rate. PLoS One, **8:** e59353 (2013).
24. Kell, D.B. and Oliver, S.G., Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era. Bioessays, **26:** 99-105 (2004).
25. Kusano, M., Baxter, I., Fukushima, A., Oikawa, A., Okazaki, Y., Nakabayashi, R., Assessing metabolomic and chemical diversity of a soybean lineage representing 35 years of breeding. Metabolomics. **11:** 261-270 (2015).
26. Le, B.H., Cheng, C., Bui, A.Q., Wagmaister, J.A., Henry, K.F., Pelletier, J., Global analysis of gene activity during Arabidopsis seed development and identification of seed-specific transcription factors. Proc. Natl. Acad. Sci. U.S.A., **107:** 8063-8070 (2010).
27. Le, B.H., Wagmaister, J.A., Kawashima, T., Bui, A.Q., Harada, J.J., Goldberg, R.B., Using genomics to study legume seed development. Plant Physiol., **144:** 562-574 (2007).
28. Li, L., Hur, M., Lee, J.Y., Zhou, W., Song, Z., Ransom, N., A systems biology approach toward understanding seed composition in soybean. BMC Genomics 16:1471-2164 (2015).
29. Li, W., Gao, Y., Xu, H., Zhang, Y., Wang, J., A proteomic analysis of seed development in Brassica campestri L. PLoS One. **7:** e50290 (2012).
30. Lin, H., Rao, J., Shi, J., Hu, C., Cheng, F., Wilson, Z.A., Zhang. D. and Quan, S., Seed metabolomic study reveals significant metabolite variations and correlations among different soybean cultivars. J Integr Plant Biol., **56:** 826-836 (2014).
31. Liu, L., Qu, C., Wittkop, B., Yi, B., Xiao, Y., He, Y., A high-density SNP map for accurate mapping of seed fibre QTL in Brassica napus L. PLoS ONE, **8(12):** e83052 (2013a).
32. Liu, H., Wang, C., Komatsu, S., He, M., Liu, G. and Shen, S., Proteomic analysis of the seed development in Jatropha curcas: from carbon flux to the lipid accumulation. J. Proteomics., **91:** 23-40 (2013b).

33. Lonneke, A. and Geest, H.M., Seed genomics: germinating opportunities. Seed Sci Res.,**12:** 145-153 (2002).
34. MASC, The Multinational Coordinated Arabidopsis thaliana Functional Genomics Project-Annual Report (2007).
35. Masuda, H., Ishimaru, Y., Aung, M.S., Kobayashi, T., Kakei, Y., Takahashi, M., Higuchi, K., Nakanishi, H. and Nishizawa, N.K., Iron biofortification in rice by the introduction of multiple genes involved in iron nutrition. Sci. rep., **2:** 543 (2012).
36. Meyer, L.J., Gao, J., Xu, D. and Thelen, J.J., Phosphoproteomic analysis of seed maturation in Arabidopsis, rapeseed, and soybean. Plant Physiol., **159:** 517-528 (2012).
37. Nambara, E. and Nonogaki, H., Seed biology in the 21st century: Perspectives and new directions. Plant Cell Physiol., **53:** 1-4 (2012).
38. Narula, K., Pandey, A., Gayali, S., Chakraborty, N. and Chakraborty, S., Birth of plant proteomics in India: a new horizon. J. Proteomics., **127:** 34-43 (2015).
39. Nemeth, C., Freeman, J., Jones, H.D., Sparks, C., Pellny, T.K., Wilkinson, M.D., Dunwell, J., Andersson, A.A.M., Aman, P. and Guillon, F., Downregulation of the CSLF6 gene results in decreased (1, 3; 1, 4)-β-d-glucan in endosperm of wheat. Plant Physiol., **152:** 1209-1218 (2010).
40. Obendorf, R.L., Oligosaccharides and galactosyl cyclitols in seed desiccation tolerance. Seed Sci. Res., **7:** 63-74 (1997).
41. Pedersen, N.G., Dam, S., Laursen, B.S., Siegumfeldt, A.L., Nielsen, K., Goffard, N., Proteome analysis of pod and seed development in the model legume Lotus japonicus. J. Proteome Res., **9:** 5715-5726 (2010).
42. Putri, S.P., Yamamoto, S., Tsugawa, H., and Fukusaki, E., Current metabolomics: technological advances. J. Biosci. Bioeng. **116:** 9-16 (2013).
43. Revilla, P., Butron, A.V., Rodriguez, M., Malvar, R.A. and Orda, A., Identification of genes related to germination in aged maize seed by screening natural variability. J. Exp. Bot., **14:** 4151-4157 (2010).
44. Riedelsheimer, C., Lisec, J., Eysenberg, A.C., Sulpice, R., Flis, A., Grieder, C., Altmann, T., Stitt, M., Willmitzer, L. and Melchinger, A.E., Genome-wide associationmapping of leafmetabolic profiles for dissecting complex traits in maize. Proc Natl Acad Sci., **109:** 8872-8877 (2012).
45. Saito, K. and Matsuda, F., Metabolomics for functional genomics, systems biology, and biotechnology. Annu Rev Plant Biol., **61:** 463-489 (2010).
46. Sanger, S., Nicklen, A.R. and Coulson, A., DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci., **74(12):** 5463-5467 (1977).
47. Sawada, H., Komatsub, S., Nanjob, Y., Khana, N.A., Kohno, Y., Proteomic analysis of rice response involved in reduction of grain yield under elevated ozone stress. Environmental and Experimental Botany. **77:** 108- 116 (2012).
48. Shewry, P.R., Improving grain quality: Wheat. Wiley-Blackwell. 159-178 (2013).
49. Wang, W.Q., Liu, s.j., Song, S.Q. and Moller, I.M., Proteomics of seed development, desiccation tolerance, germination and vigour. plant physiol. biochem., **86:** 1-15 (2015a).
50. Wang, W.Q., Ye, J.Q., Rogowska-Wrzesinska, A., Wojdyla, K.I., Jensen, O.N., Moller, I. M., Song, S.Q., Proteomic comparison between maturation drying and prematurely imposed drying of Zea mays seeds reveals a potential role of maturation drying in preparing proteins for seed germination, seedling vigor, and pathogen resistance. J. Proteome Res., **13:** 606-626 (2014b).
51. Wang, Z., Gerstein, M. and Snyder, M., RNA Seq: a revolutionary tool for transcriptomics. Nature Rev. Genet., **10:** 57-63 (2009c).
52. Xu, H., Wei, Y., Zhu, Y., Lian, L., Xie, H., Cai, Q., Chen, Q., Lin, Z., Wang, Z.,

Xie, H. and Zhang, J., Antisense suppression of LOX3 gene expression in rice endosperm enhances seed longevity. Plant Bio. J., **13:** 526-539 (2015a).

53. Xu, H., Zhang, W., Gao, Y., Zhao, Y., Guo, L., Wang, J., Proteomic analysis of embryo development in rice (Oryza sativa). Planta. **235:** 687-701 (2012b).
54. Xu, S.B., Li, T., Deng, Z.Y., Chong, K., Xue, Y.B., Wang, T., Dynamic proteomic analysis reveals a switch between central carbon metabolism and alcoholic fermentation in rice filling grains. Plant Physiol., **148:** 908-925 (2008c).
55. Yang, H., Huang, Y.P., Zhi, H.J., Yu, D.Y., Proteomics-based analysis of novel genes involved in response toward soybean mosaic virus infection. Mol Biol Rep., **38:** 511-521 (2011a).
56. Yang, H., Lu, P., Wang, Y., Ma, H., The transcriptome landscape of Arabidopsis male meiocytes from high throughput sequencing: the complexity and evolution of the meiotic process. Plant J., **65:** 503-516 (2011b).
57. Yang, L., Su, N., Wu, M., Wang, C., Hu, H., Proteomics to identify pathogenesis-related proteins in rice roots under water deficit. Biologia., **66:** 477-483 (2011c).
58. Yang, P., Li, X., Wang, X., Chen, H., Chen, F. and Shen, S., Proteomic analysis of rice (Oryza sativa) seeds during germination. Proteomics. **7:** 3358-3368 (2007).