

EDITORIAL

Functional genomics, challenges and perspectives for the future

Ron Mittler* and Vladimir Shulaev

Department of Biological Sciences, College of Arts and Sciences, University of North Texas, 1155 Union Circle #305220, Denton, TX, 76203-5017, USA

e-mail: ron.mittler@unt.edu

doi:10.1111/ppl.12060

With the advent of high throughput sequencing, proteomics and metabolomics tools, functional genomics and systems biology have become a central platform in plant sciences and beyond. The numerous plant genomes sequenced (and in progress of being sequenced) have generated a wealth of information on the composition, structure and organization of different plant genomes. This has led to the identification and annotation of 10 000s of plant genes. Nevertheless, determining the function of these genes, the networks they are involved in, and the genome as a whole remains a major challenge for modern plant research.

The functional genomics/systems biology platform is an extremely complex and powerful approach to determine the function of individual genes, pathways, networks and ultimately entire genomes (Fig. 1). The principal premise behind this approach is to evaluate and study the entire cell or organism as a system and understand how different biological processes occur within this system, how they are controlled and how they are executed. To study the system as a whole it is perturbed by, for a example, a mutation or a change in growth conditions, and is then studied using different sub-platforms that determine the response of the entire genome, transcriptome, metabolome, fluxome, ionome and any additional technique that address the entire set of components in the cell (e.g. Hrmova and Fincher 2009, Baxter 2010, Davies et al. 2010, Edwards and Batley 2010, Pedreschi et al. 2010, Saito and Matsuda 2010, Araújo et al. 2012, Berkman et al. 2012). Each sub-platform generates a large data set that is curated and stored in a database and different bioinformatics tools are then used to integrate the data, mine for specific genes, pathways and networks, annotate and generate a visualization output that attempts to link all the different components involved (e.g. Mochida and Shinozaki 2010, Higashi and Saito 2013). Modeling is then applied in an attempt to explain the dynamic behavior of the entire system and

generate new hypotheses, propose additional perturbations and assign a function to the different cellular components involved (e.g. Moreno-Risueno et al. 2010, Yin and Struik 2010). Of course assigning a function requires a phenotype and this is addressed by the phenomics platform (e.g. Kuromori et al. 2009, Kondou et al. 2010, Furbank and Tester 2011), as well as by the initiation of a new cycle of experiments using a slightly different perturbation of the system. The power of this approach lies in the fact that each individual perturbation results in a response that involves the entire network/system of the cell and encompasses 1000s of different genes, transcripts and metabolites, generating relational data sets that could be linked and produce multiple correlations that imply a function. The function is then tested via further perturbations and phenotypic analysis that could be a visually measured phenotype, such as altered growth or tolerance to a specific abiotic stress, or for example a metabolic phenotype that would be determined with one of the different platforms. The drawback of applying this approach to different biological systems is of course that it is extremely expensive and complex. Nevertheless, recent advances in technology had significantly reduced the cost of whole genome and transcriptome sequencing opening many new avenues of research.

Perhaps the most obvious advancement in technology in recent years was the development of Next Generation (NextGen) DNA sequencing platforms. These have increased our ability to sequence plant genomes (more than 100 plant genomes have already been fully sequenced), and decreased the cost of whole genome and whole transcriptome sequencing to a level that is highly accessible for individual researchers (e.g. Appleby et al. 2009, Edwards et al. 2013). The implications of this advancement are immense and far reaching in many fields including medicine, agriculture, public health, defense and more. In addition to the sequence data that generates a physical map of the genome, NextGen

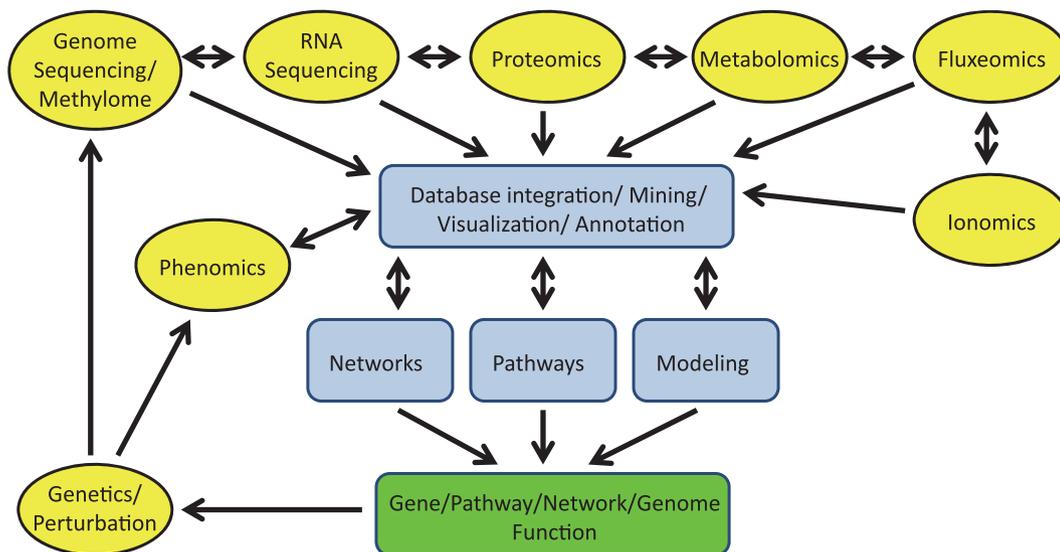


Fig. 1. The overall flow and hierarchy of functional genomics and systems biology. Flow chart showing the integration of the different components and sub-platforms of functional genomics and systems biology. Experimental platforms are shown in yellow, bioinformatics and modeling are shown in light blue, and function is shown in green.

sequencers can provide information on the methylation state of the entire genome (methylome), thereby enabling us to determine the epigenetic control of different genes and regulons within the genome. Of course the ability to sequence the entire population of cellular RNA (RNA-seq) using the NextGen sequencer enables us to study how transcripts are spliced and how many of the regulatory and structural RNAs function in the cell. Despite major advancement in the development of proteomics and metabolomics tools, these sub-platforms are still costly compared to NextGen sequencing. Another more recent advancement was the establishment of phenomics facilities and the increased availability of seed banks for insertional mutants in *Arabidopsis*, rice, and other plants (e.g. Kuromori et al. 2009, Kondou et al. 2010, Furbank and Tester 2011). As a result large systematic phenotypic screens are now being conducted not only by companies, but also by different consortiums and individual research laboratories.

One of the most notable developments in plant biology to date is a growing number of functional genomics projects focused on so-called 'nonmodel' plants, and a spike in translational research. The fast explosion in data generation and throughput techniques and the reduced cost of analysis made it possible to move the 'omics' research platform beyond model organism toward commercial crops, ornamentals and wild species (e.g. Sonah et al. 2011). This development provides a greater venue for faster integration of improved crops into agricultural practices.

Perhaps the biggest challenge facing functional genomics and systems biology to date is the development of bioinformatics tools that will integrate and analyze data obtained from multiple 'omics' platforms in a comprehensive way to generate a holistic view of cellular systems and networks. Another challenge is to integrate different relational databases, statistically dissect them and present the outcome in a way that would be comprehensive to humans. Several projects have been established to generate visualization tools and these are having varying degrees of success in presenting the complex data sets as linked networks of different cellular components (e.g. RNA/protein/metabolite) and regulatory networks (e.g. Mochida and Shinozaki 2010, Higashi and Saito 2013). It is not entirely clear however how the availability in the near future to 1000s of fully sequenced plant genomes would be integrated; and new levels of data organization and interrogation methods may need to be developed. Another challenge has to do with the fact that different cells within the organism respond differently. To comprehend how the organism as a whole responds to a perturbation would therefore require in vivo and in vitro imaging and micromanipulation techniques to visualize and sample individual cells or even organelles and subject them to functional genomics analysis using tools such as NextGen sequencing, metabolomics and proteomics. In addition, cell-to-cell communication would need to be integrated into the different modeling and analysis tools. The development of advanced screening and selection

methods, some at the single cell level, and/or using new and advanced reporter systems, would also significantly advance functional genomics and genetics studies in the near future.

When it comes to the future of functional genomics, even the sky is not the limit. The massive effort to sequence plant genomes coupled with the application of additional functional genomics tools would eventually yield an extremely high level of understanding of plant function, development and regulation. In the near future there are a number of different research avenues that would significantly benefit from the application of this platform (Fig. 2):

- (1) The development and design of functional foods could be enhanced by the application of functional genomics tools to identify beneficial pathways and engineer them into or out of specific plants and crops. This would result in the enhancement of food quality without the need to modify foods at the postharvest stage.
- (2) Most of the functional genomics studies conducted to date are performed on laboratory- or greenhouse-grown plants. The field environment is nevertheless significantly different from the conditions used in the lab and the development of better crops would be significantly enhanced if the functional genetics platform would be applied to plants/crops grown under real field conditions.
- (3) Functional genomics has been extensively used for crop improvements, typically in conjugation

with advanced breeding techniques (e.g. Tran and Mochida 2010, Kurowska et al. 2011, Langridge and Fleury 2011, Mir et al. 2012). The advent of genome and RNA sequencing is likely however to significantly facilitate the breeding process replacing many of the marker assisted breeding techniques with cheap whole genome or transcriptome studies. These could also be applied to different varieties and ecotypes of crops enhancing the availability of potential target genes for breeding.

- (4) Biofuel development is a hot topic these days and functional genomics is already applied to identify and engineer different pathways into plants and algae. The massive sequencing of different plant and algae genomes/transcriptomes, as well as the different metagenomics projects under way (e.g. Singh et al. 2009, Hettich et al. 2012), are likely to significantly enhance these efforts identifying, characterizing and engineering new sources for biofuels, some could even be from unexpected sources, or produced by proteins which currently have an unknown function.
- (5) The analysis of entire ecosystems and the potential for their management and modification would dramatically benefit from the lower cost of DNA/RNA sequencing (e.g. Friesen and Wettberg 2010, Whitehead 2012). Projects that involve the

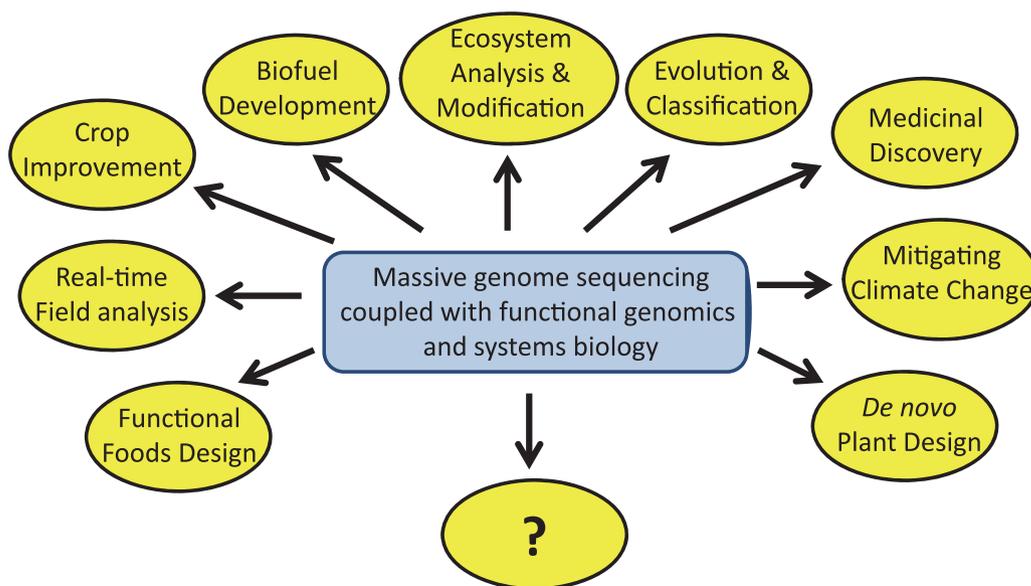


Fig. 2. Genome sequencing and functional genomics: where do we go from here? Flow chart showing some of the current and future applications of functional genomics and NextGen genome/RNA sequencing.

sequencing of 100s or even 1000s of individual plants genomes/transcriptomes from a particular ecosystem would be possible, precisely determining the genetic variability of the system. When other species will be included in this genome/transcriptome analysis, interactions between different species could be traced to particular genes and pathways, and new and novel interactions could be identified. Of course the health of the ecosystem and its response to global climatic changes could be determined as well at the genomics level.

- (6) The application of NextGen sequencing to 1000s of different plants and other organisms would create a wealth of data that would enhance our organism classification, as well as understanding of evolution (e.g. Paterson et al. 2010, Tirosh and Barkai 2011). Combined with the application of this method to fossils, ancient seeds and prehistoric animal remains, our understanding of evolutionary process at the genomics level would be significantly enhanced.
- (7) Medicinal discovery is perhaps one of the most important applications for functional genomics (e.g. Chen et al. 2011). Combined with advanced tools to detect compounds with a medicinal property, functional genomics could identify the precise pathways and genes involved in its biosynthesis. These could then be used for drug production under controlled conditions and the development of novel drugs with different properties.
- (8) Mitigating climate change, especially from the view point of increasing CO₂ consumption by plants and algae could be achieved by modifying existing plant and algal systems. Modeling and functional genomics could be used for this purpose and the resulting organisms could be grown on a mass produced based, especially in combination with biofuel production.
- (9) Perhaps the most exciting aspect of functional genomics and systems biology is the potential for de novo design of plants. De novo design of organisms is a newly emerging area of synthetic biology aimed at designing biological systems and components that do not exist in nature. Currently de novo design of organisms is limited to viruses, mycoplasma and perhaps in the near future bacteria. With the advent of DNA synthesis tools and our understanding of plant systems it is not impossible to imagine a future in which plants would be designed from scratch, put together in a tube and grown in the field.

The explosion of genomics data coupled with functional genomics, bioinformatics and systems biology is going to dominant the future of scientific research leading to the big question of what lies ahead beyond these platforms?

References

- Appleby N, Edwards D, Batley J (2009) New technologies for ultra-high throughput genotyping in plants. *Methods Mol Biol* 513: 19–39
- Araújo WL, Nunes-Nesi A, Williams TC (2012) Functional genomics tools applied to plant metabolism: a survey on plant respiration, its connections and the annotation of complex gene functions. *Front Plant Sci* 3: 210
- Baxter I (2010) Ionomics: The functional genomics of elements. *Brief Funct Genomics* 9: 149–156
- Berkman PJ, Lai K, Lorenc MT, Edwards D (2012) Next-generation sequencing applications for wheat crop improvement. *Am J Bot* 99: 365–371
- Chen S, Xiang L, Guo X, Li Q (2011) An introduction to the medicinal plant genome project. *Front Med* 5: 178–184
- Davies HV, Shepherd LV, Stewart D, Frank T, Röhlig RM, Engel KH (2010) Metabolome variability in crop plant species – when, where, how much and so what? *Regul Toxicol Pharmacol* 58: S54–S61
- Edwards D, Batley J (2010) Plant genome sequencing: applications for crop improvement. *Plant Biotechnol J* 8: 2–9
- Edwards D, Batley J, Snowdon RJ (2013) Accessing complex crop genomes with next-generation sequencing. *Theor Appl Genet* 126: 1–11
- Friesen ML, von Wettberg EJ (2010) Adapting genomics to study the evolution and ecology of agricultural systems. *Curr Opin Plant Biol* 13: 119–125
- Furbank RT, Tester M (2011) Phenomics-technologies to relieve the phenotyping bottleneck. *Trends Plant Sci* 16: 635–644
- Hettich RL, Sharma R, Chourey K, Giannone RJ (2012) Microbial metaproteomics: identifying the repertoire of proteins that microorganisms use to compete and cooperate in complex environmental communities. *Curr Opin Microbiol* 15: 373–380
- Higashi Y, Saito K (2013) Network analysis for gene discovery in plant-specialized metabolism. *Plant Cell Environ*. doi: 10.1111/pce.12069
- Hrmova M, Fincher GB (2009) Functional genomics and structural biology in the definition of gene function. *Methods Mol Biol* 513: 199–227
- Kondou Y, Higuchi M, Matsui M (2010) High-throughput characterization of plant gene functions by using gain-of-function technology. *Annu Rev Plant Biol* 61: 373–393
- Kuromori T, Takahashi S, Kondou Y, Shinozaki K, Matsui M (2009) Phenome analysis in plant species using

- loss-of-function and gain-of-function mutants. *Plant Cell Physiol* 50: 1215–1231
- Kurowska M, Daszkowska-Golec A, Gruszka D, Marzec M, Szurman M, Szarejko I, Maluszynski M (2011) TILLING: a shortcut in functional genomics. *J Appl Genet* 52: 371–390
- Langridge P, Fleury D (2011) Making the most of ‘omics’ for crop breeding. *Trends Biotechnol* 29: 33–40
- Mir RR, Zaman-Allah M, Sreenivasulu N, Trethowan R, Varshney RK (2012) Integrated genomics, physiology and breeding approaches for improving drought tolerance in crops. *Theor Appl Genet* 125: 625–645
- Mochida K, Shinozaki K (2010) Genomics and bioinformatics resources for crop improvement. *Plant Cell Physiol* 51: 497–523
- Moreno-Risueno MA, Busch W, Benfey PN (2010) Omics meet networks – using systems approaches to infer regulatory networks in plants. *Curr Opin Plant Biol* 13: 126–131
- Paterson AH, Freeling M, Tang H, Wang X (2010) Insights from the comparison of plant genome sequences. *Annu Rev Plant Biol* 61: 349–372
- Pedreschi R, Hertog M, Lilley KS, Nicolai B (2010) Proteomics for the food industry: opportunities and challenges. *Crit Rev Food Sci Nutr* 50: 680–692
- Saito K, Matsuda F (2010) Metabolomics for functional genomics, systems biology, and biotechnology. *Annu Rev Plant Biol* 61: 463–489
- Singh AH, Doerks T, Letunic I, Raes J, Bork P (2009) Discovering functional novelty in metagenomes: examples from light-mediated processes. *J Bacteriol* 191: 32–41
- Sonah H, Deshmukh RK, Singh VP, Gupta DK, Singh NK, Sharma TR (2011) Genomic resources in horticultural crops: status, utility and challenges. *Biotechnol Adv* 29: 199–209
- Tirosh I, Barkai N (2011) Inferring regulatory mechanisms from patterns of evolutionary divergence. *Mol Syst Biol* 7: 530
- Tran LS, Mochida K (2010) Functional genomics of soybean for improvement of productivity in adverse conditions. *Funct Integr Genomics* 10: 447–462
- Whitehead A (2012) Comparative genomics in ecological physiology: toward a more nuanced understanding of acclimation and adaptation. *J Exp Biol* 215: 884–891
- Yin X, Struik PC (2010) Modelling the crop: from system dynamics to systems biology. *J Exp Bot* 61: 2171–2183