



# Grand challenges in plant systems biology: closing the circle(s)

Alisdair R. Fernie\*

Department of Molecular Physiology, Max-Planck-Institute of Molecular Plant Physiology, Potsdam, Germany

\*Correspondence: fernie@mpimp-golm.mpg.de

## WHAT IS SYSTEMS BIOLOGY?

A prerequisite for any article on systems biology, despite the fact that it is no longer a novel concept, is to define what is meant by the term. Two definitions which cover what most people mean when they use the term are:

The reductionist approach has successfully identified most of the components and many of the interactions but, unfortunately, offers no convincing concepts or methods to understand how system properties emerge...the pluralism of causes and effects in biological networks is better addressed by observing, through quantitative measures, multiple components simultaneously and by rigorous data integration with mathematical models Sauer et al. (2007).

and

Systems biology...is about putting together rather than taking apart, integration rather than reduction. It requires that we develop ways of thinking about integration that are as rigorous as our reductionist programmes, but different... It means changing our philosophy, in the full sense of the term. Noble (2006).

The definition of Sauer is arguably the broadest in that it encompasses both experimental and theoretical mathematical aspects and as such is one that I personally favor. The definition of Noble, as well as that of Kitano (2004) in his seminal paper on systems biology is somewhat more controversial. Indeed the view that the systems biology approach represents a paradigm shift can be and is quite frequently challenged. For example Bothwell (2006) quite correctly argues that systems biology has had a long past, he exemplifies this stance by citing Aristotle, Wilhelm Pfeffer, Charles Darwin, the field of cybernetics and the first winner of the Nobel prize for Chemistry, Jacobus van't Hoff, as exponents of systems biology. Similarly,

Robeva (2010) argues that scientists and philosophers alike, including Kant, Goethe, Schrödinger, and Delbruck, in discussing that a living organism is more than a list of its parts essentially were already viewing life from a holistic perspective. Undoubtedly Popper (1934) could and should be added to this list. As such defining systems biology as a new kind of biology or the successor to molecular biology as has been done in the past appears rather disingenuous. To the above lists can also certainly be added Kacser and Burns (1973) and Heinrich and Rapoport (1974) for their development of metabolic control analysis – a theoretical framework describing the extent to which a given enzyme controls the metabolite flux through entire pathways. This theory was the first attempt in computational biology to establish the relation between the parts with the sum theorem clearly demonstrating that properties of biological relevance exist that depend on the system as a whole. As such it paved the way for further modeling approaches looking at metabolic pathways and networks that are of considerably greater scale (see for example Forster et al., 2003). What has changed from the times of even the most recent of these works is, however, the magnitude of available data. The implementation of post-genomic technologies and the parallel increase in computational capacities facilitates the collection of ever-expanding datasets has facilitated the collection of vast datasets from microarrays (see for example Schena et al., 1995) and mass-spectrometric analysis of proteins and metabolites (see for example Baerenfaller et al., 2008; Caldana et al., 2011). The advent of next generation sequencing has already rapidly accelerated the rate at which genome sequences are appearing and also dramatically increased the coverage and accuracy of transcript profiling methodologies (Jiménez-Gómez, 2011; Schneeberger and Weigel, 2011). These developments along-

side the large-scale web-based searchable databases of transcript, protein, and metabolite abundances have however certainly dramatically altered the scientific landscape (<https://www.genevestigator.com>; Bais et al., 2010; Joshi et al., 2011). Such massively multi-parallel approaches are of course not without their own issues with careful controls being required to ensure adequate data quality (Clarke and Zhu, 2006; Fernie et al., 2011; Tohge et al., 2011) and tightly controlled vocabularies needed to allow accurate comparisons to be made across datasets (see for example The Gene Ontology Consortium, 2012). That said the research communities focusing on such data acquisition are by and large already aware of these problems and are actively pursuing approaches to ameliorate them. Whilst this is certainly a challenge I believe that efficiently computationally mining these data to better understand metabolism and nutrition, growth and development, and responses to the environment in short *understanding of the complexities in all aspects of plant responses* represents the grand challenge of plant systems biology.

## SEMINAL PAPERS LAYING THE FOUNDATION FOR (PLANT) SYSTEMS BIOLOGY

It is all well and good throwing down the gauntlet, however, more important is to suggest possible ways in which the challenge can be addressed I will come back to this in the final section of the article but first I would like to highlight some important conceptual and experimental works which I believe represent solid foundations for the field. There are already many excellent reviews covering aspects of systems biology (Ideker et al., 2001; Oltvai and Barabasi, 2002; Kitano, 2004; Somerville et al., 2004) and although historically systems biology was applied exclusively to mathematical modeling strategies (see for example Edwards and Palsson, 1999), it is

now more widely applied particularly with respect to genomics (Baginsky and Fernie, 2007). This said as alluded to above I think that the scope of systems biology is well illustrated by several key papers only a few of which cover plant function. These examples fit into what Breitling (2010) refers to as the esthetic foundations of systems biology in addressing either simplicity, complexity, or diversity. The quest for simplicity is perhaps best illustrated in Kitano (2002) framework of systems biology in which life is described in terms of engineering principles defining system structures, system dynamics, control, and design methods. Modeling of regulatory circuitry also is rooted in this approach as is the search for recurring network motifs capable of implementing similar functions at multiple positions within biological networks (see for example Milo et al., 2002; Mangan and Alon, 2003; Tyson et al., 2003). In plants, one of the nicest examples of systems biology meeting the challenge was provided by the work of Locke et al. (2006) who experimentally validated a feedback loop predicted by modeling of the multi-oscillator clock of *Arabidopsis thaliana*. Similar efforts are currently being advanced in order to model various aspects of plant regulation including that of root, shoot, and leaf development (Bilsborough et al., 2011; Muraro et al., 2011; Vernoux et al., 2011).

Complexity is arguably also best addressed in terms of networks with the work of Barabasi and Oltvai (2004) being instrumental in popularizing the adoption of network perspectives. Such approaches analyze the networks and assign importance to individual network elements (nodes and edges) from purely topological/structural point of view. Important examples of this approach include gene and transcriptional networks, protein interaction networks signaling, and metabolic networks (van Lenne et al., 2010; Vernoux et al., 2011). A second approach is to include stoichiometry and one to one relationships which pervade biological systems. In terms of metabolism the widespread adoption of genome-scale metabolic networks represents a very good example of this approach. Models have now been published for 10s of species and several plant specific networks being published in the last few years (for a review see Sweetlove and Ratcliffe, 2011). These models are not

only finding utility in the understanding of metabolic regulation and as foundation for metabolic engineering but also can be used to solve evolutionary problems (Bekaert et al., 2011). Another approach, kinetic modeling, provides classical cases in which non-intuitive understanding has been provided by computational analysis. A recent article using this approach to decipher the regulation of plant aspartate metabolism is an excellent example of the utility of this approach (Curien et al., 2009). A further alternative approach commonly taken here is the modeling of experimental evaluation of isotope redistribution – this approach is very much in renaissance driven by the increased availability of mass-spectrometry based detection systems (Schwender, 2011). However, it is important to note that the types of modeling required in the different instances mentioned above varies greatly.

Diversity is as yet relatively untapped as a resource for systems biology despite being widely used as a tool for quantitative genetics (Koornneef et al., 2004). However, even in plants, systems biology approaches are being adopted which include the data obtained from quantitative genetic screens in a manner which embraces modeling in addition to the statistical analyses on which they are founded. A theoretical expansion of this premise was recently published by Kliebenstein (2010) Moreover, in experimental work from his group mapping of network quantitative loci was used as a method to uncover linkages between metabolic pathways and the clock in *Arabidopsis* (Kerwin et al., 2011).

As the example of Locke et al. (2006) illustrates systems biology is perhaps most effective as an iterative strategy in which properties of cellular/organismal systems are acquired and modeled as the first “half of the cycle” then hypotheses are generated and experimentally tested as the “second half of the cycle.” To date relatively few studies in plants reported to date complete such cycles. In the microbial field pioneering work of Palsson has shown how effective the layering of experimentally derived data atop of network models also represents a highly powerful validation tool (Edwards et al., 2001; Cho et al., 2011). First, examples of this approach in plants are now appearing (Williams et al., 2010) and it would seem likely that, following on from the considerable success of efforts aimed at integrating

data obtained from plants at different levels of the cellular hierarchy, such approaches will prove highly successful in the future.

## A FUTURE PERSPECTIVE

It is increasing apparent that adoption of the engineering principles of (plant) systems biology will provide a solid foundation for metabolic engineering and synthetic biology (Breitling, 2010; Frommer, 2010). The latter of these is in its infancy in plants, however, it is important to note that successful attempts to transfer the production of important secondary metabolites from plants to microbes have been reported (Facchini et al., 2012). Another challenge for systems biology in plants, as opposed to microbes, is the additional complexities that arise in a multicellular organism, given that these features are not unique to plants but are common to many eukaryotes. With this in mind initial attempts at defining tissue-specific (Mintz-Oron et al., 2012) and environment-specific (Sajitz-Hermstein and Nikoloski, 2010) genome-scale networks represent an important advance. It will be important for researchers active in this area to maintain a broad interest in the parallel developments of systems biology in non-plant systems.

The future development of plant systems certainly requires increasing capacities at theoretical and experimental levels to develop both independently and in concert. Such advances in systems biology and associated disciplines including acquisition and statistical analyses of quantitative data will ultimately be required to reach the ultimate aim of a comprehensive understanding of the complexities of plant responses. Given the inherent multidisciplinary nature of systems biology it would appear likely that this challenge will best be met via collaboration between experimental and computationally minded scientists the impact of dramatic future developments that either community makes in isolation is impossible to foresee. Another challenge that is yet to be fully addressed is the integration of networks from different levels of the system in order to connect heterogeneous data. With this in mind I see no value in precluding any of the myriad of available and emergent opportunities and approaches in addressing this challenge and work embracing computation should be equally encouraged as that facilitating the acquisition and/or databasing of information required

for computation. In addition integration of various types and models will certainly require additional theoretical developments and experimental design will likely need to be periodically reviewed in order to fully exploit developing opportunities.

## ACKNOWLEDGMENTS

I thank Lee Sweetlove and Zoran Nikoloski for their constructive criticism of the manuscript.

## REFERENCES

- Baerenfaller, K., Grossmann, J., Grobei, M. A., Hull, R., Hirsch-Hoffmann, M., Yalovsky, S., Zimmermann, P., Grossniklaus, U., Gruissem, W., and Baginsky, S. (2008). Genome-scale proteomics reveals *Arabidopsis thaliana* gene models and proteome dynamics. *Science* 320, 938–941.
- Baginsky, S., and Fernie, A. R. (2007). *Plant Systems Biology*. Basel: Birkhäuser.
- Bais, P., Moon, S. M., He, K., Leitao, R., Dreher, K., Walk, T., Sudaet, Y., Barkan, L., Wohlgenuth, G., Roth, M. R., Wurtele, E. S., Dixon, P., Fiehn, O., Lange, B. M., Shulaev, V., Sumner, L. W., Welti, R., Nikolau, B. J., Rhee, S. Y., and Dickerson, J. A. (2010). PlantMetabolomics.org: a web portal for plant metabolomics experiments. *Plant Physiol.* 152, 1807–1816.
- Barabasi, A. L., and Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nat. Rev. Genet.* 5, 101–115.
- Bekaert, M., Edger, P. P., Pires, J. C., and Conant, G. C. (2011). Two-phase resolution of polyploidy in the *Arabidopsis* metabolic network gives rise to relative and absolute dosage constraints. *Plant Cell* 23, 1719–1728.
- Bilborough, G. D., Runions, A., Barkoulas, M., Jenkins, H. W., Hasson, A., Galinha, C., Laufs, P., Hay, A., Prusinkiewicz, P., and Tsiantis, M. (2011). Model for the regulation of *Arabidopsis thaliana* leaf margin development. *Proc. Natl. Acad. Sci. U.S.A.* 108, 3424–3430.
- Bothwell, J. H. F. (2006). The long past of systems biology. *New Phytol.* 170, 6–10.
- Breitling, R. (2010). What is systems biology? *Front. Physiol.* 1:9. doi: 10.3389/fphys.2010.00009
- Caldana, C., Degenkolbe, T., Cuadros-Inostroza, A., Klie, S., Sulpice, R., Leisse, A., Steinhauser, D., Fernie, A. R., Willmitzer, L., and Hannah, M. A. (2011). High-density kinetic analysis of the metabolomic and transcriptomic response of *Arabidopsis* to eight environmental conditions. *Plant J.* 67, 869–884.
- Cho, B. K., Federowicz, S., Park, Y. S., Zengler, K., and Palsson, B. O. (2011). Deciphering the transcriptional regulatory logic of amino acid metabolism. *Nat. Chem. Biol.* 13, 65–71.
- Clarke, J. D., and Zhu, T. (2006). Microarray analysis of the transcriptome as a stepping stone towards understanding biological systems: practical considerations and perspectives. *Plant J.* 45, 630–650.
- Curien, G., Bastien, O., Robert-Genthon, M., Cornish-Bowden, A., Cárdenas, M. L., and Dumas, R. (2009). Understanding the regulation of aspartate metabolism using a model based on measured kinetic parameters. *Mol. Syst. Biol.* 5, 271.
- Edwards, J. S., Ibarra, R. U., and Palsson, B. O. (2001). In silico predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nat. Biotechnol.* 19, 125–130.
- Edwards, J. S., and Palsson, B. O. (1999). Systems properties of the *Haemophilus influenzae* Rd metabolic genotype. *J. Biol. Chem.* 274, 17410–17416.
- Facchini, P. J., Bohlmann, J., Covello, P. S., De Luca, V., Mahadevan, R., Page, J. E., Ro, D. K., Sensen, C. W., Storms, R., and Martin, V. J. J. (2012). Synthetic biosystems for the production of high-value plant metabolites. *Trends Biotechnol.* Available at: <http://dx.doi.org/10.1016/j.tibtech.2011.10.001>
- Fernie, A. R., Aharoni, A., Willmitzer, L., Stitt, M., Tohge, T., Kopka, J., Carroll, A. J., Saito, K., Fraser, P. D., and DeLuca, V. (2011). Recommendations for reporting metabolite data. *Plant Cell* 23, 2477–2482.
- Forster, J., Famili, I., Fu, P., Palsson, B. O., and Nielsen, J. (2003). Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. *Genome Res.* 13, 244–253.
- Frommer, W. B. (2010). Grand opportunities in physiology to address the grand challenges facing the planet. *Front. Physiol.* 1:11. doi: 10.3389/fphys.2010.00011
- Heinrich, R., and Rapoport, T. A. (1974). Linear steady-state treatment of enzymatic chains – general properties, control and effector strength. *Eur. J. Biochem.* 42, 89–95.
- Ideker, T., Galitski, T., and Hood, L. (2001). A new approach to decoding life: systems biology. *Annu. Rev. Genomics Hum. Genet.* 2, 343–372.
- Jiménez-Gómez, J. M. (2011). Next generation quantitative genetics in plants. *Front. Plant Sci.* 2:77. doi: 10.3389/fpls.2011.00077
- Joshi, H. J., Hirsch-Hoffmann, M., Baerenfaller, K., Gruissem, W., Baginsky, S., Schmidt, R., Schulze, W. X., Sun, Q., van Wijk, K. J., Egelhofer, V., Wienkoop, S., Weckwerth, W., Bruley, C., Rolland, N., Toyoda, T., Nakagami, H., Jones, A. M., Briggs, S. P., Castleden, I., Tanz, S. K., Millar, A. H., and Heazlewood, J. L. (2011). MASCP Gator: an aggregation portal for the visualization of *Arabidopsis* proteomics data. *Plant Physiol.* 155, 259–270.
- Kacser, H., and Burns, J. A. (1973). The control of flux. *Symp. Soc. Exp. Biol.* 28, 65–104.
- Kerwin, R. E., Jimenez-Gomez, J. M., Harmer, S. L., Maloof, J. N., and Kliebenstein, D. J. (2011). Network quantitative trait loci mapping of circadian clock outputs identifies metabolic pathway-to-clock linkages in *Arabidopsis*. *Plant Cell* 23, 471–485.
- Kitano, H. (2002). Systems biology: a brief overview. *Science* 5560, 1662–1664.
- Kitano, H. (2004). Biological robustness. *Nat. Rev. Genet.* 5, 826–837.
- Kliebenstein, D. J. (2010). Systems biology uncovers the foundation of natural genetic diversity. *Plant Physiol.* 152, 480–486.
- Koornneef, M., Alonso-Blanco, C., and Vreugdenhil, D. (2004). Naturally occurring genetic variation in *Arabidopsis thaliana*. *Annu. Rev. Plant Biol.* 55, 141–172.
- Locke, J. C. W., Kozma-Bognar, L., Gould, P. D., Feher, B., Kevei, E., Nagy, F., Turner, M. S., Hall, A., and Millar, A. J. (2006). Experimental validation of a predicted feed-back loop in the multi-oscillator clock of *Arabidopsis thaliana*. *Mol. Syst. Biol.* 2, 59.
- Mangan, S., and Alon, U. (2003). Structure and function of the feed-forward loop network motif. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11980–11985.
- Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U. (2002). Network motifs: simple building blocks for complex networks. *Science* 298, 824–827.
- Mintz-Oron, S., Meir, S., Malitsky, S., Ruppim, E., Aharoni, A., and Shlomi, T. (2012). Reconstruction of *Arabidopsis* metabolic network models accounting for subcellular compartmentalization and tissue-specificity. *Proc. Natl. Acad. Sci. U.S.A.* 109, 339–344.
- Muraro, D., Byrne, H., King, J., Voss, U., Kieber, J., and Bennett, M. (2011). The influence of cytokinin-auxin cross-regulation on cell-fate determination in *Arabidopsis thaliana* root development. *J. Theor. Biol.* 283, 152–167.
- Noble, D. (2006). *The Music of Life: Biology Beyond the Genome*. Oxford: Oxford University Press.
- Oltvai, Z. N., and Barabasi, A. L. (2002). Systems biology. Life's complexity pyramid. *Science* 298, 763–764.
- Popper, K. R. (1934). *Logic of Scientific Discovery*. London: Routledge Classics.
- Robeva, R. (2010). Systems biology – old concepts, new science, new challenges. *Front. Psychiatry* 1:1. doi: 10.3389/fpsy.2010.00001
- Sajitz-Hermstein, M., and Nikoloski, Z. (2010). A novel approach for determining environment-specific protein costs: the case of *Arabidopsis thaliana*. *Bioinformatics* 26, 582–588.
- Sauer, U., Heinemann, M., and Zamboni, N. (2007). Getting close to the whole picture. *Science* 316, 550–551.
- Schena, M., Shalon, D., Davis, R. W., and Brown, P. O. (1995). Quantitative monitoring of gene-expression patterns with a complementary-DNA microarray. *Science* 270, 467–470.
- Schneeberger, K., and Weigel, D. (2011). Fast-forward genetics enabled by new sequencing technologies. *Trends Plant Sci.* 16, 282–288.
- Schwender, J. (2011). Experimental flux measurements on a network scale. *Front. Plant Sci.* 2:63. doi: 10.3389/fpls.2011.00063
- Somerville, C. R., Bauer, S., Brininstool, G., Facette, M., Hamann, T., Milne, J., Osborne, E., Paredez, A., Persson, S., Raab, T., Vorwerk, S., and Youngs, H. (2004). Towards a systems approach to understanding plant cell walls. *Science* 306, 2206–2211.
- Sweetlove, L. J., and Ratcliffe, R. G. (2011). Flux-balance modeling of plant metabolism. *Front. Plant Sci.* 2:38. doi: 10.3389/fpls.2011.00038
- The Gene Ontology Consortium. (2012). The Gene Ontology: enhancements for 2011. *Nucleic Acids Res.* 40. doi: 10.1093/nar/gkr1028
- Tohge, T., Mettler, T., Arrivault, S., Carroll, A. J., Stitt, M., and Fernie, A. R. (2011). From models to crop species: caveats and solutions for translational metabolomics. *Front. Plant Sci.* 2:61. doi: 10.3389/fpls.2011.00061
- Tyson, J. J., Chen, K. C., and Novak, B. (2003). Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signalling pathways in the cell. *Curr. Opin. Cell Biol.* 15, 221–231.
- Van Lenne, J., Hollunder, J., Eeckhout, D., Persiau, G., Van De Slijke, E., Stals, H., Van Isterdael, G., Verkest, A., Neiryck, S., Buffel, Y., De Bodt, S., Maere, S., Laukens, K., Pharazyn, A., Ferreira, P. C., Eloy, N., Renne, C., Meyer, C., Faure, J. D., Steinbrenner, J., Beynon, J.,

- Larkin, J. C., Van de Peer, Y., Hilson, P., Kuiper, M., De Veylder, L., Van Onckelen, H., Inzé, D., Witters, E., and De Jaeger, G. (2010). Targeted interactomics reveals a complex core cell cycle machinery in *Arabidopsis thaliana*. *Mol. Syst. Biol.* 6, 397.
- Vernoux, T., Brunoud, G., Farcot, E., Morin, V., Van den Daele, H., Legrand, J., Oliva, M., Das, P., Larrieu, A., Wells, D., Guédon, Y., Armitage, L., Picard, F., Guyomarç'h, S., Cellier, C., Parry, G., Koumproglou, R., Doonan, J. H., Estelle, M., Godin, C., Kepinski, S., Bennett, M., De Veylder, L., and Traas, J. (2011). The auxin signalling network translates dynamic input into robust patterning at the shoot apex. *Mol. Syst. Biol.* 7, 508.
- Williams, T. C. R., Poolman, M. G., Howden, A. J. M., Schwarzlander, M., Fell, D. A., Ratcliffe, R. G., and Sweetlove, L. J. (2010). A genome-scale metabolic model accurately predicts fluxes in central carbon metabolism under stress conditions. *Plant Physiol.* 154, 311–323.

Received: 16 January 2012; accepted: 07 February 2012; published online: 22 February 2012.

Citation: Fernie AR (2012) Grand challenges in plant systems biology: closing the circle(s). *Front. Plant Sci.* 3:35. doi: 10.3389/fpls.2012.00035

This article was submitted to *Frontiers in Plant Systems Biology*, a specialty of *Frontiers in Plant Science*.

Copyright © 2012 Fernie. This is an open-access article distributed under the terms of the Creative Commons Attribution Non Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.