frontiers in
PLANT SCIENCE

EDITORIAL
published: 25 November 2011
doi: 10.3389/fpls.2011.00078

Editorial overview – computational approaches in aid of advancing understanding in plant physiology

Alisdair R. Fernie*

Max-Planck-Institute of Molecular Plant Physiology, Potsdam, Germany
**Correspondence: fernie@mpimp-golm.mpg.de*

The exact impact of computers on any branch of science is impossible to predict, however, as highlighted in a recent article in Science the internet has already had a profound effect on the way academics use their brains as information retrieval units (Sparrow et al., 2011). Whilst Coleridge is often cited as the last man who read (or would have been capable of reading) every article in print, the advent of the internet dwarfs even the rapid expansion of the printing presses. It's impact on science, though vast, is clearly incalculable. From a data, as opposed to a text, perspective computation has undoubtedly greatly enabled genomics – a discipline that would certainly not exist in its current form without recent advances in computational power. For example the recently sequenced potato genome (Xu et al., 2011) could easily be stored on a standard laptop computer (Usadel, personal communication).

Within this special issue of Frontiers are collected both reviews and primary research papers in which computational support played a major role. Whilst the call for papers was open to submissions from any plant biological discipline the collected papers are focused on next generation RNA sequencing, co-expression analysis, protein sub-cellular compartment prediction, sub-cellular metabolite analyses, the expansion of capacities for metabolite profiling, translational metabolomics, and metabolic flux analyses. There is thus a clear bias toward studies focused on metabolites, however, it is likely that this reflects their relative complexity both in terms of chemical structure and difficulty of analyses (Stitt and Fernie, 2003; Matsuda and Saito, 2010) as much as the interest in these problems from a biological standpoint.

The article by Jimenez-Gomez provides a detailed perspective of how computational analysis has begun, and will continue, to revolutionize the analysis of continuous phenotypic trait variation. It highlights the current state of the art in using next generation sequencing methods for the analysis of expression quantitative trait loci (eQTL) detailing recent technical, computational, and technical innovations which have facilitated the detection of molecular markers at higher resolution than previously achievable (Jimenez-Gomez, 2011). In addition to providing examples of how this works within the context of species for which a reference genome sequence is present Jimenez-Gomez also describes the utility of next generation sequencing in cases where it is absent – a technique which will prove highly useful in addressing the grand challenge of translational biology (Huber, 2011). In addition he describes complexities of next generation sequencing with respect to expression profiling and the identification of allele specific expression. Two further studies in this collection, those or Ruprecht et al. (2011) and Tohge et al. (2011), also address aspects of translational biology by means of co-expression analysis of transcript data and by addressing experimental and computational caveats of the translational application of metabolite profiling protocols, respectively. In the article by Ruprecht et al. (2011) the recently described PlaNet platform (Mutwil et al., 2011) was utilized to perform large scale condition-dependent comparisons of primary and secondary cell wall related cellulose synthase A co-expression networks. The authors used this approach to select genes from gene families that were conserved across seven species to correlate with cellulose synthase A and analyzed cell wall properties of *Arabidopsis* mutant lines of these gene families. One of these lines was demonstrated to be lignin deficient thus demonstrating the utility of this approach and suggesting that it will likely be a highly useful strategy for gene functional annotation. Also taking a cross-species approach, Tohge et al. (2011) analyzed the difficulties of using standard metabolite profiling approaches in cross-species experiments. Presented results support arguments for the need for the adoption of closely controlled empirical adaptations each time a new species or tissue is analyzed (Fernie et al., 2011). Such experiments are required because reliability of protocols for harvesting, handling, and analysis depends on biological features and chemical composition of the plant tissue. Tohge et al. (2011) provide cases studies of two different liquid chromatography mass spectrometry (LC–MS) based metabolomics platforms and four species in order to illustrate how measurement errors can be detected and circumvented.

The article of Voll et al. (2011) investigated the transcriptomic and metabolomics response of three diverse pathosystems, the barley powdery mildew fungus (*Blumeria graminis* f. sp. *hordei*), the corn smut fungus *Ustilago maydis,* and the maize anthracnose fungus *Colletotrichum graminicola.* Intriguingly, analysis of 42 water-soluble metabolites, allowed the separation of early biotrophic from late biotrophic interactions by hierarchical cluster analysis and principal component analysis, irrespective of the plant host. Both metabolome and transcript data were employed to generate models of leaf primary metabolism during early biotrophy for the three investigated interactions and these models will likely prove highly important for future studies of these pathosystems.

Sticking with metabolomics, Matsuda and co-workers present a novel framework for automated elucidation of metabolite structures in LC–MS and a co-responding metabolite ontology system. As a proof of concept the metabolome of 20 *Arabidopsis* accessions was evaluated and 704 metabolites were analyzed (Matsuda et al., 2011). Exact chemical structure determination remains one of the grand challenges of metabolomics and this strategy allowed structural estimates for an impressive 30% of these signals. In a similar vein, the paper of Hummel et al. (2011) describes a novel Ultra Performance Liquid Chromatography-based method as an alternative to widely used direct infusion based shotgun-lipidomics approaches coupled to a database search software which allows both

targeted and non-targeted lipidomic and metabolomics analysis of all kinds of mass spectral data. Widespread adoption of either the Matsuda et al. (2011) or the Hummel et al. (2011), strategy will likely greatly enhance knowledge retrieval from data acquired by similar methods.

The identification of biomarkers and complex metabolic signatures is receiving increasing attention. One of the first such searches defined the metabolite signature associated with high growth rates in *Arabidopsis* (Meyer et al., 2007) whilst recent studies have identified starch and protein levels to be key integrators of metabolism and growth (Sulpice et al., 2009, 2010). In this issue Kusano et al. (2011) describe the use of gas chromatography mass spectrometry (GC–MS) aligned with multivariate projection methods to define a metabolite signature associated with short-day induced growth cessation in aspen. In this paper the authors use this casestudy to highlight the power of statistical data analyses including principal component analysis (PCA) and orthogonal projection to latent structures (OPLS) in data interpretation.

One of the greatest challenges we currently face in plant biology is that of understanding spatial compartmentation of metabolic pathways and indeed of any other biological function. This problem is particularly acute in plants due to the myriad of cell types and sub-cellular compartments they contain (Fernie, 2007; Lunn, 2007). Whilst considerable advances have been made in determining protein location using a combination of reporter gene constructs, subcellular proteomics, and *in silico* sequence analysis (see for example Millar et al., 2009), our inventories for protein content remain incomplete and in some instances inaccurate. In the paper by Ryngajllo et al. (2011), the authors explored whether gene expression data could be harnessed to enhance bioinformatics location prediction performance. In this paper they show that utilizing their approach they could greatly enhance plastid localization prediction with notable improvements for the mitochondrion, Golgi apparatus, and plasma membrane. On the basis of these results they created the SLocX sub-cellular location predictor engine that even works in cases where only partial gene sequences are available suggesting that it may additionally have great utility for non-sequenced or poorly annotated genomes. The sub-cellular localization of metabolites is additionally currently seeing somewhat of a renaissance. Early plant studies were initiated in the 80s (Gerhardt and Heldt, 1984), however, these techniques were not commonly adopted prior to the advent of metabolomics (Farre et al., 2001). In their article Klie et al. (2011) discusses computational aspects associated with the non-aqueous fractionation method. In addition they provide a new version of the BestFit command line tool for calculation and evaluation of sub-cellular metabolite distributions and also discuss caveats and benefits of the approach.

The final two articles of Schwender (2011) and Sweetlove and Ratcliffe (2011) describe two different approaches for assessing metabolic fluxes. In the first, the ^{13}C-metabolic flux profiling approach is reviewed and Schwender describes the principle of the approach before outlining how the model boundaries are defined and the need for reaction stoichiometries for this approach. He also details computational aspects of ^{13}C-metabolic flux profiling as defined by the modeling framework of Wiechert et al. (2001), providing recent examples of network definition and validation in plants before ending with a perspective for future developments of this approach. The article of Sweetlove and Ratcliffe (2011) reviews the complementary technique of flux balance modeling. They define flux balance modeling as a constraints-based approach in which steady-state fluxes are predicted using optimization algorithms within an experimentally bounded solution space. Sweetlove and Ratcliffe argue that despite the undoubted power of the approach described by Schwender it has several limitations and postulate that these have driven to the adoption of alternate flux balance based approaches. They provide a comprehensive review of the field from its early beginnings in microbial systems to the several plant models which have been published in the last 2 years, covering modeling of specific cell types, accounting for cell maintenance energy costs, and the evaluation of metabolic efficiency via this approach. In concluding their article Sweetlove and Ratcliffe make convincing arguments for the adoption of flux balance modeling as an important complement to ^{13}C-metabolic flux profiling both for understanding metabolic regulation and ultimately as a means to determine targets for rational crop improvement.

When taken as a whole these articles cover many, although by no means all, of the ways in which computational approaches are rapidly advancing our understanding of plant function. As well as providing informative overviews of the fields defined in the opening paragraphs several of the articles also describe and provide software which should allow a relatively simple adoption of the described techniques by researchers from other laboratories. I thank all the authors for their support in putting together this special issue and hope people enjoy reading it as much as I enjoyed editing it.

REFERENCES

Farre, E. M., Tiessen, A., Roessner, U., Geigenberger, P., Trethewey, R. N., and Willmitzer, L. (2001). Analysis of the compartmentation of glycolytic intermediates, nucleotides, sugars, organic acids, amino acids and sugar alcohols in potato tubers using a nonaqueous fractionation method. *Plant Physiol.* 127, 685–700.

Fernie, A. (2007). The future of metabolic phytochemistry: larger numbers of metabolites, higher resolution, greater understanding. *Phytochemistry* 68, 2861–2880.

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.

Gerhardt, R., and Heldt, H. W. (1984). Measurement of subcellular metabolite levels in leaves by fractionation of freeze stopped material in nonaqueous media. *Plant Physiol.* 75, 542–547.

Huber, S. (2011). Grand challenges in plant physiology: the underpinning of translational research. *Front. Plant Sci.* 2:48. doi: 10.3389/fpls.2011.00048

Hummel, J., Segu, S., Irgang, S., Jueppner, J., and Giavalisco, P. (2011). Ultra performance liquid chromatography and high resolution mass spectrometry fort he analysis of plant lipids. *Front. Plant Sci.* 2:54. doi: 10.3389/fpls.2011.00054

Jimenez-Gomez, J. M. (2011). Next generation quantitative genetics in plants. *Front. Plant Sci.* 2:77. doi: 10.3389/fpls.2011.00077

Klie, S. Krueger, S., Krall, L., Giavalisco, P., Flügge, U. I., Willmitzer, L., and Steinhauser, D. (2011). Analysis of the compartmentalized metabolome – a validation oft he non-aqueous fractionation technique. *Front. Plant Sci.* 2:55. doi: 10.3389/fpls.2011.00055

Kusano, M., Jonsson, P., Fukushima, A., Gulberg, J., Sjöstrom, M., Trygg, J., and Moritz, T. (2011). Metabolite signature during short-day induced growth cessation in *Populus*. *Front. Plant Sci.* 2:29. doi: 10.3389/fpls.2011.00029

Lunn, J. E. (2007). Compartmentation in plant metabolism. *J. Exp. Bot.* 58, 35–47.

Matsuda, F., Nakabayashi, R., Sawada, Y., Suzuki, M., Hirai, M. Y., Kanaya, S., and Saito, K. (2011). Mass spectra-based framework for automated structural elucidation of metabolome data to explore phytochemical diversity.

Front. Plant Sci. 2:40. doi: 10.3389/fpls.2011.00040

Matsuda, F., and Saito, K. (2010). Metabolomics for functional genomics, systems biology and biotechnology. *Annu. Rev. Plant Biol.* 61, 463–489.

Meyer, R. C., Steinfath, M., Lisec, J., Becher, M., Witucka-Wall, H., Törjek, O., Fiehn, O., Eckardt, A., Willmitzer, L., Selbig, J., and Altmann, T. (2007). The metabolic signature related to high plant growth rate in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 13, 4759–4764.

Millar, A. H., Carrie, C., Pogson, B., and Whelan, J. (2009). Exploring the function-location nexus: using multiple lines of evidence in defining the subcellular location of plant proteins. *Plant Cell* 21, 1625–1631.

Mutwil, M., Klie, S., Tohge, T., Giorgi, F. M., Wilkins, O., and Campbell, M. M. (2011). PlaNet: Combined sequence and expression comparisons across plant networks derived from seven species. *Plant Cell* 23, 895–910.

Ruprecht, C., Mutwil, M., Saxe, F., Edler, M., Nikoloski, Z., and Persson, S. (2011). Large-scale co-expression approach to dissect secondary cell wall formation across plant species. *Front. Plant Sci.* 2:23. doi: 10.3389/fpls.2011.00023

Ryngajllo, M., Childs, L., Lohse, M., Giorgi, F. M., Lude, A., Selbig, J., and Usadel, B. (2011). SLocX: predicting subcellular localization of *Arabidopsis* proteins leveraging gene expression data. *Front. Plant Sci.* 2:43. doi: 10.3389/fpls.2011.00043

Schwender, J. (2011). Experimental flux measurements on a network scale. *Front. Plant Sci.* 2:63. doi: 10.3389/fpls.2011.00063

Sparrow, B., Liu, J., and Wegner, D. M. (2011). Google effects on memory: cognitive consequences of having information at our fingertips. *Science* 333, 776–778.

Stitt, M., and Fernie, A. R. (2003). From measurements of metabolites to metabolomics: an "on the fly" perspective illustrated by recent studies of carbon-nitrogen interactions. *Curr. Opin. Biotechnol.* 14, 136–144.

Sulpice, R., Pyl, E. T., Ishihara, H., Trenkamp, S., Steinfath, M., Witucka-Wall, H., Gibon, Y., Usadel, B., Poree, F., Piques, M. C., Von Korff, M., Steinhauser, M. C., Keurentjes, J. J., Guenter, M., Hoehne, M., Selbig, J., Fernie, A. R., Altmann, T., and Stitt, M. (2009). Starch as a major integrator in the regulation of plant growth. *Proc. Natl. Acad. Sci. U.S.A.* 106, 10348–10353.

Sulpice, R., Trenkamp, S, Steinfath, M., Usadel, B., Gibon, Y., Witucka-Wall, H., Pyl, E. T., Tschoep, H., Steinhauser, M. C., Guenther, M., Hoehne, M., Rohwer, J. M., Altmann, T., Fernie, A. R., and Stitt, M. (2010). Network analysis of enzyme activities and metabolite levels and their relationship to biomass in a large panel of *Arabidopsis* accessions. *Plant Cell* 22, 2872–2893.

Sweetlove, L. J., and Ratcliffe, R. G. (2011). Flux-balance modeling of plant metabolism. *Front. Plant Sci.* 2:38. 2:38. doi: 10.3389/fpls.2011.00038

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

Voll, L. M., Horst, R. J., Voitsik, A.-M., Zajic, D., Samans, B., Pons-Kühnemann, J., Doehlemann, G., Münch, S., Wahl, R., Molitor, A., Hofmann, J., Schmiedl, A., Waller, F., Deising, H. B., Kahmann, R., Kämper, J., Kogel, K.-H., and Sonnewald, U. (2011). Common motifs in the response of cereal primary metabolism to fungal pathogens are not based on similar transcriptional reprogramming. *Front. Plant Sci.* 2:39. doi: 10.3389/fpls.2011.00039

Wiechert, W., Möllney, M., Petersen, S., and De Graaf, A. A. (2001). 13C metabolic flux analysis. *Metab. Eng.* 3, 265–283.

Xu, X., Pan, S., Cheng, S., Zhang, B., Mu, D., Ni, P., Zhang, G., Yang, S., Li, R., Orieda, G., Guzman, F., Torres, M., Lozano, R., Ponce, O., Martinez, D., De la Cruz, G., Chakrabartu, S. K., Patil, V. U., Skrvabin, K. G., Kuznetsov, B. B., Ravin, N. V., Kolganova, T. V., Beletsky, A. V., Mardanov, A. V., Di Genova, A., Bolser, D. M., Martin, D. M., Li, G., Yang, Y., Kuang, H., Hu, Q, Xiong, X., Bishop, G. J., Sagredo, B., Majia, N., Zagorski, W., Gromadka, R., Gawor, J, Szczesny, P., Huang, S., Zhang, Z., Liang, C., He, J., Li, Y., He, Y., Xu, J., Zhang, Y., Xie, B., Du, Y., Qu, D., Bonierbale, M., Ghislain, M., Herrera Mdel, R., Giuliano, G., Pietrella, M., Perrotta, G., Facello, P., O'Brien, K., Feingold, S. E., Barreiro, L. E., Massa, G. A., Diambra, L., Whitty, B. R., Vaillancourt, B., Lin, H., Massa, A. N., Geoffroy, M., Lundback, S., Dellapenna, D., Buell, C. R., Sharma, S. K., Marshall, D. F., Waugh, R., Bryan, G. J., Destefanis, M., Nagy, I., Milbourne, D., Thomson, S. J., Fiers, M., Jacobs, J. M., Nielsen, K. L., Sonderkaer, M., Iovene, M., Torres, G. A., Jiang, J., Veilleux, R. E., Bachem, C. W., de Boer, J., Borm, T., Kloosterman, B., van Eck, H., Datema, E., Hekkert, B. L., Goverse, A., van Ham, R. C., and Visser, R. G. (2011). Genome sequence and analysis of the tuber crop potato. *Nature* 475, 189–195.

Received: 14 September 2011; accepted: 26 October 2011; published online: 25 November 2011.

Citation: Fernie AR (2011) Editorial overview – computational approaches in aid of advancing understanding in plant physiology. Front. Plant Sci. ***2****:78. doi: 10.3389/fpls.2011.00078*

This article was submitted to Frontiers in Plant Physiology, a specialty of Frontiers in Plant Science.