Available online at www.sciencedirect.com

SciVerse ScienceDirect

Current Opinion in Plant Biology

Systems-integration of plant metabolism: means, motive and opportunity

Mark Stitt

System integration of metabolism is considered in analogy to the investigation of corporate misdemeanour. Motive, or goal-oriented explanation, provides hypotheses that can guide the investigation of network structure. Opportunity can be established by correlative analysis using large-scale omics resources. However, correlative approaches on their own remain inconclusive and seldom identify all the links in a network. Establishment of means, or the ability to act on other network components and contribute to a phenotype, is therefore crucial. This requires functional information. Integration of quantitative data in the context of pathway models provides a powerful approach to establish 'means'. This is illustrated by discussing: first, how protein abundance is regulated by a network including transcript abundance, translation and protein degradation and second, how a combination of experimentation and modelling provides information about pathway flux, an emergent network property that integrates changes in proteins and metabolites and determines composition and biomass.

Address
Max Planck Institute of Molecular Plant Physiology, Am Muehlenberg 1, 14476 Potsdam-Golm, Germany

Corresponding author: Stitt, Mark (mstitt@mpimp-golm.mpg.de)

Current Opinion in Plant Biology 2013, **16**:381–388

This review comes from a themed issue on **Physiology and metabolism**

Edited by **John Browse** and **Edward Farmer**

For a complete overview see the Issue and the Editorial

Available online 2nd May 2013

1369-5266/$ – see front matter, Published by Elsevier Ltd.

http://dx.doi.org/10.1016/j.pbi.2013.02.012

Introduction

Advances in technology to measure transcripts [1,2,3•,4•], proteins [4•,5•] and metabolites [6,7•,8•] are generating daunting amounts of data that have to be integrated and linked to phenotypes in a predictive manner [9,10]. This is challenging due to the heterogenous nature of omics data [11•,12], the interactive structure of biological networks [11•,13] and their flexible remodelling during development and in response to the environment [11•,14•,15]. Recent reviews have discussed how metabolic traits can be mapped to genetic polymorphisms [15,16,17] and omics data can be integrated into networks in specialised metabolism [10,16–19]. Complex pathway topology and overlapping functionalities in primary metabolism make integration especially difficult. Conviction is especially difficult when crimes are perpetrated by many people, acting often via intermediaries. I examine systems integration in terms of the criminalistics concepts of 'means, motive and opportunity' and highlight the importance of establishing 'means' by integrating quantitative data into pathway models. A broader treatment of the importance of quantification in biology can be found in [20,21•].

Means, motive and opportunity

The lines of evidence that establish guilt in a law court can be schematised as the opportunity to commit the crime, the ability or means to commit it, and motive. Motive is not essential for conviction but makes it easier to persuade a jury, due to the human propensity to rationalise in terms of goals. In biology, goal-oriented explanations are, at the worst, teleological arguments. Used judiciously, they are useful as hypotheses to guide experimental design and data evaluation and demonstrate opportunity and means.

Opportunity — integration of large data sets by correlation

In a law court, opportunity is usually established by demonstrating that the defendant was present at the scene of crime, and disproved by an alibi.

Omics data can be used to establish opportunity via 'guilt-by-association' [18]. Web resources exist to identify sets of transcripts that change in a correlated manner in different data sets (see [10] for references). Analogous approaches are applicable to proteins and metabolites, although currently limited by amount of data in the public domain and, for metabolites, the lack of standardised protocols for sample handling, metabolite determination and data validation [7•]. Further omics resources include information on cell-specific expression patterns of transcripts and proteins [5•,22–25,26•], the subcellular location of proteins (http://suba.plantenergy.uwa.edu.au/) and large-scale analyses of protein-protein interactions [27].

Genetics can provide alibis by, for example, showing that a given phenotypic response is unaltered when a gene is deleted. However, this approach is not unambiguous because redundant genes may substitute for the deleted gene. The demonstration that deletion prevents or modifies a response is usually taken as evidence for a

gene's involvement. This conclusion depends on there being no side effects, which is unlikely in a complex network, emphasising the important of using small and, if possible, spatially restricted or temporally restricted genetic interventions.

Correlative analysis followed by functional analysis of candidates has led to important discoveries in secondary metabolism [10,16–19] but has arguably been less successful in central metabolism. Recent advances include definition of the starch degradation pathway (see [28]), the identification of novel transporters and pathways in C4 photosynthesis [3•,29•] and the identification of transporters in photorespiration [30•,31•].

Advances have been made in identifying metabolic states that are associated with faster growth. Whilst individual metabolites seldom correlate with biomass, multivariate approaches have identified sets of metabolites that predict biomass in large populations of Arabidopsis genotypes [32,33]. Higher biomass also correlated with allocation of a larger proportion of total protein to enzymes in central metabolism [34]. The recent discovery in Arabidopsis [35•,36•] and maize [37•] that the extent of heterosis can be predicted from the metabolic profiles of parental lines has practical implications. Multivariate prediction of quality traits is also possible [38]. Nonetheless, correlative studies only establish a probability that a given metabolic state is associated with faster growth, but do not explain why this is so. Genetic mapping can define genomic regions and ultimately polymorphisms that underlie variation in metabolic traits [10,16,17]. However, while this provides insights into the genetic and molecular architecture of the networks that control metabolic traits, it may not explain why certain metabolic states are associated with greater biomass production.

Several issues can interfere with correlative analyses of large omics data sets. Coordinated responses often make it difficult to identify key genes, proteins or metabolites. Analyses are complicated by complex spatial location [5•,23,25,26•] and temporal responses. Transcripts typically change within minutes to a few hours, some proteins change in a similar time frame but many change over a time frame of days, whilst metabolites show an even wider dynamic range [4•,11•,14•,39,40•]. Another problem is redundancy, which is itself often context-dependent. Sucrose synthase (*SUS*) provides a striking example. Whilst mutations in individual *SUS* genes lead to strong phenotypes in maize seeds, a quadruple *SUS* knockout had no major phenotype in Arabidopsis rosettes [41]. In fact, correlative analyses frequently fail to detect connections between entities that are known to be mechanistically linked. For example, levels of transcripts and their encoded proteins often change independently in microbes [42,43], animals [44,45•] and plants [4•,39,46]. Agreement is better at the cell-specific level [5•] and after long-term treatments, but is especially weak in short-term responses [4•,34]. Another example is a poor correlation between the levels of enzymes and the metabolites they act on [12,34].

Establishing the 'means' — using quantitative data to establish mechanistic links

Even if 'opportunity' can be established, conviction requires evidence of 'means', that is, the ability to commit a crime. Most omics data are non-quantitative, in the sense that the units are relative. Relative data are useful for building up a qualitative understanding of function via aggregation of information. However, they are unsuitable for a quantitative predictive analysis, which is often necessary to demonstrate a chain or web of events.

Absolute transcript abundance can be measured by adding external RNA standards before extraction [47] or by combining RNA-seq with information about total RNA levels [45•,48,49]. Absolute protein abundance can be determined using peptide standards or, on a larger scale, by relating the summed mass spectrometric (MS) signals to total protein content or using APEX, which estimates protein abundance from the fraction of peptide mass spectra associated with that protein after correction for prior expectation of observing the peptides [49]. For metabolites, it is essential to add standards [50], preferably also before tissue extraction to assess loss during sample handling and preparation [7•,12].

The value of large-scale quantitative data sets for transcripts, proteins and metabolites will be illustrated using the two apparent discrepancies between omics data sets mentioned earlier. They illustrate how quantitative data and models can be used to learn if events at one level can lead to a given output at a higher level in a network.

Quantitative analysis of the relation between transcript levels and protein abundance

As already mentioned, the responses of transcripts and the proteins that they encode are often poorly correlated. Knowledge about mechanistic relationships points to possible reasons for this discrepancy. First, the rate of protein synthesis depends on the rate of translation as well as transcript abundance. A recent analysis in mouse fibroblasts showed that translational regulation makes a larger contribution than transcript abundance to the determination of protein abundance [45•]. Translational regulation also plays a key role in plants [51,52•]. Second, a change in transcript abundance cannot lead to a significant increase in protein abundance unless there is enough transcript to support a high enough rate of translation to alter protein abundance in a given time interval. A study with Arabidopsis rosettes [47] used quantitative data on transcript abundance and polysome loading to model the synthesis rates of 35 enzymes. These estimates were

Download English Version:

https://daneshyari.com/en/article/10869275

Download Persian Version:

https://daneshyari.com/article/10869275

Daneshyari.com