Metabolomics (2013) 9:755–756
DOI 10.1007/s11306-013-0562-5

OPINION

Microbial metabolomics: welcome to the real world!

Edward E. K. Baidoo · Jay D. Keasling

Published online: 10 July 2013

In the post-genomics era there has been a growing emphasis on understanding the metabolic consequences of engineering cells to produce desired chemicals. Proteomics measurements have improved our understanding of gene expression, and metabolomics promises to take us a step further. A consequence of the metabolome being further down the line from gene function is that the interpretation of metabolomics data reflects more closely the activities of a cell at the functional level and potentially leads to the revelation of metabolite function. This latter point is of great importance especially when we consider that a given metabolite may come from more than one metabolic pathway and may impact the cell in multiple ways. Further, by accurately quantifying the metabolome, we can begin to understand the stoichiometric significance of metabolite production.

Since metabolomics was first coined, it was received with a great deal of fanfare. Like the earlier omics techniques, it was thought that comprehensive metabolomics data sets would be a staple of functional genomics analyses. That was until it was realized that the structural diversity that exists within the metabolome would make it virtually impossible, with current analytical technologies, to characterize the metabolomes of microorganisms fully. To make matters worse, the diverse chemical and physical properties within the metabolome, together with the structural differences in the cell walls of microorganisms, meant that it would be extremely difficult to establish a single sample extraction procedure that was suitable for all metabolites. Added to the fact that the recovery of metabolites would largely depend on the extraction protocol used, meant that the achievement of global metabolite analysis was not a trivial matter. Questions that should have been asked at the time were, how much of the metabolome was necessary to characterize metabolism and what aspects of metabolomics data were truly meaningful? While the goal of complete holistic metabolomics data sets is still what we aspire to achieve, at present we have softened our stance on metabolomics by focusing on local metabolite profiling. In addition to this, we have concentrated our efforts on looking for specific differences in metabolite levels in response to genetic modification, stress responses and environmental change. Even though this is a far cry from global metabolite profiling, these approaches combined with the latest analytical technologies are being used to successfully unravel gene function and to identify biomarkers and unknown metabolites.

The most promising techniques that are currently employed in microbial metabolomics research are nuclear magnetic resonance (NMR), array-based mass spectrometry (MS), gas chromatography–MS (GC–MS), liquid chromatography–MS (LC–MS), and capillary electrophoresis–MS (CE–MS). Though NMR can provide high throughput fingerprinting and structural elucidation, cost, sensitivity and quantification are issues that need to be addressed if it is to compete with the MS-based technologies. While array-based techniques, such as nanostructure-initiator MS (NIMS), are ideal for screening crude microbial extracts, developments in nano surface technology are essential if

E. E. K. Baidoo · J. D. Keasling (✉)
Physical Biosciences Division, Joint Bioenergy Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
e-mail: jdkeasling@lbl.gov

E. E. K. Baidoo
e-mail: eebaidoo@lbl.gov

J. D. Keasling
Departments of Chemical and Biomolecular Engineering and of Bioengineering, University of California, Berkeley, Berkeley, CA, USA

they are to be used for global metabolite profiling. Fast chromatography techniques are often coupled to fast scanning MS, such as time-of-flight (TOF), to take advantage of the highly efficient separations they produce. Recent advances in 2-dimensional GC have led to a growing number of applications in microbial metabolomics despite the reliance of this technique on volatile metabolites (i.e., naturally volatile or derivatized). In addition, rapid LC separations of are now routinely performed by ultra high performance LC (UHPLC) or by superficially porous stationary phases (fused-core^{TM}) on current analytical LC systems. Manufacturers, however, have yet to adapt all of their LC columns to the aforementioned technologies. Highly efficient CE separation, when coupled to TOF MS, has also garnered significant attention in microbial metabolomics studies. Despite the inherent problems of CE, in recent years there have been noticeable improvements in CE robustness (e.g., the use of pressure-assisted CE for anions) and sensitivity (via new developments in nanospray, electrospray and ion transmission in MS instrumentation). While it would appear that TOF is the ideal MS for metabolomics, recent advances in triple quadrupole technology have resulted in an increase in the number of metabolomics applications, especially when either the single-reaction-monitoring or multiple-reaction-monitoring modes are employed with time segments.

To complement these analytical technologies, there are now a growing number of metabolomics data analysis platforms being used, and all appear to approach data analysis in a similar fashion. That is, they first extract all the components from the spectral noise, normalize the data, then perform the relevant statistical analysis and identify the compounds of interest. Further, the identified compounds can now be linked to metabolic pathways, and the entire data set can be correlated to proteomics and transcriptomics data sets. However, some data analysis platforms can only be used with data obtained from the instrument manufacturer. In addition, the process of analysing the acquired data can be longwinded and, at times, cumbersome and often requires extremely powerful computers to do so.

At the Joint Bioenergy Institute, our strategy has been to use synthetic biology to engineer biofuel-producing microbial strains and validate them via molecular biology tools, proteomics (if necessary), and the quantification of the resulting fuel products by metabolite profiling (via GC–MS). Then, to understand the significance of production, intermediates of the biosynthetic pathway are profiled to reveal any potential bottlenecks (via LC–MS). Since the engineered biosynthetic pathways are reliant upon central carbon metabolites, the amounts of these intermediates are quantified by metabolomics measurements (via LC–MS). This has enabled us to understand how our engineering affects metabolism and, hence, the cells' physiological state. This combinatorial approach, together with the pursuit of unknown biomarkers, will eventually enable us to identify the regulatory mechanisms that cells may use to circumvent our engineering.

While global metabolite profiling in microorganisms may be possible in the future, in order achieve this it is imperative that we continue to push the envelope in development in all aspects of metabolomics technologies/research. Consequently, as metabolomics measurements become more comprehensive, they are likely to be applied to fluxomics studies, leading to predictive models for microbial metabolic networks that provide greater understanding of the regulation involved in these networks which, of course, is of great importance to the future success of both synthetic and systems biology.