



Analytics for metabolic engineering

Christopher J. Petzold^{1*}, Leanne Jade G. Chan¹, Melissa Nhan¹ and Paul D. Adams^{1,2}

¹Joint BioEnergy Institute, Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA,

²Department of Bioengineering, University of California Berkeley, Berkeley, CA, USA

Realizing the promise of metabolic engineering has been slowed by challenges related to moving beyond proof-of-concept examples to robust and economically viable systems. Key to advancing metabolic engineering beyond trial-and-error research is access to parts with well-defined performance metrics that can be readily applied in vastly different contexts with predictable effects. As the field now stands, research depends greatly on analytical tools that assay target molecules, transcripts, proteins, and metabolites across different hosts and pathways. Screening technologies yield specific information for many thousands of strain variants, while deep omics analysis provides a systems-level view of the cell factory. Efforts focused on a combination of these analyses yield quantitative information of dynamic processes between parts and the host chassis that drive the next engineering steps. Overall, the data generated from these types of assays aid better decision-making at the design and strain construction stages to speed progress in metabolic engineering research.

Keywords: metabolic engineering, RNA-seq, proteomics, metabolomics, high-throughput screening, microfluidics

OPEN ACCESS

Edited by:

Stephen Fong,
Virginia Commonwealth University,
USA

Reviewed by:

Weiqin Zhang,
Tianjin University, China
M. Kalim Akhtar,
University of Edinburgh, UK

*Correspondence:

Christopher J. Petzold,
5885 Hollis Street – 4th Floor,
Emeryville, CA, USA
cjpetzold@lbl.gov

Specialty section:

This article was submitted to
Synthetic Biology, a section of the
journal *Frontiers in Bioengineering
and Biotechnology*

Received: 07 July 2015

Accepted: 24 August 2015

Published: 07 September 2015

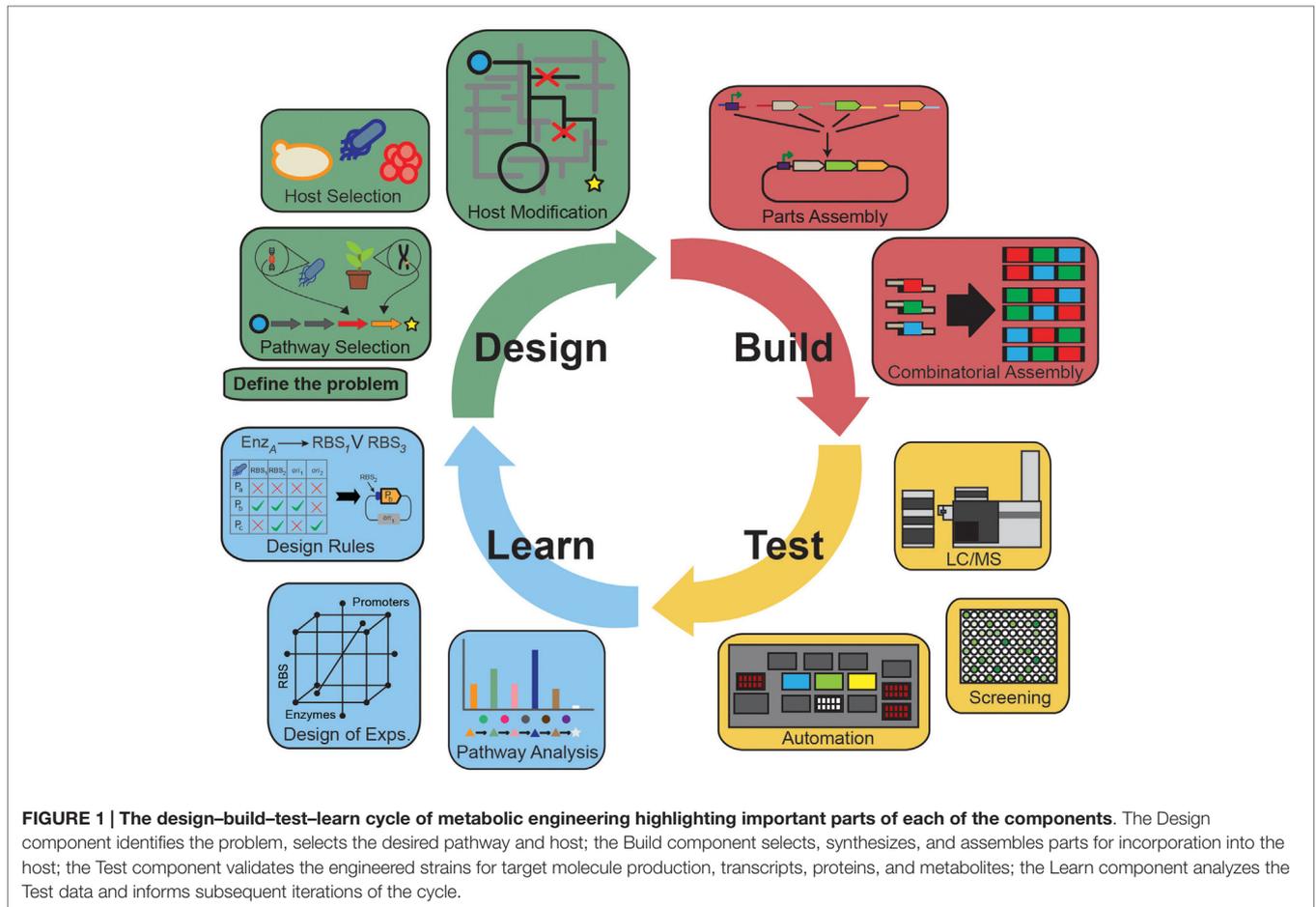
Citation:

Petzold CJ, Chan LJG, Nhan M and
Adams PD (2015) Analytics for
metabolic engineering.
Front. Bioeng. Biotechnol. 3:135.
doi: 10.3389/fbioe.2015.00135

Introduction

Over the past decade, the field of metabolic engineering has realized several significant achievements centered on production of bulk (Atsumi et al., 2008; Saxena et al., 2009; Yim et al., 2011; Jarboe et al., 2012), high-value (Yang et al., 2013; Zhang and Stephanopoulos, 2013), and therapeutic compounds (Paddon et al., 2013). Stemming from these efforts have been transformational developments in the design–build–test–learn (DBTL) paradigm to accelerate pathway discovery, construction, characterization, and understanding. Innovative tools in each component of the cycle are propelling the field from the artisanal/single researcher model of science toward engineering principles of standardization, parameterization, and robust operation (**Figure 1**). Yet, attempts to generalize knowledge from earlier studies have not led to rapid progress toward these goals. As a result, successful integration of each component into a coherent whole that can rapidly inform subsequent engineering efforts is still a challenging endeavor. True paradigm-shifting advancement in metabolic engineering requires seamless workflows that are capability matched across all components to inform subsequent cycles.

Until recently, strain design was often treated as a one-off process, relying heavily on domain expertise with no standardization. Now, many software platforms that integrate and automate design parameters, enable strain construction, and incorporate knowledge from past experiments are readily available (Kelwick et al., 2014). A variety of novel computational tools are available to identify biosynthetic gene clusters (Hatzimanikatis et al., 2005; Campodonico et al., 2014;



Medema et al., 2014; Weber et al., 2015), select pathways based on retrosynthesis of products (Hatzimanikatis et al., 2005; Campodonico et al., 2014; Carbonell et al., 2014a,b) or retro-fit enzymes to engineered pathways (Brunk et al., 2012). Likewise, genome-scale models (Feist and Palsson, 2008; Schellenberger et al., 2011; McCloskey et al., 2013) can identify beneficial host modifications *in silico* to prioritize changes based on factors related to target molecule production.

Once the pathway is chosen, there remain many questions related to which specific enzymes and how much of them are optimal, as well as which expression system should be used to achieve balanced protein levels. Consequently, a significant amount of research has focused on characterization of libraries of promoters (Alper et al., 2005; Anderson et al., 2010; Babiskin and Smolke, 2011; Davis et al., 2011; Mutalik et al., 2013) and terminators (Chen et al., 2013), as well as develop tools to predict ribosome-binding strength (Salis et al., 2009) or control protein location in the cell (Dueber et al., 2009). New technologies have advanced the Build component via low cost gene synthesis, rapid genome modifications [e.g., CAS9/CRISPR (Jinek et al., 2012)], and facile methods to rationally generate strain diversity [i.e., multiplex automated genome engineering (MAGE) (Wang et al., 2009) and trackable multiplex recombineering (TRMR) (Warner et al., 2010)]. As the scope of strain construction has grown,

combinatorial approaches, like multivariate modular metabolic engineering (Ajikumar et al., 2010), have been applied to various parts to rapidly balance flux through the pathway (Latimer et al., 2014; Smanski et al., 2014). Yet, knowledge from previous work is infrequently predictive for new pathways, genes, and gene expression levels (Cardinale and Arkin, 2012; Kittleson et al., 2012). This problem is compounded when multiple genetic circuits or pathways are combined into a single organism often leading to unintended consequences. Thus, massive over construction of a pathway is required to find the best strain.

As a result, large-scale analysis of the engineered organisms is needed from the Test component. Yet, it lags far behind the recent Design- and Build-related advancements especially with respect to throughput, robustness, and generalizability. High-throughput assays, such as screens or selections that assay the target molecule, are ideally suited to strain optimization efforts, yet fail to provide sufficient information to efficiently identify pathway bottlenecks. Alternatively, detailed analyses of transcripts, proteins, and metabolites to query strain function can provide a rich dataset to identify bottlenecks that inform subsequent strain design, albeit for only a tiny fraction of strains that can be built. Detailed analyses of strain function are needed because engineering efforts often disrupt native cell processes that compete with the intended result. It is currently possible to analyze only a tiny fraction of the

engineered strains for detailed omics analysis and quantification. The net result is a significant capability gap between the Design and Build components and that of the Test component of the DBTL cycle.

Unfortunately, learning is possibly the most weakly supported step in current DBTL cycle. Learn efforts to generalize knowledge from past experiments to inform design and build process decisions with the explicit goal of increasing the rate of successful outcomes are frequently limited. Many failure modes are possible and they are often difficult to identify and alleviate. Successful learning stems from observations from multiple iterations of the cycle, including analysis of failure modes at multiple functional levels (i.e., transcripts, proteins, metabolites). From this knowledge, improved design rules for assembling biological systems with predictable behavior can be created. With the Design and Build capabilities outpacing Test developments metabolic engineers face distinct difficulties making significant progress. This lack of actionable data prevents meaningful learning from past efforts. This review will detail the common analytical techniques to enable characterization of a broad range of components and their interactions.

Target Molecule Detection

By far the most common assay, and arguably the only necessary one, in metabolic engineering is target molecule detection. Target molecule detection takes several common forms that balance throughput and flexibility, with increased throughput typically coming at the cost of lower flexibility (Table 1). Small-scale engineering efforts often quantify the target of interest by using techniques such as gas or liquid chromatography (GC, LC) with UV absorbance or mass spectrometry (MS) detection (Figure 2A). The vast majority of older metabolic engineering studies relied heavily on these assays for initial pathway validation. Mass specific detection is particularly powerful because it permits monitoring the target molecule and, in many cases, pathway intermediates within complex matrices. Furthermore, assay development is typically fast and easy when standard compounds are available making them applicable to many types of targets. These methods produce confident target identifications with high sensitivity as well as accurate and precise quantification. There are many tools to generate rationally designed libraries or random strain diversity, which push the sample throughput capacity of these assays. More recently, these methods have been used to verify the top “hits” from high-throughput screening (HTS) assays.

TABLE 1 | Analytical attributes of common target molecule assay types.

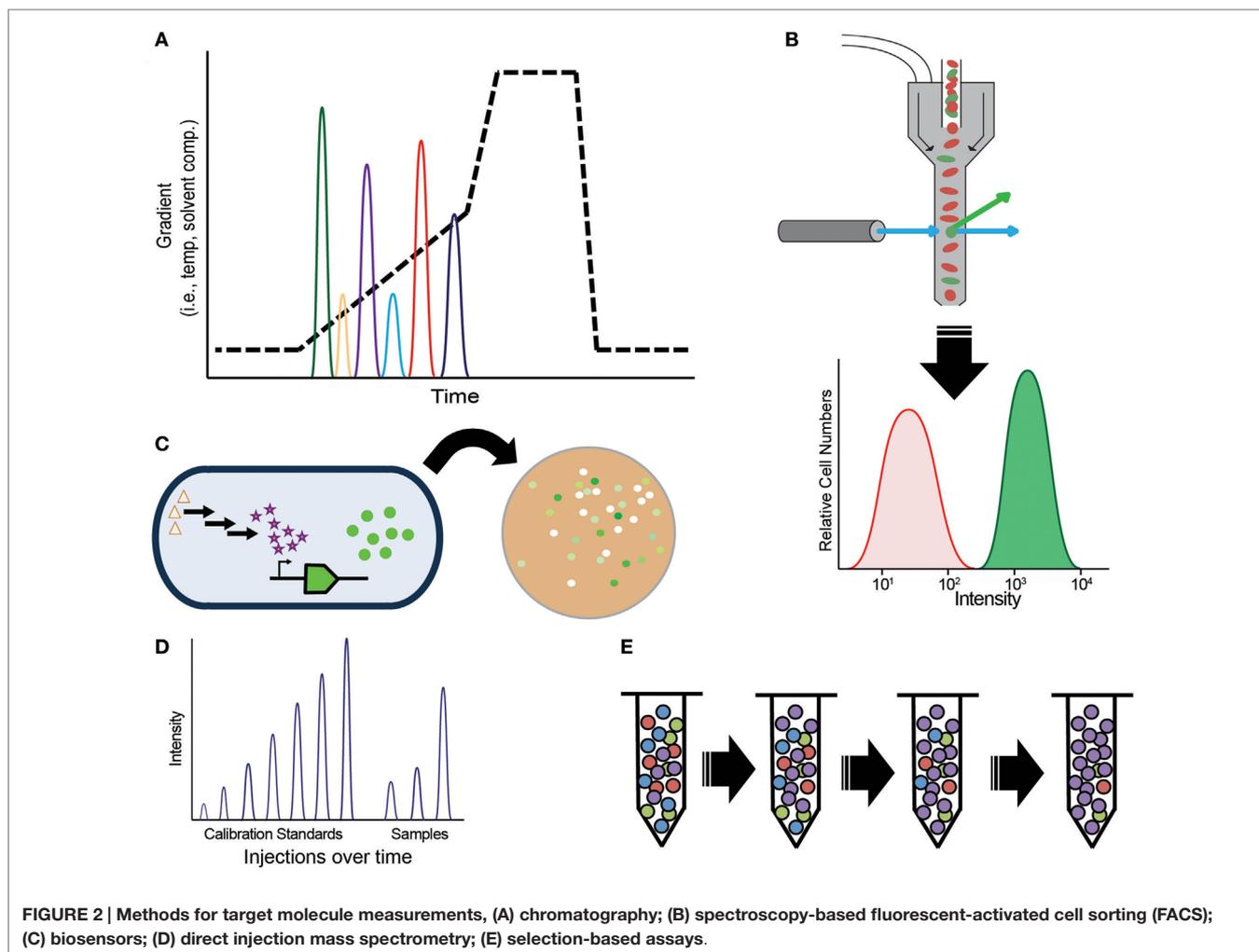
	Sample throughput (per day)	Sensitivity (LLOD)	Flexibility	Linear response	Dynamic range
Chromatography	10–100	mM	++	+++	+++
Direct mass spectrometry	100–1000	nM	+++	+++	++
Biosensors	1000–10,000	pM	+	+	+
Screens	1000–10,000	nM	+	++	++
Selection	10⁷⁺	nM	+	+	+

The optimal method for a given criteria is highlighted in bold.

While flexibility and confident identification are highly desired for proof-of-concept experiments, higher throughput assays, such as screens, selections, or biosensors, are preferred (Dietrich et al., 2010; Van Rossum et al., 2013) for titer, yield, and productivity optimization, steps required to develop economically viable strains. To achieve the necessary throughput, the vast majority of HTS assays rely on spectroscopic measurements, such as colorimetric, UV absorbance, or fluorescence in micro-well plates or via fluorescent-activated cell sorting (FACS) (Figure 2B). High-throughput measurements remain difficult for most target molecules because they lack an appropriate fluorophore, chromophore, or may not be essential for cell growth. A variety of chemical biology tools are available to overcome this complication by modifying target molecules with chemical tags. Bio-orthogonal chemistries, such as “Click” chemistry (Kolb et al., 2001), as well as protein bio-conjugation methods (Stephanopoulos and Francis, 2011) can be used to identify glycans, proteins, lipids, and nucleic acids. However, much effort is necessary to develop quantitative assays based on chemical biology tools. Ultimately, the usefulness of these assays depends on key performance features: dynamic range, sensitivity, and linear range of detection (Dietrich et al., 2010).

Most biosensors function via protein or transcript-based sensing of a target molecule coupled to some reporter (Figure 2C). By far, the most common use of biosensor assays is for spectroscopic measurement of the consumption or production of a cofactor, changes to pH, or H₂O₂. If gene expression as a response to target molecule production is known, then the expression of a reporter protein can be used. Recently, many enzymatic reporter systems have been generalized to broaden their applicability to new target molecules (Eggeling et al., 2015), but many lack suitable ligand recognition or binding elements. Consequently, engineering RNA aptamers (Win and Smolke, 2008; Babiskin and Smolke, 2011; Carothers et al., 2011; Michener et al., 2012; Zhang et al., 2012a), transcription factors (Binder et al., 2012; Dietrich et al., 2013), ligand binding proteins (Looger et al., 2003; Tang and Cirino, 2011; Shong and Collins, 2013), and protein–protein interactions (Dueber et al., 2003; Skerker et al., 2008) are fertile areas of research. Likewise, biosensors engineered to report via light/dark changes can be tuned to report on target production or various cell functions (Tabor et al., 2011). Whole cell biosensors that respond to target molecules via growth (Pfleger et al., 2006) effectively separate the reporter from the engineered microbe. These systems remove sensor components from the producer strain simplifying assay development and troubleshooting. By using synthetic transcription factors, one can produce metabolic enzyme sensors that recognize specific target molecules. Despite many examples of high-throughput assays in the literature, it is challenging to build a biosensor or screen with sufficient dynamic range, sensitivity, and linear response for a broad range of optimization conditions. To alleviate this constraint, more effort is needed to engineer genetic circuits that permit tight control of signal transduction, amplification, and response time (Kobayashi et al., 2004; Tabor et al., 2009).

Mass spectrometry is capable of the necessary throughput, sensitivity, and linear response for screening studies; however, signal repression from the complex matrix associated with



cellular analyses confounds many MS assays, requiring slow separations prior to MS detection. New technologies, such as the Agilent Rapidfire™ system, attempt to overcome this limitation by integrating solid phase extraction with HT direct infusion analyses (**Figure 2D**) at a rate of one analysis per 10–12 s (one 96-well plate in (11 min)). Initially applied to drug discovery studies (Vanderporten et al., 2013), there is a great interest in implementing the system to screen cell extracts for metabolic engineering research. This technique has the potential to determine target levels and comprehensively analyze multiple types of cell metabolites in much greater throughput than currently possible. Alternatively, development of less complex engineering hosts or cell-free systems would simplify MS measurements overcoming this obstacle.

Key to screen development is establishing conditions that produce accurate quantitative results that can be used to identify strains with improved production. Statistical methods, leveraged from drug discovery studies to verify hits, reduce false positives and false negatives (Malo et al., 2006, 2010) facilitate differentiation of subtle improvements. For industrial applications, the quality of the screen with respect to scale-up process requirements determines the ultimate success of the method.

Consequently, development of these assays typically requires significant time, resources, and testable strains to ensure that quantitative improvements determined for micro-bioreactor conditions translate into analogous gains at production scale.

Despite the high-throughput nature of screening, there remains several bottlenecks that limit the number of strains that can be tested. Heavy reliance on colony picking and liquid handling systems produce a practical limit to throughput in addition to significant needs for culturing space. Selections, which circumvent these obstacles, are powerful methods to test very large libraries (10^{10} cells) to rapidly identify the optimal genotype (**Figure 2E**). Techniques such as TRMR (Warner et al., 2010) use selections to improve host tolerance and target production, yet few target molecules meet the criteria for selection-based assays (Dietrich et al., 2010). It can be very difficult to engineer a chosen host organism to be auxotrophic for a specific target molecule; consequently, there is a great interest in methods to adapt existing selections to new targets via biosensors. Biosensors combined with feedback-regulated evolution of phenotype (FREPE) (Chou and Keasling, 2013) are a powerful way to evolve traits. Fluorescent-activated cell sorting of strains engineered with biosensor reporters offers

significant increase in throughput without the associated culturing and colony-picking bottlenecks. Yet, variance is often broad requiring follow-up screens or chromatography-based assays to validate strain improvement. In most situations, achieving well-defined performance criteria for high-throughput assays is a long arduous process.

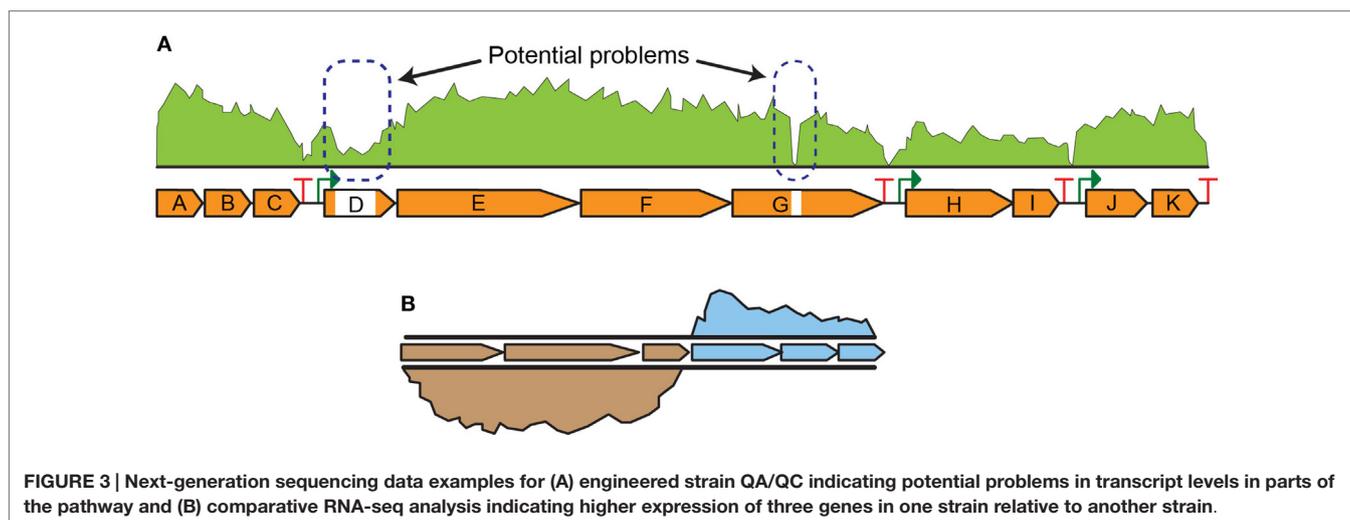
Transcriptomic Analysis

Since the primary mechanism of change via metabolic engineering is DNA, many studies have focused on techniques and methods that reduce the cost of gene synthesis, strain construction, and other aspects of molecular biology. Through analysis of transcript levels and mapping the outcomes of failed systems, new constraints can be applied to successive iterations of the cycle. Traditional methods like real-time quantitative PCR (RT-qPCR) and microarray analyses are routinely used to verify that the host has been engineered correctly and to query regulatory and stress-related effects under production conditions. Combining microarray analysis with pathway intermediate detection, Kizer et al. (2008) identified stress response in *Escherichia coli* due to an imbalanced mevalonate pathway that accumulated 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA).

Transcript analysis has been particularly helpful for host-engineering efforts. Oh and Liao (2000) examined *E. coli* grown on glucose, acetate, and glycerol media and identified transcripts that were up- or down-regulated during protein overexpression. Alternatively, transcriptome analysis aided the identification of promoters responsive to specific growth conditions, such as the presence of inhibitors, pathway intermediates, and oxygenation state of the cell (Rutherford et al., 2010; Zhang et al., 2012a). For instance, to increase *E. coli* tolerance to the presence of ionic liquids microarray analysis was used to identify native *E. coli* promoters responsive to sub-lethal concentrations of ionic liquid (Frederix et al., 2014). Dynamic control of a newly discovered pump was engineered to alleviate ionic liquid toxicity. Designing dynamic control of pathway enzymes based on accumulation of

pathway intermediates is greatly desired for commercial scale fermentations to eliminate the need for costly inducer compounds. Microarrays have also been used to identify promoters that respond to pathway intermediate accumulation. Amorpho-4,11-diene production in *E. coli* was improved by engineering pathway enzyme production in response to increased farnesyl pyrophosphate (FPP) levels (Dahl et al., 2013).

Next-generation sequencing (NGS) technologies (Smith et al., 2009) are emerging as the preferred approach for transcript analysis. NGS assays are suited to a broad range of applications, including quality assurance/quality control (QA/QC) of DNA construction efforts and quantitative whole genome transcript analysis (i.e., RNA-seq experiments) (Figure 3), due to high sensitivity, multiplex advantages, and dynamic range (Robles et al., 2012; Ghaffari et al., 2015). The most direct use of NGS is for QA/QC experiments of engineering pathways and strains. Tracking genomic changes during pathway and host engineering is crucial to determining the true (and testable) source of production improvements. NGS permits comprehensive validation of engineered strains to identify unintended mutations or other types of transcriptional failures (Figure 3A). For instance, identification of transcript read-through of terminators facilitated optimization of nitrogen fixation pathway in *E. coli* (Smanski et al., 2014). Genomic characterization of engineered strains via NGS is especially useful to rapidly identify single nucleotide changes that contribute to the observed phenotype. Comparative RNA-seq analysis was used to characterize the transcriptional response of *E. coli* to aromatic inhibitors from pretreated biomass and discover a D-galacturonic acid transporter candidate gene in *Neurospora crassa* that was then expressed in yeast (Benz et al., 2014). Similarly, it was used to identify xylose utilization genes and their regulation in *Saccharomyces cerevisiae* (Feng and Zhao, 2013). Additionally, RNA-seq is foundational to multi-omic analyses where integrated datasets are used to elucidate complex cellular interactions and identify differences between strains. Data generated from RNA-seq are also useful for genome-scale metabolic models (GSMs) as it is complementary to flux analysis (Gowen and Fong, 2010).



Proteomic Analysis

Transcriptomic analysis readily queries the success or failure of strain construction efforts, yet it is often a poor proxy for protein abundance. Consequently, proteomic analysis is valuable for characterization of the functional aspects of engineered strains. Proteomic analysis is inherently more challenging than transcriptomic analysis because proteins are not readily amplified nor are they easily separated from each other; factors that dramatically impact sample throughput. Protein detection and quantification are frequently achieved via immunoblot assays because they are selective, established, fast, and easily analyzed in parallel. A variety of techniques are available for quantification of proteins via immunoblot assays or with fluorescent protein surrogates; however, these are also often inaccurate (Cardinale and Arkin, 2012). Furthermore, it can be challenging to obtain accurate quantitative information from these methods when assaying many different proteins in the same strain as is common for multi-step pathway engineering.

Current shotgun proteomic methods based on LC-MS/MS are useful for identification and quantification of thousands of proteins. Differential relative quantification is frequently achieved by culturing strains in media containing isotopically labeled substrates or by using functional group-specific chemical labels [e.g., iTRAQ (Shadforth et al., 2005), TMT (McAlister et al., 2012)] during sample preparation, or by so-called “label-free” techniques (Arike et al., 2012), such as data-independent analysis (DIA) (Gillet et al., 2012). Shotgun proteomic analyses are often combined with transcriptomic analyses to characterize cell stress responses to high levels of pathway intermediates or the final product (Rutherford et al., 2010) and to identify metabolic sinks that reduce carbon flux through the pathway.

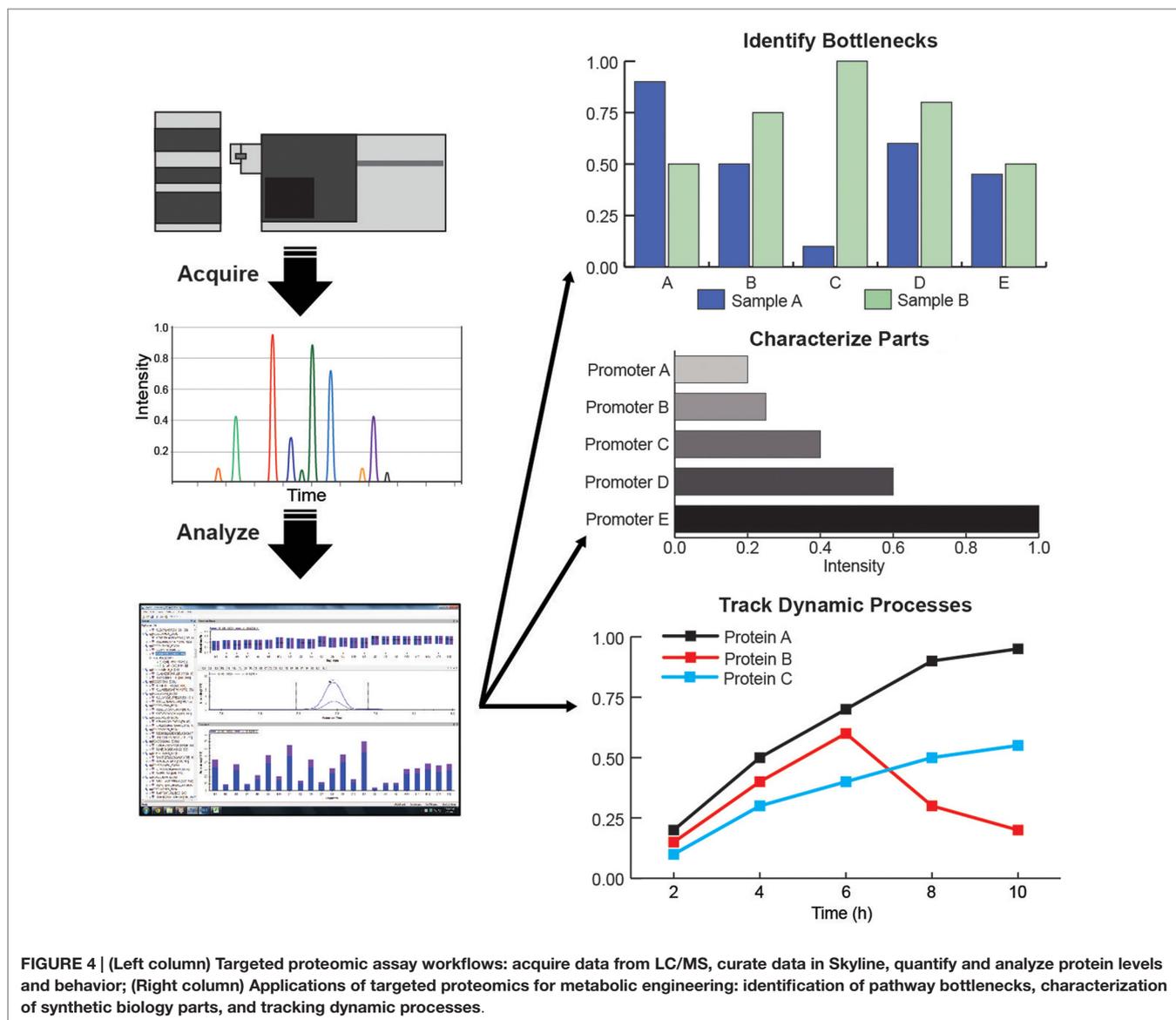
For more specific hypotheses, a targeted proteomics approach, via selected-reaction monitoring (SRM) MS can be used to accurately quantify a select group of proteins (Picotti et al., 2009; Picotti and Aebersold, 2012). The targeted proteomics approach has been applied in a variety of ways to test engineered microbes. Redding-Johanson et al. (2011) used targeted proteomics to identify protein-associated bottlenecks in the mevalonate pathway expressed in *E. coli*, resulting in over threefold improvement in the final product. This technique has been used to quantify protein levels for promoter and ribosome-binding site (RBS) variants (Nowroozi et al., 2014), comparison of enzyme homologs (Ma et al., 2011), and to track dynamic regulation of protein levels (Zhang et al., 2012b; Dahl et al., 2013) (Figure 4). More recently, targeted proteomic methods have been optimized for greater throughput (Bath et al., 2014) as well as used to characterize and quantify stable post-translational modifications for engineered microbes. For instance, on polyketide synthases (PKS), by targeting the active site peptide the degree to which it is modified by acyl precursors can be monitored via a phosphopantetheinyl-ejection assay (Dorrestein et al., 2006; Meier et al., 2011). Bottlenecks in PKS function can be determined from this assay similar to pathway metabolite analysis for modular pathways. It has been used to identify competing reactions and characterize the mechanism of PKS function (Hagen et al., 2014; Poust et al., 2015). Pairing bioinformatics tools to identify gene clusters with assays specific

to PKS and NRPS enzymes will stimulate discovery of novel natural products (Bumpus et al., 2009). Adaptation of this technique to other PTMs will facilitate characterization of other classes of enzymes for metabolic engineering applications.

Metabolite Analysis

Measuring protein levels in the cell goes a long way toward characterizing a microbial cell factory, yet, assuming that all of the protein is functional is often incorrect. Consequently, metabolite analyses at the pathway and organism level provide functional information for both pathway and host-engineering research. Tools for strain improvement, such as metabolic flux analysis and constraint-based reconstruction and analysis (COBRA), rely heavily on accurate metabolite data and carbon-flux measurements to restrict parameters for model predictions. Monitoring metabolites that are part the engineered pathway as well as central carbon metabolism aids identification of bottlenecks, and helps identify where increasing specific protein levels can yield dramatic improvements to the product titer or where allosteric regulation is limiting flux through the pathway. Metabolite analysis is commonly carried out as a part of GC-MS and LC-MS target molecule detection assays since pathway intermediates often have similar chemical structures to the target. Fortunately, LC-MS methods that were developed to study central metabolism (Bajad et al., 2006; Lu et al., 2008; Reaves and Rabinowitz, 2011) are directly applicable to metabolic engineering work. Various methods to quantify changes in secondary metabolites or specific pathway intermediates can be readily implemented by labs with the necessary instrumentation. Yet, developing assays for each pathway intermediate is often challenging due to the lack of available standards, intermediates that degrade rapidly, or ones that are isomers. These complications often result in incomplete information regarding the pathway metabolite levels. Designing specific methods for each pathway or class of target molecule is a time- and resource-consuming process that typically results in long LC-MS methods, severely limiting sample throughput. One way to circumvent low throughput LC-MS methods is by using flow-injection acquisition (FIA) (Fuhrer et al., 2011) that omits the chromatography step, relying on high mass accuracy, high-resolution MS for confident metabolite identification. FIA is heavily dependent on reproducible extraction and sample preparation conditions for quantification. A potential compromise between direct injection and long chromatographic methods is the combination of an on-line solid-phase extraction (SPE) method with direct MS detection (Vanderporten et al., 2013).

Non-targeted, discovery-based metabolomics experiments offer great opportunities for metabolic engineering based on comprehensive metabolome analysis. Extensions of genome-scale models to fully utilize the metabolome and integrate multiple omic data-types (Schellenberger et al., 2011; Lerman et al., 2012) have recently been developed to provide greater predictive power for engineered microbes. Coupling multiple extraction, sample preparation, and chromatography methods enables near complete characterization of the small molecule component of a microbe. Yet, sample complexity necessitates long chromatography gradients to adequately separate metabolites. Despite recent



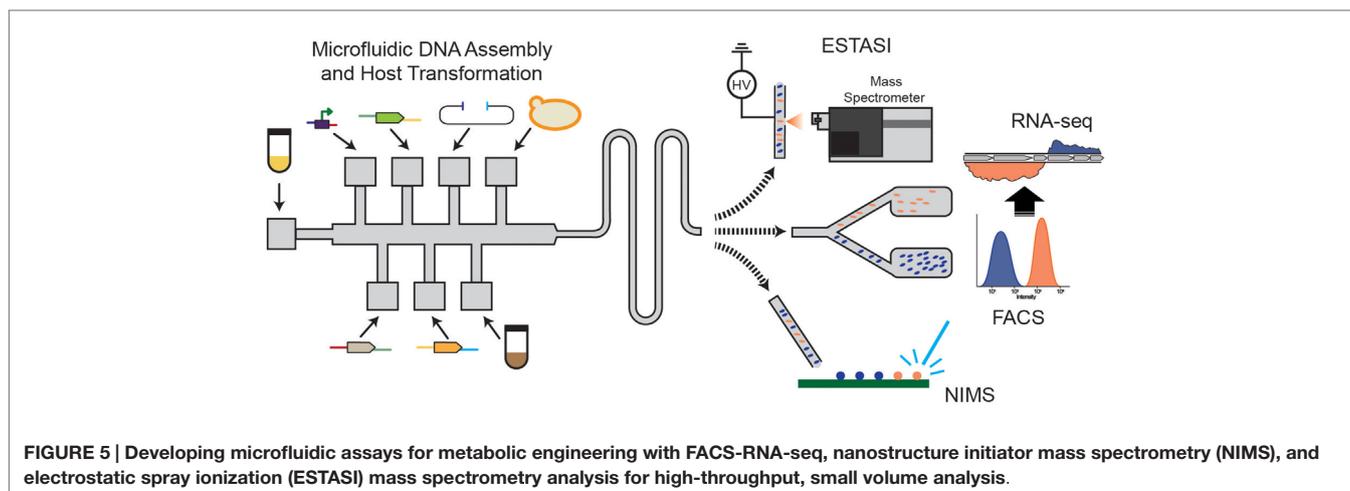
improvements in metabolite identification (Pan et al., 2011), non-targeted metabolomics is still a significant challenge requiring many control samples, metabolite standards, or extensive follow-up experiments. As a result, non-targeted metabolomics methods are most applicable for discovery experiments where the identity of the desired product is not well known, but are challenging to implement for high-throughput quantitative analyses.

Developing Tools

Beyond these technologies are new tools that have the potential to be catalysts for metabolic engineering research. Some of the most promising of these technologies are microfluidic (Liu and Singh, 2013; Wang et al., 2014; Shih et al., 2015) or droplet-based (Abate et al., 2013; Lim and Abate, 2013; Basova and Foret, 2015) systems to build, culture, and analyze many thousands of strains (Figure 5). These systems offer the tantalizing possibility of

testing more than 10,000 unique genotypes per day while overcoming bottlenecks associated with colony picking and culturing limitations. Coupling microfluidic strain construction with FACS and single cell RNA-seq analysis (Abate et al., 2013; Saliba et al., 2014) holds great promise for ultra-high-throughput metabolic engineering. The strengths of these systems revolve around time and reagent cost savings associated with nanoliter and picoliter scale experiments coupled with sensitive spectroscopic assays. Expansion of the spectroscopic toolbox beyond UV/vis and fluorescence assays will depend heavily on adaptation of near-IR, mid-IR, and Raman-based methods that have been successfully implemented in other research fields (i.e., food science, health, materials science), but need further development to be robust for this scale.

There is also great interest in coupling microfluidic and droplet-based systems with mass spectrometric detection to broaden the type of information that can be produced in this



manner (Figure 5). The prospective understanding that this level of strain characterization provides could be transformative. Attempts to couple direct infusion or desorption-based MS are underway and show promise (Kelly et al., 2009; Liu et al., 2010; Gao et al., 2013). For instance, a new ionization method, electrostatic spray ionization (ESTASI), that is compatible with droplet-based methods without dilution or an oil removal step demonstrated high sensitivity at 10 Hz sampling rate (Gasilova et al., 2014). However, issues associated with sample complexity, sensitivity, and acquisition time need to be overcome to enable broad application. The nanostructure initiator mass spectrometry (NIMS) assay combines the versatility of rapid MS analysis with specificity associated with initiators that can be synthesized for a given target (Northen et al., 2008) or are compatible with common chemical biology tools.

Complementing microfluidic and droplet-based systems are innovative assays based on cell-free systems and electrical current that yield very high-throughput advantages. Cell-free systems (Jewett et al., 2008; Hodgman and Jewett, 2012) simplify analysis due to lower complexity matrices, increase engineering flexibility, and enable rapid tuning of parameters, such as proteins or substrate levels (Wang et al., 2012). Similarly, engineering of the electron transfer pathway in *E. coli* provides the foundation for biosensor assays that produce electrical readouts (Jensen et al., 2010) to simplify and generalize detection. Likewise, new detection methods for biosensors based on carbon nanotubes and elemental-tag based antibodies have the potential to significantly broaden biosensor dynamic range and sensitivity (Selvaraju et al., 2008; Yang, 2012). Characterizing these methods for complex backgrounds, such as engineered microbes across a large target molecule concentration range, is necessary to establish robustness of the method and minimize false positives and false negatives.

Ribosomal profiling is an emerging method to quantify the fraction of mRNA transcripts that are being actively translated by ribosomes (Ingolia et al., 2009). This method offers the comprehensiveness of RNA-seq measurements with greater correlation to protein levels. Broad applicability is somewhat hindered by variable translation rates, variability in tRNA abundance, and

technical challenges associated with polysome fractionation. Yet, all of these concerns are surmountable and with the widespread availability of NGS tools the potential to quantify actively translated protein is on verge of becoming a standard workflow in metabolic engineering research. Reverse phase protein arrays (RPPA) (Tibes et al., 2006) are an alternate protein quantification method that has the potential to greatly increase the throughput of proteomic data acquisition. Initial studies, focused on biomedical applications, indicate that it is very sensitive and reproducible but the biotechnology application landscape is unexplored. The time and resources required to develop antibodies for many protein targets and the subsequent experiments needed to validate the method result in slow progress. A large-scale investment for select industrial host proteins could establish a useful quantitative tool for many metabolic engineering applications.

An integrated omics platform has long been the goal of systems biology research across the health and biotechnology fields, yet challenges associated with low data quality, difficulty comparing different types of omics data, and a general lack of datasets for multiple omics methods have kept it from becoming reality. The recent advances to omics techniques described above relating to increased sample throughput, higher data quality, greater method robustness, along with their broader adoption, are encouraging from the data acquisition standpoint. Comprehensive datasets inform statistical analyses and rule-based approaches for subsequent designs that increase the success rate for achieving production of a target molecule. Computational tools, such as correlation analysis, PCA, and machine-learning methods, are just emerging and rely on accurate information to make quality predictions (George et al., 2014; Alonso-Gutierrez et al., 2015). And, development of tools to relate information regarding the genetic background as well as factors associated with cell culturing, sample preparation, acquisition, and analysis will be needed to extract the greatest benefit.

Conclusion

The on-going cost and efficiency improvements occurring to the Design and Build components of the metabolic engineering

cycle are placing significant strain on the Test component because many constructs are needed to find the optimal parameters. This presents a classic “chicken-or-the-egg” dilemma, whereby many test samples are needed to refine parameters for designs while well-defined parameters are needed to reduce the number of samples to test. Currently, a massive amount of resources must be directed to the development of analytical technologies to match the capabilities further upstream in the cycle with the intent to comprehensively cover the design space. Standardized procedures and data-type reporting for metabolic engineering test measurements will enable greater applicability of data across the field. Efforts focused on a combination of these analyses will yield detailed understanding of dynamic processes between multiple parts or systems. Efficient metabolic engineering will be possible with design-of-experiments methods to intelligently

sample strains from large libraries of engineered strains. Large datasets, statistical analyses, and rule-based approaches for subsequent designs will inform models to increase the success rate for achieving production and optimization of a target molecule. In the end, with a rapidly turning DBTL cycle significant steps toward realizing the promise of metabolic engineering will be achievable.

Acknowledgments

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U. S. Department of Energy.

References

- Abate, A. R., Hung, T., Sperling, R. A., Mary, P., Rotem, A., Agresti, J. J., et al. (2013). DNA sequence analysis with droplet-based microfluidics. *Lab. Chip* 13, 4864–4869. doi:10.1039/c3lc50905b
- Ajikumar, P. K., Xiao, W. H., Tyo, K. E., Wang, Y., Simeon, F., Leonard, E., et al. (2010). Isoprenoid pathway optimization for taxol precursor overproduction in *Escherichia coli*. *Science* 330, 70–74. doi:10.1126/science.1191652
- Alonso-Gutierrez, J., Kim, E. M., Batth, T. S., Cho, N., Hu, Q., Chan, L. J., et al. (2015). Principal component analysis of proteomics (PCAP) as a tool to direct metabolic engineering. *Metab. Eng.* 28, 123–133. doi:10.1016/j.ymben.2014.11.011
- Alper, H., Fischer, C., Nevoigt, E., and Stephanopoulos, G. (2005). Tuning genetic control through promoter engineering. *Proc. Natl. Acad. Sci. U.S.A.* 102, 12678–12683. doi:10.1073/pnas.0504604102
- Anderson, J. C., Dueber, J. E., Leguia, M., Wu, G. C., Goler, J. A., Arkin, A. P., et al. (2010). BglBricks: a flexible standard for biological part assembly. *J. Biol. Eng.* 4, 1. doi:10.1186/1754-1611-4-1
- Arike, L., Valgepea, K., Peil, L., Nahku, R., Adamberg, K., and Vilu, R. (2012). Comparison and applications of label-free absolute proteome quantification methods on *Escherichia coli*. *J. Proteomics* 75, 5437–5448. doi:10.1016/j.jprot.2012.06.020
- Atsumi, S., Hanai, T., and Liao, J. C. (2008). Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* 451, 86–89. doi:10.1038/nature06450
- Babiskin, A. H., and Smolke, C. D. (2011). A synthetic library of RNA control modules for predictable tuning of gene expression in yeast. *Mol. Syst. Biol.* 7, 471. doi:10.1038/msb.2011.4
- Bajad, S. U., Lu, W., Kimball, E. H., Yuan, J., Peterson, C., and Rabinowitz, J. D. (2006). Separation and quantitation of water soluble cellular metabolites by hydrophilic interaction chromatography-tandem mass spectrometry. *J. Chromatogr. A* 1125, 76–88. doi:10.1016/j.chroma.2006.05.019
- Basova, E. Y., and Foret, F. (2015). Droplet microfluidics in (bio)chemical analysis. *Analyst* 140, 22–38. doi:10.1039/c4an01209g
- Batth, T. S., Singh, P., Ramakrishnan, V. R., Sousa, M. M., Chan, L. J., Tran, H. M., et al. (2014). A targeted proteomics toolkit for high-throughput absolute quantification of *Escherichia coli* proteins. *Metab. Eng.* 26C, 48–56. doi:10.1016/j.ymben.2014.08.004
- Benz, J. P., Protzko, R. J., Andrich, J. M., Bauer, S., Dueber, J. E., and Somerville, C. R. (2014). Identification and characterization of a galacturonic acid transporter from *Neurospora crassa* and its application for *Saccharomyces cerevisiae* fermentation processes. *Biotechnol. Biofuels* 7, 20. doi:10.1186/1754-6834-7-20
- Binder, S., Schendzielorz, G., Stabler, N., Krumbach, K., Hoffmann, K., Bott, M., et al. (2012). A high-throughput approach to identify genomic variants of bacterial metabolite producers at the single-cell level. *Genome Biol.* 13, R40. doi:10.1186/gb-2012-13-5-r40
- Brunk, E., Neri, M., Tavernelli, I., Hatzimanikatis, V., and Rothlisberger, U. (2012). Integrating computational methods to retrofit enzymes to synthetic pathways. *Biotechnol. Bioeng.* 109, 572–582. doi:10.1002/bit.23334
- Bumpus, S. B., Evans, B. S., Thomas, P. M., Ntai, I., and Kelleher, N. L. (2009). A proteomics approach to discovering natural products and their biosynthetic pathways. *Nat. Biotechnol.* 27, 951–956. doi:10.1038/nbt.1565
- Campononico, M. A., Andrews, B. A., Asenjo, J. A., Palsson, B. O., and Feist, A. M. (2014). Generation of an atlas for commodity chemical production in *Escherichia coli* and a novel pathway prediction algorithm, GEM-Path. *Metab. Eng.* 25, 140–158. doi:10.1016/j.ymben.2014.07.009
- Carbonell, P., Parutto, P., Baudier, C., Junot, C., and Faulon, J. L. (2014a). Retropath: automated pipeline for embedded metabolic circuits. *ACS Synth. Biol.* 3, 565–577. doi:10.1021/sb4001273
- Carbonell, P., Parutto, P., Herisson, J., Pandit, S. B., and Faulon, J. L. (2014b). XTMS: pathway design in an eXTended metabolic space. *Nucleic Acids Res.* 42, W389–W394. doi:10.1093/nar/gku362
- Cardinale, S., and Arkin, A. P. (2012). Contextualizing context for synthetic biology – identifying causes of failure of synthetic biological systems. *Biotechnol. J.* 7, 856–866. doi:10.1002/biot.201200085
- Carothers, J. M., Goler, J. A., Juminaga, D., and Keasling, J. D. (2011). Model-driven engineering of RNA devices to quantitatively program gene expression. *Science* 334, 1716–1719. doi:10.1126/science.1212209
- Chen, Y. J., Liu, P., Nielsen, A. A., Brophy, J. A., Clancy, K., Peterson, T., et al. (2013). Characterization of 582 natural and synthetic terminators and quantification of their design constraints. *Nat. Methods* 10, 659–664. doi:10.1038/nmeth.2515
- Chou, H. H., and Keasling, J. D. (2013). Programming adaptive control to evolve increased metabolite production. *Nat. Commun.* 4, 2595. doi:10.1038/ncomms3595
- Dahl, R. H., Zhang, F., Alonso-Gutierrez, J., Baidoo, E., Batth, T. S., Redding-Johanson, A. M., et al. (2013). Engineering dynamic pathway regulation using stress-response promoters. *Nat. Biotechnol.* 31, 1039–1046. doi:10.1038/nbt.2689
- Davis, J. H., Rubin, A. J., and Sauer, R. T. (2011). Design, construction and characterization of a set of insulated bacterial promoters. *Nucleic Acids Res.* 39, 1131–1141. doi:10.1093/nar/gkq810
- Dietrich, J. A., Mckee, A. E., and Keasling, J. D. (2010). High-throughput metabolic engineering: advances in small-molecule screening and selection. *Annu. Rev. Biochem.* 79, 563–590. doi:10.1146/annurev-biochem-062608-095938
- Dietrich, J. A., Shis, D. L., Alikhani, A., and Keasling, J. D. (2013). Transcription factor-based screens and synthetic selections for microbial small-molecule biosynthesis. *ACS Synth. Biol.* 2, 47–58. doi:10.1021/sb300091d
- Dorrestein, P. C., Bumpus, S. B., Calderone, C. T., Garneau-Tsodikova, S., Aron, Z. D., Straight, P. D., et al. (2006). Facile detection of acyl and peptidyl intermediates on thiotemplate carrier domains via phosphopantetheinyl elimination reactions during tandem mass spectrometry. *Biochemistry* 45, 12756–12766. doi:10.1021/bi061169d
- Dueber, J. E., Wu, G. C., Malmirchegini, G. R., Moon, T. S., Petzold, C. J., Ullal, A. V., et al. (2009). Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol.* 27, 753–759. doi:10.1038/nbt.1557
- Dueber, J. E., Yeh, B. J., Chak, K., and Lim, W. A. (2003). Reprogramming control of an allosteric signaling switch through modular recombination. *Science* 301, 1904–1908. doi:10.1126/science.1085945

- Eggeling, L., Bott, M., and Marienhagen, J. (2015). Novel screening methods-bio-sensors. *Curr. Opin. Biotechnol.* 35C, 30–36. doi:10.1016/j.copbio.2014.12.021
- Feist, A. M., and Palsson, B. O. (2008). The growing scope of applications of genome-scale metabolic reconstructions using *Escherichia coli*. *Nat. Biotechnol.* 26, 659–667. doi:10.1038/nbt1401
- Feng, X., and Zhao, H. (2013). Investigating host dependence of xylose utilization in recombinant *Saccharomyces cerevisiae* strains using RNA-seq analysis. *Biotechnol. Biofuels* 6, 96. doi:10.1186/1754-6834-6-96
- Frederix, M., Hutter, K., Leu, J., Bath, T. S., Turner, W. J., Ruegg, T. L., et al. (2014). Development of a native *Escherichia coli* induction system for ionic liquid tolerance. *PLoS ONE* 9, e101115. doi:10.1371/journal.pone.0101115
- Fuhrer, T., Heer, D., Begemann, B., and Zamboni, N. (2011). High-throughput, accurate mass metabolome profiling of cellular extracts by flow injection-time-of-flight mass spectrometry. *Anal. Chem.* 83, 7074–7080. doi:10.1021/ac201267k
- Gao, D., Liu, H., Jiang, Y., and Lin, J. M. (2013). Recent advances in microfluidics combined with mass spectrometry: technologies and applications. *Lab. Chip* 13, 3309–3322. doi:10.1039/c3lc50449b
- Gasilova, N., Yu, Q., Qiao, L., and Girault, H. H. (2014). On-chip spyhole mass spectrometry for droplet-based microfluidics. *Angew. Chem. Int. Ed. Engl.* 53, 4408–4412. doi:10.1002/anie.201310795
- George, K. W., Chen, A., Jain, A., Bath, T. S., Baidoo, E. E., Wang, G., et al. (2014). Correlation analysis of targeted proteins and metabolites to assess and engineer microbial isopentenol production. *Biotechnol. Bioeng.* 111, 1648–1658. doi:10.1002/bit.25226
- Ghaffari, P., Mardinoglu, A., Asplund, A., Shoaie, S., Kampf, C., Uhlen, M., et al. (2015). Identifying anti-growth factors for human cancer cell lines through genome-scale metabolic modeling. *Sci. Rep.* 5, 8183. doi:10.1038/srep08183
- Gillet, L. C., Navarro, P., Tate, S., Rost, H., Selevsek, N., Reiter, L., et al. (2012). Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol. Cell. Proteomics* 11, O111016717. doi:10.1074/mcp.O111.016717
- Gowen, C. M., and Fong, S. S. (2010). Genome-scale metabolic model integrated with RNAseq data to identify metabolic states of *Clostridium thermocellum*. *Biotechnol. J.* 5, 759–767. doi:10.1002/biot.201000084
- Hagen, A., Poust, S., De Rond, T., Yuzawa, S., Katz, L., Adams, P. D., et al. (2014). In vitro analysis of carboxyacyl substrate tolerance in the loading and first extension modules of borrelidin polyketide synthase. *Biochemistry* 53, 5975–5977. doi:10.1021/bi500951c
- Hatzimanikatis, V., Li, C., Ionita, J. A., Henry, C. S., Jankowski, M. D., and Broadbelt, L. J. (2005). Exploring the diversity of complex metabolic networks. *Bioinformatics* 21, 1603–1609. doi:10.1093/bioinformatics/bti213
- Hodgman, C. E., and Jewett, M. C. (2012). Cell-free synthetic biology: thinking outside the cell. *Metab. Eng.* 14, 261–269. doi:10.1016/j.jymben.2011.09.002
- Ingolia, N. T., Ghaemmaghami, S., Newman, J. R., and Weissman, J. S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* 324, 218–223. doi:10.1126/science.1168978
- Jarboe, L. R., Liu, P., Kautharapu, K. B., and Ingram, L. O. (2012). Optimization of enzyme parameters for fermentative production of biorenewable fuels and chemicals. *Comput. Struct. Biotechnol. J.* 3, e201210005. doi:10.5936/csbj.201210005
- Jensen, H. M., Albers, A. E., Malley, K. R., Londer, Y. Y., Cohen, B. E., Helms, B. A., et al. (2010). Engineering of a synthetic electron conduit in living cells. *Proc. Natl. Acad. Sci. U.S.A.* 107, 19213–19218. doi:10.1073/pnas.1009645107
- Jewett, M. C., Calhoun, K. A., Voloshin, A., Wu, J. J., and Swartz, J. R. (2008). An integrated cell-free metabolic platform for protein production and synthetic biology. *Mol. Syst. Biol.* 4, 220. doi:10.1038/msb.2008.57
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821. doi:10.1126/science.1225829
- Kelly, R. T., Page, J. S., Marginean, I., Tang, K., and Smith, R. D. (2009). Dilution-free analysis from picoliter droplets by nano-electrospray ionization mass spectrometry. *Angew. Chem. Int. Ed. Engl.* 48, 6832–6835. doi:10.1002/anie.200902501
- Kelwick, R., Macdonald, J. T., Webb, A. J., and Freemont, P. (2014). Developments in the tools and methodologies of synthetic biology. *Front Bioeng. Biotechnol.* 2, 60. doi:10.3389/fbioe.2014.00060
- Kittleson, J. T., Wu, G. C., and Anderson, J. C. (2012). Successes and failures in modular genetic engineering. *Curr. Opin. Chem. Biol.* 16, 329–336. doi:10.1016/j.copbio.2012.06.009
- Kizer, L., Pitera, D. J., Pfleger, B. F., and Keasling, J. D. (2008). Application of functional genomics to pathway optimization for increased isoprenoid production. *Appl. Environ. Microbiol.* 74, 3229–3241. doi:10.1128/AEM.02750-07
- Kobayashi, H., Kaern, M., Araki, M., Chung, K., Gardner, T. S., Cantor, C. R., et al. (2004). Programmable cells: interfacing natural and engineered gene networks. *Proc. Natl. Acad. Sci. U.S.A.* 101, 8414–8419. doi:10.1073/pnas.0402940101
- Kolb, H. C., Finn, M. G., and Sharpless, K. B. (2001). Click chemistry: diverse chemical function from a few good reactions. *Angew. Chem. Int. Ed. Engl.* 40, 2004–2021. doi:10.1002/1521-3773(20010601)40:11<2004::AID-ANIE2004>3.0.CO;2-5
- Latimer, L. N., Lee, M. E., Medina-Cleghorn, D., Kohnz, R. A., Nomura, D. K., and Dueber, J. E. (2014). Employing a combinatorial expression approach to characterize xylose utilization in *Saccharomyces cerevisiae*. *Metab. Eng.* 25, 20–29. doi:10.1016/j.jymben.2014.06.002
- Lerman, J. A., Hyduke, D. R., Latif, H., Portnoy, V. A., Lewis, N. E., Orth, J. D., et al. (2012). In silico method for modelling metabolism and gene product expression at genome scale. *Nat. Commun.* 3, 929. doi:10.1038/ncomms1928
- Lim, S. W., and Abate, A. R. (2013). Ultrahigh-throughput sorting of microfluidic drops with flow cytometry. *Lab. Chip* 13, 4563–4572. doi:10.1039/c3lc50736j
- Liu, J., Wang, H., Manicke, N. E., Lin, J. M., Cooks, R. G., and Ouyang, Z. (2010). Development, characterization, and application of paper spray ionization. *Anal. Chem.* 82, 2463–2471. doi:10.1021/ac902854g
- Liu, Y., and Singh, A. K. (2013). Microfluidic platforms for single-cell protein analysis. *J. Lab. Autom.* 18, 446–454. doi:10.1177/2211068213494389
- Looger, L. L., Dwyer, M. A., Smith, J. J., and Hellinga, H. W. (2003). Computational design of receptor and sensor proteins with novel functions. *Nature* 423, 185–190. doi:10.1038/nature01556
- Lu, W., Bennett, B. D., and Rabinowitz, J. D. (2008). Analytical strategies for LC-MS-based targeted metabolomics. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 871, 236–242. doi:10.1016/j.jchromb.2008.04.031
- Ma, S. M., Garcia, D. E., Redding-Johanson, A. M., Friedland, G. D., Chan, R., Bath, T. S., et al. (2011). Optimization of a heterologous mevalonate pathway through the use of variant HMG-CoA reductases. *Metab. Eng.* 13, 588–597. doi:10.1016/j.jymben.2011.07.001
- Malo, N., Hanley, J. A., Carlile, G., Liu, J., Pelletier, J., Thomas, D., et al. (2010). Experimental design and statistical methods for improved hit detection in high-throughput screening. *J. Biomol. Screen.* 15, 990–1000. doi:10.1177/1087057110377497
- Malo, N., Hanley, J. A., Cerquozzi, S., Pelletier, J., and Nadon, R. (2006). Statistical practice in high-throughput screening data analysis. *Nat. Biotechnol.* 24, 167–175. doi:10.1038/nbt1186
- McAlister, G. C., Huttlin, E. L., Haas, W., Ting, L., Jedrychowski, M. P., Rogers, J. C., et al. (2012). Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. *Anal. Chem.* 84, 7469–7478. doi:10.1021/ac301572t
- McCloskey, D., Palsson, B. O., and Feist, A. M. (2013). Basic and applied uses of genome-scale metabolic network reconstructions of *Escherichia coli*. *Mol. Syst. Biol.* 9, 661. doi:10.1038/msb.2013.18
- Medema, M. H., Cimermancic, P., Sali, A., Takano, E., and Fischbach, M. A. (2014). A systematic computational analysis of biosynthetic gene cluster evolution: lessons for engineering biosynthesis. *PLoS Comput. Biol.* 10, e1004016. doi:10.1371/journal.pcbi.1004016
- Meier, J. R., Patel, A. D., Niessen, S., Meehan, M., Kersten, R., Yang, J. Y., et al. (2011). Practical 4'-phosphopantetheine active site discovery from proteomic samples. *J. Proteome Res.* 10, 320–329. doi:10.1021/pr100953b
- Michener, J. K., Thodey, K., Liang, J. C., and Smolke, C. D. (2012). Applications of genetically-encoded biosensors for the construction and control of biosynthetic pathways. *Metab. Eng.* 14, 212–222. doi:10.1016/j.jymben.2011.09.004
- Mutalik, V. K., Guimaraes, J. C., Cambay, G., Lam, C., Christoffersen, M. J., Mai, Q. A., et al. (2013). Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat. Methods* 10, 354–360. doi:10.1038/nmeth.2404
- Northen, T. R., Lee, J. C., Hoang, L., Raymond, J., Hwang, D. R., Yannone, S. M., et al. (2008). A nanostructure-initiator mass spectrometry-based enzyme activity assay. *Proc. Natl. Acad. Sci. U.S.A.* 105, 3678–3683. doi:10.1073/pnas.0712332105
- Nowroozi, F. F., Baidoo, E. E., Ermakov, S., Redding-Johanson, A. M., Bath, T. S., Petzold, C. J., et al. (2014). Metabolic pathway optimization using ribosome binding site variants and combinatorial gene assembly. *Appl. Microbiol. Biotechnol.* 98, 1567–1581. doi:10.1007/s00253-013-5361-4

- Oh, M. K., and Liao, J. C. (2000). DNA microarray detection of metabolic responses to protein overproduction in *Escherichia coli*. *Metab. Eng.* 2, 201–209. doi:10.1006/mben.2000.0149
- Paddon, C. J., Westfall, P. J., Pitera, D. J., Benjamin, K., Fisher, K., McPhee, D., et al. (2013). High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* 496, 528–532. doi:10.1038/nature12051
- Pan, C., Fischer, C. R., Hyatt, D., Bowen, B. P., Hettich, R. L., and Banfield, J. F. (2011). Quantitative tracking of isotope flows in proteomes of microbial communities. *Mol. Cell. Proteomics* 10, M110006049. doi:10.1074/mcp.M110.006049
- Pfleger, B. F., Pitera, D. J., Smolke, C. D., and Keasling, J. D. (2006). Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. *Nat. Biotechnol.* 24, 1027–1032. doi:10.1038/nbt1226
- Picotti, P., and Aebersold, R. (2012). Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. *Nat. Methods* 9, 555–566. doi:10.1038/nmeth.2015
- Picotti, P., Bodenmiller, B., Mueller, L. N., Dörmann, B., and Aebersold, R. (2009). Full dynamic range proteome analysis of *S. cerevisiae* by targeted proteomics. *Cell* 138, 795–806. doi:10.1016/j.cell.2009.05.051
- Poust, S., Phelan, R. M., Deng, K., Katz, L., Petzold, C. J., and Keasling, J. D. (2015). Divergent mechanistic routes for the formation of gem-dimethyl groups in the biosynthesis of complex polyketides. *Angew. Chem. Int. Ed. Engl.* 54, 2370–2373. doi:10.1002/anie.201410124
- Reaves, M. L., and Rabinowitz, J. D. (2011). Metabolomics in systems microbiology. *Curr. Opin. Biotechnol.* 22, 17–25. doi:10.1016/j.copbio.2010.10.001
- Redding-Johanson, A. M., Bath, T. S., Chan, R., Krupa, R., Szmids, H. L., Adams, P. D., et al. (2011). Targeted proteomics for metabolic pathway optimization: application to terpene production. *Metab. Eng.* 13, 194–203. doi:10.1016/j.ymben.2010.12.005
- Robles, J. A., Qureshi, S. E., Stephen, S. J., Wilson, S. R., Burden, C. J., and Taylor, J. M. (2012). Efficient experimental design and analysis strategies for the detection of differential expression using RNA-sequencing. *BMC Genomics* 13, 484. doi:10.1186/1471-2164-13-484
- Rutherford, B. J., Dahl, R. H., Price, R. E., Szmids, H. L., Benke, P. I., Mukhopadhyay, A., et al. (2010). Functional genomic study of exogenous n-butanol stress in *Escherichia coli*. *Appl. Environ. Microbiol.* 76, 1935–1945. doi:10.1128/AEM.02323-09
- Saliba, A. E., Westermann, A. J., Gorski, S. A., and Vogel, J. (2014). Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res.* 42, 8845–8860. doi:10.1093/nar/gku555
- Salis, H. M., Mirsky, E. A., and Voigt, C. A. (2009). Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.* 27, 946–950. doi:10.1038/nbt.1568
- Saxena, R. K., Anand, P., Saran, S., and Isar, J. (2009). Microbial production of 1,3-propanediol: recent developments and emerging opportunities. *Biotechnol. Adv.* 27, 895–913. doi:10.1016/j.biotechadv.2009.07.003
- Schellenberger, J., Que, R., Fleming, R. M., Thiele, I., Orth, J. D., Feist, A. M., et al. (2011). Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nat. Protoc.* 6, 1290–1307. doi:10.1038/nprot.2011.308
- Selvaraju, T., Das, J., Han, S. W., and Yang, H. (2008). Ultrasensitive electrochemical immunosensing using magnetic beads and gold nanocatalysts. *Biosens. Bioelectron.* 23, 932–938. doi:10.1016/j.bios.2007.09.010
- Shadforth, I. P., Dunkley, T. P., Lilley, K. S., and Bessant, C. (2005). i-Tracker: for quantitative proteomics using iTRAQ. *BMC Genomics* 6, 145. doi:10.1186/1471-2164-6-145
- Shih, S. C., Gach, P. C., Sustarich, J., Simmons, B. A., Adams, P. D., Singh, S., et al. (2015). A droplet-to-digital (D2D) microfluidic device for single cell assays. *Lab. Chip* 15, 225–236. doi:10.1039/c4lc00794h
- Shong, J., and Collins, C. H. (2013). Engineering the esaR promoter for tunable quorum sensing-dependent gene expression. *ACS Synth. Biol.* 2, 568–575. doi:10.1021/sb4000433
- Skerker, J. M., Perchuk, B. S., Siryaporn, A., Lubin, E. A., Ashenberg, O., Goulian, M., et al. (2008). Rewiring the specificity of two-component signal transduction systems. *Cell* 133, 1043–1054. doi:10.1016/j.cell.2008.04.040
- Smanski, M. J., Bhatia, S., Zhao, D., Park, Y., and Woodruff, L. (2014). Functional optimization of gene clusters by combinatorial design and assembly. *Nat. Biotechnol.* 32, 1241–1249. doi:10.1038/nbt.3063
- Smith, A. M., Heisler, L. E., Mellor, J., Kaper, F., Thompson, M. J., Chee, M., et al. (2009). Quantitative phenotyping via deep barcode sequencing. *Genome Res.* 19, 1836–1842. doi:10.1101/gr.093955.109
- Stephanopoulos, N., and Francis, M. B. (2011). Choosing an effective protein bio-conjugation strategy. *Nat. Chem. Biol.* 7, 876–884. doi:10.1038/nchembio.720
- Tabor, J. J., Levskaya, A., and Voigt, C. A. (2011). Multichromatic control of gene expression in *Escherichia coli*. *J. Mol. Biol.* 405, 315–324. doi:10.1016/j.jmb.2010.10.038
- Tabor, J. J., Salis, H. M., Simpson, Z. B., Chevalier, A. A., Levskaya, A., Marcotte, E. M., et al. (2009). A synthetic genetic edge detection program. *Cell* 137, 1272–1281. doi:10.1016/j.cell.2009.04.048
- Tang, S. Y., and Cirino, P. C. (2011). Design and application of a mevalonate-responsive regulatory protein. *Angew. Chem. Int. Ed. Engl.* 50, 1084–1086. doi:10.1002/anie.201006083
- Tibes, R., Qiu, Y., Lu, Y., Hennessy, B., Andreeff, M., Mills, G. B., et al. (2006). Reverse phase protein array: validation of a novel proteomic technology and utility for analysis of primary leukemia specimens and hematopoietic stem cells. *Mol. Cancer Ther.* 5, 2512–2521. doi:10.1158/1535-7163.MCT-06-0334
- Van Rossum, T., Kengen, S. W., and Van Der Oost, J. (2013). Reporter-based screening and selection of enzymes. *FEBS J.* 280, 2979–2996. doi:10.1111/febs.12281
- Vanderporten, E., Frick, L., Turincio, R., Thana, P., Lamarr, W., and Liu, Y. (2013). Label-free high-throughput assays to screen and characterize novel lactate dehydrogenase inhibitors. *Anal. Biochem.* 441, 115–122. doi:10.1016/j.ab.2013.07.003
- Wang, B. L., Ghaderi, A., Zhou, H., Agresti, J., Weitz, D. A., Fink, G. R., et al. (2014). Microfluidic high-throughput culturing of single cells for selection based on extracellular metabolite production or consumption. *Nat. Biotechnol.* 32, 473–478. doi:10.1038/nbt.2857
- Wang, H. H., Huang, P. Y., Xu, G., Haas, W., Marblestone, A., Li, J., et al. (2012). Multiplexed in vivo His-tagging of enzyme pathways for in vitro single-pot multienzyme catalysis. *ACS Synth. Biol.* 1, 43–52. doi:10.1021/sb3000029
- Wang, H. H., Isaacs, F. J., Carr, P. A., Sun, Z. Z., Xu, G., Forest, C. R., et al. (2009). Programming cells by multiplex genome engineering and accelerated evolution. *Nature* 460, 894–898. doi:10.1038/nature08187
- Warner, J. R., Reeder, P. J., Karimpour-Fard, A., Woodruff, L. B., and Gill, R. T. (2010). Rapid profiling of a microbial genome using mixtures of barcoded oligonucleotides. *Nat. Biotechnol.* 28, 856–862. doi:10.1038/nbt.1653
- Weber, T., Charusanti, P., Musiol-Kroll, E. M., Jiang, X., Tong, Y., Kim, H. U., et al. (2015). Metabolic engineering of antibiotic factories: new tools for antibiotic production in actinomycetes. *Trends Biotechnol.* 33, 15–26. doi:10.1016/j.tibtech.2014.10.009
- Win, M. N., and Smolke, C. D. (2008). Higher-order cellular information processing with synthetic RNA devices. *Science* 322, 456–460. doi:10.1126/science.1160311
- Yang, H. (2012). Enzyme-based ultrasensitive electrochemical biosensors. *Curr. Opin. Chem. Biol.* 16, 422–428. doi:10.1016/j.cbpa.2012.03.015
- Yang, J. E., Choi, S. Y., Shin, J. H., Park, S. J., and Lee, S. Y. (2013). Microbial production of lactate-containing polyesters. *Microb. Biotechnol.* 6, 621–636. doi:10.1111/1751-7915.12066
- Yim, H., Haselbeck, R., Niu, W., Pujol-Baxley, C., Burgard, A., Boldt, J., et al. (2011). Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol. *Nat. Chem. Biol.* 7, 445–452. doi:10.1038/nchembio.580
- Zhang, F., Carothers, J. M., and Keasling, J. D. (2012a). Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nat. Biotechnol.* 30, 354–359. doi:10.1038/nbt.2149
- Zhang, F., Ouellet, M., Bath, T. S., Adams, P. D., Petzold, C. J., Mukhopadhyay, A., et al. (2012b). Enhancing fatty acid production by the expression of the regulatory transcription factor FadR. *Metab. Eng.* 14, 653–660. doi:10.1016/j.ymben.2012.08.009
- Zhang, H., and Stephanopoulos, G. (2013). Engineering *E. coli* for caffeic acid biosynthesis from renewable sugars. *Appl. Microbiol. Biotechnol.* 97, 3333–3341. doi:10.1007/s00253-012-4544-8

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Petzold, Chan, Nhan and Adams. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.