Recent Applications of Metabolomics to Advance Microbial Biofuel Production

Authors: Julia I Martien¹, Daniel Amador-Noguez¹

¹Department of Bacteriology, University of Madison Wisconsin, United States

Abstract

Biofuel production from plant biomass is a promising source of renewable energy [1]. However, efficient biofuel production involves the complex task of engineering high-performance microorganisms, which requires detailed knowledge of metabolic function and regulation. This review highlights the potential of mass-spectrometry-based metabolomic analysis to guide rational engineering of biofuel-producing microbes. We discuss recent studies that apply knowledge gained from metabolomic analyses to increase the productivity of engineered pathways, characterize the metabolism of emerging biofuel producers, generate novel bioproducts, enable utilization of lignocellulosic feedstock, and improve the stress tolerance of biofuel producers.

Introduction

Maximizing microbial biofuel production from plant biomass (i.e. lignocellulosic biomass or plant dry matter) requires reprogramming metabolism to ensure a seamless supply of carbon, energy (e.g. ATP), and reducing power (e.g. NAD(P)H) towards engineered biofuel pathways. This must be accomplished without eliciting unintended metabolic inefficiencies that frequently accompany engineering efforts. Furthermore, engineered microbes must be capable of catabolizing the diverse sugars (i.e. hexoses and pentoses) present in plant biomass and be tolerant to inhibitory chemicals generated during biomass pretreatment and fermentation (Fig. 1). Over the last decade, advancements in chromatography, mass spectrometry, and computational analysis have enabled accurate systems-level quantification of intracellular metabolites and metabolic fluxes, helping to inform rational bioengineering by revealing the complex processes driving cellular metabolism [2,3]. This review highlights recent studies that have used knowledge gained from mass-spectrometry-based metabolomic analyses to improve the

efficiency of engineered pathways, enable utilization of lignocellulosic sugars, and increase stress tolerance in biofuel-producing microbes (Table 1).

1. Improving activity of engineered pathways and increasing product yields

At the heart of efficient biofuel production lies the challenge of metabolic engineering. Thanks to recent developments in synthetic biology and genetic engineering, manipulating heterologous enzyme expression to construct biofuel-producing pathways in a microbial host is no longer the roadblock it once was [4,5]. However, optimal pathway activity is rarely achieved by simple expression (or overexpression) of required enzymes; product formation can be affected by many other factors including consumption of substrates by competing pathways, energetic and redox imbalances caused by engineered pathway activity, and inhibition due to product accumulation. Metabolomic analyses can guide pathway optimization by identifying sources of metabolic inefficiency and revealing strategies to increase activity of engineered pathways.

1.1 Modeling central metabolism

Characterizing and modeling central carbon metabolism is a critical step in metabolic engineering as it gives a holistic picture of carbon, energy, and redox sources and sinks. Although the central metabolism of model organisms such as *E. coli* and *S. cerevisiae* has been extensively characterized, many organisms with less studied metabolism have emerged as promising alternatives to traditional microbes due to attractive qualities such as high heat tolerance, rapid carbon assimilation, or diverse substrate utilization [6–8]. Metabolomics data can be used to build computational models of carbon flux through metabolic networks in order to develop strategies for metabolically engineering new and promising microbes. Recently, ¹³C Metabolic Flux Analysis (¹³C MFA), a method that models flux based on steady-state isotopic labeling data, has been employed to model metabolism in emerging biofuel producers such as *Thermus thermophilus*, *Clostridium acetobutylicum*, and *Geobacillus sp.* [9–11]. These studies verify genome annotations, characterize non-canonical pathways, and determine major sources and sinks of energy and reduced cofactors.

Although several organisms have already been identified as promising biofuel-producers, the vast diversity of microbial life has not yet been fully explored in the search for optimal organisms. A study by Hollinshead et al. develops a method to improve the rate at which central

carbon metabolism can be mapped in novel organisms [12]. This study describes a ^{13}C fingerprinting kit which allows for metabolic analysis of central pathways including both classical Embden-Meyerhof-Parnas (EMP) and Entner-Doudoroff (ED) glycolysis, the pentose phosphate (PP) pathway, and the TCA cycle via quantification of only a few amino acids. Low cost metabolic analysis such as this will greatly increase the rate at which new organisms with desirable native features can be characterized for further development as biofuel producers.

Central metabolism can be dramatically altered by genetic modifications aimed at increasing bioproduct yields. By quantifying the metabolic responses elicited by increased activity of biofuel-producing pathways, ^{13}C MFA of central metabolism can be used to propose new methods for supporting the carbon, energy and redox demands required for excessive biofuel production [13,14]. For example, He et al. used ^{13}C MFA to quantify the metabolic impact of increased fatty acid production on central metabolism in *E. coli* [13]. Their study found that the PP pathway and TCA cycle were flexible nodes of central metabolism: flux through these pathways adjusted to accommodate carbon and redox demands from fatty acid overproduction, while ED glycolysis was a rigid node: flux through the ED pathway remained low despite its ability to increase supply of both NADPH and acetyl-coA for fatty acid synthesis. Based on these flux dynamics, He et al. identified upregulation of ED glycolysis as a potential strategy to further improve fatty acid production by *E. coli*.

1.2 Altering redox balancing to support biofuel production

Production of highly reduced fuels such as alcohols, fatty acids and hydrocarbons requires an adequate supply of reduced cofactors such as NAD(P)H. Since the availability of reduced cofactors is determined by the combined activity of many catabolic and biosynthetic pathways in primary and secondary metabolism, systems-level metabolic analysis can provide critical insight into redox regulation. Additionally, it has been shown that cellular responses to increased NAD(P)H demands can be metabolically rather than transcriptionally regulated, making metabolomics an important complement to transcriptomic analysis [15].

Recent studies have used metabolomics to identify changes in redox balance caused by metabolic engineering and evaluate the implications for biofuel production [13–16]. Anasontzis et al. used systems-level metabolite quantification to determine the mechanism by which constitutive expression of phosphoglucomutase in upper glycolysis and transaldolase in the PP

pathway resulted in increased ethanol production by *Fusarium oxysporum*. They found that increased expression of these two enzymes improved NADPH regeneration, leading to a shift in carbon flux away from acetate and towards ethanol production [16]. He et al. employed ^{13}C MFA to show that flux through the oxidative PP pathway and conversion of NADH to NADPH by transhydrogenase increased to support NADPH demands from fatty acid overproduction [13]. A complementary study by Wada et al. found that during overproduction of mevalonate, a bioproduct requiring less reducing power than fatty acids, PP pathway flux did not increase, and enhanced transhydrogenase activity was sufficient to support NADPH demands [14]. Taken together, these studies reveal the potential for improving biofuel production by manipulating activity of the oxidative PP pathway and transhydrogenase reactions.

In addition to biofuel production, availability of reducing equivalents is essential for cellular maintenance and protein synthesis. The added energetic and redox demands required for heterologous enzyme expression often limits the efficiency of engineered pathways [17]. Nocon et al. used ^{13}C MFA to compare the flux profiles of several *Pichia pastoris* strains overexpressing different combinations of enzymes in the PP pathway [18]. They found that strains with higher flux through the oxidative PP pathway were better able to express human superoxide dismutase, the best performing strain exhibiting 4-fold increase in heterologous enzyme production. Metabolomics has therefore been demonstrated to be a useful tool in understanding how cellular metabolism can be adjusted to support the redox demands of engineered pathways, both directly for product generation, and indirectly for expression of recombinant enzymes.

1.3 Identifying metabolic bottlenecks and competing pathways

Metabolomics is an important complement to other -omic analysis in assessing pathway efficiency as it directly measures pathway intermediates rather than relying on more indirect indicators such as transcript or enzyme abundance. Accumulation of intermediate metabolites in an engineered pathway is suggestive of a bottleneck caused by enzymatic inefficiency, which may occur even in cases of high protein abundance (Fig 2). George et al. observed this when they sought to improve production of 3-methyl-1-butanol via a heterologous mevalonate pathway in *E. coli* [19]. Although they measured weak expression of upstream enzyme HMG-CoA reductase, metabolite profiling showed that downstream product isopentenyl pyrophosphate (IPP) accumulated to a level higher than any other intermediate in the pathway. This suggested

that the activity of NudB, the enzyme responsible for conversion of IPP to 3-methyl-3-buten-1-ol, was a rate-limiting step. As predicted by this, NudB overexpression resulted in decreased IPP accumulation and a 60% increase in 3-methyl-3-buten-1-ol yields. Pathway activity and metabolic bottlenecks can also be evaluated using time-dependent metabolite quantification or dynamic isotopic labeling to measure the rate of production of pathway intermediates [20,21].

Metabolomics has also proved useful for identifying competing pathways that contribute to inefficiencies in biofuel production. For example, Barton et al. used ^{13}C steady-state isotope labeling to identify competing reactions to 1,4-butanediol production in their engineered *E. coli* strain [21]. Carbon labeling indicated that succinate semialdehyde, the first intermediate in their engineered pathway, was being diverted to succinate and consumed in the TCA cycle. This information directed transcriptomic analysis, leading to the identification of two highly expressed succinate semialdehyde dehydrogenases, whose deletion resulted in improved 1,4-butanediol yields.

1.4 Production of advanced biofuels

Although ethanol remains the most common biofuel, it is considered a suboptimal fuel choice due to its low energetic density and high miscibility in water. Alternatives to ethanol are emerging as target products for microbial engineering, including higher alcohols (e.g. butanol, butanediol and isobutanol), medium- and long-chain hydrocarbons, fatty acids, and isoprenoids [22,23]. These bioproducts are high-grade fuels and have the added potential for use as commodity molecules or replacements for fossil-fuel-derived polymers. Metabolomic analysis has been useful in advancing microbial production of several of these advanced biofuels [19–21,24,25].

Medium-chain hydrocarbon biosynthesis was investigated by Shaw et al. using steady-state labeling [24]. ^{13}C labeling patterns showed that medium-chain hydrocarbon (3E,5E,7E)-nona-1,3,5,7-tetraene was produced from acetate via a pathway related to polyketide synthesis. Characterization of this biosynthetic pathway opens the door for further development of hydrocarbon bioproducts through manipulation of polyketide metabolism. In another study, Yang et al. examined isoprene production by quantifying flux through two isoprene synthesis pathways, the mevalonate (MVA) pathway and the methylerythritol phosphate (MEP) pathway

in *E. coli* [25]. Upregulation of the native MEP pathway and expression of a recombinant MVA pathway resulted in a 20-fold increase in isoprene yield. Flux analysis showed that the two pathways were synergistic: flux through each increased more when the two were overexpressed simultaneously than when each was manipulated individually. Identifying this interplay between the MVA and MEP pathways sets the stage for new advances in microbial engineering for isoprenoid production.

2. Lignocellulose utilization: xylose assimilation and microbial pre-processing

Carbohydrate-rich crops such as sugar-cane and corn have traditionally been used as feedstock for biofuel production as they provide easily accessible sugar substrate. However, lignocellulosic feedstocks such as corn stover, rice straw, or switchgrass are preferable alternatives to edible crops as they lower production cost and reduce competition with global food supply [26]. The challenge in using these lignocellulosic feedstocks is that sugars are present as hard-to-degrade cellulose and hemicellulose polymers that are tightly bound to lignin (a complex aromatic polymer). Due to an inability to liberate simple sugars from lignocellulosic biomass, an inability to metabolize the five-carbon sugars present in lignocellulose, or both, most commonly-used biofuel producers are poorly suited to lignocellulose utilization, making feedstock diversification another critical area for metabolic improvement.

2.1 Xylose assimilation

One of the primary challenges associated with lignocellulosic feedstock is its high pentose content, primarily in the form of xylose which makes up to 25-45% of the sugar content of lignocellulose [27]. Many commonly used biofuel producers including *S. cerevisiae*, *Zymomonas mobilis*, *Clostridium thermocellum*, and *Clostridium cellulolyticum* are unable to utilize xylose natively. Metabolic engineering of some of these organisms has produced strains with heterologous xylose assimilation pathways but this has often led to metabolic imbalances that interfere with xylose to biofuel conversion [28,29]. Recent studies have used ¹³C MFA to identify sources of these inefficiencies and suggest methods for improvement [29–31]. Feng et al. found that in yeast expressing non-native xylose reductase, xylitol dehydrogenase, and xylulose kinase enzymes, flux through the TCA cycle was inversely correlated with biomass yield, indicating that the TCA cycle was used for energy generation to support heterologous

protein expression rather than to support growth or ethanol production [30]. This study also measured high flux through NADH-consuming reactions, indicating an imbalance between NADPH and NADH levels in the recombinant strain. Wasylenko et al. showed that expression of a xylose-isomerase-dependent assimilation pathway relieves these redox imbalances, but results in accumulation of glyceraldehyde-3-phosphate, indicating a bottleneck in lower glycolysis. They suggest increasing activity of enzymes in lower glycolysis to improve ethanol yields [31].

In addition to improving heterologous xylose assimilation pathways, efforts have been made to improve sugar co-utilization by organisms capable of native or evolved xylose assimilation. A study by Aristilde et al. used ^{13}C MFA to examine native sugar co-utilization by *C. acetobutylicum* [32]. This study found that pentoses were assimilated via the phosphoketolase pathway in addition to the PP pathway, thereby identifying overexpression of phosphoketolase enzymes as a potential method by which to increase efficiency of sugar co-utilization. Cordova et al. studied co-utilization of glucose and xylose by a laboratory-evolved strain of *T. thermophilus*, finding that NADPH production was a growth-limiting factor during sugar co-utilization [33]. They suggested incorporation of recombinant oxidative PP pathway enzymes as a way to improve NADPH availability and strain performance.

2.2 Microbial pre-processing

Despite advances in xylose utilization, the complex and recalcitrant structure of lignocellulose makes extracting fermentable sugars difficult. Current methods of lignocellulosic feedstock processing utilize chemical and enzymatic pretreatment to liberate cellulose and hemicellulose from lignin and convert them to simple sugars. Such pre-treatment is a costly step in the biofuel production process, prompting a growing interest in microbial-based pretreatments. Metabolomic analysis of native ligninolytic pathways in wood-rot fungus has contributed to the development of fungal pre-treatment as a replacement for chemical lignin degradation [34]. Additionally, metabolic profiling and engineering of cellulolytic organisms such as *C. cellulolyticum* and *C. thermocellum* has resulted in strains capable of converting crystalline cellulose to biofuels without the addition of hydrolase enzymes [35,36]. These developments offer promising opportunities for significantly reducing the cost of biofuel production, and may also help alleviate growth inhibition by toxic compounds produced during chemical pre-treatment.

3. Understanding stress responses and improving stress tolerance

Breakdown of lignocellulosic biomass during pretreatment is accompanied by the generation of a variety of inhibitory byproducts that lead to inefficient conversion of sugars into biofuels. For example, acid-based pretreatments convert sugars to toxic furans such as furfural and hydroxymethylfurfural, and ammonia-based pretreatments generate a variety of phenolic inhibitors [37]. In many instances, the chemical agents used during pretreatment are themselves inhibitors of microbial metabolism [38–40]. Recent metabolomics studies have started to elucidate the mechanisms underlying the toxicity of this diverse set of microbial inhibitors in order to establish systematic methods for improving stress tolerance of biofuel-producing microbes [39–46].

3.1 Chemical pre-treatment and feedstock-derived inhibitors

Metabolomic analysis has provided valuable information as to which stress responses are induced by pre-treatment-associated inhibitors such as ions (e.g. ammonia, chloride and sulfate), organic weak acids (e.g. acetic and formic acid), furans, and aromatics [39–46]. These studies have identified genetic modifications or nutrient supplementations that improve microbial tolerance to such inhibitors. For example, Hasunuma et al. showed that intermediates of the non-oxidative PP pathway accumulate in xylose-consuming yeast during exposure to acetic acid. Based on this information, they developed a strain with over-expressed transaldolase, leading to improved ethanol yields and titers in the presence of acetic acid [41]. Wang et al. used metabolite quantification and regression modeling to investigate the inhibitory mechanisms of furfural, acetic acid, and phenol in *S. cerevisiae* [42]. They identified proline and myo-inositol synthesis as key pathways associated with tolerant phenotypes and further showed that deletion of genes involved in the synthesis of either metabolite lead to increased sensitivity, while increased biosynthesis or external supplementation with either proline or myo-inositol increased tolerance. A recent study by Pisithkul et al. used dynamic ¹³C-labeling to investigate the toxicity mechanism of phenolic amides (e.g. feruloyl amide and coumaroyl amide) predominant in ammonia-pretreated biomass hydrolysates [43]. They discovered that these phenolic amides disrupt *de novo* nucleotide biosynthesis by competitive inhibition of glutamine

amidotransferases, and that external nucleoside supplementation prevents growth inhibition by allowing nucleotide biosynthesis via salvage pathways.

3.2 Product inhibition

Product inhibition from biofuel accumulation imposes an upper limit on biofuel titer. Metabolomic analysis is at the forefront of understanding the metabolic mechanisms by which tolerant organisms are able to grow under high biofuel concentrations. Regression models that use metabolite profiles as a predictor variable and stress tolerance as a response variable have been a successful tool for identifying candidate pathways, metabolites, and genes involved in biofuel tolerance [42,44–47]. Teoh et al. successfully applied a regression model to engineer a strain of *S. cerevisiae* with improved 1-butanol tolerance based on the identification of threonine as a key tolerance-conferring metabolite [45].

Directed evolution has been a successful method for improving microbial stress tolerance [48]. However, one disadvantage of directed evolution is that the mechanisms behind improved phenotypes are not always apparent, even when genetic alterations are known. Metabolomics can reveal these mechanisms by characterizing the metabolic differences between parent and evolved strains. This was done by Hu et al. to identify mechanisms for butanol tolerance in *Methylobacterium extorquens* [47]. They found a lower abundance of carotenoid precursors in the evolved, tolerant strain despite a lack of genetic mutation in any genes associated with carotenoid synthesis. Membrane fluidity due to decreased carotenoid levels was therefore identified as one mechanism contributing to butanol tolerance. Metabolomic analysis was essential for this discovery, as changes in carotenoid synthesis could not be detected through genomic analysis.

Many recent metabolomics-based studies of ethanol and butanol tolerance have identified the same groups of metabolites involved in stress response, including amino acids, lipids, and disaccharides [42,44–47]. Although the mechanism of action behind some of these metabolites is thought to be understood, most require further study. Most notably, tryptophan has been consistently found by independent studies to be associated with stress tolerance across many taxa [42,44–47]. However, the role of tryptophan in stress response remains unclear, making it a potential area of study for further improvement of stress tolerance in biofuel-producers.

Conclusion

Metabolomic analysis is an indispensable tool for metabolic engineering of biofuel producers. Metabolomics techniques such as metabolite quantification, isotopic labeling, ^{13}C MFA, and regression modeling have been successfully used to guide methods for increasing product yield, enabling utilization of lignocellulosic feedstock, and improving stress tolerance in biofuel producing microbes. However, despite this demonstrated utility, only a small fraction of metabolic engineering projects fully incorporate metabolomics-based analysis into their methods. Further integration of metabolomic analysis into the standard bioengineering tool-set will facilitate advances in biofuel production, helping to meet the growing demand for renewable energy.

References

1. Liao JC, Mi L, Pontrelli S, Luo S: **Fuelling the future: microbial engineering for the production of sustainable biofuels.** *Nat. Publ. Gr.* 2016, **14**.
2. Wang Y, Liu S, Hu Y, Li P, Wan J-B: **Current state of the art of mass spectrometry-based metabolomics studies – a review focusing on wide coverage, high throughput and easy identification .** *RSC Adv.* 2015, **5**:78728–78737.
3. Antoniewicz MR: **Methods and advances in metabolic flux analysis: a mini-review.** *J. Ind. Microbiol. Biotechnol.* 2015, **42**:317–325.
4. Liu R, Bassalo MC, Zeitoun RI, Gill RT: **Genome scale engineering techniques for metabolic engineering .** *Metab. Eng.* 2015, **32**:143–154.
5. Liu Y-J, Zhang J, Cui G-Z, Cui Q: **Current progress of targetron technology: Development, improvement and application in metabolic engineering .** *Biotechnol. J.* 2015, **10**:855–865.
6. Scully SM, Orlygsson J: **Recent advances in second generation ethanol production by thermophilic bacteria.** *Energies* 2015, **8**:1–30.
7. He M, Wu B, Qin H, Ruan Z, Tan F, Wang J, Shui Z, Dai L, Zhu Q, Pan K, et al.: **Zymomonas mobilis: a novel platform for future biorefineries .** *Biotechnol. Biofuels* 2014, **7**:101.
8. Cho C, Jang Y-S, Moon HG, Lee J, Lee SY: **Metabolic engineering of clostridia for the production of chemicals .** *Biofuels, Bioprod. Biorefining* 2015, **9**:211–225.
9. * Swarup A, Lu J, DeWoody KC, Antoniewicz MR: **Metabolic network reconstruction, growth characterization and ¹³C-metabolic flux analysis of the extremophile Thermus thermophilus HB8 .** *Metab. Eng.* 2014, **24**:173–180.
10. ** Au J, Choi J, Jones SW, Venkataramanan KP, Antoniewicz MR: **Parallel labeling experiments validate Clostridium acetobutylicum metabolic network model for ¹³C metabolic flux analysis .** *Metab. Eng.* 2014, **26**:23–33.
11. ** Cordova LT, Antoniewicz MR: **¹³C metabolic flux analysis of the extremely**

thermophilic, fast growing, xylose-utilizing Geobacillus strain LC300 . *Metab. Eng.* 2016, **33**:148–157.

12. ** Hollinshead WD, Henson WR, Abernathy M, Moon TS, Tang YJ: **Rapid metabolic analysis of *Rhodococcus opacus* PD630 via parallel ¹³ C-metabolite fingerprinting** . *Biotechnol. Bioeng.* 2016, **113**:91–100.

13. ** He L, Xiao Y, Gebreselassie N, Zhang F, Antoniewicz MR, Tang YJ, Peng L: **Central metabolic responses to the overproduction of fatty acids in *Escherichia coli* based on ¹³ C-metabolic flux analysis** . *Biotechnol. Bioeng.* 2014, **111**:575–585.

14. * Wada K, Toya Y, Banno S, Yoshikawa K, Matsuda F, Shimizu H: **(13)C-metabolic flux analysis for mevalonate-producing strain of *Escherichia coli*** . *J. Biosci. Bioeng.* 2016, doi:10.1016/j.jbiosc.2016.08.001.

15. Celton M, Sanchez I, Goelzer A, Fromion V, Camarasa C, Dequin S, Nielsen J, Bro C, Regenber B, Forster J, et al.: **A comparative transcriptomic, fluxomic and metabolomic analysis of the response of *Saccharomyces cerevisiae* to increases in NADPH oxidation** . *BMC Genomics* 2012, **13**:317.

16. * Anasontzis GE, Kourtoglou E, Mamma D, Villas-Boas SG, Hatzinikolaou DG, Christakopoulos P: **Constitutive homologous expression of phosphoglucumutase and transaldolase increases the metabolic flux of *Fusarium oxysporum*** . *Microb. Cell Fact.* 2014, **13**:43.

17. Rosano GL, Ceccarelli EA: **Recombinant protein expression in *Escherichia coli*: advances and challenges** . *Front. Microbiol.* 2014, **5**:172.

18. * Nocon J, Steiger M, Mairinger T, Hohlweg J, Rußmayer H, Hann S, Gasser B, Mattanovich D: **Increasing pentose phosphate pathway flux enhances recombinant protein production in *Pichia pastoris*** . *Appl. Microbiol. Biotechnol.* 2016, doi:10.1007/s00253-016-7363-5.

19. * George KW, Thompson MG, Kang A, Baidoo E, Wang G, Chan LJG, Adams PD, Petzold CJ, Keasling JD, Soon Lee T: **Metabolic engineering for the high-yield production of isoprenoid-based C5 alcohols in *E. coli*** . *Sci. Rep.* 2015, **5**:11128.

- 344 20. * Liu H, Lu T: **Autonomous production of 1,4-butanediol via a de novo biosynthesis**
345 **pathway in engineered Escherichia coli** . *Metab. Eng.* 2015, **29**:135–141.
- 346 21. ** Barton NR, Burgard AP, Burk MJ, Crater JS, Osterhout RE, Pharkya P, Steer BA, Sun
347 J, Trawick JD, Dien SJ Van, et al.: **An integrated biotechnology platform for**
348 **developing sustainable chemical processes**. *J Ind Microbiol Biotechnol* 2015, **42**:349–
349 360.
- 350 22. Lee SY, Kim HM, Cheon S: **Metabolic engineering for the production of hydrocarbon**
351 **fuels** . *Curr. Opin. Biotechnol.* 2015, **33**:15–22.
- 352 23. Beller HR, Lee TS, Katz L: **Natural products as biofuels and bio-based chemicals:**
353 **fatty acids and isoprenoids** . *Nat. Prod. Rep.* 2015, **0**:1–19.
- 354 24. * Shaw JJ, Spakowicz DJ, Dalal RS, Davis JH, Lehr NA, Dunican BF, Orellana EA,
355 Narváez-Trujillo A, Strobel SA: **Biosynthesis and genomic analysis of medium-chain**
356 **hydrocarbon production by the endophytic fungal isolate Nigrograna mackinnonii**
357 **E5202H** . *Appl. Microbiol. Biotechnol.* 2015, **99**:3715–3728.
- 358 25. ** Yang C, Gao X, Jiang Y, Sun B, Gao F, Yang S: **Synergy between methylerythritol**
359 **phosphate pathway and mevalonate pathway for isoprene production in Escherichia**
360 **coli** . *Metab. Eng.* 2016, **37**:79–91.
- 361 26. Taha M, Foda M, Shahsavari E, Aburto-Medina A, Adetutu E, Ball A: **Commercial**
362 **feasibility of lignocellulose biodegradation: possibilities and challenges** . *Curr. Opin.*
363 *Biotechnol.* 2016, **38**:190–197.
- 364 27. Collins SR, Wellner N, Martinez Bordonado I, Harper AL, Miller CN, Bancroft I,
365 Waldron KW: **Variation in the chemical composition of wheat straw: the role of tissue**
366 **ratio and composition.** . *Biotechnol. Biofuels* 2014, **7**:121.
- 367 28. Chen R, Dou J: **Biofuels and bio-based chemicals from lignocellulose: metabolic**
368 **engineering strategies in strain development** . *Biotechnol. Lett.* 2016, **38**:213–221.
- 369 29. Matsushika A, Nagashima A, Goshima T, Hoshino T, Hong K, Nielsen J, Gong C, Cao N,
370 Du J, Tsao G, et al.: **Fermentation of Xylose Causes Inefficient Metabolic State Due to**
371 **Carbon/Energy Starvation and Reduced Glycolytic Flux in Recombinant Industrial**

372 **Saccharomyces cerevisiae** . *PLoS One* 2013, **8**:e69005.

373 30. Feng X, Zhao H: **Investigating xylose metabolism in recombinant Saccharomyces**
374 **cerevisiae via 13C metabolic flux analysis** . *Microb. Cell Fact.* 2013, **12**:114.

375 31. * Wasylenko TM, Stephanopoulos G: **Metabolomic and ¹³ C-metabolic flux analysis of**
376 **a xylose-consuming *Saccharomyces cerevisiae* strain expressing xylose isomerase** .
377 *Biotechnol. Bioeng.* 2015, **112**:470–483.

378 32. ** Aristilde L, Lewis IA, Park JO, Rabinowitz JD: **Hierarchy in Pentose Sugar**
379 **Metabolism in *Clostridium acetobutylicum*** . *Appl. Environ. Microbiol.* 2015, **81**:1452–
380 1462.

381 33. ** Cordova LT, Lu J, Cipolla RM, Sandoval NR, Long CP, Antoniewicz MR: **Co-**
382 **utilization of glucose and xylose by evolved *Thermus thermophilus* LC113 strain**
383 **elucidated by 13C metabolic flux analysis and whole genome sequencing** . *Metab.*
384 *Eng.* 2016, **37**:63–71.

385 34. Bak JS: **Lignocellulose depolymerization occurs via an environmentally adapted**
386 **metabolic cascades in the wood-rotting basidiomycete *Phanerochaete chrysosporium***.
387 *Microbiologyopen* 2015, **4**:151–166.

388 35. Gaida SM, Liedtke A, Jentges AHW, Engels B, Jennewein S, Ajanovic A, Panoutsou C,
389 Bauen A, Duffield J, Hill J, et al.: **Metabolic engineering of *Clostridium cellulolyticum***
390 **for the production of n-butanol from crystalline cellulose** . *Microb. Cell Fact.* 2016,
391 **15**:6.

392 36. Dash S, Ng CY, Maranas CD, Amador-Noguez D, Brasg I, Feng X, Au J, Choi J, Jones S,
393 Beguin P, et al.: **Metabolic modeling of clostridia: current developments and**
394 **applications** . *FEMS Microbiol. Lett.* 2016, **363**:23–33.

395 37. Jönsson LJ, Martín C: **Pretreatment of lignocellulose: Formation of inhibitory by-**
396 **products and strategies for minimizing their effects**. *Bioresour. Technol.* 2016,
397 **199**:103–112.

398 38. Dickinson Q, Bottoms S, Hinchman L, McIlwain S, Li S, Myers CL, Boone C, Coon JJ,
399 Hebert A, Sato TK, et al.: **Mechanism of imidazolium ionic liquids toxicity in**

Saccharomyces cerevisiae and rational engineering of a tolerant, xylose-fermenting strain. . *Microb. Cell Fact.* 2016, **15**:17.

39. Mehmood N, Husson E, Jacquard C, Wewetzer S, Büchs J, Sarazin C, Gosselin I: **Impact of two ionic liquids, 1-ethyl-3-methylimidazolium acetate and 1-ethyl-3-methylimidazolium methylphosphonate, on *Saccharomyces cerevisiae*: metabolic, physiologic, and morphological investigations.** *Biotechnol. Biofuels* 2015, doi:10.1186/s13068-015-0206-2.

40. Casey E, Mosier NS, Adamec J, Stockdale Z, Ho N, Sedlak M, Mussatto S, Dragone G, Guimarães P, Silva J, et al.: **Effect of salts on the Co-fermentation of glucose and xylose by a genetically engineered strain of *Saccharomyces cerevisiae*.** *Biotechnol. Biofuels* 2013, **6**:83.

41. Hasunuma T, Sanda T, Yamada R, Yoshimura K, Ishii J, Kondo A: **Metabolic pathway engineering based on metabolomics confers acetic and formic acid tolerance to a recombinant xylose-fermenting strain of *Saccharomyces cerevisiae*.** . *Microb. Cell Fact.* 2011, **10**:2.

- * Wang X, Bai X, Chen D-F, Chen F-Z, Li B-Z, Yuan Y-J: **Increasing proline and myo-inositol improves tolerance of *Saccharomyces cerevisiae* to the mixture of multiple lignocellulose-derived inhibitors.** *Biotechnol. Biofuels* 2015, **8**:142.

- * Pisithkul T, Jacobson TB, O'Brien TJ, Stevenson DM, Amador-Noguez D: **Phenolic amides are potent inhibitors of de novo nucleotide biosynthesis.** *Appl. Environ. Microbiol.* 2015, **81**:5761–5772.

44. Hashim Z, Fukusaki E, **Metabolomics-based prediction models of yeast strains for screening of metabolites contributing to ethanol stress tolerance.** *IOP Conf. Ser. Earth Environ. Sci.* 2016, **36**:12046.

- ** Teoh ST, Putri S, Mukai Y, Bamba T, Fukusaki E: **A metabolomics-based strategy for identification of gene targets for phenotype improvement and its application to 1-butanol tolerance in *Saccharomyces cerevisiae*.** . *Biotechnol. Biofuels* 2015, **8**:144.

46. Zhu H, Ren X, Wang J, Song Z, Shi M, Qiao J, Tian X, Liu J, Chen L, Zhang W:

428 **Integrated OMICS guided engineering of biofuel butanol-tolerance in photosynthetic**
429 **Synechocystis sp. PCC 6803.** . *Biotechnol. Biofuels* 2013, **6**:106.

- 430 47. * Hu B, Yang Y-M, Beck DAC, Wang Q-W, Chen W-J, Yang J, Lidstrom ME, Yang S:
431 **Comprehensive molecular characterization of Methylobacterium extorquens AM1**
432 **adapted for 1-butanol tolerance** . *Biotechnol. Biofuels* 2016, **9**:84.

- 433 48. Cobb RE, Sun N, Zhao H: **Directed evolution as a powerful synthetic biology tool.**
434 *Methods* 2013, **60**:81–90.

435

436

437 Acknowledgements:

438 Thank you to Tyler Jacobson for reading the manuscript.

439 This work was supported by the Department of Energy, Great Lakes Bioenergy Research Center
440 (DOE BER Office of Science DE-FC02-07ER64494).

441 s/o SBE

442

443 Figure Legends:

444

445 Figure 1. The major goals of metabolic engineering for microbial biofuel production are (1) to
446 direct metabolic flux towards maximum biofuel generation, (2) to enable use of economical
447 feedstock such as lignocellulose, and (3) to improve stress tolerance to inhibitors produced
448 during pre-processing or biofuel production. The studies featured in this review apply knowledge
449 gained from metabolomics-based methods to achieve these goals.

450

451 Figure 2. Metabolomics is an essential complement to proteomics and transcriptomics for
452 identifying targets for pathway improvement. In this scheme, metabolite D is the desired
453 bioproduct. Based on proteomics or transcriptomics alone, gene 1 might be targeted for pathway
454 improvement due to its low transcript and enzyme abundance. However, accumulation of
455 metabolite B indicates that the conversion of metabolite B to D by enzyme 3 constitutes the
456 major bottleneck in the pathway, making gene 3 a better target for increased expression rather
457 than gene 1. Flux analysis also identifies conversion of metabolite B to C by enzyme 2 as a
458 major competing flux, making gene 2 a target for downregulation or deletion.

459

460

461

462

463

Area	Major findings	Metabolomics Methods	Bioproduct(s)	Organism	Reference
Pathway Engineering	Metabolite quantification detected accumulation of isopentenyl pyrophosphate, indicating that NudB was a bottleneck enzyme in engineered heterologous MVA pathway. Increased NudB expression resulted in a 60% increase in methylbutenol production.	LC-MS	Isoprene-derived alcohols: 3-methyl-2-butenol, 3-methyl-3-butenol, and 3-methyl-1-butanol	<i>E. coli</i>	[19]*
	Time-dependent metabolite profiling was used to verify enzymatic activity of engineered xylose to butanediol pathway before regulation of the pathway was optimized.	GC and GC-MS	1,4-butanediol	<i>E. coli</i>	[20]*
	Integrated –omics were used to design and optimize a succinyl-CoA to butanediol pathway. Dynamic labeling showed that aldehyde dehydrogenase was a rate-limiting step, guiding targeted enzyme engineering that resulted in a 20% increase in titer. Competing TCA cycle reactions were identified using targeted transcriptomics, directed by isotopic labeling.	LC-MS, dynamic isotopic labeling	1,4-butanediol	<i>E. coli</i>	[21]**
	The mevalonate (MVA) and methylerythritol-phosphate (MEP) pathways were found to be synergistic in isoprene production. Dual overexpression of MAV and MEP pathways resulted in enhanced productivity beyond the expected additive effect of individual overexpression, with a final titer of 24 g/L.	LC-MS, steady-state isotopic labeling, ¹³ C MFA	Isoprene	<i>E. coli</i>	[25]**
	Medium-chain hydrocarbon nonatetraene was found to be synthesized via a polyketide-synthesis-related pathway whereby head to tail condensation of acetate is followed by decarboxylation. Candidate polyketide synthases were identified.	GC-MS, steady-state isotopic labeling	(3E,5E,7E)-nona-1,3,5,7-tetraene	<i>Nigrograna mackinonnii</i>	[24]*
Redox Balancing	Yeast responded to increased NADPH demand by increasing acetate production and rerouting flux towards the PPP. These responses were metabolically or transcriptionally controlled depending on the severity of NADPH demand.	GC-MS and IE-MS/MS, steady-state isotopic labeling, ¹³ C MFA	any	<i>Saccharomyces cerevisiae</i>	[15]
	Constitutive expression of phosphoglucosmutase and transaldolase increased ethanol yield. Metabolomic analysis identified increased NADPH availability leading to reduced acetate production as a major source of improvement.	GC-MS	Ethanol	<i>Fusarium oxysporum</i>	[16]*
	Increased carbon and redox demands of fatty acid overproduction resulted in increased flux through the oxidative PPP and increased conversion of NADH to NADPH by transhydrogenases. ED glycolysis increased only slightly despite its ability to generate both NADPH and acetyl-coA for fatty acid synthesis, making ED glycolysis a potential target for metabolic improvement.	GC-MS, steady-state isotopic labeling, ¹³ C MFA	Fatty Acids	<i>E. coli</i>	[13]**

	Increased carbon and redox demands of mevalonate overproduction were met by conversion of NADH to NADPH via transhydrogenase. This study did not see increases in oxidative PPP flux, suggesting that the moderate NADPH demands for mevalonate production could be met by transhydrogenase reactions alone.	GC-MS, steady-state isotopic labeling, ^{13}C MFA	Mevalonate	<i>E. coli</i>	[14]*
	A correlation was observed between increased PPP flux and increased expression of recombinant protein. Combined overexpression of glucose-6-phosphate dehydrogenase and 6-phosphogluconolactonase resulted in the highest PPP flux and the highest expression levels of recombinant protein.	GC-MS, steady-state isotopic labeling, ^{13}C MFA	any	<i>Pichia pastori</i>	[18]*
Central Metabolism	Flux modeling of central carbon metabolism verified the absence of ED glycolysis and oxidative PPP, and showed high TCA cycle flux.	GC-MS, steady-state isotopic labeling, ^{13}C MFA	any	<i>Thermus thermophilus</i>	[9]*
	Flux modeling of central carbon metabolism revealed non-canonical TCA cycle reactions, generation of C_1 from pyruvate, and isoleucine production via citramalate synthase.	GC-MS, parallel steady-state isotopic labeling, ^{13}C MFA	Butanol, ethanol and acetone	<i>C. acetobutylicum</i>	[10]**
	Flux modeling of central carbon metabolism showed that the TCA cycle and oxidative PPP are responsible for NADPH production during growth on xylose.	GC-MS, parallel steady-state isotopic labeling, ^{13}C MFA	any	<i>Geobacillus sp.</i>	[11]**
	^{13}C fingerprinting based on labeling patterns of only a few amino acids was used to assess the metabolic activity of EMP and ED glycolysis, gluconeogenesis, the glyoxylate shunt, anaplerotic pathways, and amino acid synthesis in a non-model organism.	GC-MS, parallel steady-state isotopic labeling, ^{13}C fingerprinting	any	<i>Rhodococcus opacus</i> PD630	[12]**
Xylose Utilization	Expression of heterologous xylose reductase (XR), xylitol dehydrogenase (XDH), and xylulose kinase enzymes leads to increased flux through the oxidative PPP and TCA cycle to meet increased NADPH and energy demands, limiting ethanol production.	GC-MS, steady-state isotopic labeling, ^{13}C MFA	Ethanol	<i>S. cerevisiae</i> (XR-XDH xylose assimilation)	[30]
	Yeast strain with xylose-isomerase (XI) based xylose assimilation did not exhibit high flux through oxidative PPP suggesting XI ameliorates the redox imbalances seen in XR-HDH strains. However, bottlenecks in lower glycolysis limit ethanol production.	GC-MS, steady-state isotopic labeling, ^{13}C MFA	Ethanol	<i>S. cerevisiae</i> (XI xylose assimilation)	[31]*
	During sugar co-utilization, hexoses were assimilated via glycolysis while pentoses were incorporated into the PPP. The phosphoketolase pathway plays an important role in pentose metabolism and could be targeted for strain improvement.	GC-MS, parallel steady-state isotopic labeling, ^{13}C MFA	Butanol, ethanol, and acetone	<i>C. acetobutylicum</i> BOH3	[32]**
	In xylose-utilizing strain developed via directed evolution, NADPH production was identified as a limiting factor during growth on xylose, suggesting that expression of heterologous oxidative PPP enzymes may improve strain performance.	GC-MS, parallel steady-state isotopic labeling, ^{13}C MFA	any	<i>T. thermophilus</i> developed via adaptive evolution of strain HB8	[33]**

Stress Tolerance	Acetic acid was found to inhibit xylose fermentation due to an accumulation of intermediates of the non-oxidative PPP. Overexpression of transaldolase relived this bottleneck and improved ethanol yields.	GC-MS and CE-MS	Ethanol	<i>S. cerevisiae</i> (XR-XDH xylose assimilation)	[41]
	Proline and myo-inositol were identified as key metabolites in tolerance to furfural, acetic acid, and phenol. Accumulation or supplementation of these metabolites mitigates growth inhibition.	GC-MS, Regression modeling (PLS-DA)	any	<i>S. cerevisiae</i>	[42]*
	Phenolic amides inhibit nucleotide biosynthesis via competitive inhibition of glutamine amidotransferase. External supplementation with nucleosides mitigates growth inhibition.	LC-MS, dynamic isotopic labeling	any	<i>E. coli</i>	[43]*
	Threonine was identified as a key metabolite contributing to butanol tolerance based on metabolomics-based regression modeling. Gene deletions aimed at increasing threonine accumulation resulted in improved butanol tolerance, providing a proof of concept for semi-rational engineering based on metabolomics data.	GC-MS, Regression modeling (OPLS)	1-butanol	<i>S. cerevisiae</i>	[45]**
	Directed evolution for improved butanol tolerance resulted in increased abundance of disaccharides and saturated fatty acids and decreased levels of carotenoids and carotenoid precursors, suggesting membrane fluidity and osmotic control are important factors in butanol tolerance.	GC-MS and LC-MS	1-butanol	<i>Methylobacterium extorquens</i> AM1	[47]*

Table 1. Application of metabolomic-based analyses to biofuel production. Studies marked with * are of special interest due to their use of in-depth diagnostic metabolomic analysis to propose or develop systematic methods for improved biofuel production. Studies marked with ** are of outstanding interest due to their development of widely applicable methods, work with emerging biofuel-producers and/or bioproducts, extremely thorough use of metabolomics tools, or some combination of these three aspects.