A concerted systems biology analysis of phenol metabolism in *Rhodococcus opacus* PD630

Garrett W. Roell[1], Rhiannon R. Carr[1], Tayte Campbell[2], Zeyu Shang[1], William R. Henson[1], Jeffrey J. Czajka[1], Hector García Martín[3,4,5,6], Fuzhong Zhang[1], Marcus Foston[1], Gautam Dantas[2,7,8,9], Tae Seok Moon[1]*, and Yinjie J. Tang[1]*

[1] Department of Energy, Environmental and Chemical Engineering, Washington University in St. Louis, St. Louis, MO, 63130, USA.

[2] The Edison Family Center for Genome Sciences and Systems Biology, Washington University in St. Louis School of Medicine, St. Louis, MO, 63110, USA

[3] DOE, Joint BioEnergy Institute, Emeryville, CA, 94608, USA

[4] DOE, Agile BioFoundry, Emeryville, CA, 94608, USA

[5] Biological Systems and Engineering Division, Lawrence Berkeley National Lab, Berkeley, CA, 94720, USA

[6] BCAM, Basque Center for Applied Mathematics, Bilbao, Spain

[7] Department of Pathology and Immunology, Washington University in St. Louis School of Medicine, St. Louis, MO, 63108, USA

[8] Department of Biomedical Engineering, Washington University in St. Louis, St Louis, MO, 63130, USA

[9] Department of Molecular Microbiology, Washington University in St. Louis School of Medicine, St. Louis, MO, 63108, USA

* Corresponding authors: tsmoon@wustl.edu (TSM) and yinjie.tang@seas.wustl.edu (YT)

Additional authors: garrettroell@wustl.edu (GR), rrcarr@wustl.edu (RC), tayte.campbell@gmail.com (TC), Zeyu.Shang@wustl.edu (ZS), williamrhenson@gmail.com (WH), jjczajka@umich.edu (JC), hgmartin@ldl.gov (HGM), fzhang29@wustl.edu (FZ), mfoston@wustl.edu (MF), dantas@wustl.edu (GD)

Abstract

Rhodococcus opacus PD630 metabolizes aromatic substrates and naturally produces branched-chain lipids, which are advantageous traits for lignin valorization. To provide insights into its lignocellulose hydrolysate utilization, we performed ^{13}C pathway tracing, ^{13}C-pulse-tracing, transcriptional profiling, biomass compositional analysis, and metabolite profiling in conjunction with ^{13}C-metabolic flux analysis (^{13}C-MFA) of phenol metabolism. We found that 1) phenol is metabolized mainly through the ortho–cleavage pathway; 2) phenol-utilization requires a highly active TCA cycle; 3) NADPH is generated mainly via NADPH-dependent isocitrate dehydrogenase; 4) active cataplerotic fluxes increase plasticity in the TCA cycle; 5) gluconeogenesis occurs partially through the reversed Entner–Doudoroff pathway (EDP). We also found that phenol-fed *R. opacus* PD630 generally has lower sugar phosphate concentrations (e.g., fructose 1,6-bisphosphatase) compared to metabolite pools in ^{13}C-glucose-fed *Escherichia coli* (set as internal standards), while its TCA metabolites (e.g., malate, succinate, and α-ketoglutarate) accumulate intracellularly with measurable succinate secretion. In addition, we found that phenol utilization was inhibited by benzoate, while catabolite repressions by other tested carbon substrates (e.g., glucose and acetate) were absent in *R. opacus* PD630. Three adaptively-evolved strains display very different growth rates when fed with phenol as a sole carbon source, but they maintain a conserved flux network. These findings improve our understanding of *R. opacus*' metabolism for future lignin valorization.

Key words: ^{13}C-MFA, ^{13}C pulse-tracing, Entner–Doudoroff pathway, Gluconeogenesis, Lignin

Introduction

Engineering bacterial cells to utilize lignocellulosic biomass as a feedstock could offer cost-competitive production of renewable fuels, chemicals, and materials (Ragauskas et al., 2006). Lignocellulosic biomass is composed mainly of cellulose, hemicellulose, and lignin. Lignocellulosic carbohydrates, cellulose and hemicellulose, can be deconstructed into fermentable sugars for biological conversion into a range of bioproducts. Lignocellulose is a recalcitrant feedstock for biological conversion, in part due to its up to 40% lignin content (Gani and Naruse, 2007). Pretreatments are employed prior to carbohydrate deconstruction to increase sugar and thus product yield. However, lignin degradation during pretreatment often produces aromatic compounds that strongly inhibit the growth of most bacteria (Beckham et al., 2016). Moreover, conventional fermentative biorefineries treat lignin as by-product, and burn it for low-value electricity or process heat (Ragauskas et al., 2014). Techno-economic analyses have reported that the use of lignin for co-product generation is a key to profitable lignocellulose valorization (Valdivia et al., 2016). Lignin cannot be rapidly broken down by biological methods (Boerjan et al., 2003; Studer et al., 2011). Thus, an approach for lignin utilization has been proposed that combines the rapid depolymerization kinetics of chemical processing with the metabolic funneling and selective abilities of microbial systems (Linger et al., 2014; Wheeldon et al., 2017). Recently, research groups are attempting to develop microbial hosts with improved tolerance to aromatic compounds for fermentation of lignocellulose-derived sugars and lignin-derived aromatic compounds (Abdelaziz et al., 2016; Ragauskas et al., 2014; Janusz et al. 2017). In particular, lignin-rich streams have been utilized to generate lipid-based biofuels (i.e., waste-to-fuel applications) (Le et al., 2017; Yaguchi et al., 2017).

70 *Rhodococcus opacus* PD630 (hereafter *R. opacus*) is a Gram-positive actinobacterium. It is
71 natively able to utilize sugars, aromatics, furans, and organic acids (Holder et al., 2011; Kurosawa
72 et al., 2015). Isolated from a gas works plant, this strain is capable of accumulating triacylglycerol
73 (TAG), the precursor for biodiesel up to 80% of its dry cell weight (Alvarez et al., 1996). Studies
74 of its mechanisms of aromatic tolerance and TAG accumulation have found promising results for
75 its use in the conversion of lignin-derived aromatic substrates (DeLorenzo et al., 2017; Henson et
76 al., 2018a; Henson et al., 2018b; Yoneda et al., 2016). *R. opacus* has been extensively studied for
77 its growth kinetics and transcriptional activities, yet key knowledge gaps from genotype to
78 phenotype still remain. Specifically, there is little knowledge about the distribution of the
79 metabolic fluxes in this biofuel producer when consuming aromatic substrates. Functional
80 characterization of the central metabolic network is necessary to develop this nonmodel platform
81 for metabolic engineering applications, particularly to facilitate genome-scale modeling for
82 rational strain design.

83 In this work, we have characterized phenol-fed cultures of *R. opacus* based on integrated omics
84 analyses. ^{13}C-Metabolic Flux Analysis (^{13}C-MFA) can measure *in vivo* reaction rates of central
85 metabolism and confirm global gene regulation patterns inferred from transcriptomic data
86 (Chubukov et al., 2013). In the past, ^{13}C-MFA mostly focused on cellular metabolisms fed with
87 sugars, organic acids, and CO_2 (photosynthesis). To our knowledge, this work is the first report of
88 a full MFA completed in any organism using phenol as the sole carbon source. Additionally, we
89 performed ^{13}C-pathway tracing, metabolite analysis, flux balance analysis, and biomass
90 composition analysis. The flux-based insights were further connected to recently reported
91 transcriptional profiles and proteomics data. We also examined how adaptive laboratory evolution

affects the central flux network. These systems analyses provide a deeper understanding of *R. opacus* metabolism of lignin-derived substrates at different layers of cellular processes.

Materials and Methods

Chemicals, strains, and growth conditions. 1-^{13}C phenol, 2,6-^{13}C phenol, and U-^{13}C phenol were purchased from Cambridge Isotope Laboratories Inc. (Tewksbury, MA, USA). All other chemicals were purchased from Sigma-Aldrich (St. Louis, MO, USA). All cells were grown in minimal media, previously described as media B (DeLorenzo et al., 2017). Unless otherwise noted, the sole carbon source was 0.5 g/L phenol, and the nitrogen source was 1 g/L ammonium sulfate. For all cultures, a single colony of *R. opacus* from a tryptic soy broth plate was inoculated into 2 mL minimal medium, which was used as a seed culture for 10 mL cultures. Isotopomer measurements were conducted when cell cultures reached an optical density at 600 nm (OD_{600}) of 0.2-0.5, corresponding to exponential growth.

Analysis of proteinogenic amino acid labeling. 10 mL labeled biomass from 2,6-^{13}C phenol or 1-^{13}C phenol growing cultures (biological duplicates, n=2) were centrifuged down then hydrolyzed at 100 ^{o}C for 24 hours. The resulting amino acids were analyzed by GC-MS via the TBDMS (N-(tert-butyldimethylsilyl)-N-methyltrifluoroacetamide) method as described previously (You et al., 2012). In brief, the $[m-57]^+$ and $[m-159]^+$ fragments for amino acids were mainly used for ^{13}C-MFA, while the $[m-15]^+$, $[m-85]^+$, and $f(302)^+$ fragments were also analyzed to provide additional labeling information. The m+0, m+1, m+2, etc. data corresponds to the fraction of fragments that

were unlabeled, singly labeled, doubly labeled, etc. Labeling data were corrected for naturally-occurring isotopes.

Dynamic labeling (pulse-tracing), free metabolite extraction, and LC-MS analysis. For phenol dynamic labeling experiments, 20 mL unlabeled phenol cultures were grown to OD_{600} ~ 0.2 and then pulsed with 2 mL of a 5.25 g/L U-^{13}C phenol stock solution for a final labeled phenol concentration of ~0.5 g/L. The cultures were quenched at 5 seconds, 1 minute, and 1 hour using the Fast-Cooling procedure (Hollinshead et al. 2016; Abernathy et al., 2017). To study the effect of benzoate on phenol utilization, 40 mL cultures with 0.5 g/L fully labeled phenol were grown to OD_{600} ~ 0.35, and then the labeled cultures were mixed with 4 mL of 4 g/L unlabeled benzoate. After the benzoate pulse, the cultures were quenched at 20 seconds, 1 minute, and 3 hours.

At each metabolite quenching time, 10 mL of culture was poured into ice cold carbon-free media B, and then placed into a liquid nitrogen bath. The samples were stirred briefly and then centrifuged (8000 g) at 0 °C for 5 minutes. The pellets were stored at -80 °C until metabolite extraction. To determine labeling in free metabolites, cell pellets were resuspended in 1 mL of 7:3 methanol:chloroform, and shaken at 250 rpm at 4 °C for 6 hours with hourly vortexing. The methanol layer was separated by the addition of 500 µL water followed by centrifugation after which the aqueous layer was collected. This separation process was repeated twice. The collected samples were frozen at -80 °C, then lyophilized at -50 °C. The samples were reconstituted with water and analyzed using an Ion Pairing LC-MS method at the Donald Danforth Plant Science Center (Creve Coeur, MO, USA) (Abernathy et al., 2017). To measure metabolite pool size, an isotope ratio approach was used (Bennett et al., 2008; Abernathy et al., 2017). More specifically, *E. coli* was grown on U-^{13}C glucose to produce fully-labeled internal standards for each metabolite,

while *R. opacus* culture was grown on unlabeled substrates. The cultures were quenched and mixed together at a 1:1 biomass content ratio for metabolite extraction and LC-MS measurements. The relative metabolite pool size was calculated based on ^{13}C ratios.

Flux calculations. Steady-state flux analysis calculations were done using the software INCA (Isotopomer Network Compartmental Analysis) (Young, 2014). The basis of the *R. opacus* flux model is the KEGG genome annotation, including the β–ketoadipate pathway, the TCA cycle, the anaplerotic reactions, glycolysis, and the EDP. Carbon mappings for both ortho and meta pathways are shown in **Supplemental Figure 1**. The *R. opacus* central flux network is largely similar to the model bacterial species *E. coli* with major differences in the inclusion of phenol → acetyl-CoA + succinyl-CoA, a modified biomass equation (Crown et al., 2015), and the reversibility of glycolysis and the EDP to allow for gluconeogenesis. Labeling data of key proteinogenic amino acids, biomass growth, and substrate consumption rate were used for flux calculation. Based on measurements errors from parallel labeling experiments, 90% confidence intervals of fluxes were calculated via INCA's parameter continuation function (See supplementary Excel file for metabolic reactions, flux constraints, amino acid labeling fitting, and statistical analysis).

Flux balance analysis (FBA) of the central metabolic network was done using INCA to further compare actual metabolic fluxes with optimal flux topologies. In addition to maximizing biomass or fatty acid productions, three other objective functions were also tested with fixed biomass growth rate: maximizing NAD(P)H, ATP, or minimizing sum of fluxes (i.e., enzyme usage). The FBA results were compared to ^{13}C-MFA by the sum of squared residuals between optimal flux and measured flux in central pathways.

Biomass composition analyses and substrate/product concentration measurements. Cell cultures were grown to exponential phase, pelleted, washed with sterile water and then lyophilized. Amino acid composition analysis was performed by the Molecular Structure Facility, University of California (Davis, CA, USA). Fatty acid composition analysis was performed by Microbial ID (Newark, DE, USA). Overflow metabolites were analyzed using enzyme kits from R-Biopharm (Germany). Phenol and benzoate in the culture supernatant were derivatized using methyl chloroformate (Madsen et al., 2016) and quantified using a custom GC-MS-FID system (gas chromatography-mass spectrometry-flame ionization detector). Briefly, the culture supernatant was centrifuged at 16,000 g for 5 min, and 200 µL of the culture supernatant was mixed with 40 µL of 5.0% (w/w) sodium hydroxide solution, 200 µL of methanol, and 50 µL of pyridine. Methyl chloroformate (50 µL) was added to the mixture in two 25 µL aliquots. Next, 400 µL of chloroform containing a decane internal standard was added to the sample, followed by the addition of 400 µL of 50 mM sodium bicarbonate solution to induce phase separation of the aqueous and organic layers. Samples were vortexed between each step to ensure complete mixing. After phase separation, the organic phase was transferred to a GC vial with a 350 µL glass insert (Agilent), and samples were analyzed using an Agilent 7890A GC coupled to both an Agilent 5975C mass spectrometer containing a triple-axis detector and an Agilent G3461A FID detector with a methanizer (Activated Research Company; Eden Prarie, MN, USA). The Agilent 7890A GC was equipped with a Restek fused silica RTX-50 capillary column (30 m by 0.25 mm, 0.5 µm film thickness), and helium was used as the carrier gas. 1 µL of the organic phase was injected with a splitting ratio of 10:1 using the autosampler. For GC runs, the inlet was maintained at 250 °C, and the oven was held for 2 min at 40 °C, heated to 300 °C using a 5 °C/min ramp, and held at 300 °C

for 5 min. All data was exported and analyzed using the Agilent ChemStation Software, and peak intensities were normalized to the decane internal standard. Phenol and benzoate were identified based on retention time and concentrations were determined using external standard curves.

RNA-seq analysis. Raw transcriptomic sequences were downloaded from the Sequence Read Archive (SRA) accession number SRP131196 (Henson et al., 2018a). Sequences were quality trimmed using Trimmomatic (Bolger et al., 2014) before mapping to the Chinese Academy of Science reference genome of *R. opacus* (ASM59954v1) (Chen et al., 2014) using Bowtie2 (Langmead and Salzberg, 2012). The resulting SAM files were converted to the BAM format and indexed using SAMtools (Li et al., 2009). Expression counts were then calculated using featureCounts (Liao et al., 2014) and normalized using DESeq2 (Love et al., 2014). **Supplemental Table 1** summarizes RNA-seq data of *R. opacus* strains growing with glucose and phenol.

Results and Discussion

Cell growth and biomass composition. The glucose metabolism of *R. opacus* has been studied previously (Hollinshead et al., 2015). Here, we investigated the growth rates of *R. opacus* on acetate, phenol, phenol with glucose, and glucose (**Table 1**). The fastest growth rate (0.24 hr^{-1}) for wild type *R. opacus* was observed when glucose was provided as the sole carbon source. Growth on phenol was, as expected, slower (0.17 hr^{-1}, P=0.013; two-tailed Student's t test), likely owing to a combination of its toxic effects and less efficient phenol degradation processes. When phenol is provided in combination with glucose, these disadvantages are largely mitigated – the rapidity with which glucose can be processed apparently compensated for the kinetic disadvantages of

phenol metabolism (e.g., requirement of gluconeogenesis), which allowed the cultures to grow nearly as quickly as the glucose-only case (P=0.46). Moreover, the co-utilization of both substrates produced more biomass at the end of the cultivations (P<0.01). Interestingly, acetate-fed cultures showed similar growth performance to phenol-fed cultures (P=0.65). The TCA cycle is the entering point of the central flux network for both acetate and phenol, but phenol utilization requires more conversion steps than acetate consumption. Therefore, the similar growth rates indicate that the β–ketoadipate pathway in *R. opacus* is as efficient as the acetate conversion route (acetate→acetyl-CoA).

Table 1. A summary of the growth rate, corresponding doubling times (t_D), and maximum optical density of *R. opacus* on four different carbon substrates. All data presented is the average of biological triplicates ± standard deviation.

Carbon Substrate	Growth Rate (hr^{-1})	t_D (hr)	Maximum OD_{600}
2.0 g/L Acetate	0.16 ± 0.02	4.30	1.08 ± 0.001
0.5 g/L Phenol	0.17 ± 0.02	4.10	0.82 ± 0.02
0.3 g/L Phenol + 1.0 g/L Glucose	0.24 ± 0.01	2.96	1.53 ± 0.01
1.0 g/L Glucose	0.25 ± 0.02	2.81	1.20 ± 0.02

A biomass composition analysis was conducted on cultures grown in phenol, glucose, and a phenol-glucose mixture. The data from the phenol culture was used to modify the biomass equation in the metabolic flux model. **Figure 1A** shows the proteinogenic amino acid composition of *R. opacus* from three tested samples, where glutamate/glutamine and alanine were most abundant. Proteinogenic amino acid synthesis in *R. opacus* undergoes minimal changes when grown on different carbon sources. *R. opacus* is known to store high concentrations of lipids in late growth

phases, and reportedly can accumulate lipids up to 80% of its dry cell weight under low-nitrogen stress (Alvarez et al., 1996; Henson et al., 2018a). In the exponential growth phase when nitrogen is still available, ~0.3 g lipid per gram dry cell weight was accumulated in biomass samples. We further probed the fatty acid profiles of *R. opacus* when fed the three previously-described carbon sources with sufficient nitrogen (**Figure 1B**). In general, the most abundant fatty acid (in the form of acyl groups of lipids) is palmitic acid (16:0), accounting for approximately 30% of total lipid fatty acids. Compared to the model bacterium *E. coli*, *R. opacus* contains higher proportions of long-chain fatty acids (**Supplemental Figure 2**). Specifically, 47% of fatty acids in *R. opacus* were found to contain 17 carbons or more, whereas only 12% of fatty acids in *E. coli* have over 16 carbons (Neidhardt et al., 1990). Additionally, *R. opacus* natively produces C15 fatty acids and branched-chain fatty acids (BCFAs) containing 16, 17, and 18 carbon atoms. Unlike most Gram-positive bacteria and some engineered Gram-negative strains which only produce terminal BCFAs (Bentley et al., 2016; Jiang et al., 2015; Kaneda, 1966), *R. opacus*' BCFAs have internal branches, which are expected to produce biofuels with better cold-flow properties than terminally-branched fuels (Pond and Langworthy, 1987). The carbon source affected the abundance of some fatty acids. Notably, ~20% of the total fatty acids in phenol-fed culture were C18 BCFAs.

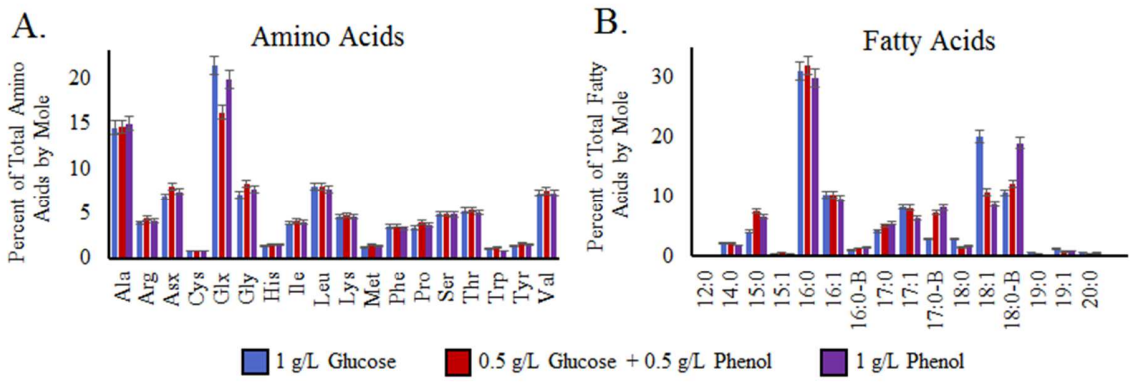


Figure 1. The distribution of amino acids (A) and fatty acids (B) in *R. opacus* with three different carbon sources. Asx = aspartate/asparagine and Glx = glutamate/glutamine; fatty acids are denoted as carbon chain length: number of unsaturated bonds, with a B indicating branching. Error bars are estimates of technical uncertainty.

Phenol degradation via the β-ketoadipate pathway. The enzymatic pathways to degrade aromatic substrates can be grouped into categories based on the position of bond cleavage (Fuchs et al., 2011). The KEGG annotation for *R. opacus* contains every gene for the meta cleavage of catechol, while one of the genes in the ortho pathway, 3-oxoadipate CoA-transferase, is unannotated. Both pathways begin with a two-component flavoprotein monooxygenase, phenol hydroxylase, which oxidizes the phenol ring to form catechol (**Figure 2**). Saa et al. studied this enzyme in *Rhodococcus erythropolis* and found that it can accept electrons from both NADH and NADPH to cleave O_2 molecules. However, phenol hydroxylase's affinity for NADH was measured to be 5-10 times greater than its affinity for NADPH (Saa et al., 2010). This may be an evolutionary adaptation to avoid competition for NADPH between phenol consumption and biomass formation. The ortho and meta pathways diverge with the ortho branch employing catechol-1,2-dioxygenase and the meta branch using catechol-2,3-dioxygenase. The intermediates in the ortho pathway include cis-cis muconic acid and β-ketoadipate, and its end products are acetyl-CoA and succinyl-CoA. In contrast, the meta pathway's intermediates include 2-

hydroxymuconic semialdehyde, and its final products are pyruvate, formate, and acetyl-CoA (Sridevi et al., 2012). Meta-cleavage has been found in *Pseudomonas* species to degrade phenol and phenol derivatives (Kukor and Olsen, 1991). To determine which route is used by *R. opacus*, we used 1-^{13}C phenol labeling (**Figure 2**). The symmetry of catechol results in two different alanine labeling patterns for each branch. In the ortho pathway, 50% of labeled carbon will end up as the 4th position carbon of succinyl-CoA and the remaining 50% will occupy the 1st position of acetyl-CoA. In the TCA cycle, succinyl-CoA is converted to succinate, another symmetric molecule. The reactions of cataplerosis, e.g., the phosphoenolpyruvate (PEP) carboxykinase-catalyzed reaction, release the 4th position carbon as CO_2, and convert the remaining 3-carbon fragment to pyruvate. The expected labeling pattern for pyruvate derived from the ortho branch is therefore 25% first-position labeled and 75% unlabeled. In contrast, the meta-cleavage will create pyruvate molecules with an even split between those labeled in the first and the second positions. Alanine's sole precursor is pyruvate, and thus pyruvate's labeling can be determined by measuring alanine. Our analysis showed that ~23% of alanine was labeled only at the 1st position, and ~77% of alanine was unlabeled. This provides strong evidence that the ortho pathway is the exclusive route for phenol degradation. The labeling result is consistent with our previous gene knockout analyses (Henson et al., 2018a). Specifically, when muconate cycloisomerase, an ortho pathway enzyme, was disrupted, *R. opacus* was no longer able to grow on phenol. Additionally, proteomic data from this species has shown that several enzymes in the ortho pathway were abundant during aromatic metabolism (Xie et al., 2019). The missing ortho pathway reaction (3-oxoadipate → 3-oxoadipyl CoA) could be catalyzed by succinyl-CoA:3-ketoacid-CoA transferase (PD630_RS25340 + PD630_RS25345) (Yoneda et al., 2016). We also found that the presence of low concentration formate (0.5 g/L) was inhibitory to cell growth, implying that the meta pathway

with byproduct formate is unlikely to be beneficial for growth. It is worth noting that the genes coding the ortho pathway enzymes are found in two distinct clusters on the chromosome, while the genes for the meta pathway are scattered across the genome and two endogenous plasmids. All these analyses prove the value of ^{13}C tracing to delineate functional pathways and to fill the annotation gaps.

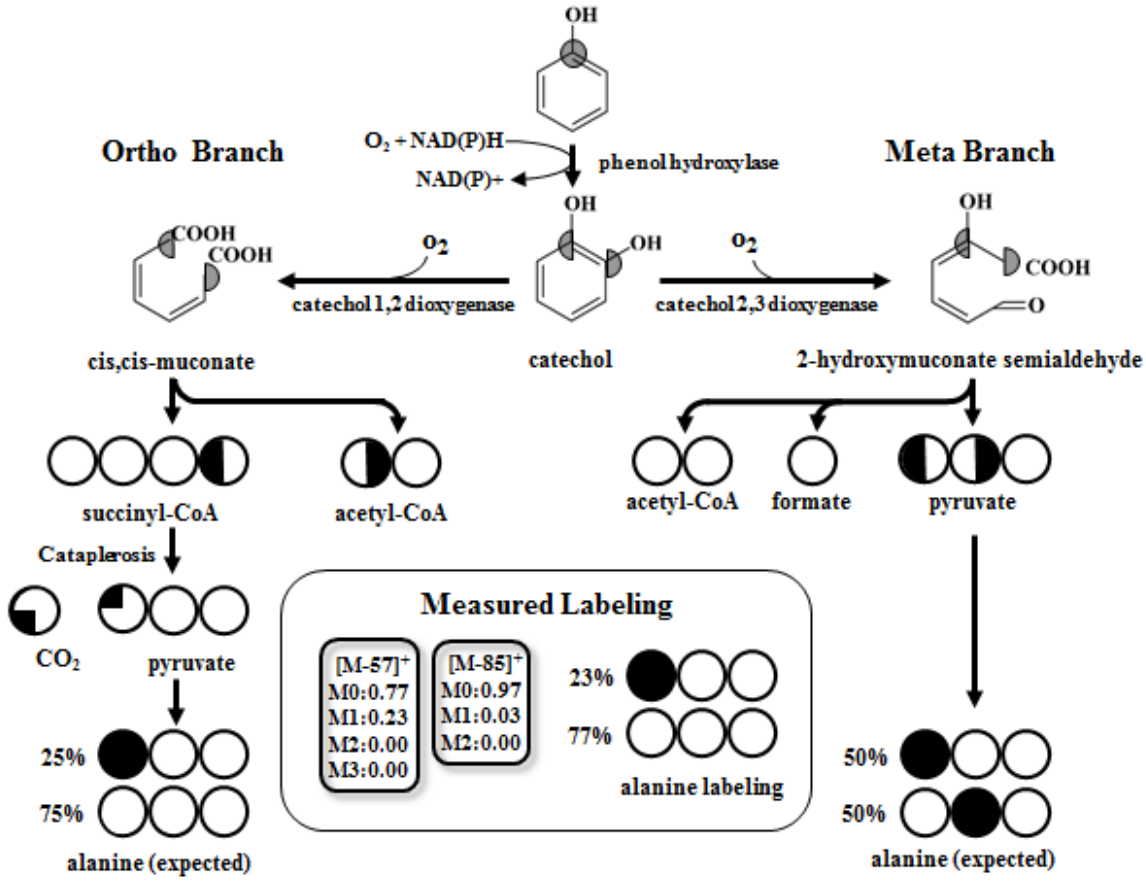


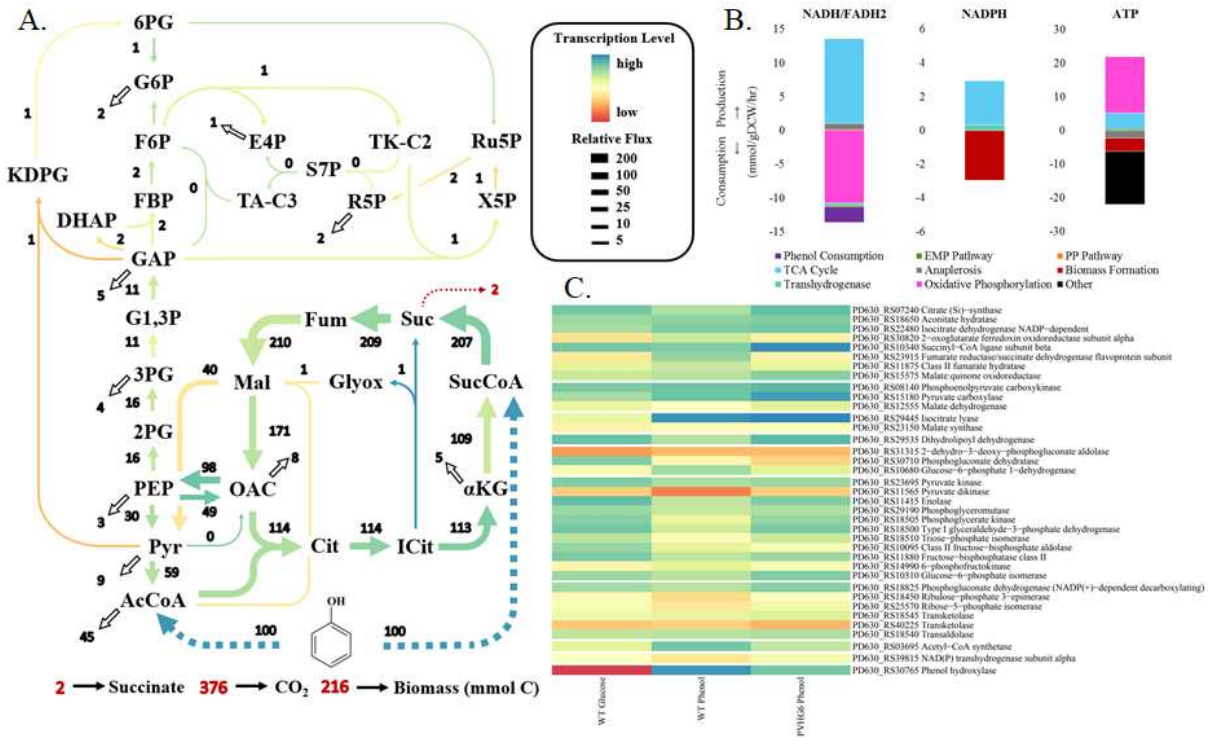
Figure 2. 1-^{13}C phenol catabolism in *R. opacus* is through the ortho-cleavage. GC-MS data of alanine gives labeling data for pyruvate. The $[M-57]^+$ and $[M-85]^+$ fragments were used to determine the labeling of each carbon of pyruvate. The labeling of pyruvate closely matches the expected labeling for phenol catabolism through the ortho branch. A detailed tracing is shown in **Supplemental Figure 1**.

***R. opacus* flux map.** ^{13}C-MFA was performed to comprehensively investigate *R. opacus*' phenol metabolism (**Figure 3A**). The metabolic model had 71 reactions and 33 independent parameters, and was constrained by 174 independent mass isotopomer measurements, resulting in 141 degrees of freedom. Amino acid labeling data with 0.75 mol % measurement error from parallel tracer experiments (1-^{13}C Phenol and 2,6-^{13}C Phenol) was used to generate the flux map. The SSR (sum of square residuals) for the presented fit was 119.4, which is inside the 90% confidence bounds for the SSR [114.6 169.7]. Experimentally, the rate of phenol uptake (i.e., the amount of phenol entering the central metabolism per cell mass within a given time) at exponential phase was measured to be 2.3 mmol/hr/g dry cell weight (**Supplemental Figure 3**). ^{13}C-MFA shows that per 100 mmol of phenol consumed ~2 mmol extracellular overflow metabolites (i.e., succinate), ~370 mmol CO_2, and ~220 mmol C of biomass (~4.7 g) are produced, resulting in a ~50% yield of biomass from phenol (**Figure 3A**).

The TCA cycle of *R. opacus* in phenol culture has very high fluxes to sustain aerobic growth. The citrate entrance node has a relative flux over 100, which is much higher than reported fluxes in glucose-fed *E. coli* (Crown et al., 2015). The direct infusion of succinyl-CoA from phenol degradation drives high fluxes through succinyl-CoA ligase, succinate dehydrogenase, and fumarate hydratase in the TCA cycle and causes succinate overflow (**Supplemental Figure 4**). In the supernatant, α-ketoglutarate could also be detected while acetate secretion was not observed. The overflow of the TCA cycle works to balance excessive succinyl-CoA flux from phenol degradation. In contrast, excessive acetyl-CoA (i.e, not forming citrate in the TCA cycle) enters into biomass pools (lipid and amino acids). Further, the glyoxylate shunt is measurable, offering an alternative route for acetyl-CoA consumption. On the other hand, several anaplerotic reactions are also active to direct flux outside the TCA cycle (note: the confidence intervals of these fluxes

318 are relatively wide due to limited labeling resolution). For instance, the malic enzyme is active to
319 pump the TCA fluxes for pyruvate synthesis. Moreover, both ^{13}C-MFA and gene expression data
320 identified the GTP-dependent PEP carboxykinase (PD630_RS08140) as the driving force for
321 gluconeogenesis (OAC→PEP) and flux towards the upper pathways (the EDP and the pentose
322 phosphate pathways). Interestingly, the model shows that the EDP, operating in reverse, can serve
323 as a secondary route to the upper pathways. Fluxes in the upper pathways of phenol-fed cultures
324 are small and primarily used for biomass synthesis.

325 To offer a deeper understanding of *R. opacus* flux network optimality, the ^{13}C-MFA model
326 was compared to flux balance analysis using five different objective functions (**Supplemental**
327 **Figure 5**) (Schuetz et al., 2007). Based on the SSRs between the MFA fluxes and FBA fluxes, the
328 metabolism of *R. opacus* is most closely optimized to maximize the production of energy
329 molecules rather than the production of biomass or fatty acids or the minimization of enzyme use
330 (**Supplemental Figure 6**). The high TCA cycle fluxes elucidated by ^{13}C-MFA facilitate NADH,
331 and subsequently ATP production. The elevated priority of energy molecules is the tradeoff against
332 biomass growth with stress inducing aromatic substrates. Moreover, *R. opacus* flux topology does
333 not gear towards minimal usage of enzymes, allowing for metabolic flexibility to utilize
334 unconventional substrates.



335

336 **Figure 3. Flux and transcriptional analyses of *R. opacus* strains.** (A) *R. opacus* PD630's flux
337 map when phenol is the sole carbon source. The flux values are relative flux distributions based
338 on 100 mmol of phenol consumed by the cell to generate 100 mmol of influx toward both acetyl-
339 CoA and succinyl-CoA. Arrow color corresponds to the relative gene expression levels of the
340 enzyme that catalyzes the reaction (yellow/orange=low and green/blue=high). Arrow size
341 corresponds to relative flux. White arrows are fluxes to biomass formation. The dashed red arrow
342 is flux out of the cell. (B) The major sources and sinks of NADH, NADPH, and ATP in *R.*
343 *opacus*'s phenol metabolism (note: pyruvate dehydrogenase reaction was lumped with the EMP
344 pathway). (C) The transcript levels from wild type and an adapted strain (PVHG6) *R. opacus*
345 grown on glucose or phenol were compared for the following metabolisms: TCA cycle, anaplerotic
346 pathways, the glyoxylate shunt, the EDP, the EMP pathway, the pentose phosphate pathway,
347 transhydrogenation, and phenol consumption (Henson et al., 2018a).

348

349

350 **Cofactor balance and energy metabolism.** Phenol metabolism carries very low fluxes through
351 the oxidative pentose phosphate pathway, which is the major NADPH source for common glucose-
352 utilizing microbes (such as *E. coli*). Instead, significant TCA fluxes are responsible for generating
353 large amounts of NADH, NADPH, and ATP (**Figure 3B**). Unlike NADPH, which is primarily

consumed in the formation of biomass, NADH is used for the oxidation of phenol to catechol and the production of ATP. *R. opacus* can use NAD-dependent malic enzyme (PD630_RS12555, EC 1.1.1.38), but lacks NADP-dependent malic enzyme (EC 1.1.1.40). Consequently, most NADPH generation occurs in the TCA cycle by NADPH-dependent isocitrate dehydrogenase (PD630_RS13780), which explains the necessity to maintain high flux through the TCA cycle. A search of *R. opacus*'s genome identified NAD(P) transhydrogenase subunit alpha (PD630_RS39815). The transhydrogenase might further fine-tune the cofactor balance. On the other hand, ATP can be generated directly by the TCA cycle (as GTP) and indirectly through oxidative phosphorylation. ATP is needed mainly for biomass formation (in particular for lipid synthesis) and other cellular processes (e.g., cell maintenance, substrate active transport, and gluconeogenesis). Additionally, we calculated the intracellular energy charge based on LC-MS measurements to determine the relative concentrations of ATP, ADP, and AMP (*Energy Charge* $= \frac{[ATP]+½\,[ADP]}{[ATP]+[ADP]+[AMP]}$) (**Supplemental Figure 7**). Phenol-fed cultures have an energy charge of ~0.9, indicating that the concentration of ATP is much greater than that of ADP and AMP combined. Glucose-fed cultures have a comparable energy charge (~0.85). An explanation for this similarity is that, during aerobic growth, the NADH synthesized from either substrate can be effectively used via oxidative phosphorylation to generate ATP (**Figure 3B**). These results along with FBA results demonstrate that *R. opacus* has sufficient energy carrying molecules to support the production of high energy biofuels from different feedstocks.

RNA-seq analysis of flux network. RNA-seq was used to compare transcriptional regulation between conditions (**Figure 3, Supplemental Table 1**). It is important to note that while

transcriptomics can illuminate how genes are regulated, it does not strictly correlate to enzyme activities or flux values due to variations in translational rate and allosteric regulation. Using data generated in the previous report, we compared the transcription levels of central pathway enzymes in glucose- and phenol-fed cultures of wild type *R. opacus* and *R. opacus* PVHG6, a mutant strain which was adaptively evolved for improved growth on a mixture of aromatic compounds (**Figure 3C**) (Henson et al., 2018a). When consuming glucose, *R. opacus* expressed phenol hydroxylase at low levels, while its key enzymes for glycolysis are mostly highly expressed. During phenol utilization, *R. opacus*' isocitrate lyase (PD630_RS29445) was highly expressed, but the glyoxylate shunt flux was low.

Gluconeogenesis, which is required when phenol is the sole carbon source, is typically thought to occur through the EMP pathway operating in reverse. However, the ^{13}C-MFA model suggests that the EDP enzymes (in the reverse direction) also play a role in *R. opacus*. Transcriptomic data was used to investigate this finding. Despite its low transcription levels in glucose cultures, the EDP was found to be the main route for glucose consumption in *R. opacus* (Hollinshead et al., 2015). The EDP has comparable transcription levels in phenol-fed cultures and glucose cultures, and this suggests that it is active in the reverse direction when phenol is the carbon source. When running in reverse, the EMP pathway uses most of the same enzymes as it does when it runs forward, with the exception of conversion between FBP to F6P and pyruvate to phosphoenolpyruvate (PEP). In the glycolytic direction, F6P is converted to FBP by 6-phosphofructokinase (PD630_RS14990); in the gluconeogenic direction, FBP → F6P occurs via fructose-1,6-bisphosphatase (PD630_RS11880). Interestingly, the number of transcripts of fructose-1,6-bisphosphatase is lower in phenol cultures than in glucose cultures (**Supplemental Table 1**), implying fructose-1,6-bisphosphatase activity is reduced during phenol metabolism. Additionally, the gene encoding

pyruvate dikinase (PD630_RS11565, pyruvate → PEP) has very low expression levels during phenol metabolism. Therefore, the anaplerotic reaction via phosphoenolpyruvate carboxykinase (PD630_RS08140, oxaloacetate → PEP) is indispensable for supporting gluconeogenic fluxes. Intriguingly, the aromatic mixture adapted strain, PVHG6, was found to express pyruvate dikinase and succinyl-CoA ligase at higher rates than the wild type in phenol culture, which may facilitate aromatic utilization and tolerance.

Pulse-Trace labeling and metabolite pool measurement. The gluconeogenesis pathways were further investigated with a ^{13}C-pulse experiment (**Figure 4A, 4B, Supplemental Figure 8)**. Conceptually, metabolites with the highest flux are expected to incorporate ^{13}C carbon most rapidly. The labeling order of glycolysis metabolites observed in phenol-fed cultures is the reverse of that in glucose-fed metabolism: phosphoenolpyruvate (PEP) was labeled first followed closely by 3-phosphoglycerate (3PG). Fructose-bis-phosphate (FBP) was found to be labeled significantly faster than fructose-6-phosphate (F6P) (P=0.03). This may be due to a combination of the thermodynamic barrier between FBP and F6P (i.e., reaction Gibbs free energy favors flux from F6P to FBP), the small pool size of FBP (**Figure 4C)**, and the low transcription levels of fructose-1,6-bisphosphatase (**Figure 3C, Supplemental Table 1**).

The intracellular metabolite concentrations, or pool sizes, were measured for *R. opacus* cultures grown on glucose and phenol at stationary and exponential time points **(Figure 4C**). Cells growing exponentially on phenol have very high quantities of TCA cycle intermediates. However, their intracellular concentrations of fructose-bis-phosphate (FBP), dihydroxyacetone phosphate (DHAP), and glyceraldehyde-3-phosphate (GAP) are at minimum 10-fold lower than those of *E. coli*. This offers thermodynamic evidence for high TCA cycle fluxes paired with low reverse EDP

and gluconeogenic fluxes. Compared to glucose-fed *E.coli*, phenol-fed *R. opacus* has relatively low citrate concentrations. This may be the result of a highly active isocitrate dehydrogenase, an enzyme needed for both αKG and NADPH production. Also, *R. opacus* has relatively high concentrations of α-ketoglutarate (αKG), especially during exponential growth. An explanation is that succinyl-CoA/succinate accumulation from phenol degradation inhibits α-ketoglutarate dehydrogenase (αKG → succinyl-CoA), leading to excess amounts of αKG (note: both succinate and αKG were detected in the supernatant). Therefore, succinyl-CoA is likely a bottleneck node causing metabolite overflow during phenol metabolism.

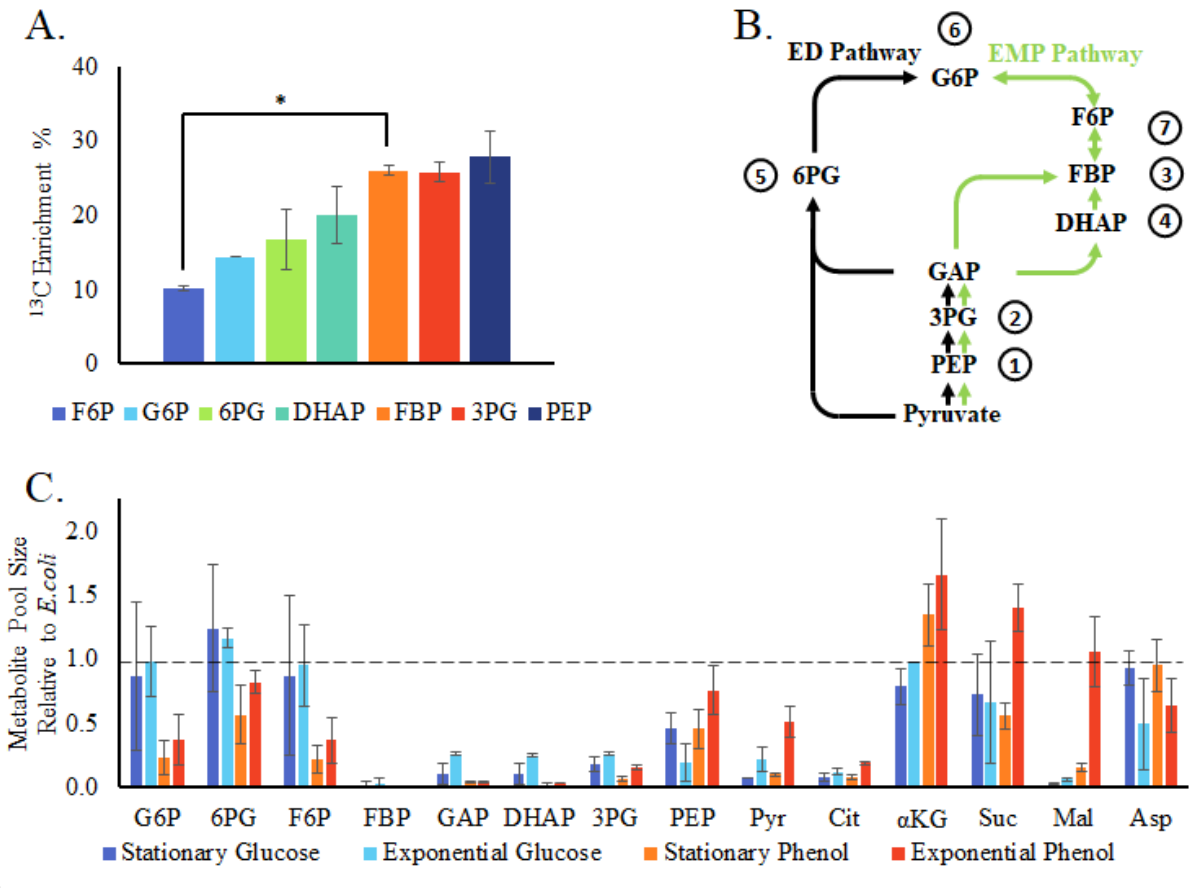


Figure 4. Dynamic labeling and metabolite pool size measurement. (A) Pulse-trace labeling data from cultures that were quenched within one minute after ^{13}C phenol was added. Asterisk (*) indicates that the difference between F6P and FBP values is statistically significant ($P < 0.05$, one mean, two-tailed Student's t-test). (B) A simplified metabolic map of the ED and EMP pathways annotated with 1 being the fastest labeled metabolite and 7 being the slowest. (C) Intracellular metabolite concentrations relative to those in glucose-fed *E. coli* (used as internal standards).

Phenol co-utilization with other substrates. Industrial applications of *R. opacus* will require growth on lignin hydrolysates, which contain a heterogeneous mixture of compounds (Beckham et al., 2016). Notably, *R. opacus* can consume aromatic compounds simultaneously with glucose without exhibiting catabolite repression (Hollinshead et al., 2015). However, it was found that *R. opacus* consumes some aromatic compounds sequentially, rather than concurrently (Henson et al., 2018a). Here we further investigated co-metabolism of phenol with competing substrates (glucose, acetate, succinate, and benzoate) (**Figure 5A, 5B, 5C**). Our findings for phenol-glucose cultures are consistent with the previous report, finding no interaction between the two substrates. A separate study observed that phenol degradation enzyme transcripts in *Rhodococcus erythropolis* CCM2595 were downregulated in the presence of succinate (Szőköl et al., 2014). Yet, when phenol and succinate were co-fed in our experiment, phenol was dominantly used (>90% of carbon originated from phenol). Since phenol metabolism produces excessive succinate, *R. opacus* may repress the utilization of external succinate in order to preferentially consume the toxic compound. Similarly, *R. opacus* can co-utilize phenol and acetate, and phenol is preferred (~70% of proteinogenic carbon originated from phenol). On the other hand, benzoate was found to be utilized faster than phenol (**Figure 5B**, **5C**).

Three experiments were performed to elucidate the interaction of phenol and benzoate. In the first experiment, amino acid labeling was used to quantify the relative uptake rate between benzoate and phenol (**Figure 5B**). We found that the ratio of benzoate/phenol uptake rate

459 (calculated via ^{13}C labeling data) during early growth phase showed a linear relation to the
460 concentration ratio of the two substrates in the media (**Supplemental Figure 9**). When cells grew
461 in media containing equal moles of fully labeled phenol and unlabeled benzoate, the proteinogenic
462 amino acids contained ~90% ^{12}C carbon from benzoate. This result indicates that the wild type
463 strain consumes both substrates, but its benzoate utilization is much faster than its phenol
464 utilization. The second experiment examined how the concentrations of benzoate and phenol in
465 the media decreased over time (**Figure 5C**). We validated that benzoate serves as the main carbon
466 source when both benzoate and phenol are present: phenol was consumed slowly in the presence
467 of benzoate, and its consumption increased once benzoate was nearly exhausted. Lastly, a benzoate
468 pulse was given into fully-labeled phenol-fed, exponentially growing cultures to investigate the
469 rate at which unlabeled benzoate enters into the central metabolism of *R. opacus* (**Figure 5D**). To
470 consider contribution of CO_2 fixation to the unlabeled carbon percentage (Hollinshead et al.,
471 2015), we included no-pulse controls with which the benzoate-pulsed cultures were compared
472 (using one mean, two-tailed Student's t-test). When the phenol growing *R. opacus* was quenched
473 within 20 seconds of a benzoate pulse, little amount of carbon from benzoate had entered the
474 central metabolism (e.g., *P* values = 0.11 for αKG). In contrast, at 10 minutes, a significant amount
475 of carbon had been incorporated into central metabolites from the added benzoate ($P < 0.05$ for
476 αKG, PEP, and G6P), which indicate phenol catabolism in *R. opacus* could be quickly shifted to
477 benzoate utilization without long lag phase.

478 The downstream products of phenol are acetyl-CoA and succinyl-CoA. To investigate how
479 these two substrates interact, we grew *R. opacus* in media containing fully labeled acetate and
480 unlabeled succinate (1 g/L each) (**Figure 5E**). We hypothesized that acetate would be preferred
481 over succinate since succinate is not fully consumed in phenol metabolism. GC-MS measurements

showed that *R. opacus* prefers acetate, as the carbon in its amino acids was mostly derived from acetate (~70%). This observation further confirmed that the metabolite nodes succinate/succinyl-CoA might be the rate limiting steps during phenol metabolism, leading to flux congestion.

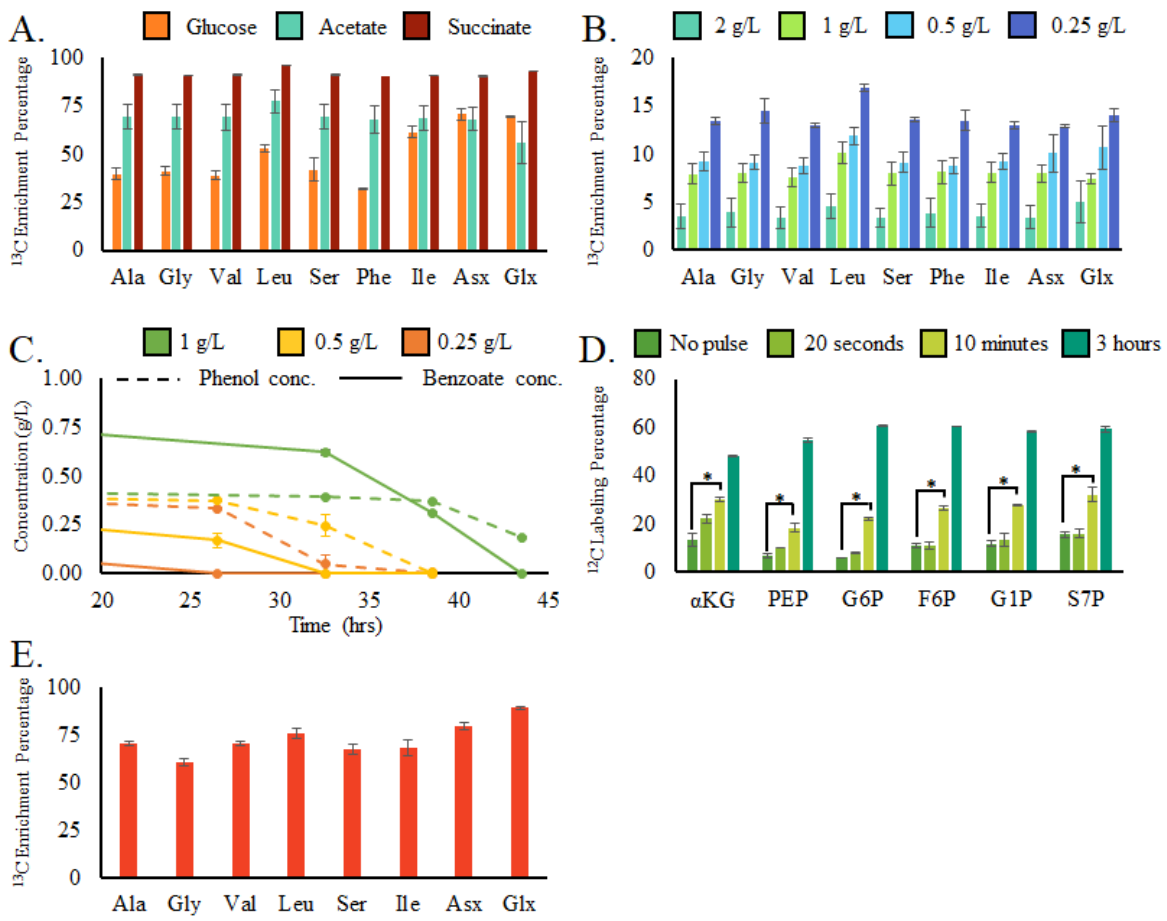


Figure 5. Mixotrophic phenol metabolism. (A) The contribution of phenol-derived carbon to amino acids in the presence of 0.5 g/L U-^{13}C phenol and 1 g/L of unlabeled glucose, acetate, or succinate. Samples were taken at early exponential growth (OD_{600}=0.4~0.5) when both substrates remained in the media. (B) The contribution of phenol to the production of amino acids in the presence of 0.5 g/L U-^{13}C phenol and decreasing concentrations of unlabeled benzoate. Samples were taken at early exponential growth phase (OD_{600}=0.25~0.3) when both substrates remained in the media. The bars correspond to the relative uptake rate of phenol compared to that of benzoate.

(C) The supernatant concentrations of phenol and benzoate over time when the starting concentration of U-^{13}C phenol was 0.5 g/L and the starting concentration of benzoate was varied. (D) The contribution of benzoate-derived carbon to central metabolites at 20 seconds, 10 minutes, and 3 hours after a benzoate pulse into phenol-fed cultures. Asterisk (*) indicates that the difference between labeling is statistically significant ($P < 0.05$, one mean, two-tailed Student's t-test) (E) The contribution of acetate to amino acids in the presence of 1 g/L U-^{13}C acetate and 1 g/L unlabeled succinate. Figures 5A, 5B, 5C, and 5D represent data from biological duplicates. Figure 5E represents data from technical triplicates.

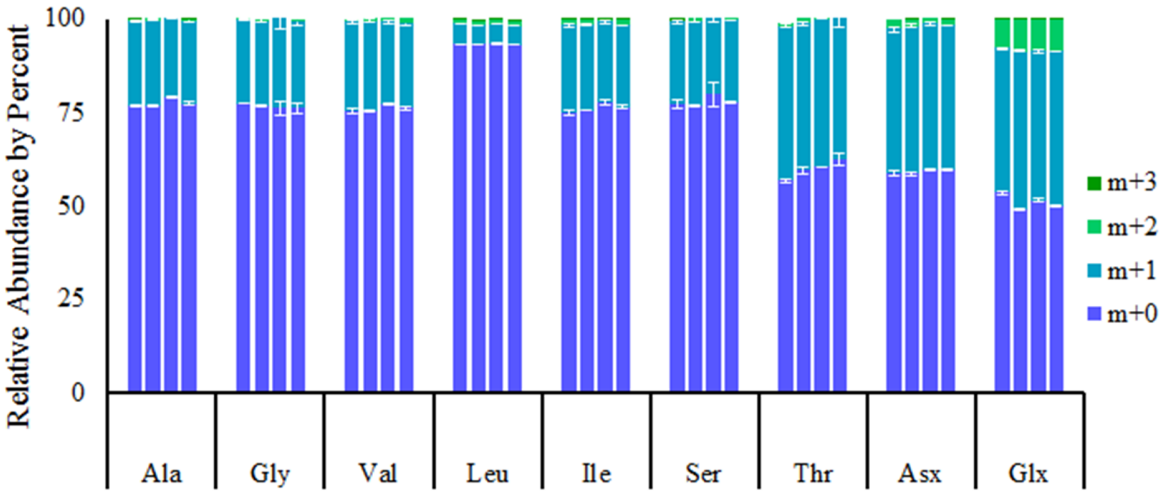


Figure 6. The amino acid labeling pattern of wild type *R. opacus* and three mutant strains when grown on 1-^{13}C phenol (From left to right: wild type, phenol-adapted strain P1, benzoate-adapted strain B2, and aromatic-mixture-adapted strain PVHG6; m+n = unlabeled mass m plus n labeled carbons).

Mutant flux network after adaptive evolution. In a previous report, *R. opacus* was adaptively evolved to tolerate high concentrations of various aromatic compounds, including phenol, benzoate, vanillate, 4-hydroxybenzoate, and guaiacol (Henson et al., 2018a). Three mutant strains (P1, B2, and PVHG6) were obtained via adaptive evolution on increasing concentrations of phenol (P), benzoate (B), or a mixture of phenol, vanillate, 4-hydroxybenzoate, and guaiacol (PVHG), respectively. These mutant strains were found to have distinct mutations and different growth

profiles when provided with a mixture of phenol, vanillate, 4-hydroxybenzoate, guaiacol, and benzoate as a carbon source. To determine whether these mutations altered flux through central pathways, these evolved strains were grown on 0.5 g/L of 1-^{13}C phenol, and the resulting amino acid labeling patterns were compared to that of the wild type strain (**Figure 6**). Adapted stains, P1 and PVHG6 showed similar growth rates on phenol as that of the wild type (P=0.70 and P=0.11, respectively) (**Supplemental Table 2**). The benzoate-adapted strain showed reduced growth on phenol compared to the wild type strain, only reaching an OD_{600} value of 0.2 after 50 hours (P=3.7×10^{-5}), and a maximum growth rate of 0.10 hr^{-1} (P=0.01). This supports previous findings that while adaptive evolution can help improve specific microbial traits, its untargeted nature can result in unexpected negative consequences or tradeoffs in other traits (Yi and Dean, 2016). Despite the growth defect in the benzoate-adapted strain, all three adapted strains produced amino acid labeling similar to the wild type, where the variances in labeling data were generally less than 2% (suggesting the invariability of relative flux distributions for biomass synthesis). This provides strong evidence that substrate adaptive evolution did not rewire the central flux network, which is a testament to the rigidity of fluxes during adaptive evolutions (Long et al., 2017).

Conclusion and Perspectives

Recently, non-model organisms are gaining attention for addressing challenges in bio-manufacturing from renewable feedstock (Czajka et al., 2017). Researchers leverage innate metabolic strengths of non-model species, including their tolerance and utilization abilities for toxic lignin-derived aromatics. However, rational strain engineering, which is routinely done for model organisms, cannot be done until the complex metabolic networks of these microbes are better understood. In this work, we analyzed the aromatic metabolism in *R. opacus*, one of the

539 promising non-model organisms for lignin bioconversion, using a concerted systems biology
540 approach. Importantly, this is the first reported ^{13}C-metabolic flux analysis completed in any
541 organism using phenol as the sole carbon source. This study also shows that ^{13}C tracing is a
542 powerful tool to precisely elucidate these functional pathways with annotation gaps and to
543 determine the time-scale of metabolic responses. Moreover, we established key links between the
544 transcriptome, metabolome, and fluxome of this organism, and uncovered or confirmed some of
545 its mechanisms for aromatic utilization.

546 By detailing *R. opacus*' metabolic network, this work offers guidelines for developing this
547 species as a bio-manufacturing host. First, phenol-fed *R. opacus* shows strong fluxes through its
548 TCA cycle. Thus, this host is ideal for producing chemicals derived from TCA cycle intermediates
549 (e.g., acetyl-CoA and α-ketoglutarate). In contrast, its fluxes from phenol through gluconeogenesis
550 are small, making it undesirable to redirect sugar phosphate metabolites as precursors for bio-
551 manufacturing. Second, we investigated the carbon source utilization hierarchy of *R. opacus*.
552 Interestingly, this strain utilizes glucose and phenol simultaneously without catabolite repression,
553 preferentially consumes a toxic aromatic compound (phenol) over succinate and acetate, and uses
554 benzoate faster than phenol. Since lignocellulosic degradation products are mixtures of
555 compounds, *R. opacus*'s co-utilization patterns will help to develop predictive metabolic models
556 as well as guide future metabolic engineering efforts toward more-efficient conversion of
557 lignocellulose to biofuels and biochemicals. Combining our findings from this work with our
558 previous reports (Hollinshead et al., 2015; Henson et al., 2018a; Henson et al., 2018b, Yoneda et
559 al., 2016), we are developing a genome-scale metabolic model that enables prediction of metabolic
560 fluxes for diverse carbon sources. Quantitative flux profiles obtained through ^{13}C-MFA provide
561 an accurate description of internal metabolism that is invaluable when deciding future metabolic

562 engineering approaches. For example, these flux profiles can be used in conjunction with methods
563 like ROOM or MOMA to predict the effect of knockouts on metabolism and to offer computer
564 aided strain design (Martín et al., 2015; Ando and Martín, 2018). Third, ^{13}C analysis of the
565 adaptively-evolved strains shows that their genetic differences do not lead to different flux
566 networks. Recently, reliable genetic parts and engineering tools have been developed for *R.*
567 *opacus*, allowing for gene expression control and genome modification in this strain (Delorenzo
568 et al., 2017; Delorenzo et al., 2018). Future work will investigate how model-guided perturbations
569 of its metabolism (e.g., knockouts and overexpression), as opposed to adaptive evolution, affects
570 *R. opacus*'s flux network. With the ^{13}C-MFA model and detailed information regarding *R. opacus*'
571 metabolic network provided from this work, future metabolic engineering efforts will be
572 prediction-driven, speeding up its development as a bio-manufacturing host.
573
574

575 Data Availability

576 1. Supplemental tables and figures

577 2. Flux model construction and isotopomer data

578

579 Abbreviations

580 3PG = 3-phosphoglycerate
581 6PG = 6-phosphogluconic acid
582 AcCoA = acetyl-CoA

583 αKG = α-ketoglutarate

584 Asx = aspartate and asparagine

585 Cit = citrate

586 DHAP = dihydroxyacetone phosphate

587 E4P = erythrose-4-phosphate

588 Fum = fumarate

589 F6P = fructose-6-phosphate

590 FBP = fructose-1,6-phosphate

591 G1P = glucose-1-phosphate

592 G1,3P = glycerate-1,3-phosphate

593 G2P = glycerate-2-phosphate

594 G6P = glucose-6-phosphate

595 GAP = glyceraldehyde-3-phosphate

596 Glx = glutamate and glutamine

597 ICit= isocitrate

598 KDPG = 2-dehydro-3-deoxy-phosphogluconate

599 Mal = malate

600 OAC = oxaloacetate

601 PEP = phosphoenolpyruvate

602 Pyr = pyruvate

603 R5P = ribose-5-phosphate

604 Ru5P = ribulose-5-phosphate

605 Suc = succinate

606 SucCoA = succinyl-CoA

607 TA-C3 = 3-carbon fragment transferred by transaldolase

608 TK-C2 = 2-carbon fragment transferred by transketolase

609 X5P=xylose-5-phosphate

610

Acknowledgement

This project is supported by the US DOE (DE-SC0018324). For LC-MS services, The Donald Danforth Plant Science Center QTRAP LC-MS/MS acquisition was supported by National Science Foundation (NSF) Grant DBI-1427621. We would like to thank Dr. Mary Abernathy for her help with sample preparation and data analysis of LC-MS experiments.

Table and Figure Captions

Table 1. A summary of the growth rate, corresponding doubling times (t_D), and maximum optical density of *R. opacus* on four different carbon substrates. All data presented is the average of biological triplicates ± standard deviation.

Figure 1. The distribution of amino acids (A) and fatty acids (B) in *R. opacus* with three different carbon sources. Asx = aspartate/asparagine and Glx = glutamate/glutamine; fatty acids are denoted as carbon chain length: number of unsaturated bonds, with a B indicating branching. Error bars are estimates of technical uncertainty.

Figure 2. 1-^{13}C phenol catabolism in *R. opacus* is through the ortho-cleavage. GC-MS data of alanine gives labeling data for pyruvate. The $[M-57]^+$ and $[M-85]^+$ fragments were used to determine the labeling of each carbon of pyruvate. The labeling of pyruvate closely matches the expected labeling for phenol catabolism through the ortho branch. A detailed tracing is shown in **Supplemental Figure 1**.

Figure 3. Flux and transcriptional analyses of *R. opacus* strains. (A) *R. opacus* PD630's flux map when phenol is the sole carbon source. The flux values are relative flux distributions based on 100 mmol of phenol consumed by the cell to generate 100 mmol of influx toward both acetyl-CoA and succinyl-CoA. Arrow color corresponds to the relative gene expression levels of the enzyme that catalyzes the reaction (yellow/orange=low and green/blue=high). Arrow size corresponds to relative flux. White arrows are fluxes to biomass formation. The dashed red arrow is flux out of the cell. (B) The major sources and sinks of NADH, NADPH, and ATP in *R. opacus*'s phenol metabolism (note: pyruvate dehydrogenase reaction was lumped with the EMP pathway). (C) The transcript levels from wild type and an adapted strain (PVHG6) *R. opacus* grown on glucose or phenol were compared for the following metabolisms: TCA cycle, anaplerotic pathways, the glyoxylate shunt, the EDP, the EMP pathway, the pentose phosphate pathway, transhydrogenation, and phenol consumption (Henson et al., 2018a).

Figure 4. Dynamic labeling and metabolite pool size measurement. (A) Pulse-trace labeling data from cultures that were quenched within one minute after ^{13}C phenol was added. Asterisk (*)

indicates that the difference between F6P and FBP values is statistically significant ($P < 0.05$, one mean, two-tailed Student's t-test). (B) A simplified metabolic map of the ED and EMP pathways annotated with 1 being the fastest labeled metabolite and 7 being the slowest. (C) Intracellular metabolite concentrations relative to those in glucose-fed *E. coli* (used as internal standards).

Figure 5. Mixotrophic phenol metabolism. (A) The contribution of phenol-derived carbon to amino acids in the presence of 0.5 g/L U-^{13}C phenol and 1 g/L of unlabeled glucose, acetate, or succinate. Samples were taken at early exponential growth (OD_{600}=0.4~0.5) when both substrates remained in the media. (B) The contribution of phenol to the production of amino acids in the presence of 0.5 g/L U-^{13}C phenol and decreasing concentrations of unlabeled benzoate. Samples were taken at early exponential growth phase (OD_{600}=0.25~0.3) when both substrates remained in the media. The bars correspond to the relative uptake rate of phenol compared to that of benzoate. (C) The supernatant concentrations of phenol and benzoate over time when the starting concentration of U-^{13}C phenol was 0.5 g/L and the starting concentration of benzoate was varied. (D) The contribution of benzoate-derived carbon to central metabolites at 20 seconds, 10 minutes, and 3 hours after a benzoate pulse into phenol-fed cultures. Asterisk (*) indicates that the difference between labeling is statistically significant ($P < 0.05$, one mean, two-tailed Student's t-test) (E) The contribution of acetate to amino acids in the presence of 1 g/L U-^{13}C acetate and 1 g/L unlabeled succinate. Figures 5A, 5B, 5C, and 5D represent data from biological duplicates. Figure 5E represents data from technical triplicates.

Figure 6. The amino acid labeling pattern of wild type *R. opacus* and three mutant strains when grown on 1-^{13}C phenol (From left to right: wild type, phenol-adapted strain P1, benzoate-adapted strain B2, and aromatic-mixture-adapted strain PVHG6; m+n = unlabeled mass m plus n labeled carbons).

References

Abdelaziz, O. Y., Brink, D. P., Prothmann, J., Ravi, K., Sun, M., García-Hidalgo, J., Sandahl, M., Hulteberg, C. P., Turner, C., Lidén, G., Gorwa-Grauslund, M. F., 2016. Biological valorization of low molecular weight lignin. Biotechnology Advances. 34, 1318-1346.

Abernathy, M.H., Yu, J., Ma, F., Liberton, M., Ungerer, J., Hollinshead, W.D., Gopalakrishnan, S., He, L., Maranas, C.D., Pakrasi, H.B., Allen, D.K., Tang, Y.J., 2017. Deciphering cyanobacterial phenotypes for fast photoautotrophic growth via isotopically nonstationary metabolic flux analysis. Biotechnology for Biofuel. 10, 273

Alvarez, H., Mayer, F., Fabritius, D., Steinbüchel, A., 1996. Formation of intracytoplasmic lipid inclusions by *Rhodococcus opacus* strain PD630. Archives of Microbiology. 165**,** 377-386.

Ando, D., García Martín, H., 2019. Genome-scale ^{13}C fluxomics modeling for metabolic engineering of *Saccharomyces cerevisiae*. Microbial Metabolomics. Methods in Molecular Biology. 1859, 317-345.

Beckham, G. T., Johnson, C. W., Karp, E. M., Salvachua, D., Vardon, D. R., 2016. Opportunities and challenges in biological lignin valorization. Curr Opin Biotechnol. 42**,** 40-53.

Bennett, B. D., Yuan, J., Kimball, E. H., Rabinowitz, J. D., 2008. Absolute quantitation of intracellular metabolite concentrations by an isotope ratio-based approach. Nature protocols. 3**,** 1299-1311.

Bentley, G. J., Jiang, W., Guamán, L. P., Xiao, Y., Zhang, F., 2016. Engineering *Escherichia coli* to produce branched-chain fatty acids in high percentages. Metabolic Engineering. 38**,** 148-158.

Boerjan, W., Ralph, J., Baucher, M., 2003. Lignin Biosynthesis. Annual Review of Plant Biology. 54**,** 519-546.

Bolger, A. M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 30**,** 2114-2120.

Chen, Y., Ding, Y., Yang, L., Yu, J., Liu, G., Wang, X., Zhang, S., Yu, D., Song, L., Zhang, H., Zhang, C., Huo, L., Huo, C., Wang, Y., Du, Y., Zhang, H., Zhang, P., Na, H., Xu, S., Zhu, Y., Xie, Z., He, T., Zhang, Y., Wang, G., Fan, Z., Yang, F., Liu, H., Wang, X., Zhang, X., Zhang, M. Q., Li, Y., Steinbüchel, A., Fujimoto, T., Cichello, S., Yu, J., Liu, P., 2014. Integrated omics study delineates the dynamics of lipid droplets in *Rhodococcus opacus* PD630. Nucleic acids research. 42**,** 1052-1064.

Chubukov, V., Uhr, M., Le Chat, L., Kleijn, R. J., Jules, M., Link, H., Aymerich, S., Stelling, J., Sauer, U., 2013. Transcriptional regulation is insufficient to explain substrate-induced flux changes in *Bacillus subtilis*. Molecular systems biology. 9**,** 709.

Crown, S. B., Long, C. P., Antoniewicz, M. R., 2015. Integrated ^{13}C-metabolic flux analysis of 14 parallel labeling experiments in *Escherichia coli*. Metabolic Engineering. 28**,** 151-158.

Czajka, J., Wang, Q., Wang, Y., Tang, Y. J., 2017 Synthetic biology for manufacturing chemicals: constraints drive the use of non-conventional microbial platforms. Applied microbiology and biotechnology. 101, 7427-7434.

DeLorenzo, D. M., Henson, W. R., Moon, T. S., 2017. Development of Chemical and Metabolite Sensors for *Rhodococcus opacus* PD630. ACS Synthetic Biology. 6**,** 1973-1978.

DeLorenzo, D. M., Rottinghaus, A. G., Henson, W. R., Moon, T. S., 2018. Molecular Toolkit for Gene Expression Control and Genome Modification in *Rhodococcus opacus* PD630. ACS Synthetic Biology. 7, 727-738.

Fu, Y., Yoon, J. M., Jarboe, L., Shanks, J. V., 2015. Metabolic flux analysis of *Escherichia coli* MG1655 under octanoic acid (C8) stress. Applied Microbiology and Biotechnology. 99**,** 4397-4408.

Fuchs, G., Boll, M., Heider, J., 2011. Microbial degradation of aromatic compounds — from one strategy to four. Nature Reviews Microbiology. 9**,** 803.

Gani, A., Naruse, I., 2007. Effect of cellulose and lignin content on pyrolysis and combustion characteristics for several types of biomass. Renewable Energy. 32**,** 649-661.

García Martín, H., Kumar, V. S., Weaver, D., Ghosh, A., Chubukov, V., Mukhopadhyay, A., Arkin, A., Keasling, J. D., 2015. A method to constrain genome-scale models with ^{13}C labeling data. PLOS Computational Biology. 11, e1004363.

Henson, W. R., Campbell, T., DeLorenzo, D. M., Gao, Y., Berla, B., Kim, S. J., Foston, M., Moon, T. S., Dantas, G., 2018a. Multi-omic elucidation of aromatic catabolism in adaptively evolved *Rhodococcus opacus*. Metabolic Engineering. 49**,** 69-83.

Henson, W. R., Hsu, F.-F., Dantas, G., Moon, T. S., Foston, M., 2018b. Lipid metabolism of phenol-tolerant *Rhodococcus opacus* strains for lignin bioconversion. Biotechnology for biofuels. 11**,** 339.

Holder, J. W., Ulrich, J. C., DeBono, A. C., Godfrey, P. A., Desjardins, C. A., Zucker, J., Zeng, Q., Leach, A. L. B., Ghiviriga, I., Dancel, C., Abeel, T., Gevers, D., Kodira, C. D., Desany, B., Affourtit, J. P., Birren, B. W., Sinskey, A. J., 2011. Comparative and Functional Genomics of *Rhodococcus opacus* PD630 for Biofuels Development. PLOS Genetics. 7**,** e1002219.

Hollinshead, W. D., Henson, W. R., Abernathy, M., Moon, T. S., Tang, Y. J., 2015. Rapid metabolic analysis of *Rhodococcus opacus* PD630 via parallel 13C-metabolite fingerprinting. Biotechnology and Bioengineering. 113**,** 91-100.

Hollinshead, W. D., Rodriguez, S., Martin, H. G., Wang, G., Baidoo, E. E. K., Sale, K. L., Keasling, J. D., Mukhopadhyay, A., Tang, Y. J., 2016. Examining *Escherichia coli* glycolytic pathways, catabolite repression, and metabolite channeling using Δpfk mutants. Biotechnology for biofuels. 9**,** 212.

Janusz, G., Pawlik, A., Sulej, J., Swiderska-Burek, U., Jarosz-Wilkolazka, A., Paszczynski, A., 2017. Lignin degradation: microorganisms, enzymes involved, genomes analysis and evolution. FEMS Microbiology Reviews. 41, 941-962.

Jiang, W., Jiang, Y., Bentley, G. J., Liu, D., Xiao, Y., Zhang, F., 2015. Enhanced production of branched-chain fatty acids by replacing β-ketoacyl-(acyl-carrier-protein) synthase III (FabH). Biotechnology and Bioengineering. 112**,** 1613-1622.

Kaneda, T., 1966. Biosynthesis of branched-chain fatty acids: IV. Factors affecting relative abundance of fatty acids produced by *Bacillus subtilis*. Canadian Journal of Microbiology. 12**,** 501-514.

Kukor, J.J., Olsen, R.H., 1991. Genetic organization and regulation of a meta-cleavage pathway for catechols produced from catabolism of toluene, benzene, phenol, and cresols by *Pseudomonas pickettii* PKO1. J Bacteriol 173:4587–4594

Kurosawa, K., Laser, J., Sinskey, A., 2015. Tolerance and adaptive evolution of triacylglycerol-producing *Rhodococcus opacus* to lignocellulose-derived inhibitors. Biotechnology for biofuels. 8**,** 76.

Langmead, B., Salzberg, S. L., 2012. Fast gapped-read alignment with Bowtie 2. Nature methods. 9**,** 357-359.

Le, R. K., Das, P., Mahan, K. M., Anderson, S. A., Wells, T., Jr., Yuan, J. S., Ragauskas, A. J., 2017. Utilization of simultaneous saccharification and fermentation residues as feedstock for lipid accumulation in *Rhodococcus opacus*. AMB Express. 7**,** 185.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 25**,** 2078-2079.

Liao, Y., Smyth, G. K., Shi, W., 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 30**,** 923-930.

Linger, J. G., Vardon, D. R., Guarnieri, M. T., Karp, E. M., Hunsinger, G. B., Franden, M. A., Johnson, C. W., Chupka, G., Strathmann, T. J., Pienkos, P. T., Beckham, G. T., 2014.

772 Lignin valorization through integrated biological funneling and chemical catalysis.
773 Proceedings of the National Academy of Sciences. 111**,** 12013-12018.
774 Long, C. P., Gonzalez, J. E., Feist, A. M., Palsson, B. O., Antoniewicz, M. R., 2017. Fast growth
775 phenotype of *E. coli* K-12 from adaptive laboratory evolution does not require
776 intracellular flux rewiring. Metabolic Engineering. 44**,** 100-107.
777 Love, M. I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion
778 for RNA-seq data with DESeq2. Genome Biology. 15**,** 550.
779 Madsen, R. B., Jensen, M. M., Mørup, A. J., Houlberg, K., Christensen, P. S., Klemmer, M.,
780 Becker, J., Iversen, B. B., Glasius, M., 2016. Using design of experiments to optimize
781 derivatization with methyl chloroformate for quantitative analysis of the aqueous phase
782 from hydrothermal liquefaction of biomass. Anal Bioanal Chem. 408**,** 2171-2183.
783 Neidhardt, F. C., Ingraham, J. L., M Schaechter, M., 1990. Physiology of the bacterial cell. A
784 molecular approach. Sinauer Associates Inc, Sunderland, MA.
785 Pond, J. L., Langworthy, T. A., 1987. Effect of growth temperature on the long-chain diols and
786 fatty acids of *Thermomicrobium roseum*. Journal of Bacteriology. 169**,** 1328-1330.
787 Ragauskas, A. J., Beckham, G. T., Biddy, M. J., Chandra, R., Chen, F., Davis, M. F., Davison, B.
788 H., Dixon, R. A., Gilna, P., Keller, M., Langan, P., Naskar, A. K., Saddler, J. N.,
789 Tschaplinski, T. J., Tuskan, G. A., Wyman, C. E., 2014. Lignin Valorization: Improving
790 Lignin Processing in the Biorefinery. Science. 344. 6185, 1246843.
791 Ragauskas, A. J., Williams, C. K., Davison, B. H., Britovsek, G., Cairney, J., Eckert, C. A.,
792 Frederick, W. J., Hallett, J. P., Leak, D. J., Liotta, C. L., Mielenz, J. R., Murphy, R.,
793 Templer, R., Tschaplinski, T., 2006. The Path Forward for Biofuels and Biomaterials.
794 Science. 311**,** 484-489.
795 Saa, L., Jaureguibeitia, A., Largo, E., Llama, M. J., Serra, J. L., 2010. Cloning, purification and
796 characterization of two components of phenol hydroxylase from *Rhodococcus*
797 *erythropolis* UPV-1. Applied Microbiology and Biotechnology. 86**,** 201-211.
798 Schuetz, R., Kuepfer, L., Sauer, U., 2007. Systematic evaluation of objective functions for
799 predicting intracellular fluxes in *Escherichia coli*. Mol Syst Biol. 3, 119.
800 Sridevi, V., V V Chandana Lakshmi, M., Manasa, M., Sravani, M., 2012. Metabolic pathways
801 for the biodegradation of phenol. Int. J. Eng. Sci. Adv. Technol. 2**,** 695-705.
802 Studer, M. H., DeMartini, J. D., Davis, M. F., Sykes, R. W., Davison, B., Keller, M., Tuskan, G.
803 A., Wyman, C. E., 2011. Lignin content in natural Populus variants affects sugar release.
804 Proceedings of the National Academy of Sciences. 108**,** 6300-6305.
805 Szőköl, J., Rucká, L., Šimčíková, M., Halada, P., Nešvera, J., Pátek, M., 2014. Induction and
806 carbon catabolite repression of phenol degradation genes in *Rhodococcus erythropolis*
807 and *Rhodococcus jostii*. Applied Microbiology and Biotechnology. 98**,** 8267-8279.
808 Valdivia, M., Galan, J. L., Laffarga, J., Ramos, J.-L., 2016. Biofuels 2020: Biorefineries based
809 on lignocellulosic materials. Microbial biotechnology. 9**,** 585-594.
810 Wheeldon, I., Christopher, P., Blanch, H., 2017. Integration of heterogeneous and biochemical
811 catalysis for production of fuels and chemicals from biomass. Curr Opin Biotechnol. 45**,**
812 127-135.
813 Xie, S., Sun, S., Lin, F., Li, M., Pu, Y., Cheng, Y., Xu, B., Liu, Z., da Costa Sousa, L., Dale, B.
814 E., Ragauskas, A. J.,Dai, S. Y., Yuan, J. S., 2019. Mechanism-Guided Design of Highly
815 Efficient Protein Secretion and Lipid Conversion for Biomanufacturing and
816 Biorefining. Adv. Sci. 1801980.

Yaguchi, A., Robinson, A., Mihealsick, E., Blenner, M., 2017. Metabolism of aromatics by *Trichosporon oleaginosus* while remaining oleaginous. Microbial Cell Factories. 16**,** 206.
Yi, X., Dean, A. M., 2016. Phenotypic plasticity as an adaptation to a functional trade-off. eLife. 5**,** e19307.
Yoneda, A., Henson, W. R., Goldner, N. K., Park, K. J., Forsberg, K. J., Kim, S. J., Pesesky, M. W., Foston, M., Dantas, G., Moon, T. S., 2016. Comparative transcriptomics elucidates adaptive phenol tolerance and utilization in lipid-accumulating *Rhodococcus opacus* PD630. Nucleic Acids Research. 44**,** 2240-2254.
You, L., Page, L., Feng, X., Berla, B., Pakrasi, H. B., Tang, Y. J., 2012. Metabolic Pathway Confirmation and Discovery Through (13)C-labeling of Proteinogenic Amino Acids. Journal of Visualized Experiments : JoVE. 3583.
Young, J. D., 2014. INCA: a computational platform for isotopically non-stationary metabolic flux analysis. Bioinformatics. 30**,** 1333-1335.