Title
Excavating the genome mine of Pseudomonas putida KT2440

Author
Incha, Matthew Ryan

Publication Date
2023

Peer reviewed|Thesis/dissertation

Excavating the genome mine of *Pseudomonas putida* KT2440

by

Matthew Ryan Incha

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Microbiology

in the

Graduate Division

of the

University of California, Berkeley

Committee in Charge:

Professor Jay D. Keasling, Chair
Professor Adam P. Arkin
Professor John D. Coates

Spring 2023

Excavating the genome mine of *Pseudomonas putida* KT2440

Abstract

Excavating the genome mine of *Pseudomonas putida* KT2440

by

Matthew Ryan Incha

Doctor of Philosophy in Microbiology

University of California, Berkeley

Professor Jay D. Keasling, Chair

Pseudomonas putida is an important chassis strain for the valorization of lignin. While we know much about this host and have developed numerous tools to engineer it, we still lack a complete understanding of its accessory genes that may be important for a variety of metabolisms and stress responses. Using a functional genomics method called barcode abundance sequencing (BarSeq), thousands of conditionally essential genes can be assayed in a pooled manner to elucidate their functions. With these assays novel metabolisms and novel gene functions can be identified and used to inform genetic engineering efforts. To validate the practicality of this method, we first assayed the aromatic metabolic pathways of *Pseudomonas putida* KT2440, and then applied the initial step of hydroxycinnamate catabolism toward the synthesis of a curcuminoid. Following this success, we sought to characterize the other diverse metabolisms of *P. putida*. Fatty acids and alcohols are highly sought after bioproducts for use as fuels and commodity chemicals; however, *P. putida* is known to metabolize most of these compounds. We therefore used BarSeq to elucidate the metabolism of 23 of these compounds with applications ranging from surfactants to plastic monomers. The results of these assays could then be used to inform engineering efforts for the production of these molecules and their derivatives. Given the diversity of metabolisms, next we sought to analyze the nitrogen metabolisms of *P. putida* as these again have applications as commodity chemicals. To this end, 52 different nitrogen-containing compounds were assayed via BarSeq as sole nitrogen sources. These molecules have many overlapping gene requirements in their metabolism, thus requiring a method for effectively analyzing and interpreting the data. The machine learning method known as t-stochastic neighbor embedding (tSNE) was employed to generate an interactive map of clustered genes with the distances between genes determined by the similarity of each gene's BarSeq data.

For Eric

Table of Contents

List of Figures

List of Tables

Acknowledgements

My graduate student experience was an interesting one. I owe a lot of credit to my friend and mentor, Mitchell Thompson, as well as my close friends in my department, Sebastian Fernandez, Pierre Joubert, Kevin Yin. I'm grateful for the support and love from my old friends who have been by my side since childhood, Dylan Hunter, Kevin Mckenna, and Nathan Bolen, and my friends from college Ellie Suhendra, Brandon Humboldt, Dilpreet Singh Randhawa, and Nicholas Sieb. Additionally, this PhD would not be possible without the work of my close collaborators and friends Allison Pearson and Matthias Schmidt. My parents, Mark Incha and Laura Waldvogel, and sisters Morgan Grant, Jillian Incha, and Elyse Incha have been the cornerstone of love and support through this journey.

Chapter 1. *Pseudomonas putida*: Beginning of a great adventure

1.1. Preface

Pseudomonas alloputida (née *P. putida* KT2440 and hereafter referred to as *P. putida* KT2440) has gained promise in biotechnology for its ability to catabolize a wide range of carbon sources and its genetic tractability. It can utilize inexpensive feedstocks such as lignin monomers and convert them into alcohols, fatty acids and other desirable biomolecules. Due to these characteristics, it is essential to have a strong fundamental understanding of the genes, pathways, and proteins involved. While there exists general annotations for most of the genes in the genome, these still lack a specific functional annotation due to the prevalence of paralogous genes. This presents an issue to the metabolic engineer as it can affect the predictive power of models which can direct their efforts. Because of this, it becomes necessary to utilize high-throughput methods to generate a clearer picture of this microbe. To this end, we use a functional genomics method barcode abundance sequencing (BarSeq) to identify conditionally essential genes that are involved in the metabolism of feedstocks and desired bioproducts. Guided by this data I generated an engineered pathway that employed some native genes to produce a curcuminoid. Then, using the same method we elucidated genes involved in the metabolism of diverse fatty acids, fatty alcohols and nitrogen sources. This became an immense dataset containing over 1000 genes with significant phenotypes. To better visualize the data I present the use of a manifold learning method, t-stochastic neighbor embedding.

1.2. Organization

In **Chapter 2** we provide an overview of the fatty acid and alcohol metabolisms of *P. putida* KT2440 identified via random barcode transposon abundance sequencing, RB-TnSeq, assays. These include short chain alcohols with industrial importance like isoprenol and 1,4-butanediol, and fatty acids such as octanoate and decanoate, which have importance as potential biofuel precursors. Finally we discuss the utility of this method in engineering *P. putida* for the production of these compounds.

In **Chapter 3** we illustrate an overview of the nitrogen assimilation pathways of *P. putida* KT2440 identified via RB-TnSeq assays. Using RB-TnSeq we identified several important 'nodes' in the assimilation of non-standard nitrogen-containing compounds and demonstrate the use of a manifold learning algorithm (t-SNE) in the compression and interpretation of these large functional genomics datasets

In **Chapter 4** we employ RB-TnSeq to investigate the catabolic pathways of *Pseudomonas putida* KT2440 for lignin monomers and related aromatic compounds. We then identified coumaroyl-CoA as a toxic intermediate to *P. putida* and propose pathways that may be the mechanism of detoxification. The data from functional genomics assays on aromatic carbon sources provides a roadmap for the heterologous production of type-III PKS in *P. putida.* First, the red pigment, flaviolin, is produced to

provide insight into the media composition and how it affects the biosynthetic performance of *P. putida*. Then we produced the plant polyketide, bisdemethoxycurcumin, which can be derived from the lignin monomer coumarate.

Chapter 2. - Pseudomonas putida, consumer of the bioprivileged

Including material from published work:
Mitchell G. Thompson, **Matthew R. Incha**, Allison N. Pearson, Matthias Schmidt, William A. Sharpless, Christopher B. Eiben, Pablo Cruz-Morales, Jacquelyn M. Blake-Hedges, Yuzhong Liu, Catharine A. Adams, Robert W. Haushalter, Rohith N. Krishna, Patrick Lichtner, Lars M. Blank, Aindrila Mukhopadhyay, Adam M. Deutschbauer, Patrick M. Shih, Jay D. Keasling, "Functional analysis of the fatty acid and alcohol metabolism of *Pseudomonas putida* using RB-TnSeq", AEM, 2020

2.0. Chapter preface

Pseudomonas putida KT2440 is rapidly emerging as a vital chassis organism for metabolic engineering due to its versatile carbon source catabolism and an expanding engineering toolkit. However, despite significant advancements in understanding this organism, considerable gaps remain in our knowledge of its genetic basis and metabolic capabilities. This chapter will address these gaps, focusing specifically on the complexities of fatty acid and alcohol catabolism where numerous paralogous genes encoding seemingly similar enzymes coexist, complicating biochemical assignment through sequence homology.

2.1. Abstract

With its ability to catabolize a wide variety of carbon sources and a growing engineering toolkit, *Pseudomonas putida* KT2440 is emerging as an important chassis organism for metabolic engineering. Despite advances in our understanding of this organism, many gaps remain in our knowledge of the genetic basis of its metabolic capabilities. These gaps are particularly noticeable in our understanding of both fatty acid and alcohol catabolism, where many paralogs putatively coding for similar enzymes co-exist making biochemical assignment via sequence homology difficult. To rapidly assign function to the enzymes responsible for these metabolisms, we leveraged Random Barcode Transposon Sequencing (RB-TnSeq). Global fitness analyses of transposon libraries grown on 13 fatty acids and 10 alcohols produced strong phenotypes for hundreds of genes. Fitness data from mutant pools grown on varying chain length fatty acids indicated specific enzyme substrate preferences, and enabled us to hypothesize that DUF1302/DUF1329 family proteins potentially function as esterases. From the data we also postulate catabolic routes for the two biogasoline molecules isoprenol and isopentanol, which are catabolized via leucine metabolism after initial oxidation and activation with CoA. Because fatty acids and alcohols may serve as both feedstocks or final products of metabolic engineering efforts, the fitness data presented here will help guide future genomic modifications towards higher titers, rates, and yields.

2.2. Introduction

Pseudomonas putida KT2440 is an important metabolic engineering chassis, which

can readily metabolize compounds derived from lignocellulosic and plastic derived feedstocks[1–3], and has an ever-growing repertoire of advanced tools for genome modification[4–7]. Its upper glycolytic pathway architecture enables *P. putida* to natively generate large amounts of reducing equivalent[8], and it more robustly withstands metabolic burdens than many other frequently used host organisms[9]. To date, a wide variety of products have been produced through metabolic engineering of *P. putida*, including valerolactam[10], curcuminoids[11], diacids[12], methyl-ketones[13], rhamnolipids[14], cis,cis-muconic acid[15], and many others[16]. Recent advances in genome-scale metabolic modeling of *P. putida* make engineering efforts more efficient[7,17]. However, a large gap still exists between genes predicted to encode enzymatic activity and functional data to support these assumptions. Recent characterizations of enzymes and transporters involved in the catabolism of lysine[12,18], levulinic acid[19], and aromatic compounds[20] highlight the need to continue functionally probing the metabolic capabilities of *P. putida*, because its native catabolism can consume many target molecules and dramatically impact titers.

Amongst the most important metabolisms not yet rigorously interrogated via omics-level analyses are fatty acid and alcohol degradation. Recently, fatty acids have been shown to be a non-trivial component of some feedstock streams[1] and, depending on their chain length, serve as high-value target molecules[21]. Furthermore, intermediates in beta-oxidation can be channeled towards mega-synthases to produce more complex molecules[22], or used in reverse beta-oxidation to produce compounds such as medium chain n-alcohols[23]. However, assigning the genetic basis of fatty acid degradation is complicated by the presence of multiple homologs of the individual *fad* genes encoded in the genome of *P. putida* KT2440[17,24]. Although work has been done to either biochemically or genetically demonstrate the substrate specificity of some individual *fad* genes, the majority of these homologs still have no functional data associated with them.

P. putida is also able to oxidize and catabolize a wide variety of alcohols. Much work has focused on the unique biochemistry and regulation of two pyrroloquinoline quinone (PQQ)-dependent alcohol dehydrogenases (ADH), *pedE* and *pedH*, which exhibit broad substrate specificity for both alcohols and aldehydes[25,26]. Specific work has also investigated the suitability of *P. putida* for the production of ethanol[27] and the genetic basis for its ability to catabolize butanol and 1,4-butanediol[28–30]. *P. putida* is also known for its ability to tolerate solvents and alcohols, making it an attractive host for their industrial production[31,32]. Tolerance to these compounds is a product of both robust efflux pumps[31] and the ability of some strains, such as *P. putida* mt-2, to catabolize a range of organic compounds[33]. Metabolic engineering has biologically produced a diverse range of alcohols with a wide array of industrial and commercial uses[34–36]. As more alcohol synthesis pathways are engineered into *P. putida*, a more complete understanding of the molecular basis of its catabolic capacities will be required to achieve high-titers.

A recent surge in omics-level data has revealed much about the metabolism of *P. putida*, with adaptive evolution[30], proteomics[10,28,29], and ^{13}C flux analysis[37–39] yielding valuable insights. An approach that has proven to be particularly powerful is Random Barcode Transposon Sequencing (RB-TnSeq)[40,41]. RB-TnSeq allows rapid and inexpensive genome-wide profiling of individual gene fitness in various conditions, and has been used in *P. putida* to identify numerous novel metabolic pathways and aid in increasing titers of the polymer precursor valerolactam[10,11,18–20]. RB-TnSeq improves over other TnSeq-based techniques by introducing a random nucleotide barcode into the

transposon which is flanked by conserved primer binding sites[40]. After one initial round of TnSeq to map the transposon insertion within the genome which also associates that insertion site with a barcode, all subsequent mutant abundance quantification can be performed using standard barcode counting via Illumina sequencing of PCR product from the conserved priming sites[40]. This advance reduces the cost per experiment, as well as greatly speeding up processing time required to conduct genome-wide fitness experiments[40]. Here, we leverage RB-TnSeq to interrogate the genetic basis for the catabolism of multiple fatty acids and alcohols to develop an evidence-based understanding of the enzymes and pathways utilized in these metabolisms.

2.3. Results and Discussion

2.3.1. Global Analysis of Fatty Acid Metabolism.

To characterize the genetic basis of fatty acid metabolism in *P. putida*, barcoded transposon mutant libraries were grown in minimal media with straight chain fatty acids (C3-C10, C12, and C14), fatty esters (Tween20 and butyl stearate), and an unsaturated fatty acid (oleic acid) as sole carbon sources. An overview of sample collection is provided in **Figure 2.1A.** Pearson correlation analyses of global fitness patterns revealed that the metabolisms of straight chain fatty acids between C7 and C14 clade together, suggesting similar overall catabolic routes (**Figure 2.1B**). Oleic acid, an 18-carbon monounsaturated fatty acid, also grouped within this clade. Shorter chain fatty acids (<C7) did not show high correlation to one another based on global fitness analyses, suggesting more independent routes of catabolism (**Figure 2.1B**). Annotations in the BioCyc database, functional assignment from a recent metabolic model of *P. putida* KT440 (*i*JN1462), and previous *in vitro* biochemical work predict the existence of several enzymes in the genome of the bacterium that may be putatively involved in fatty acid catabolism: six acyl-CoA ligases, seven acyl-CoA dehydrogenases, seven enoyl-CoA hydratases, four hydroxyacyl-CoA dehydrogenases, and five thiolases (**Figure 2.2**)[17,24,42]. Our data show discrete fitness patterns for the steps of beta-oxidation that appear to be largely dictated by chain length (**Figure 2.2**).

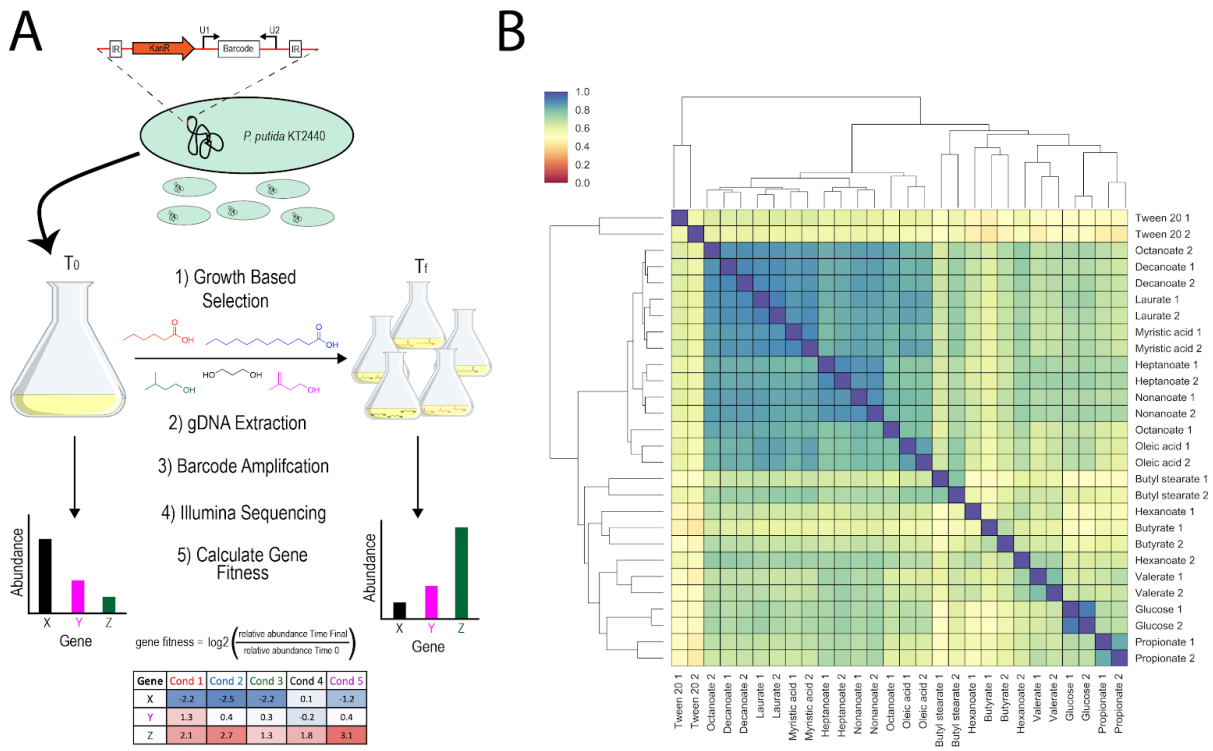


Figure 2.1. Overview of RB-TnSeq experimental design and Cladogram correlation matrix of genome-wide fitness data of *P. putida* grown on fatty acids.

A) Pooled barcoded mutants of *P. putida* KT2440 are first grown in rich media then 1) washed and used to inoculate minimal media with a single carbon source. Samples from both Time 0 and from the endpoints of grown on selective conditions then 2) have their gDNA extracted 3) have barcodes amplified from conserved priming sites 4) barcode abundance is calculated via Illumina sequencing and 5) gene fitness per condition can be calculated by comparing relative abundance of mutants before and after selection. B) The matrix shows pairwise comparisons of Pearson correlations of fitness data from *P. putida* KT2440 RB-TnSeq libraries grown on fatty acids as well as glucose. The legend in top left shows Pearson correlation between two conditions with blue showing $r = 1$, and red showing $r = 0$. The conditions were tested in duplicate and the data from each are numbered (1 & 2).

When grown on fatty acids, many bacteria require the anaplerotic glyoxylate shunt to avoid depleting TCA cycle intermediates during essential biosynthetic processes. In *P. putida*, the two steps of the glyoxylate shunt are encoded by PP_4116 (*aceA* - isocitrate lyase) and PP_0356 (*glcB* - malate synthase). Transposon mutants in both of these genes showed serious fitness defects (fitness score < -3) when grown on nearly all of the fatty acids tested (**Figure 2.2**). However, the glyoxylate shunt genes appeared dispensable for growth on valerate (C5), and showed a more severe fitness defect when grown on heptanoate (C7). Complete beta-oxidation of valerate and heptanoate results in ratios of propionyl-CoA to acetyl-CoA of 1:1 and 1:2, respectively. This higher ratio of 3-carbon to 2-carbon production presumably offers an alternate

means to replenish TCA cycle intermediates in the absence of a glyoxylate shunt (**Figure 2.2**).

In order to utilize the propionyl-CoA generated by beta-oxidation of odd-chain fatty acids, bacteria often employ the methylcitrate cycle (MCC), producing succinate and pyruvate from oxaloacetate and propionyl-CoA. In *P. putida*, the MCC is catalyzed via methylcitrate synthase (*prpC* - PP_2335), 2-methylcitrate dehydratase (*prpD* - PP_2338, or *acnB* - PP_2339), aconitate hydratase (*acnB* - PP_2339, or *acnA2* - PP_2336), and 2-methylisocitrate lyase (*mmgF* - PP_2334) (**Figure 2.3**). Unsurprisingly, the MCC appeared to be absolutely required for growth on propionate (C3), valerate (C5), heptanoate (C7), and nonanoate (C9), with PP_2334, PP_2335, and PP_2337 (a putative AcnD-accessory protein) showing severe fitness defects (**Figure 2.2 & 2.3**). While PP_2338 (*prpD*) encodes for a 2-methylcitrate dehydratase, transposon mutants showed no fitness defects when grown on odd-chain fatty acids. This reaction is likely carried out by PP_2339 (*acnB* - a bifunctional 2-methylcitrate dehydratase/aconitase hydratase B); however, there were no mapped transposon insertions for this gene (**Figures 2.2 & 2.3**). This suggests that PP_2339 was essential during the construction of the RB-TnSeq library. Furthermore, PP_2336 showed relatively modest fitness defects when grown on propionate and other odd-chain fatty acids, suggesting that PP_2339 likely accounts for much of the methylaconitate hydratase activity as well (**Figures 2.2 & 2.3**).

2.3.2. Long and Medium Chain Fatty Acid Catabolism.

Pearson correlation analysis of fitness data indicated that both long and medium chain fatty acids are likely catabolized via similar pathways. Fitness data suggests that FadD1 (PP_4549) catalyzes the initial CoA-ligation of C7 to C18 fatty acids, and may potentially act on C6 as well (**Figure 2.2**). Disruption of *fadD2* (PP_4550) did not cause fitness defects as severe as those seen in *fadD1* mutants, although it did result in moderate fitness defects when grown on C8-C10 fatty acids. These data are consistent with the biochemical characterization of FadD1 from *P. putida* CA-3, which showed greater activity on longer chain alkanoic and phenylalkanoic acids than on shorter chain substrates[43]. For fatty acids with chain lengths of C10 and greater, the data suggest that the *fadE* homolog PP_0368 is the primary acyl-coA dehydrogenase, while the nearby *fadE* homolog PP_0370 appears to be preferred for C6-C8 fatty acids (**Figure 2.2**). A relatively even fitness defect for these two *fadE* homologs indicates that PP_0368 and PP_0370 may have equal activity on nonanoate (**Figure 2.2**). These data are supported by a previous biochemical characterization of PP_0368, in which it showed greater activity on chain lengths longer than C9[44]. The *fadB* homolog PP_2136 showed severe fitness defects when grown on all fatty acids with chain lengths of C6 and longer, implicating it as the primary enoyl-CoA hydratase/3-hydroxy-CoA dehydrogenase for those substrates (**Figure 2.2**). *P. putida* was able to grow on the unsaturated substrate oleic acid, and is likely able to isomerize the position of the unsaturated bond via the enoyl-CoA isomerase PP_1845, which showed specific fitness defects when grown on oleic acid (**Figure 2.2**). *P. putida*'s primary long chain thiolase appears to be the *fadA* homolog PP_2137, which showed severe to moderate fitness defects when grown on fatty acids with chain lengths C8 or longer (**Figure 2.2**). Fitness data for mutant pools grown on heptanoate showed minor fitness defects for both PP_2137 and PP_3754 (*bktB*), suggesting that both thiolases may work on C7 substrates (**Figure 2.2**).

Both long chain fatty esters tested (Tween 20 and butyl stearate) appeared to utilize the same *fad* homologs as the long chain fatty acids. However, before either molecule can be directed towards beta-oxidation, Tween 20 and butyl stearate must be hydrolyzed to generate a C12 or C18 fatty acid, respectively. To date, no such hydrolase has been identified in *P. putida* KT2440. Comparing the mutant fitness scores between Tween 20 and laurate (C12) carbon source experiments revealed six genes (PP_0765, PP_0766, PP_0767, PP_0914, PP_2018, and PP_2019) that had significant and severe fitness defects specific to Tween 20 (fitness score < -2, $t > |4|$) in both biological replicates (**Figure 2.4**). The same comparison between butyl stearate and myristate (C14) revealed four genes specific to the fatty ester (PP_0765, PP_0766, PP_2018, and PP_4058) that had significant severe fitness defects (fitness score < -2, $t > |4|$) in both biological replicates (**Figure 2.4**). As PP_0765-6 and PP_2018 appear to have specific importance in both of the ester conditions tested, it may be possible that they contribute to the hydrolysis of the fatty ester bonds. However, it is also possible that the esterase is secreted or associated with the outer membrane[45], in which case its enzymatic activity would be shared amongst the library and it would not have the associated fitness defect expected[10].

The genes PP_2018 and PP_2019 encode a BNR-domain containing protein and a RND-family efflux transporter, respectively, and are likely co-expressed in an operon that also includes PP_2020 and PP_2021. Interestingly, although PP_2021 codes for a putative lactonase, transposon mutants had no apparent fitness defect with either of the fatty esters as the carbon source. PP_0765 and PP_0766 encode a DUF1302 family protein and DUF1329 family protein, respectively. Given their similar fitness scores, they are likely coexpressed in an operon positively regulated by the LuxR-type regulator PP_0767 (**Figure 2.4**). Previous work in multiple other species of *Pseudomonas* has observed cofitness of DUF1302/DUF1329 family genes with BNR-domain and RND-family efflux genes when grown on Tween 20[41]. The authors proposed that these genes may work together in order to export a component of the cell wall. However, an alternative hypothesis could be that PP_0765 and PP_0766 contribute to catalyzing the hydrolysis of fatty esters, accounting for the missing catabolic step of butyl stearate and Tween 20. This hypothesis is bolstered somewhat by the co-localization of PP_0765/PP_0766 near fatty acid catabolic genes in *P. putida* KT2440 and many other *Pseudomonads* (**Figure 2.5**). Future work will need to be done to biochemically characterize the potential enzymatic activity of these proteins.

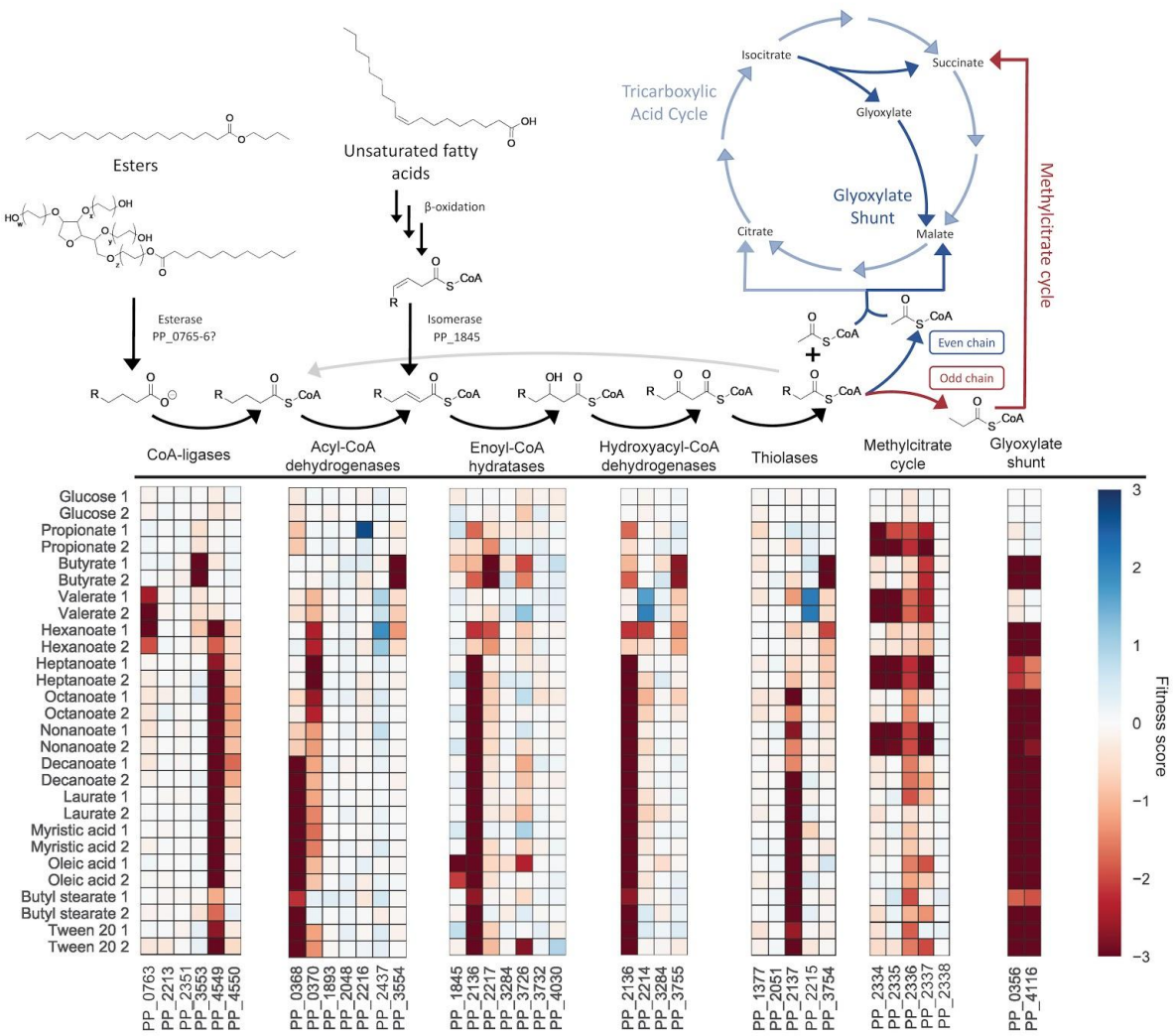


Figure 2.2. Overview of fatty acid catabolic pathways of *P. putida* KT2440.

The above diagram shows the catabolic steps of fatty ester and saturated/unsaturated fatty acid catabolism in *P. putida* KT2440, in addition to their connections to the glyoxylate shunt and the methylcitrate cycle. The heatmaps below show fitness scores when grown on fatty acids or glucose for the specific genes proposed to catalyze individual chemical reactions. Colors represent fitness scores, with blue representing positive fitness and red representing negative fitness.

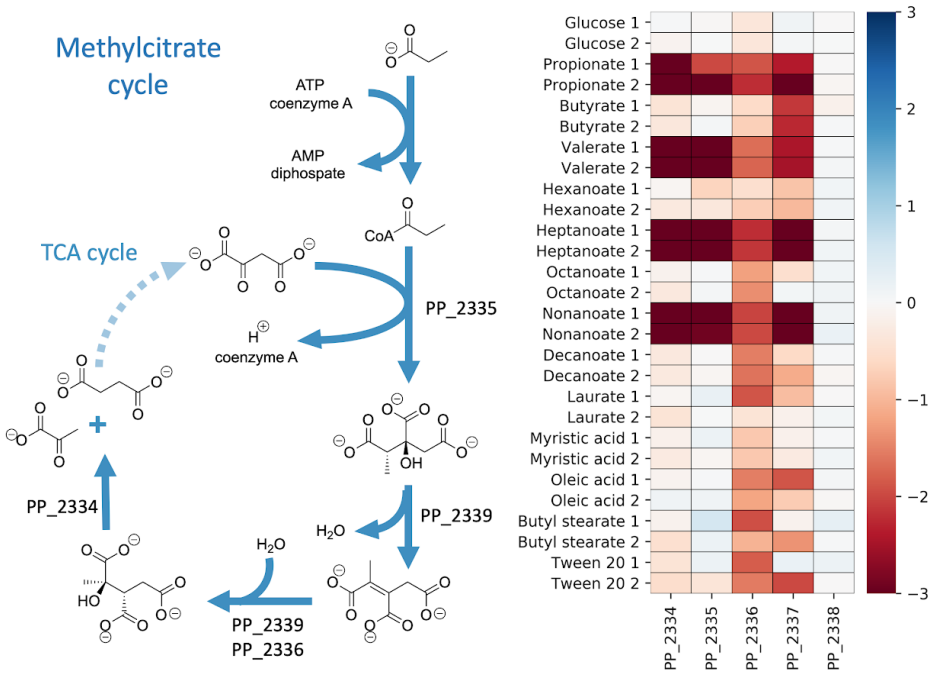


Figure 2.3. Fitness scores for genes catalyzing the methylcitrate cycle of *P. putida* KT2440.

Metabolic pathway for the MCC of *P. putida* KT2440. Heatmap to the right shows fitness scores for MCC genes when grown on fatty acids or glucose. PP_2337 is predicted to function as an aconitate isomerase not shown in the depiction (left).

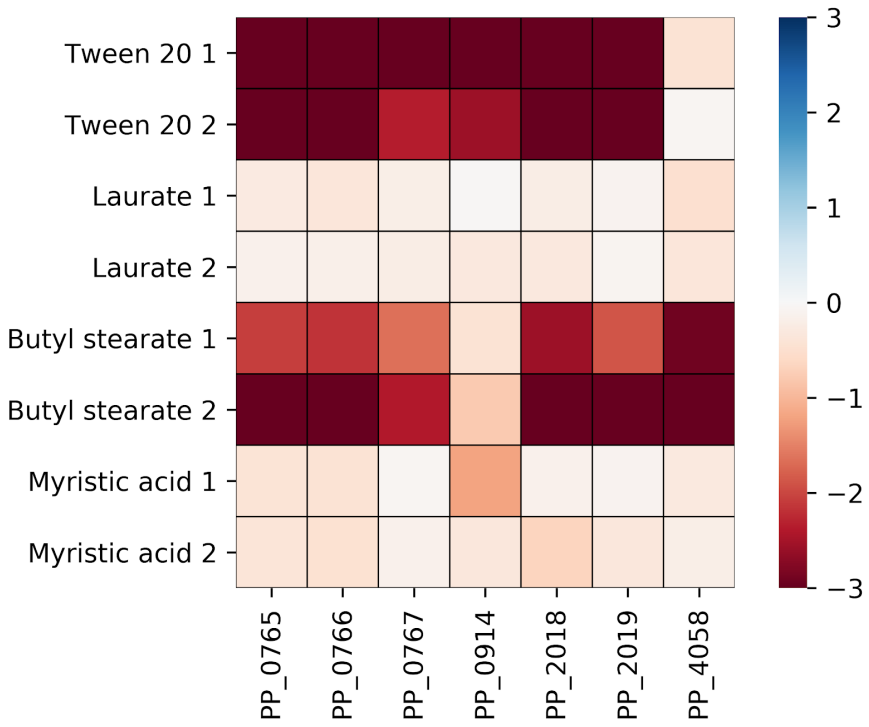


Figure 2.4. Genes in *P. putida* specifically defective for growth on long chain fatty esters compared to long chain fatty acids.

Heatmap shows fitness values of genes that were found to be specifically defective on fatty esters. Conditions shown are the two fatty esters Tween 20 and butyl stearate, as well as long chain fatty acids laurate and myristate.

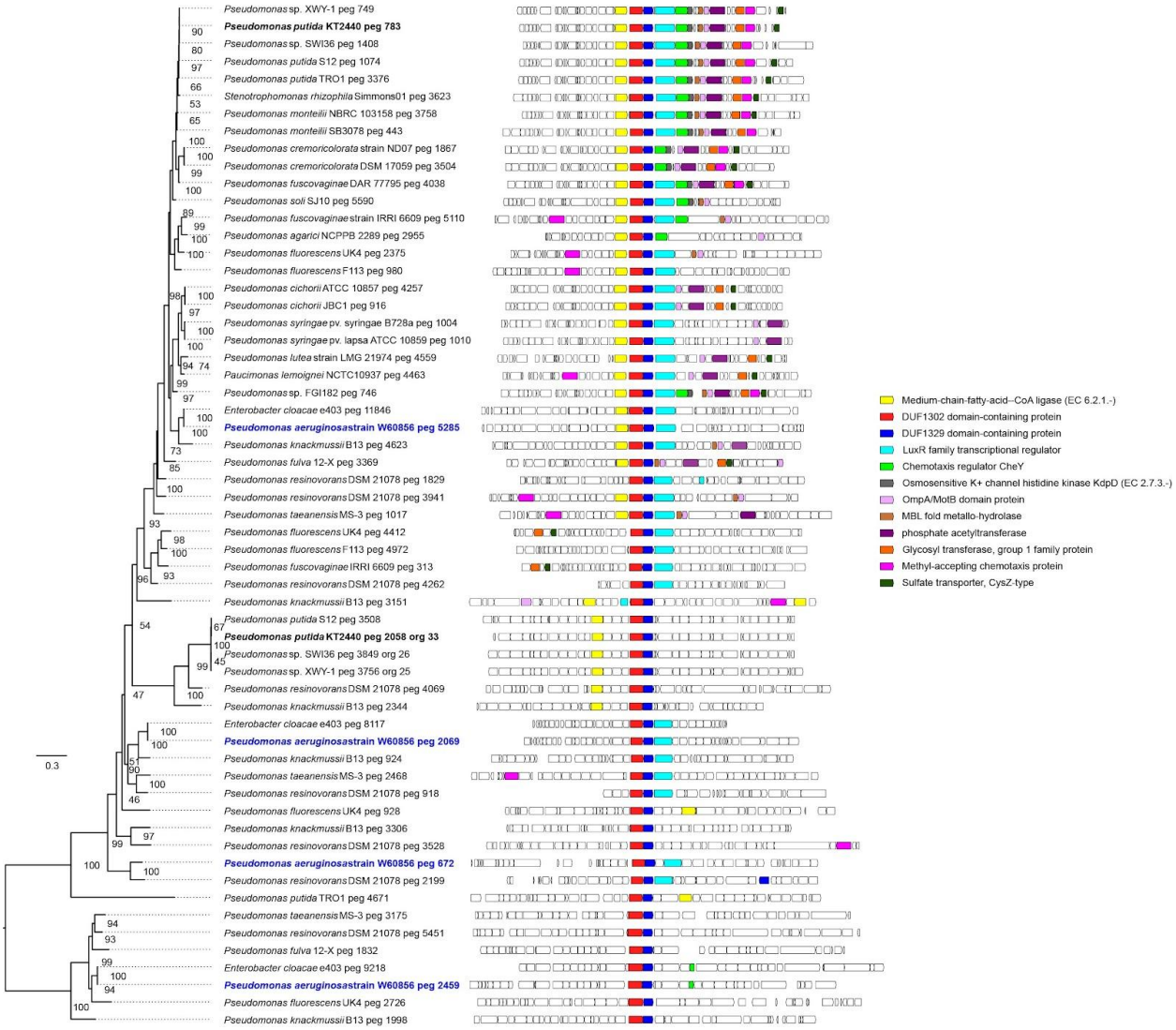


Figure 2.5. Genomic context analysis of PP_0765/PP_0766 homologs across related *Pseudomonas* species.

The genomic contexts of PP_0765/PP_0766 orthologs across multiple related *Pseudomonas* species are shown. Phylogenetic tree shows relatedness of the conserved gene pair. Genes that are commonly found near PP_0765/PP_0766 are colored, medium chain fatty acid coA-ligases are shown in yellow.

2.3.3. Short Chain Fatty Acid Catabolism.

In our genome-wide fitness assays, the mutant fitness patterns of C6 or shorter fatty acid carbon sources had lower Pearson correlation between one another than the correlations within long and medium-chain fatty acids (**Figure 2.1**). These global differences reflect what appear to be discrete preferences in beta-oxidation enzymes for growth on short chain fatty acids. Fitness data suggest that while both CoA-ligases PP_0763 and PP_4559 are required for growth on hexanoate, only PP_0763 is required

for growth on valerate (**Figure 2.2**). Furthermore, the putative positive regulator of PP_0763, LuxR-family transcription factor PP_0767, also showed a significant fitness defect (-2.0) when grown on both valerate and hexanoate (**Figure 2.2**). PP_0370 seems to be the acyl-CoA dehydrogenase largely responsible for hexanoate catabolism, though PP_3554 mutants also have minor fitness defects. The dehydrogenation of valeryl-coA appears to be distributed between the activities of PP_0368, PP_0370, and PP_3554, with no single acyl-CoA dehydrogenase mutant demonstrating a strong fitness defect when grown on valerate (**Figure 2.2**). Interestingly, though previous biochemical analysis had demonstrated that PP_2216 has activity on C4-C8 acyl-CoA substrates with a preference for shorter chain lengths[46], we observed no fitness defects for PP_2216 mutants when grown on any fatty acid carbon source (**Figure 2.2**).

It appears that the role of enoyl-CoA hydratase or hydroxyacyl-CoA dehydrogenase may be distributed across multiple enzymes for both hexanoate and valerate. Growth on hexanoate resulted in moderate fitness defects in mutants disrupted in the predicted enoyl-CoA hydratases PP_2136, PP_2217, and PP_3726; however, for mutants grown on valerate, there were almost no observable fitness defects for any of the enoyl-CoA hydratase enzymes examined in the study, suggesting that for this chain length significant functional complementation exists between the *fadB* homologs (**Figure 2.2**). Fitness data suggest that PP_2136 (*fadB*), PP_2214 (a predicted type II 3-hydroxyacyl-CoA dehydrogenase), and PP_3755 (a 3-hydroxybutyryl-CoA dehydrogenase) may all be involved in the dehydrogenation of 3-hydroxyhexanoyl-CoA (**Figure 2.2**), while there appears to be a distribution of *fadB*-like activity when it comes to the dehydrogenation of 3-hydroxyvaleryl-CoA, with PP_3755 showing only a slight fitness defect on valerate. Intriguingly, mutants disrupted in the predicted type-2 acyl-CoA dehydrogenase PP_2214 showed apparent increased fitness when grown on valerate (**Figure 2.2**). As with heptanoate, fitness data from mutant pools grown on valerate or hexanoate suggest that both PP_2137 and PP_3754 may catalyze the terminal thiolase activity of these substrates. The lack of pronounced fitness phenotypes for the beta-oxidation steps of both valerate and hexanoate underscores the necessity for further *in vitro* biochemical interrogation of these pathways.

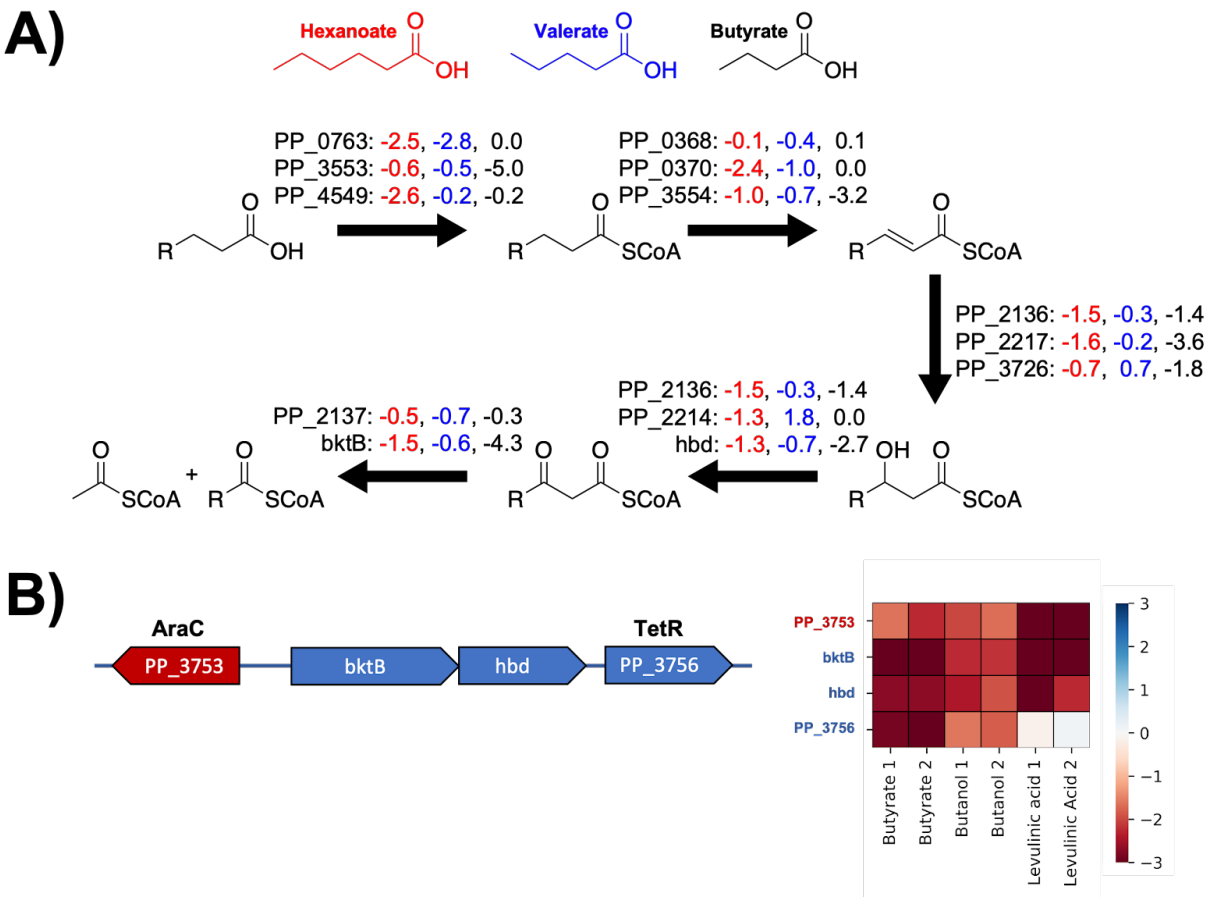


Figure 2.6. Putative pathways for short chain fatty acid catabolism in *P. putida* KT2440.

A) Individual enzymatic steps that potentially catalyze the steps of beta-oxidation for short chain fatty acids, fitness scores are listed to the right of each enzyme when grown on either butyrate, valerate, or hexanoate. B) The operonic structure of *btkB* and *hdb* flanked by an AraC-family (PP_3753) and TetR-family (PP_3756). The heatmap shows fitness scores of the genes when grown on butyrate, butanol, or levulinic acid.

Both the butanol and butyrate metabolism of *P. putida* have been studied in detail through omics-level interrogation across multiple strains[28,29]. Previous work showed that during growth on n-butanol, which is later oxidized to butyrate, three CoA-ligases are up-regulated: PP_0763, PP_3553, and PP_4487 (*acsA-1* - an acyl-CoA synthase)[29]. However, our butyrate carbon source experiments only revealed strong fitness defects in PP_3553 mutants (**Figure 2.2**, **Figure 2.6A**). The same work found that PP_3554 was the only upregulated acyl-CoA dehydrogenase, which agrees with the strong fitness defect we observed in mutants of that gene[29]. That prior work did not find upregulation of any enoyl-CoA hydratase in *P. putida* grown on butanol, but this is likely reflective of redundancy in this step; we observed fitness defects in multiple genes, including PP_2136, PP_2217, and PP_3726, with mutants in PP_2217 demonstrating the most severe fitness defect (**Figure 2.2, Figure 2.6A**). Hydroxyacyl-CoA dehydrogenase PP_2136 and 3-hydroxybutyryl-CoA dehydrogenase PP_3755 (*hbd*) have both been

shown to be upregulated during growth on butanol[29]. While our data showed fitness defects in both of these genes, the defect of PP_3755 mutants was much more severe. Three different thiolases (PP_2215, PP_3754, and PP_4636) and the 3-oxoacid CoA-transferase *atoAB* were previously observed to be upregulated during growth on butanol, but of these genes, only PP_3754 (*bktB*) had a strong fitness defect, implying that it is the main thiolase for the terminal step of butyrate catabolism (**Figure 2.2, Figure 2.6A**).

The inability of the RB-TnSeq data to clearly show which enzymes are likely responsible for specific beta-oxidation reactions suggest multiple enzymes may catalyze these steps. In addition to the lack of genotype to phenotype clarity in the enzymes responsible for the catabolic steps, we observed additional phenotypes within our fitness data that portray a complex picture of short chain fatty acid metabolism in *P. putida*. The TetR-family repressor *paaX* (PP_3286) was shown to have a negative fitness score when mutant pools were grown on fatty acids with chain lengths C7 or below (**Figure 2.7**). PaaX negatively regulates the *paa* gene cluster encoding the catabolic pathway for phenylalanine[47,48], implying that presence of phenylalanine catabolism impedes growth on short chain fatty acids. It is therefore somewhat surprising that no individual mutant within the *paa* gene cluster shows a fitness increase when grown on short chain fatty acids, though no robust fitness data exists for *paaJ* (PP_3275 - a 3-oxo-5,6-didehydrosuberyl-coA thiolase) (**Figure 2.7**).

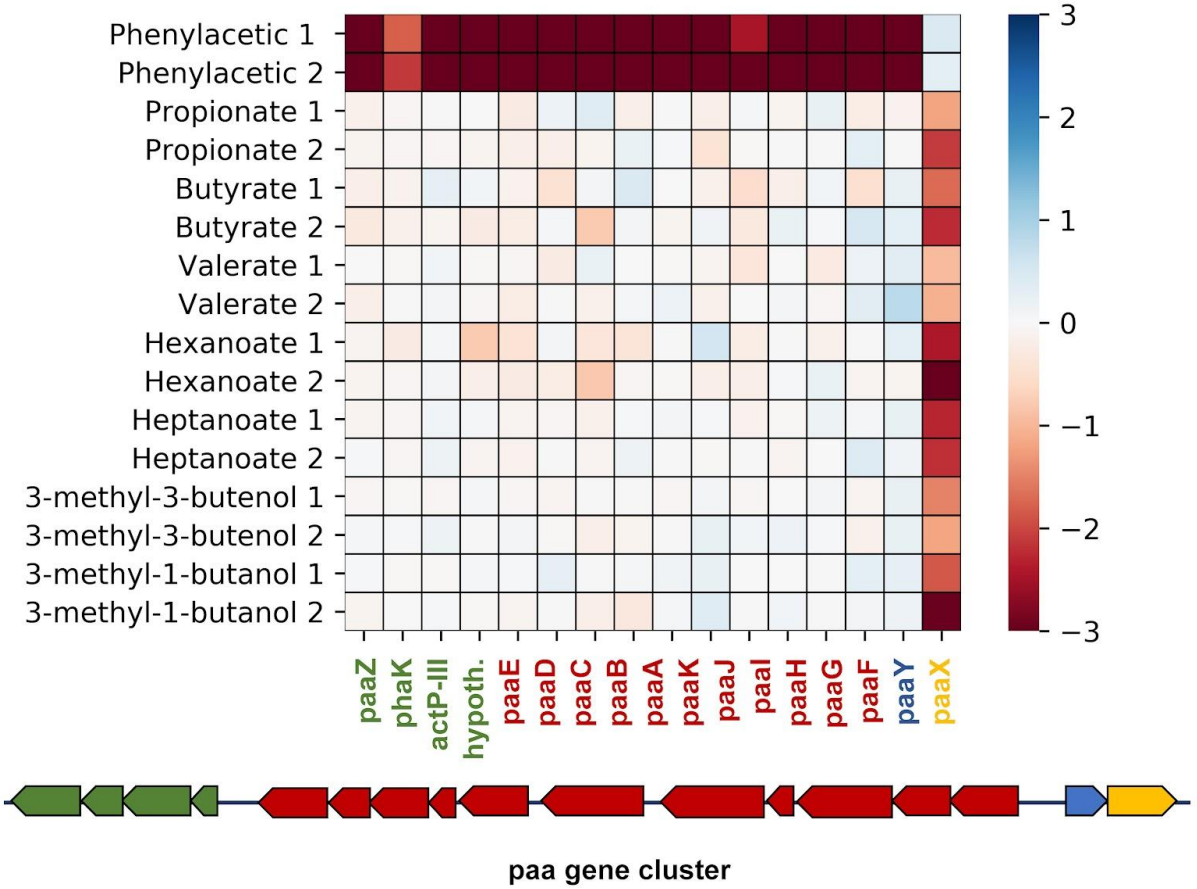


Figure 2.7. Detrimental fitness effects on short chain fatty acid catabolism by disrupting *paaX*.

Heatmap shows fitness values for the *paa* gene cluster when grown on short chain fatty acids, isoprenol, isopentanol, and phenylacetic acid.

Mutants in MerR-family regulator PP_3539 showed very high fitness benefits (fitness scores of 3.8 and 4.7 in two biological replicates) when grown on valerate. PP_3539 likely increases expression of *mvaB* (PP_3540 - hydroxymethyl-glutaryl-CoA lyase), thus suggesting that decreased levels of MvaB activity may benefit *P. putida* valerate catabolism. Unfortunately, there are no fitness data available for *mvaB*, likely because it is essential under the conditions in which the initial transposon library was constructed. The genes *hdb* and *bktB*, encoding the terminal two steps of butyrate metabolism, are flanked upstream by an AraC-family regulator (PP_3753) and downstream by a TetR-family regulator (PP_3756); the latter is likely co-transcribed with the butyrate catabolic genes (**Figure 2.6B**). When grown on butyrate, mutants in both PP_3753 and PP_3756 show decreased fitness; however, previous work to evaluate global fitness of *P. putida* grown on levulinic acid showed negative fitness values only for PP_3753, *htb*, and *btkB* (**Figure 2.6B**). These results suggest that the TetR repressor may be responding to a butyrate specific metabolite. Finally, across multiple fitness experiments, the TonB siderophore receptor PP_4994 and the TolQ siderophore transporter PP_1898 showed fitness advantages when grown on fatty acids, especially on hexanoate (**Figure 2.8**). Together, these results suggest that a wide range of environmental signals impact how *P. putida* is able to metabolize short chain fatty acids.

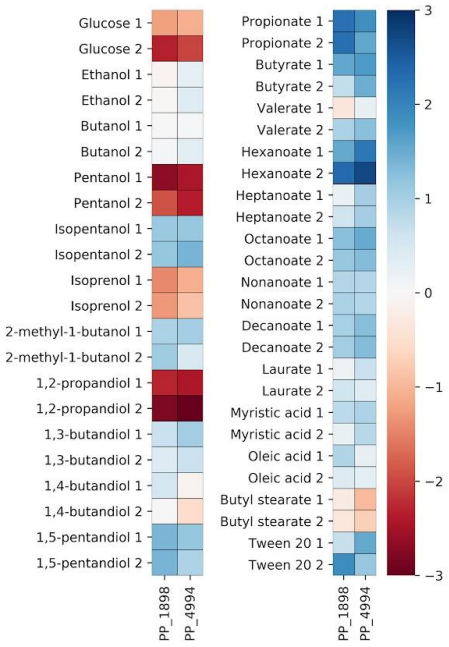


Figure 2.8: Fitness profiles of tonB siderophore transporter when grown on fatty acids or alcohols.

Heatmap shows fitness scores for PP_1898, and PP_4494 on all conditions tested in this work.

2.3.4. Global Analysis of Alcohol Catabolism

In addition to its ability to robustly catabolize a wide range of fatty acid substrates, *P. putida* is also capable of oxidizing and catabolizing a wide variety of alcohols into central metabolism through distinct pathways. To further our understanding of these pathways, transposon libraries were grown on a number of short n-alcohols (ethanol, butanol, and pentanol), diols (1,2-propanediol, 1,3-butanediol, 1,4-butanediol, and 1,5-pentanediol), and branched chain alcohols (isopentanol, isoprenol, and 2-methyl-1-butanol). Relative to growth on fatty acids, fitness experiments of *P. putida* grown on various alcohols showed less correlation to one another, reflecting the more diverse metabolic pathways used for their catabolism (**Figure 4A)**. The initial step of the catabolism of many primary alcohols is the oxidation of the alcohol to its corresponding carboxylic acid. The BioCyc database features 14 genes annotated as alcohol dehydrogenases (PP_1720, PP_1816, PP_2049, PP_2492, PP_2674, PP_2679, PP_2682, PP_2827, PP_2953, PP_2962, PP_2988, PP_3839, PP_4760, and PP_5210)[24]. Fitness data showed that the majority of these alcohol dehydrogenases had no fitness defects when grown on the alcohols used in this study (**Figure 2.9B**).

The alcohol dehydrogenases that showed the most consistent fitness defects across multiple conditions were the two PQQ-dependent alcohol dehydrogenases PP_2674 (*pedE*) and PP_2679 (*pedH*), as well as the Fe-dependent alcohol dehydrogenase PP_2682 (*yiaY*) (**Figure 2.9B**). Both *pedE* and *pedH* have been extensively studied in *P. putida* and other related bacteria, and are known to have broad substrate specificities for alcohols and aldehydes[25,26,49]. Their activity is dependent on the activity of *pedF* (PP_2675), a cytochrome *c* oxidase that regenerates the PQQ cofactor[25]. In *P. aeruginosa*, a homolog of *yiaY* (*ercA*) was shown to have a regulatory role in the expression of the *ped* cluster, and was not believed to play a direct catabolic role[50]. Recent work has validated that this function is conserved in *P. putida*[51]. In most conditions tested, disruption of *pedF* caused more severe fitness defects than either *pedE* or *pedH* individually, suggesting they can functionally complement one another in many cases. However, growth on 2-methyl-1-butanol and 1,5-pentanediol both showed more severe fitness defects in *pedE* mutants compared to *pedF* (**Figure 2.9B**). In many other alcohols, including ethanol and butanol, even disruption of *pedF* did not cause extreme fitness defects, suggesting the presence of other dehydrogenases able to catalyze the oxidation (**Figure 2.9B**).

The transcriptional regulatory systems that activate expression of various genes in the *ped* cluster could also be identified from these data. Mutants in either member of the sensory histidine kinase/response regulator (HK/RR) two component system, *pedS2*/*pedR2*, showed significant fitness defects when 2-methyl-1-butanol was supplied as the sole carbon source. This HK/RR signaling system has been shown to activate the transcription of *pedE* and repress *pedH* in the absence of lanthanide ions[52]. Since lanthanides were not supplied in the medium, this likely explains the fitness defect observed in *pedS2*/*pedR2*. The transcription factor *pedR1* (*agmR*) was also found to affect host fitness when grown on various alcohols (**Figure 2.10**). This gene has been identified

in *P. putida* U as an activator of long chain (C6+) n-alcohol and phenylethanol catabolism[53]. In *P. putida* KT2440, *pedR1* has been associated with the host response to chloramphenicol, and its regulon has been elucidated previously[54]. Our data reflect the literature, indicating that *pedR1* functions as a transcriptional activator of the *ped* cluster and *pedR2* functions as a specific regulator of *pedE* and *pedH*.

Unsurprisingly, the genes required for the biosynthesis of the PQQ cofactor were also amongst the most co-fit (cofitness is defined as Pearson correlation between fitness scores of two genes over many independent experimental conditions) with both *pedF* and *yiaY*. *P. putida* synthesizes PQQ via a well-characterized pathway, starting with a peptide encoded by the gene *pqqA* (PP_0380) which is then processed by *pqqE*, *pqqF*, and *pqqC* to generate the final cofactor (**Figure 2.9C**). The three synthetic genes (*pqqEFC*) all showed significant fitness defects on the same alcohols as the *pedF* mutants, while *pqqA* showed a less severe fitness phenotype (**Figure 2.9C**). However, the small size of *pqqA* resulted in few transposon insertions, making it difficult to draw confident conclusions. Two genes showed similar defective fitness patterns on select alcohols: *pqqB*, which has been proposed to be an oxidoreductase involved in PQQ biosynthesis; and *pqqD*, a putative PQQ carrier protein. Previous work regarding a PqqG homolog from *Methylorubrum extorquens* suggested that it forms a heterodimeric complex with PqqF that proteolytically processes PqqA peptides, although PqqF was sufficient to degrade PqqA on its own[55]. Fitness data from *P. putida* may support this hypothesis, as there was no observed fitness defect in *pqqG* mutants when grown on any alcohol, suggesting that the bacterium is still able to process PqqA with PqqF alone (**Figure 2.9C**).

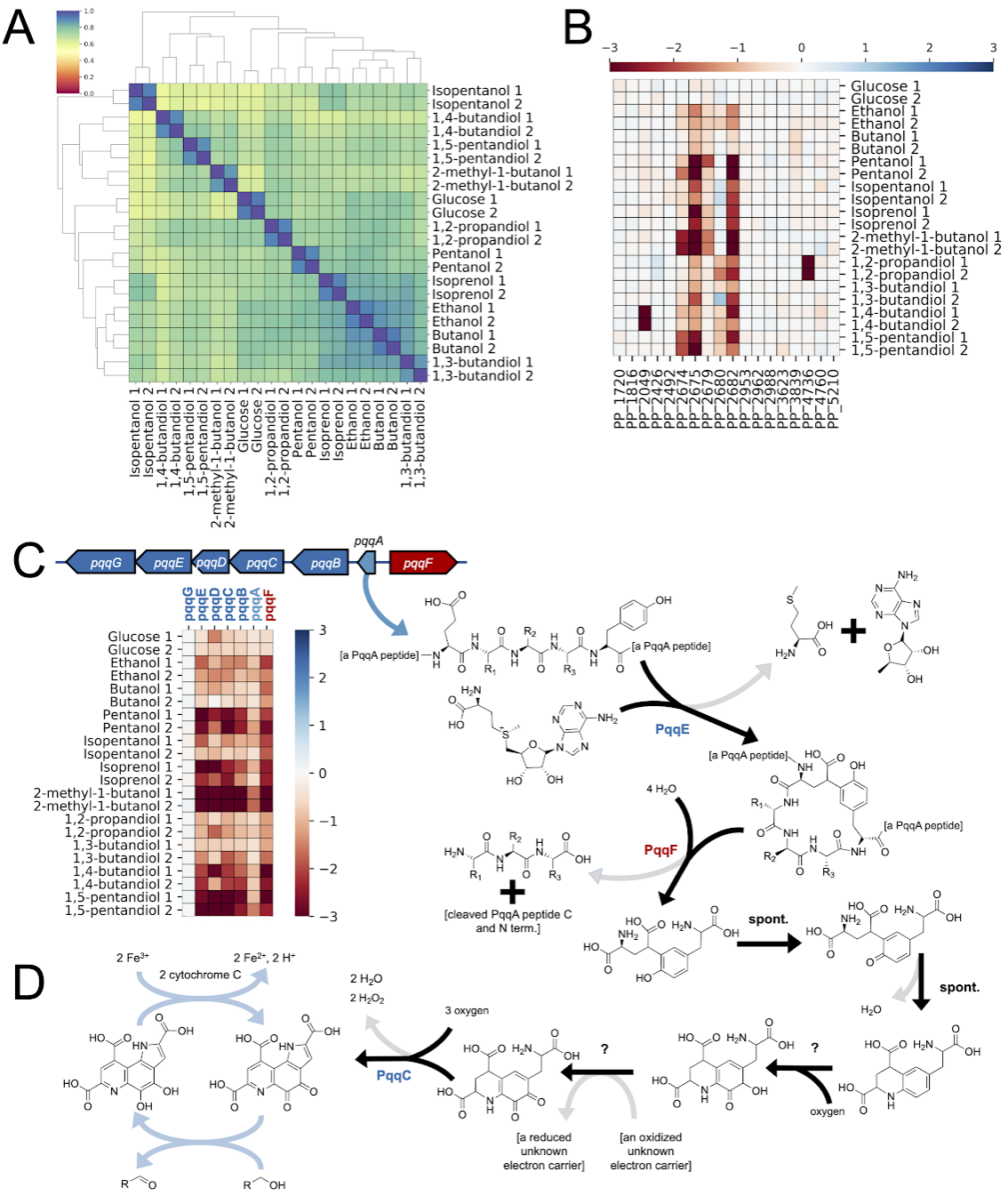


Figure 2.9. Global analysis of alcohol metabolism in *P. putida*.

A) Pairwise comparisons of Pearson correlations of fitness data from *P. putida* KT2440

RB-TnSeq libraries grown on alcohols as well as glucose grouped by overall similarity. Colors bar at top left shows the Pearson coefficient with 1 indicating greater similarity and 0 indicating greater dissimilarity. B) Heatmap shows the fitness scores of all alcohol dehydrogenases annotated on the BioCyc database as well as the cytochrome C PP_2675 when grown on various alcohols and glucose. C) Operonic diagram of the *pqq* cluster in *P. putida* and the corresponding biosynthetic pathway for the PQQ cofactor and D) How PQQ cofactors are regenerated by cytochrome C. Heatmap shows fitness scores for individual *pqq* cluster genes when grown on alcohols and glucose.

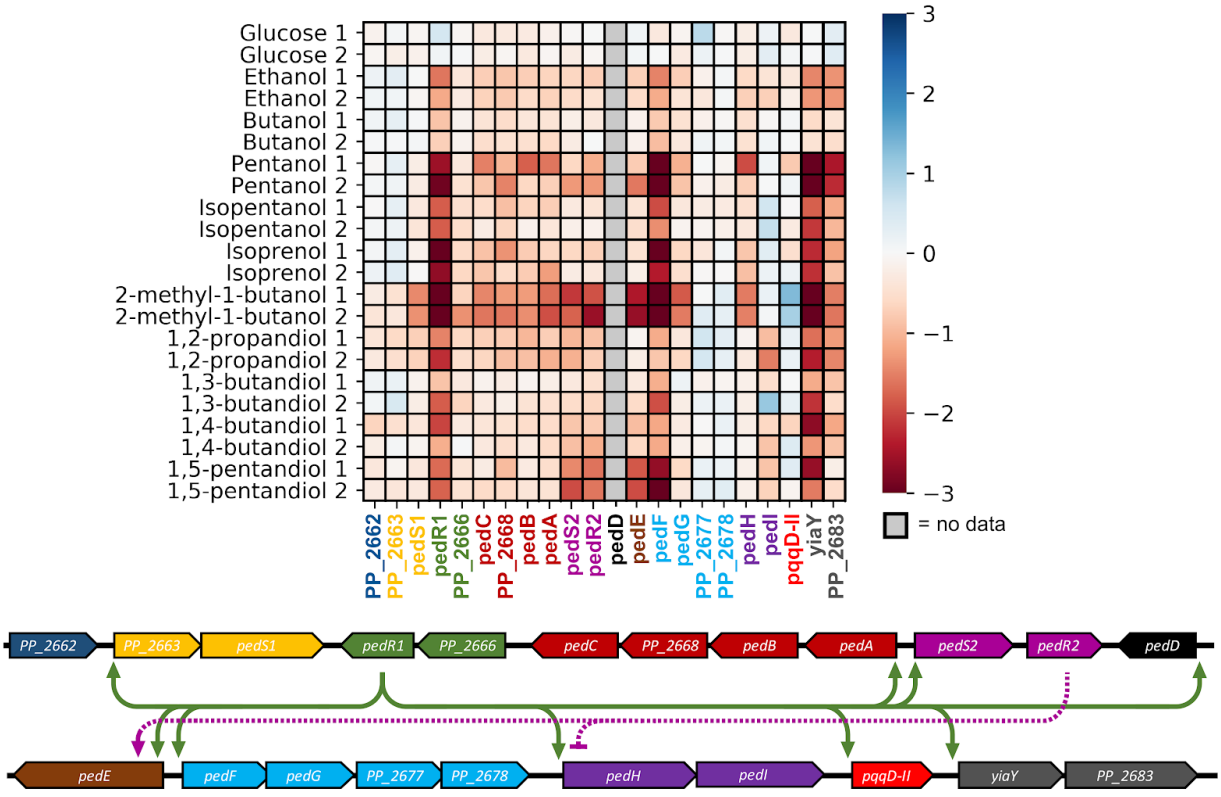


Figure 2.10: Essentiality and regulation of the *ped* cluster.

(Top) Heatmap depicting the fitness scores for genes in the *ped* cluster (PP_2662 to PP_2683) during growth on various short chain alcohols. (Bottom) Genomic context for the *ped* cluster in *P. putida* KT2440. Arrows depict transcriptionally upregulated genes of *pedR1* and *pedR2*. Blunt arrows point to genes predicted to be transcriptionally repressed in the condition tested.

2.3.5. Short chain alcohol metabolism

The metabolism of n-alcohols almost certainly proceeds through beta-oxidation using the same enzymatic complement as their fatty acid counterparts. This relationship is reflected in the high correlation in global fitness data between alcohols and fatty acids of the same chain length (ethanol and acetate - $r = 0.72$, butanol and butyrate - $r = 0.66$,

pentanol and valerate - $r = 0.72$). However, given previous work and our fitness data, the initial oxidation of these alcohols appears to be quite complex. Biochemical characterization of both PedE and PedH have shown that both have activity on ethanol, acetaldehyde, butanol, butyraldehyde, hexanol, and hexaldehyde[25]. When grown on n-pentanol, mutants disrupted in *pedF* show severe fitness defects, suggesting that PedH and PedE are the primary dehydrogenases responsible for pentanol oxidation (**Figure 2.9B**, **Figure 2.10**). However, when grown on either ethanol or n-butanol, both the PQQ-dependent alcohol dehydrogenases (PQQ-ADHs) and *pedF* show less severe fitness defects compared to when they are grown on pentanol (**Figure 2.10**). This implies that other dehydrogenases are also capable of these oxidations. One likely candidate may be PP_3839, which shows a minor fitness defect when grown on n-butanol and has been biochemically shown to oxidize coniferyl alcohol (**Figure 2.9B**)[56]. Individual gene deletion mutants of either *pedF* (PP_2675) or PP_3839 showed only minor growth defects when grown on either ethanol, butanol, or pentanol as a sole carbon source (**Figure 2.12**). However, when both genes were deleted, no growth was observed on these substrates, suggesting that the PQQ-ADHs and PP_3839 are the primary dehydrogenases responsible for the oxidation of short chain n-alcohols (**Figure 2.12**).

It is ambiguous from our data and from previous work which enzymes are oxidizing the aldehyde to the corresponding carboxylic acid. As mentioned previously, both PQQ-ADHs have been biochemically shown to act on aldehydes and could catalyze the reaction, but the lack of a strong fitness phenotype for both ethanol and n-butanol suggest they are not the only enzymes capable of catalyzing this reaction. The genomically proximal aldehyde dehydrogenase *pedI* (PP_2680) showed minor fitness defects when grown on ethanol and several other alcohols (**Figure 2.10**, **Figure 2.11A**), but showed no fitness defects when libraries were grown on butanol or pentanol. Another aldehyde dehydrogenase, *aldB-I* (PP_0545), showed virtually no fitness defects when grown on any of the short chain n-alcohols tested here (**Figure 2.11A**). The lack of any one obvious enzyme with a distinct fitness defect supports the notion that multiple enzymes are present and able to catalyze the oxidation of these aldehydes.

While the metabolism of alcohols and their corresponding fatty acids are similar, their fitness patterns showed distinct differences. When grown on acetate, mutants in *gacS* or *gacA* (PP_1650 and PP_4099 - a two-component (TCS) system), *sigS* (PP_1623 - the stationary phase sigma factor sigma S), and *ptsH* (PP_0948 - a component of the sugar phosphotransferase system (PTS)) showed large and significant fitness benefits, which were not apparent when grown on ethanol (**Figure 2.11B**). The GacS/GacA TCS is widespread across many gram-negative bacteria, and is believed to exert transcriptional control over a wide variety of functions, sometimes in concert with a small RNA binding protein (CsrA) that exerts post-transcriptional control[57]. In *Pseudomonads*, the GacA/GacS TCS has been implicated in positively controlling *sigS* expression in multiple species[58]. In *P. putida* specifically, *gacS* mutations in strains engineered to produce muconic acid have resulted in higher titers[59], but disruption of the gene was also shown to completely abolish production of medium-chain length polyhydroxyalkanoates (PHAs)[60]. Growth on butyrate also showed that *gacS*, *gacR*, *sigS*, and another component of the PTS (*ptsP*) had significant fitness benefits if disrupted, which was not observed when the library was grown on butanol (**Figure 2.11C**). Interestingly, mutants in *gacA* and *gacS* seemed to have fitness benefits when grown on either pentanol or valerate (**Figure 2.11D**). Further work is necessary to precisely characterize the nature of the benefits that occur when these genes are disrupted.

When grown on ethanol compared to acetate, relatively few genes not involved

in the oxidation of the short chain alcohols were found to be specifically and significantly unfit; however, specific phenotypes for acetate catabolism were observed (**Figure 2.11B**). Mutants in PP_1635 (a two-component system response regulator), PP_1695 (variously annotated as a sodium-solute symporter, sensory box histidine kinase, or response regulator), and *tal* (PP_2168 - transaldolase) all showed fitness defects on acetate that were not observed when libraries were grown on ethanol. The high cofitness between PP_1635 and PP_1695 observed across all publicly available fitness data ($r = 0.88$) and share homology to *crbSR* systems of other bacteria where it is known to regulate acetyl-coA synthetase[61].

Much like ethanol and acetate, there were relatively few genes that showed specific fitness defects when grown on butanol that were not also observed in butyrate. However, the genes *glgB* (PP_4058 - a 1,4-alpha-glucan branching enzyme), and the co-transcribed PP_2354 and PP_2356 (annotated as a histidine kinase/response regulator (HK/RR), and histidine kinase respectively) showed specific fitness defects when grown on butyrate relative to butanol. PaperBLAST analysis of PP_2356 and PP_2354 did not reveal any publications that had explored the function of this system, and thus further work will be needed to better characterize its regulon[62]. Mutants of genes encoding for three TCSs were found to be specifically unfit when grown on pentanol when compared to valerate. PP_2683 (a two component HK/RR), and *pedR1* (PP_2665 - RR) were both specifically unfit and, as previously described, are involved in the regulation of the *ped* cluster (**Figure 2.11D**). The gene *cbrB* (PP_4696 - sig54-dependent RR) also showed pentanol-specific defects, and is known to regulate central carbon metabolism and amino acid uptake in the *Pseudomonads*[63,64].

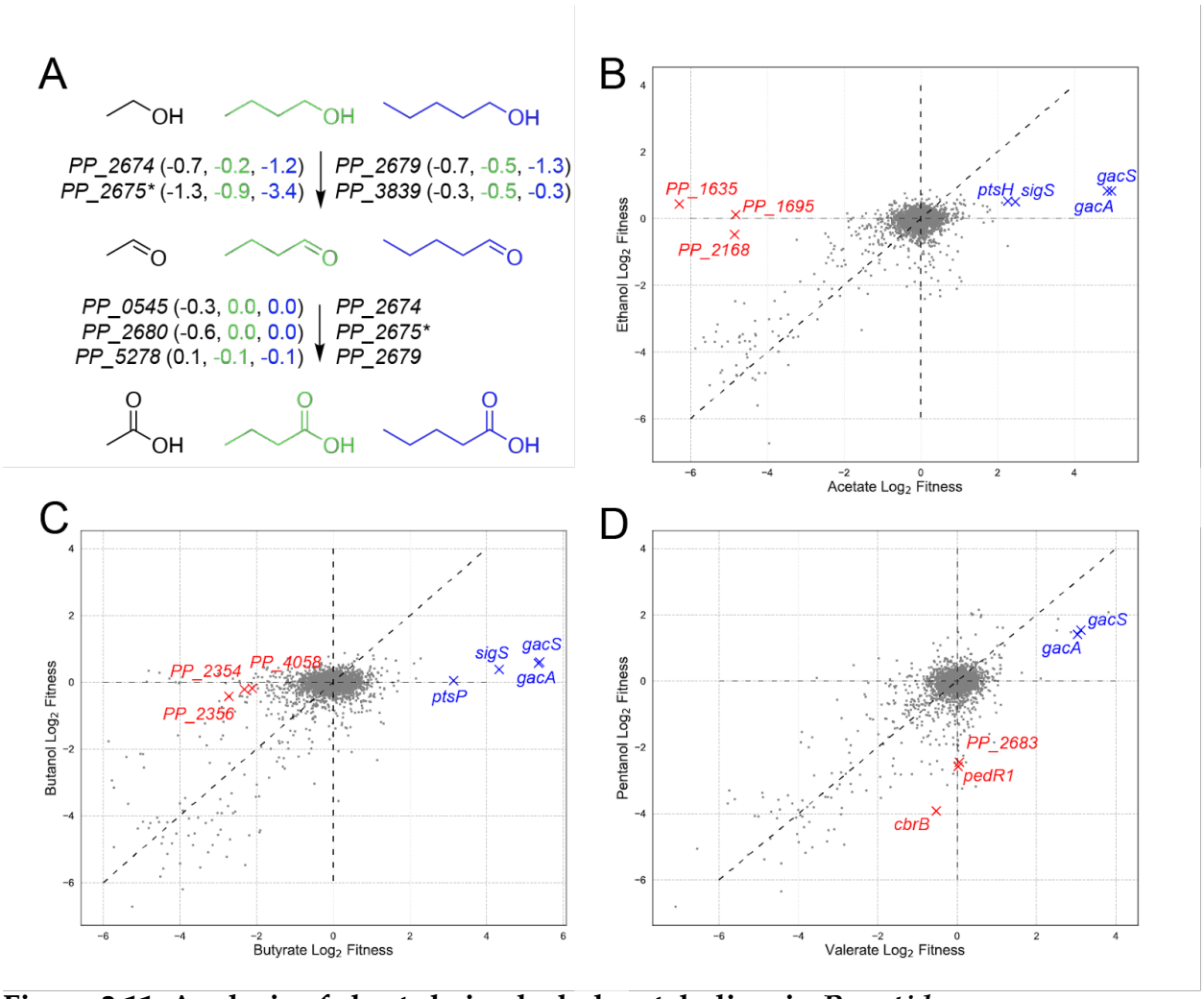


Figure 2.11. Analysis of short chain alcohol metabolism in *P. putida*

A) Putative genes involved in the initial oxidation steps of short chain alcohol assimilation in *P. putida*. PP_2675 (PedF) is involved in the regeneration of the PQQ cofactor predicted to be necessary for these oxidation reactions of PP_2674PP_2764 (PedE) and PP_2679PP_2769 (PedH). Average fitness scores for two biological reps are shown next to each gene for ethanol (black), butanol (green), and pentanol (blue). Scatter plots show global fitness scores for ethanol versus acetate (B), butanol versus butyrate (C), and pentanol versus valerate (D).

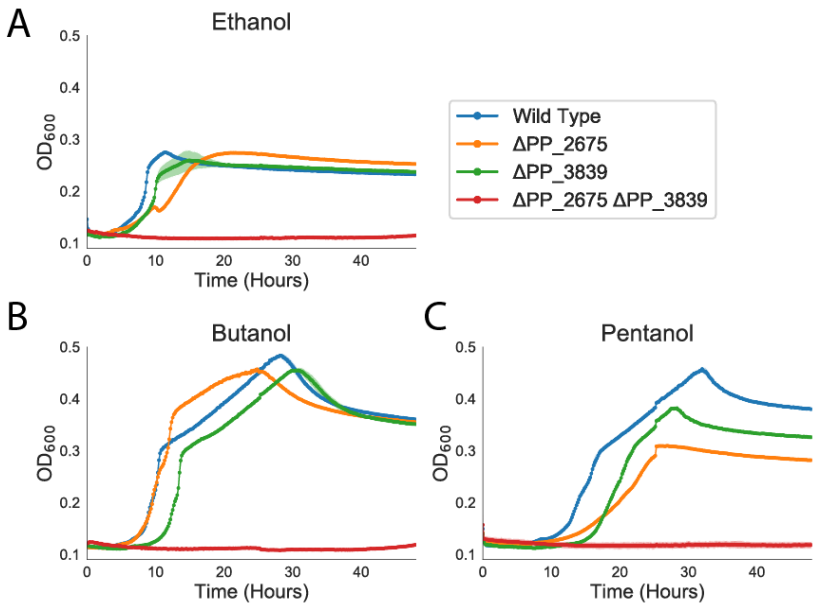


Figure 2.12. Validation of alcohol dehydrogenases involved in short chain alcohol metabolism

Growth curves of wild type (blue), ΔPP_2675 (orange), ΔPP_3839 (green), and ΔPP_2675ΔPP_3839 (red) strains of *P. putida* KT2440 on 10 mM ethanol (A), 10 mM n-butanol (B), and 10 mM n-pentanol (C). Shaded area represents 95% confidence intervals (cI), n=3.

2.3.6. Short chain diol catabolism

Another group of industrially relevant alcohols with potential for biotechnological production are short chain diols. These compounds have broad utility ranging from plasticizers to food additives[65]. As shown in **Figure 2.10**, most of the tested short chain diols result in significant fitness defects in *pedR1*, indicating that some of the genes involved in these metabolisms are in the PedR1 regulon. However, only 1,5-pentanediol had a strong fitness defect in *pedF*, indicating that multiple dehydrogenases may act on the shorter chain diols. Additionally, both 1,2-propanediol and 1,3-butanediol cause slight defects in mutants of the aldehyde dehydrogenase PP_0545. Although there is some ambiguity as to which enzymes initially oxidize the diols to their corresponding acids, the remaining steps in 1,2-propanediol, 1,3-butanediol, and 1,5-pentanediol catabolism are much more straightforward.

Oxidation of 1,2-propanediol yields lactate, and mutants in the L-lactate permease PP_4735 (*lldP*) have a fitness of -4.3 when grown on 1,2-propanediol. Furthermore, under this condition, mutants of the L- and D- lactate dehydrogenases PP_4736 (*lldD*) and PP_4737 (*lldE*) have fitness defects of -5.0 and -1.5, respectively. Since we provided a rac-1,2-propanediol as a substrate, this likely explains the fitness defects observed in both dehydrogenases[66,67]. Given these results, it appears that 1,2-propanediol is assimilated into central metabolism via oxidation to pyruvate (**Figure 2.13**).

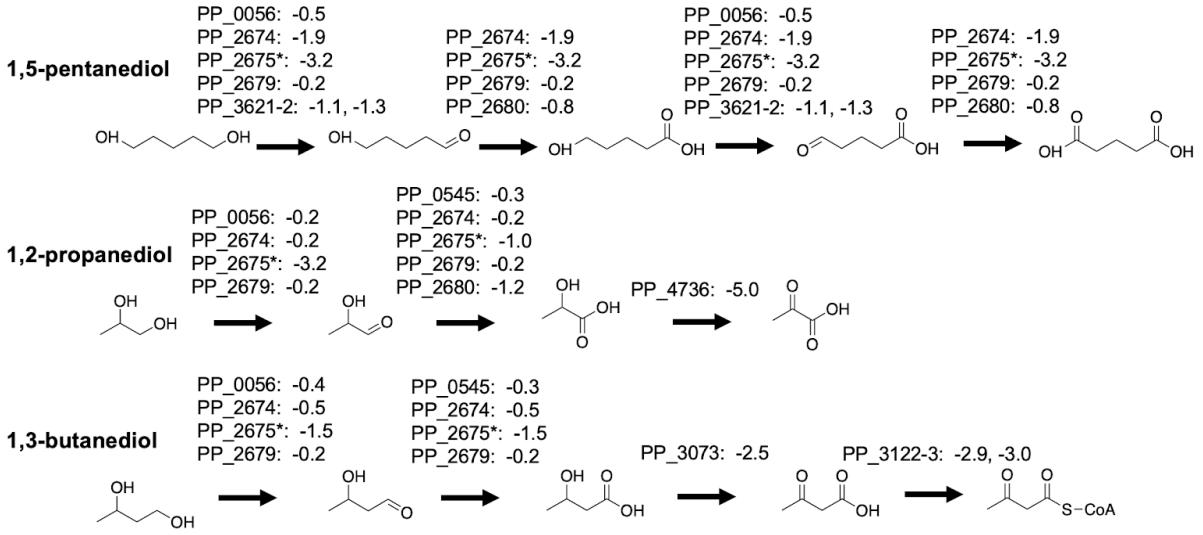


Figure 2.13. Putative catabolic pathways for 1,5-pentanediol, 1,2-propanediol, and 1,3-butanediol.

Fitness scores for two biological replicates of genes proposed to code for responsible enzymes can be found next to genes.

When grown on 1,3-butanediol, two oxidations of 1,3-butanediol result in 3-hydroxybutyrate, and we observe fitness defects of -2.5 in the D-3-hydroxybutyrate dehydrogenase PP_3073 and -1.8 in the neighboring sigma-54 dependent regulator PP_3075[68]. Dehydrogenation of 3-hydroxybutyrate results in acetoacetate, and we see a fitness defect of -2.9 and -3.0 for the subunits of the predicted 3-oxoacyl-CoA transferase PP_3122-3 (*atoAB*). This enzyme likely transfers a CoA from either succinyl-CoA or acetyl-CoA in order to generate acetoacetyl-CoA. Regarding transport, mutants in the D-beta-hydroxybutyrate permease PP_3074, located in the same operon as the 3-hydroxybutyrate dehydrogenase, have a fitness defect of -0.9, while mutants in the RarD permease PP_3776 have a fitness of -1.2.

Following oxidation by the aforementioned PQQ-dependent dehydrogenases and aldehyde dehydrogenases in the periplasm, an oxidized intermediate is likely transported into the cell for the next steps in the catabolism. This is supported by the observation that mutants of the predicted dicarboxylate MFS transporter PP_1400 and its two-component regulator PP_1401-2 have strong fitness defects on both alpha-ketoglutarate and 1,5-pentanediol. Furthermore, there is a -4.7 fitness defect in mutants of the L-2-hydroxyglutarate oxidase PP_2910, which catalyzes the second step in the glutarate hydroxylation pathway of glutarate catabolism. The glutarate hydroxylase PP_2909, which catalyzes the first step of this pathway, has a much slighter negative fitness of -0.6. This is expected, because glutarate can also be catabolized through a glutaryl-CoA dehydrogenation pathway, so mutants in PP_2909 can simply divert flux through the other catabolic route[12]. Mutants in PP_2910 are unable to oxidize L-2-hydroxyglutarate to alpha-ketoglutarate, and likely accumulate L-2-hydroxyglutarate as a dead-end metabolite.

1,4-butanediol catabolism has been previously studied; based on the results of expression data and adaptive laboratory evolution, Li et al. proposed three potential catabolic pathways for 1,4-butanediol, including a beta-oxidation pathway (**Figure 2.14**)[30]. Their evolved strains had mutations in the LysR activator PP_2046 that resulted in overexpression of the beta-oxidation operon PP_2047-51[30]. Interestingly, we found that when grown on 1,4 butanediol, transposon mutants of the acyl-CoA dehydrogenase PP_2048 had significant fitness benefits and no CoA-ligase mutants showed significant fitness defects. However, a fitness defect of -1.0 in PP_0356 (malate synthase) mutants suggests that there may be flux through the beta-oxidation pathway to glycolic acid and acetyl-CoA. A possible explanation for the positive fitness of PP_2048 mutants is that the beta-oxidation pathway is suboptimal in the wild type, and it may be beneficial to divert flux through the other pathway(s). This same reasoning could also explain the absence of CoA-ligases with fitness defects; however, this also could be due to the presence of multiple CoA-ligases capable of catalyzing that step. Mutants of the 3-hydroxyacyl-CoA dehydrogenase PP_2047, a *fadB* homolog which likely catalyzes the hydration and dehydrogenation steps to produce 3-oxo-4-hydroxybutyryl-CoA, had a strong fitness defect. When PP_2047 is non-functional, 4-hydroxycrotonyl-CoA likely accumulates as a deadend metabolite resulting in decreased fitness. Li et al. also showed that deletion mutants of PP_2046 are unable to grow on 1,4-butanediol[30]. Our data suggests that this is because PP_2049 appears to be the main alcohol dehydrogenase acting on either 1,4-butanediol or 4-hydroxybutyrate, and is in the operon under the control of PP_2046. Although our fitness data suggests that both the oxidation to succinate and beta-oxidation pathways occur, further work is necessary to determine if the pathway to succinyl-CoA is involved in the catabolism.

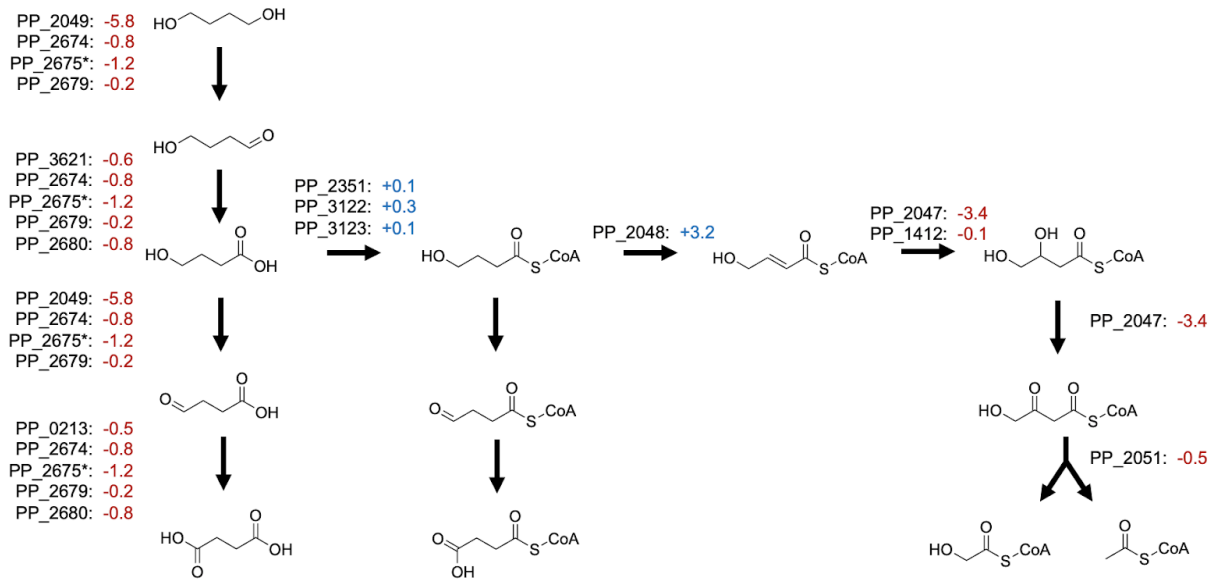


Figure 2.14. Putative routes of 1,4-butanediol catabolism in *P. putida*.

Putative genes involved in catabolism of 1,4-butanediol in *P. putida*. Average fitness scores for two biological reps are shown next to each gene. The three CoA-ligases shown were proposed by Li et al.; there were no CoA-ligases that showed significant fitness defects on 1,4-butanediol. *PP_2675 (PedF) is involved in the regeneration of the PQQ cofactor predicted to be necessary for these oxidation reactions of PP_2674PP_2764 (PedE) and PP_2679PP_2769 (PedH).

2.3.7. Branched chain alcohol metabolism

Due to their superior biofuel properties, branched chain alcohols have been targets for metabolic engineering as potential alternatives to ethanol and butanol[69]. Our fitness data suggest that *pedE* and/or *pedH* oxidize 2-methyl-1-butanol to 2-methylbutyrate, which then undergoes one round of beta-oxidation to produce acetyl-CoA and propionyl-CoA (**Figure 2.15**). Most of the genes involved in 2-methylbutyrate beta-oxidation are located in the operon PP_2213-PP_2217. With mutants having a fitness defect of -3.2, PP_2213 appears to be the main acyl-CoA ligase acting on 2-methylbutyrate. Mutants in two predicted acyl-CoA dehydrogenases, PP_2216 and PP_0358, show fitness defects of -1.1 and -2.6, respectively. The enoyl-CoA hydratase PP_2217 has a fitness defect of -5.7 and the 3-hydroxyacyl-CoA dehydrogenase PP_2214 has a fitness defect of -5.6. Finally, the acetyl-CoA acetyltransferase appears to be PP_2215, with mutants having a fitness defect of -4.8. We also observed fitness defects of -1.8 and -1.6 in mutants of the ABC transporters, PP_5538 and PP_2667, respectively. Since 2-methylbutyrate is a known intermediate in the catabolism of isoleucine, we found that the genetic data presented here closely mirror the previous biochemical characterization of this system[70,71].

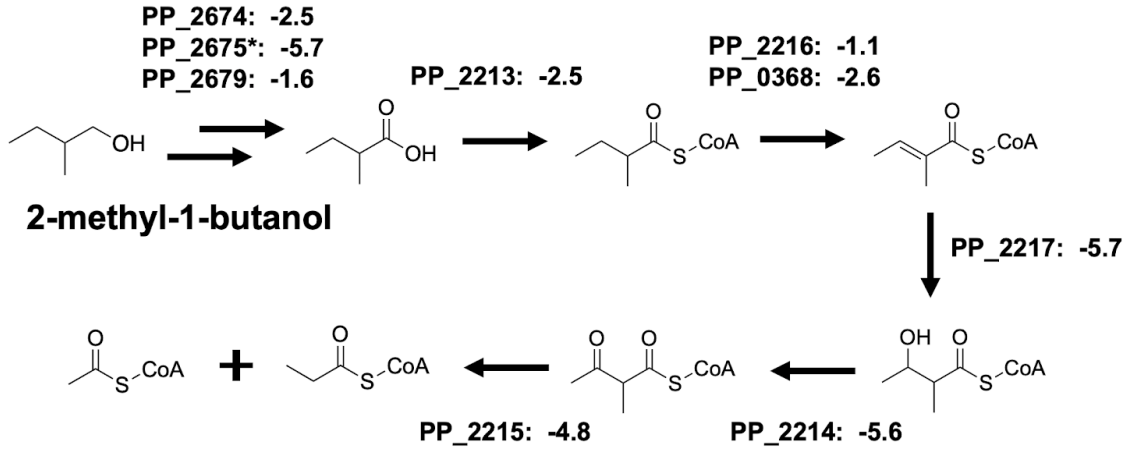


Figure 2.15. Putative catabolic pathway for 2-methyl-1-butanol

Fitness scores for two biological replicates of genes proposed to code for responsible enzymes can be found next to genes.

P. putida can readily grow on isopentanol and isoprenol but not prenol (**Figure 2.16A**). All three of these alcohols have been produced in high titers in *Escherichia coli* and other bacteria because of their potential to be suitable replacements for gasoline[72,73]. RB-TnSeq data for isopentanol and isoprenol showed severe fitness defects in genes of the leucine catabolic pathway (**Figure 2.17**). This is unsurprising, as isopentanol can be generated from the leucine biosynthetic pathway[74]. Deletion of the PP_4064-PP_4067 operon, which contains the genes that code for the conversion of isovaleryl-CoA to 3-

hydroxy-3-methylglutaryl-CoA, abolished growth on both isopentanol and isoprenol (**Figure 2.18**). Deletion of PP_3122 (acetoacetyl CoA-transferase subunit A) also abolished growth on isopentanol, and greatly reduced growth on isoprenol (**Figure 2.18**). Taken together, these results validate that both of these alcohols are degraded via the leucine catabolic pathway. Transposon insertion mutants in *pedF* showed strong fitness defects on both isopentanol and isoprenol, suggesting that *pedH* (PP_2679) and *pedE* catalyze (PP_2674) the oxidation of the alcohols. Deletion mutants in *pedH* showed only a minor delay in growth compared to wild-type when grown on either isopentanol or isoprenol, while mutants in *pedE* showed a more substantial growth defect on both alcohols (**Figure 2.16A**). Deletion of *pedF* (PP_2675) prevented growth on both isopentanol and nearly abolished growth on isoprenol when provided as a sole carbon source in minimal media (**Figure 2.16A**). When wild-type *P. putida* was grown in minimal media with 10 mM glucose and 4 mM of either isopentanol, prenol, or isoprenol, each alcohol was shown to be readily degraded with concurrent observation of increasing levels of the resultant acid (**Figure 2.16B**). Though *P. putida* was unable to utilize prenol as a sole carbon source, it was still able to readily oxidize prenol to 3-methyl-2-butenoic acid, suggesting there is no CoA-ligase present in the cell able to activate this substrate and channel it towards leucine catabolism (**Figure 2.17**). When wild-type *P. putida* was grown in LB medium supplemented with 4 mM of each alcohol individually, all alcohols were completely degraded by 24 hours post-inoculation (**Figure 2.16C**). In *pedF* deletion mutants grown under the same conditions, the rate at which the alcohols were degraded was significantly slowed; however after 48 hours ~50% of isopentanol, ~75% of isoprenol, and 100% of prenol were degraded (**Figure 2.16C**). Uninoculated controls showed that no alcohol was lost at greater than 5% on account of evaporation (data not shown). Future efforts to produce any of these alcohols in *P. putida* will be heavily impacted by this degradation, and greater effort will need to be made to identify other enzymes involved in the oxidation of these alcohols or other metabolic pathways that consume them.

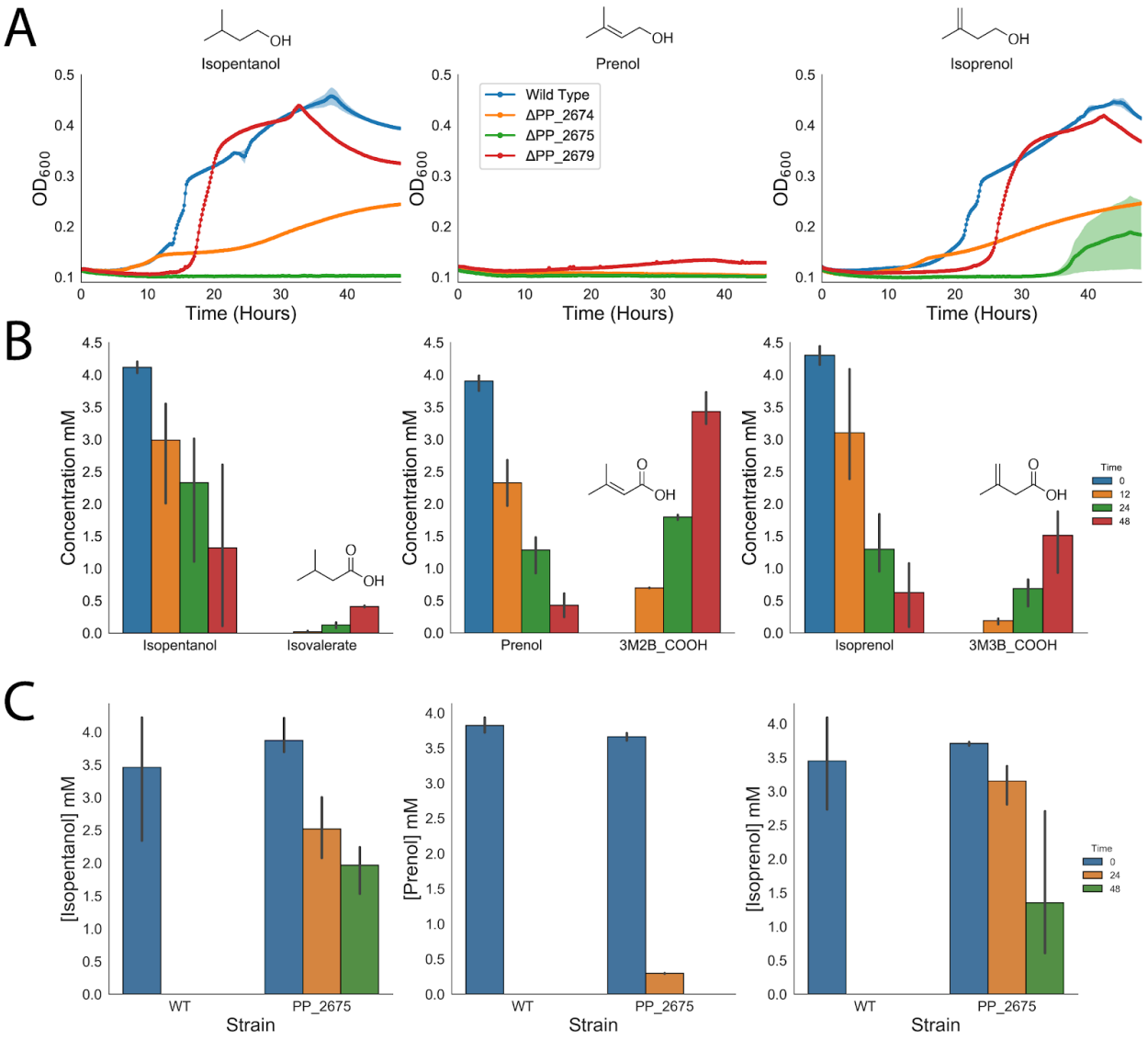


Figure 2.16. Isopentanol, Prenol, and Isoprenol consumption by *P. putida*.

A) Growth curves of wild type (blue), and ΔPP_2674 (orange), ΔPP_2675 (green), and ΔPP_2679 (red) strains of *P. putida* on isopentanol (left), prenol (middle), and isoprenol (right). Structure of alcohols are shown above graphs. Shaded area represents 95% confidence intervals (cI), n=3. B) Concentrations of alcohols consumed and their corresponding carboxylic acids produced over time by wild type. Left panel shows isopentanol and isovalerate, middle panel shows prenol and 3-methyl-2-butenoic acid, and the right panel shows isoprenol and 3-methyl-3-butenoic acid. Structures of corresponding carboxylic acids derived from alcohol are shown in graphs. Error bars represent 95% cI, n=3. C) Consumption of isopentanol (left), prenol (middle), and isoprenol (right) by wild type and ΔPP_2675 *P. putida* over time. Error bars represent 95% cI, n=3.

One mystery that remains is how isoprenol enters into leucine catabolism. GC-MS analysis confirmed oxidation of the alcohol to 3-methyl-3-butenoic acid, but it is

unclear what the next step entails. Fitness data suggests that either PP_4063 or PP_4549 may attach the CoA to isovalerate, but neither of these genes have strong phenotypes when mutant libraries are grown on isoprenol (**Figure 2.17**). That PP_4064 (isovaleryl-CoA dehydrogenase) shows strong negative fitness values when libraries are grown on isoprenol implies that its degradation goes through an isovaleryl-CoA intermediate, however this fitness defect may be the result of polar effects that disrupt the downstream steps (**Figure 2.17**). One possibility is that 3-methyl-3-butenoic acid is reduced to isovalerate in the cell; however, this seems unlikely since no isovalerate was observed via GC-MS when *P. putida* was fed isoprenol and glucose. Two other possible routes could result from the activation of 3-methyl-3-butenoic acid by an undetermined CoA-ligase. If this CoA-ligase exists, it is interesting that it would have activity on 3-methyl-3-butenoic acid but not 3-methyl-2-butenoic acid, which accumules when *P. putida* is grown in the presence of prenol. Once formed, the 3-methyl-3-butenyl-CoA could be directed into leucine catabolism via either an isomerization to 3-methylcrotonyl-CoA or a reduction to isovaleryl-CoA. Future work that leverages metabolomics to identify compounds that accumulate in leucine catabolic mutants may reveal the missing steps and help narrow the search for their enzymes.

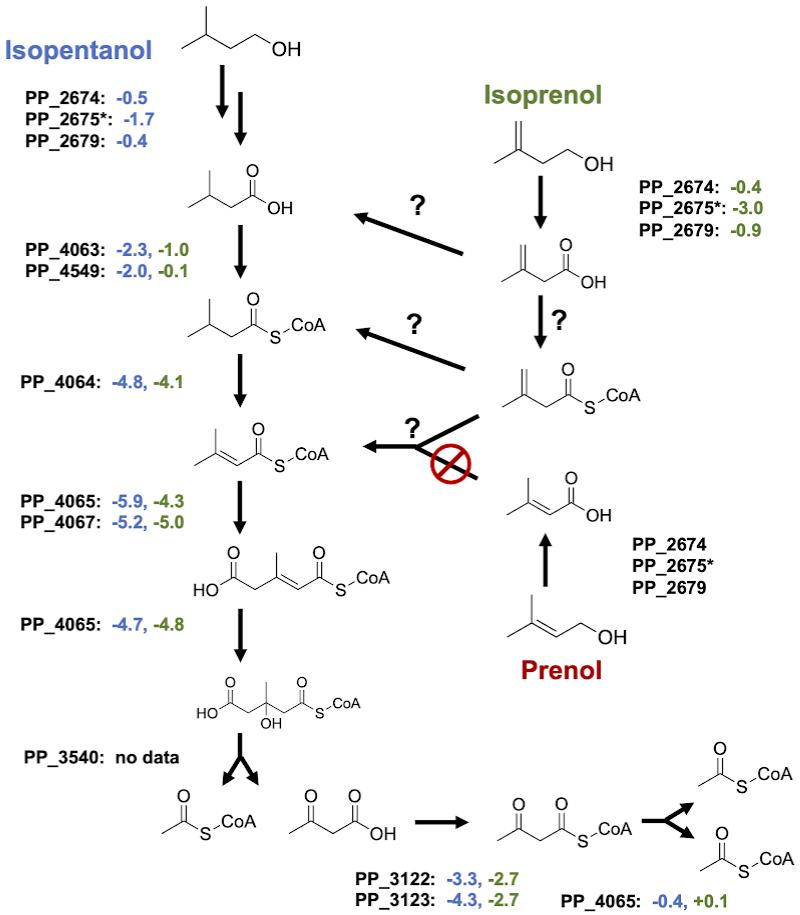


Figure 2.17. Putative routes of isopentanol and isoprenol catabolism in P. putida.

Diagram shows the proposed pathways for the catabolism of isopentanol and isoprenol. Average fitness scores of two biological replicates for individual genes can be found next to each gene. Fitness values for isopentanol are shown in blue, while fitness values for isoprenol and shown in green. Potential reactions that would bring isoprenol into leucine catabolism are marked with a question mark.

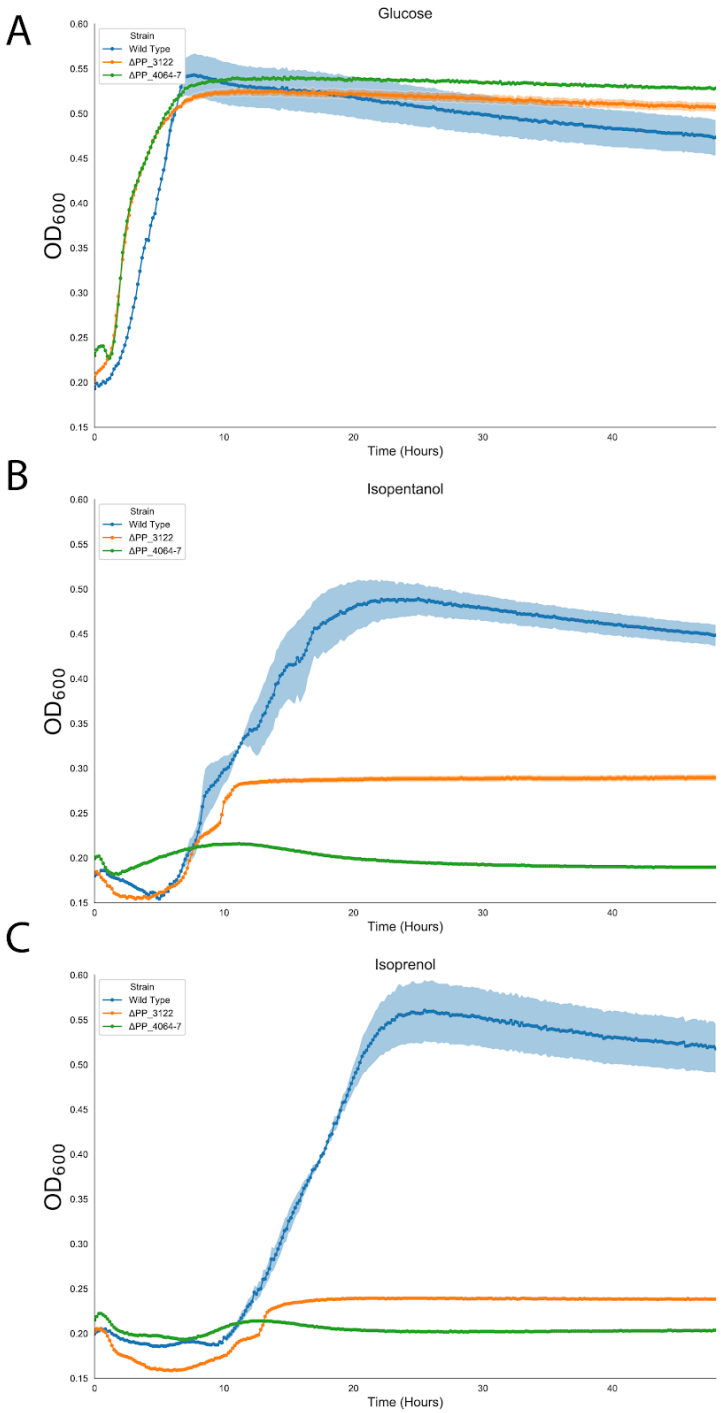


Figure 2.18. Growth of leucine catabolism deletion mutants on branched chain alcohols.

Growth curves of wild-type (blue), ΔPP_3122 (orange) , and ΔPP_4064-4067 (green) strains of *P. putida* on glucose (A), isopentanol (B), and isoprenol (C). Shaded area represents 95% cI, n=4.

2.4. Future Directions

The large set of global fitness data generated in this study provide an extensive and global overview on the putative pathways of alcohol and fatty acid degradation in *P. putida*. Overall, our fitness data agree with previously published biochemical data that explored enzymes in both fatty acid and alcohol metabolism. However, there are still many questions that our data leave unanswered. Further investigation will be required to untangle and elucidate which specific enzymes are biologically relevant in the beta-oxidation of short chain fatty acids. It is likely that biochemical characterization of individual enzymes will be required to determine which of the *fad* homologs catalyze these reactions. Another intriguing question is the function of PP_0765 and PP_0766. Biochemical interrogation and mutational analysis of the DUF1302 and DUF1329 family proteins are needed to determine whether these proteins indeed function as an esterase or, as previously predicted, play some other role in outer membrane biogenesis[41]. Additional work is also warranted to ascertain which of the proposed 1,4-butanediol catabolic routes the wild-type organism actually uses and determine whether the beta-oxidation pathway is indeed less preferable than the pathway to succinate.

To our knowledge, our finding that *P. putida* can consume both isopentanol and isoprenol are the first observations of this metabolism. If metabolic engineers wish to produce these chemicals in *P. putida*, these pathways will need to be removed. Critically, researchers will need to identify other enzymes that result in the oxidation of these alcohols or other routes of degradation within *P. putida*. How *P. putida* is able to utilize isoprenol, but not prenol, as a sole carbon source is metabolically intriguing. One of our proposed pathways of isoprenol catabolism requires the existence of a CoA-ligase that shows surprising specificity towards 3-methyl-3-butenoic acid with little to no activity on 3-methyl-2-butenoic acid. More work should be done to leverage other omics-levels techniques to try to identify this hypothetical enzyme and biochemically verify its activity. Finally, this data set as a whole will likely strengthen the assumptions made by genome-scale metabolic models. Previous models of *P. putida* metabolism have incorporated RB-TnSeq data to improve their predictions[17]. This work nearly doubles the number of available RB-TnSeq datasets in *P. putida* that are publicly available and will likely contribute greatly to further model refinement. Ultimately, large strides in our understanding of *P. putida* metabolism leveraging functional genomic approaches will provide the foundation for improved metabolic engineering efforts in the future.

2.5. Methods

2.5.1. Media, chemicals, and culture conditions

General *E. coli* cultures were grown in lysogeny broth (LB) Miller medium (BD Biosciences, USA) at 37 °C while *P. putida* was grown at 30 °C. When indicated, *P. putida* and *E. coli* were grown on modified MOPS minimal medium, which is comprised of 32.5 μM $CaCl_2$, 0.29 mM K_2SO_4, 1.32 mM K_2HPO_4, 8 μM $FeCl_2$, 40 mM MOPS, 4 mM tricine, 0.01 mM $FeSO_4$, 9.52 mM NH_4Cl, 0.52 mM $MgCl_2$, 50 mM NaCl, 0.03 μM $(NH_4)_6Mo_7O_{24}$, 4 μM H_3BO_3, 0.3 μM $CoCl_2$, 0.1 μM $CuSO_4$, 0.8 μM $MnCl_2$, and 0.1 μM $ZnSO_4$[75]. Cultures were supplemented with kanamycin (50 mg/L, Sigma Aldrich, USA), gentamicin (30 mg/L, Fisher Scientific, USA), or carbenicillin (100mg/L, Sigma Aldrich, USA), when indicated. All other compounds were purchased through Sigma Aldrich (Sigma Aldrich, USA). 3-methyl-3-butenoic acid was not available commercially and

required synthesis which is described below.

2.5.2. Strains and plasmids

All bacterial strains and plasmids used in this work are listed in Table 1. All strains and plasmids created in this work are available through the public instance of the JBEI registry. (public-registry.jbei.org/folders/456). All plasmids were designed using Device Editor and Vector Editor software, while all primers used for the construction of plasmids were designed using j5 software[76–78]. Plasmids were assembled via Gibson Assembly using standard protocols[79], or Golden Gate Assembly using standard protocols[80]. Plasmids were routinely isolated using the Qiaprep Spin Miniprep kit (Qiagen, USA), and all primers were purchased from Integrated DNA Technologies (IDT, Coralville, IA). Construction of *P. putida* deletion mutants was performed as described previously[18].

Table 2.1. Strains and plasmids used in this study.

Strain	**Description**	**Reference**
E. coli XL1 Blue		Agilent
P. putida KT2440	Wild-Type	ATCC 47054
P. putida ΔPP_2674	Strain with complete internal in-frame deletion of PP_2674	This study
P. putida ΔPP_2675	Strain with complete internal in-frame deletion of PP_2675	This study
P. putida ΔPP_2679	Strain with complete internal in-frame deletion of PP_2679	This study
P. putida ΔPP_3839	Strain with complete internal in-frame deletion of PP_3839	This study

P. putida ΔPP_2675 ΔPP_3839	A double knockout of PP_2675 and PP_3839	This study
P. putida ΔPP_4064-PP_4067	Strain with complete internal in-frame deletion of the PP_4064-4067 operon	This study
P. putida ΔPP_3122	Strain with complete internal in-frame deletion of PP_3122	This study
Plasmids		
pMQ30	Suicide vector for allelic replace Gm^r, SacB	[81]
pMQ30 ΔPP_2674	pMQ30 derivative harboring 1kb flanking regions of PP_2674	This study
pMQ30 ΔPP_2675	pMQ30 derivative harboring 1kb flanking regions of PP_2675	This study
pMQ30 ΔPP_2679	pMQ30 derivative harboring 1kb flanking regions of PP_2679	This study
pMQ30 ΔPP_3839	pMQ30 derivative harboring 1kb flanking regions of PP_3839	This study

pMQ30 ΔPP_4064-PP_4067	pMQ30 derivative harboring 1kb flanking regions of PP_4064 and PP_4067	This study
pMQ30 ΔPP_3122	pMQ30 derivative harboring 1kb flanking regions of PP_3122	This study

2.5.3. Plate-based growth assays

Growth studies of bacterial strains were conducted using microplate reader kinetic assays as described previously[82]. Overnight cultures were inoculated into 10 mL of LB medium from single colonies, and grown at 30 °C. These cultures were then washed twice with MOPS minimal media without any added carbon and diluted 1:100 into 500 μL of MOPS medium with 10 mM of a carbon source in 48-well plates (Falcon, 353072). Plates were sealed with a gas-permeable microplate adhesive film (VWR, USA), and then optical density and fluorescence were monitored for 48 hours in an Biotek Synergy 4 plate reader (BioTek, USA) at 30 °C with fast continuous shaking. Optical density was measured at 600 nm.

2.5.4. RB-TnSeq

RB-TnSeq experiments utilized *P. putida* library JBEI-1 which has been described previously with slight modification[18]. Libraries of JBEI-1 were thawed on ice, diluted into 25 mL of LB medium with kanamycin and then grown to an OD_{600} of 0.5 at 30 °C at which point three 1-mL aliquots were removed, pelleted, and stored at -80 °C. Libraries were then washed once in MOPS minimal medium with no carbon source, and then diluted 1:50 in MOPS minimal medium with 10 mM of each carbon source tested. Cells were grown in 10 mL of medium in test tubes at 30 °C shaking at 200 rpm. One 500-μL aliquot was pelleted, and stored at -80 °C until BarSeq analysis, which was performed as previously described[19,40]. The fitness of a strain is defined here as the normalized $\log_2$ ratio of barcode reads in the experimental sample to barcode reads in the time zero sample. The fitness of a gene is defined here as the weighted average of the strain fitness for insertions in the central 10% to 90% of the gene. The gene fitness values are normalized such that the typical gene has a fitness of zero. The primary statistic *t* value represents the form of fitness divided by the estimated variance across different mutants of the same gene. Statistic *t* values of $>|4|$ were considered significant. A more detailed explanation of calculating fitness scores can be found in Wetmore et al.[40]. All experiments described here passed testing using the quality metrics described previously unless noted otherwise[40]. All experiments were conducted in biological duplicate, and all fitness data are publically available at http://fit.genomics.lbl.gov.

2.5.5. GC-MS and GC-FID Analysis of Branched Alcohol Consumption

To examine the oxidation of isopentanol, prenol, and isoprenol to their corresponding acids 10mL of MOPS minimal medium supplemented with 10 mM glucose and 4mM of one of each alcohol added were inoculated with a 1:100 dilution of overnight *P. putida* culture and incubated at 30 °C with 200 rpm shaking. At 0, 12, 24, and 48-hours post-inoculation 200 µL of media were sampled and stored at - 80 °C. Alcohols and fatty acids were extracted by acidifying media with 10 µL of 10N HCl, followed by addition of an 200 µL of ethyl-acetate. To detect alcohols and their corresponding carboxylic acids via GC-MS an Agilent 6890 system equipped with a DB-5ms column (30- m×0.25 mm×0.25 µm) and an Agilent 5973 MS were used. Helium (constant flow 1 mL/min) was used as the carrier gas. The temperature of the injector was 250 °C and the following temperature program was applied: 40 °C for 2 min, increase of 10 °C/min to 100 °C then increase of 35 °C/min to 300 °C , temperature was then held at 300 °C for 1 min. Authentic standards were used to quantify analytes. Determination of isopentanol, prenol, and isoprenol consumption was conducted in 10mL LB medium with 4mM of either alcohol added. Cultures were inoculated with a 1:100 dilution of overnight *P. putida* culture and incubated at 30 °C with 200 rpm shaking. At 0, 24, and 48 hours post-inoculation 200 µL of media were sampled and stored at - 80 °C. The remaining concentration of each alcohol was determined by GC-FID as previously described[83].

2.5.6. Synthesis of 3-Methyl-3-Butenoic Acid

To a 25-mL round bottom flask was added chromium(VI) oxide (0.69 g, 6.9 mmol) and distilled water (1 mL). The reaction mixture was then cooled to 0 °C before concentrated sulfuric acid (0.6 mL, 10.5 mmol) was added dropwise, thus forming Jones reagent. The solution of Jones reagent was then diluted to a total volume of 5 mL with distilled water. To a stirred solution of 3-methyl-3-buten-1-ol (0.59 g, 6.9 mmol) in acetone (7 ml) was added dropwise the Jones reagent at 0 °C. After being stirred for 8 h at room temperature, the mixture was quenched with ethanol. The mixture was then diluted with water, and acetone was evaporated *in vacuo*. The residue was extracted with DCM, and organic layers were combined and washed three times with saturated *aq.* $NaHCO_3$ solution. The aqueous phase was acidified with a 2 M *aq.* HCl solution to pH 2-3, which was then extracted again with DCM. The extract was successively washed with water and brine, dried over $MgSO_4$, and concentrated *in vacuo*. The residue was distilled (90 °C, 100 mTorr) to yield 3-methyl-3-butenoic acid as a clear oil. ^{1}H NMR (300 MHz, Chloroform-*d*) δ 4.92 (d, *J* = 19.1 Hz, 2H), 3.08 (s, 2H), 1.84 (s, 3H) (**Figure 2.19**).

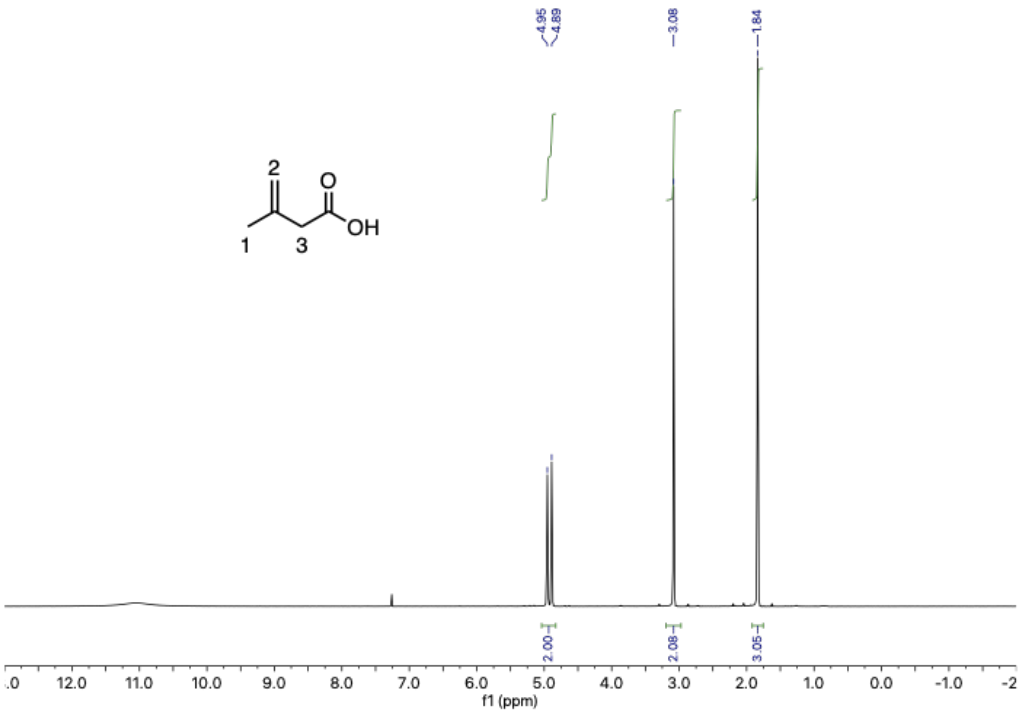


Figure 2.19. NMR validation of 3-methyl-3-butenoic acid.

2.5.7. Bioinformatic Analyses

PaperBLAST was routinely used to search for literature on proteins of interest and related homologs[62]. All statistical analyses were carried out using either the Python Scipy or Numpy libraries[84,85]. For the phylogenetic reconstructions, the best amino acid substitution model was selected using ModelFinder as implemented on IQ-tree [86] phylogenetic trees were constructed using IQ-tree, nodes were supported with 10,000 bootstrap replicates. The final tree figures were edited using FigTree v1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/). Orthologous syntenic regions were identified with CORASON-BGC[87] and manually colored and annotated.

2.6. Acknowledgements

We would like to thank Morgan Price for assistance in analyzing RB-TnSeq data. The laboratory of LMB is partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy within the Cluster of Excellence FSC 2186 'The Fuel Science Center'. This work was part of the DOE Joint BioEnergy Institute (https://www.jbei.org) supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, supported by the U.S. Department of Energy, Energy Efficiency and Renewable Energy, Bioenergy Technologies Office, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy. The views and opinions of the authors expressed herein do not necessarily

2.7. Contributions

Conceptualization, M.G.T., M.R.I., A.N.P.; Methodology, M.G.T., M.R.I., A.N.P., J.M.B, P.C.M., A.M.D.; Investigation, M.G.T., M.R.I., A.N.P, M.S., W.A.S., C.B.E., P.C.M., J.M.B., Y.L., R.W.H., C.A.A, R.N.K, P.L.; Writing – Original Draft, M.G.T., M.R.I., A.N.P.; Writing – Review and Editing, All authors.; Resources and supervision, L.M.B., A.M., A.M.D., P.M.S, J.D.K. M.G.T., M.R.I., and A.N.P. contributed equally to this work. Author order was determined by the outcome of a MarioKart 64 tournament.

2.8. Competing Interests

J.D.K. has financial interests in Amyris, Lygos, Demetrix, Napigen, Maple Bio, and Apertor Labs. C.B.E has a financial interest in Perlumi Chemicals.

Chapter 3. *Pseudomonas putida*, Nitrogen metabolism extraordinaire

Including material from published work:
Matthias Schmidt, Allison N. Pearson, **Matthew R. Incha**, Mitchell G. Thompson, Edward E. K. Baidoo, Ramu Kakumanu, Aindrila Mukhopadhyay, Patrick M. Shih, Adam M. Deutschbauer, Lars M. Blank, Jay D. Keasling, "Nitrogen metabolism in *Pseudomonas putida*: functional analysis using random barcode transposon sequencing", AEM, 2022

3.0. Chapter preface

After identifying fatty acid and alcohol metabolisms, we delved into the complex and diverse nitrogen metabolism of Pseudomonas putida KT2440. We investigated assimilatory pathways for 52 nitrogen sources using a similar approach as our previous project. In this chapter, we demonstrate the application of t-SNE for visualizing large datasets, examine the global regulatory impacts of GltBD and NtrC on these pathways, and discuss data from specific metabolisms involving nitrate, nitrite, quaternary amines, and ethanolamine.

3.1. Abstract

Pseudomonas putida KT2440 has long been studied for its diverse and robust metabolisms, yet many genes and proteins imparting these growth capacities remain uncharacterized. Using pooled mutant fitness assays, we identified genes and proteins involved in the assimilation of 52 different nitrogen containing compounds. To assay amino acid biosynthesis, 19 amino acid drop-out conditions were also tested. From these 71 conditions, significant fitness phenotypes were elicited in 672 different genes including 100 transcriptional regulators and 112 transport-related proteins. We divide these conditions into 6 classes and propose assimilatory pathways for the compounds based on this wealth of genetic data. To complement these data, we characterize the substrate range of three promiscuous aminotransferases relevant to metabolic engineering efforts *in vitro*. Furthermore, we examine the specificity of five transcriptional regulators, explaining some fitness data results and exploring their potential to be developed into useful synthetic biology tools. In addition, we use manifold learning to create an interactive visualization tool for interpreting our BarSeq data, which will improve the accessibility and utility of this work to other researchers.

3.2. Importance

Understanding the genetic basis of *P. putida's* diverse metabolism is imperative for us to reach its full potential as a host for metabolic engineering. Many target molecules of the bioeconomy and their precursors contain nitrogen. This study provides functional evidence linking hundreds of genes to their roles in the metabolism of nitrogenous compounds, and provides an interactive tool for visualizing these data. We further characterize several aminotransferases, lactamases, and regulators--which are of particular interest for metabolic engineering.

3.3. Introduction

As a free-living soil bacterium, *P. putida* encounters many different organic and inorganic nitrogen sources, and its responses to these conditions have been the target of recent study. During rhizosphere colonization, *P. putida* displays chemotaxis towards plant root exudates, which contain nitrogenous compounds such as benzylamines, polyamines, pyrrole derivatives, nucleotide derivatives, amino acids, and phenylpropanoids [88,89]. It also demonstrates chemotaxis towards and degradation of the phytotoxic and insecticidal benzoxazinoids exuded from the roots of maize seedlings [90,91]. Not only is *P. putida* able to withstand and metabolize these varied compounds in a nitrogen-rich rhizosphere, it is also capable of adapting to nitrogen-scarce conditions by increasing polyhydroxyalkanoate (PHA) production, repressing carbon catabolism, and increasing expression of transporters for nitrogen uptake [92].

Given its ability to adapt to the varied environment of the rhizosphere and its burgeoning role as a host for sustainable chemical bioproduction, it is not surprising that *P. putida*'s nitrogen metabolism has also been examined in the context of metabolic engineering. Due to the utility of PHAs as a next-generation bioplastic, the transcriptomic, proteomic, and metabolomic response of *P. putida* strains to nitrogen-limited growth conditions has been analyzed in order to better understand how PHA synthesis is triggered by nitrogen scarcity [93,94]. *P. putida*'s nitrogen metabolism and its regulation have also been a source of metabolic engineering parts. Some of its 39 predicted aminotransferases have been heterologously expressed as part of benzylamine derivative and glutaric acid production pathways [95,96]. The transcription factor regulating capro- and valerolactam degradation has been developed into a highly sensitive biosensor with the potential to be applied to increasing lactam production titers and used as an inducible system in pathway engineering [82]. Moreover, many nitrogenous compounds are relevant building blocks for commodity chemicals, and understanding *P. putida*'s metabolism of them can enable more flux to be directed towards the desired product [97].

Despite its relevance to both metabolic engineering and basic scientific research, our understanding of *P. putida*'s nitrogen metabolism is far from complete. Many gene functions have been assigned through homology predictions with limited functional evidence, which can hamper the accuracy of metabolic modeling [17]. Furthermore, *P. putida* has multiple paralogs of many of its enzymes, each potentially with different substrate preferences. Functional genomics can provide evidence for assigning gene functions and illuminate the specific roles of genes that have multiple paralogs [98].

In previous studies, we have employed barcoded transposon sequencing (BarSeq) to interrogate lysine, fatty acid, alcohol, and aromatic degradation in *P. putida* [11,18,98]. In this work, we use BarSeq to study *P. putida* KT2440's metabolism of 52 different nitrogen-containing compounds, nearly doubling the amount of publicly available BarSeq data for this bacterium. We provide evidence for many known nitrogen assimilation pathways and their regulatory systems, and also assign function to genes whose exact roles in nitrogen metabolism were not previously known. Due to their relevance in recent metabolic engineering efforts, we further examine the substrate specificity of *P. putida*'s 5-oxoprolinases and aminotransferases [10,95,96]. To understand the regulation of these enzymes *in vivo*, we further characterized their cognate regulators.

This work will bolster *P. putida*'s utility both as a host microorganism and as a source of metabolic engineering parts for sustainable chemical production.

3.4. Results and Discussion

3.4.1. BarSeq reveals the genetic bases of diverse nitrogen metabolisms

Genes involved in nitrogen utilization from natural and unnatural compounds were identified using BarSeq. In these assays, a library of barcoded transposon insertion mutants was cultured in minimal media with glucose and a variety of sole nitrogen sources (Figure 3.1). Wild-type growth on these substrates over 96 hours is shown in Figure 3.2. These assays revealed 672 genes with strong ($|\text{fitness}| > 1$) and significant ($|t| > 5$) fitness phenotypes. These include 100 transcription factors, 112 transport proteins, and numerous other enzymes with applications to engineered systems, such as aminotransferases [99,100] (Figure 3.3A, Figure 3.4). Additionally, 529 of these genes encode proteins that are currently unreviewed in the Uniprot database, and 256 have not shown significant phenotypes in previous BarSeq studies. To visualize the fitness data, the manifold learning method, t-distributed stochastic neighbor embedding (t-SNE), was employed to cluster genes based on their fitness values in the tested conditions (Figure 3.3B, Figure 3.4) [101]. The clusters were named based on the condition that elicited the largest and most frequent changes in fitness scores for genes within the cluster. In this visualization, we can identify genes that may take part in similar metabolisms. An interactive version that contains embedded hyperlinks to the Fitness Browser [41] can be accessed via this link: https://ppnitrogentsne.lbl.gov. Although genes with specific phenotypes to one condition are easy to identify, genes essential across many conditions are more challenging to assign a specific function. This is particularly problematic for genes that are essential in the majority of the tested conditions, like those involved in amino acid biosynthesis. To our surprise, clusters of genes involved in tryptophan, arginine, methionine, and branched chain amino acid biosynthesis were resolved by t-SNE (Figure 3.I.1-2). While not as exhaustive as pairwise comparisons of conditions, t-SNE provides a useful visualization of this diverse dataset.

Figure 3.I.1-2. ppnitrogentsne.lbl.gov

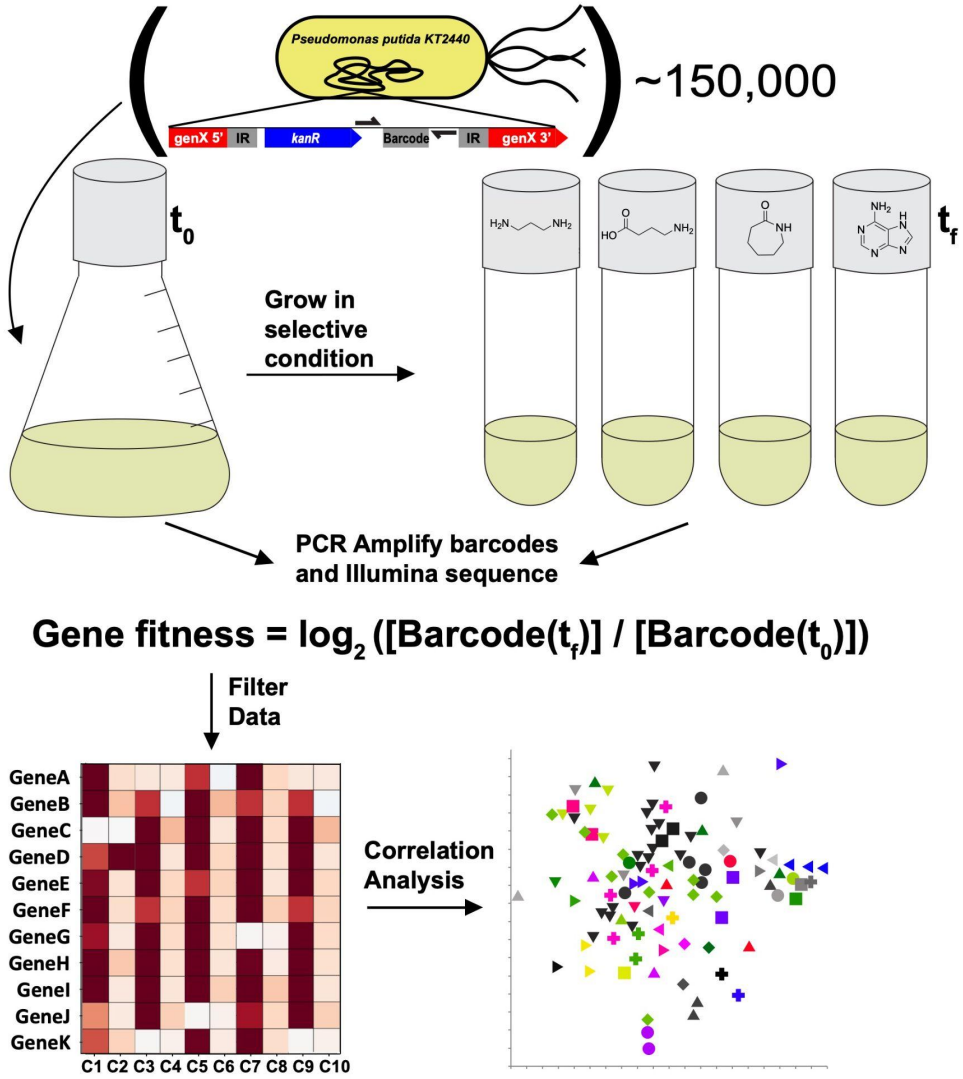


Figure 3.1. Overview of the barcode abundance sequencing (BarSeq) and data processing workflow.

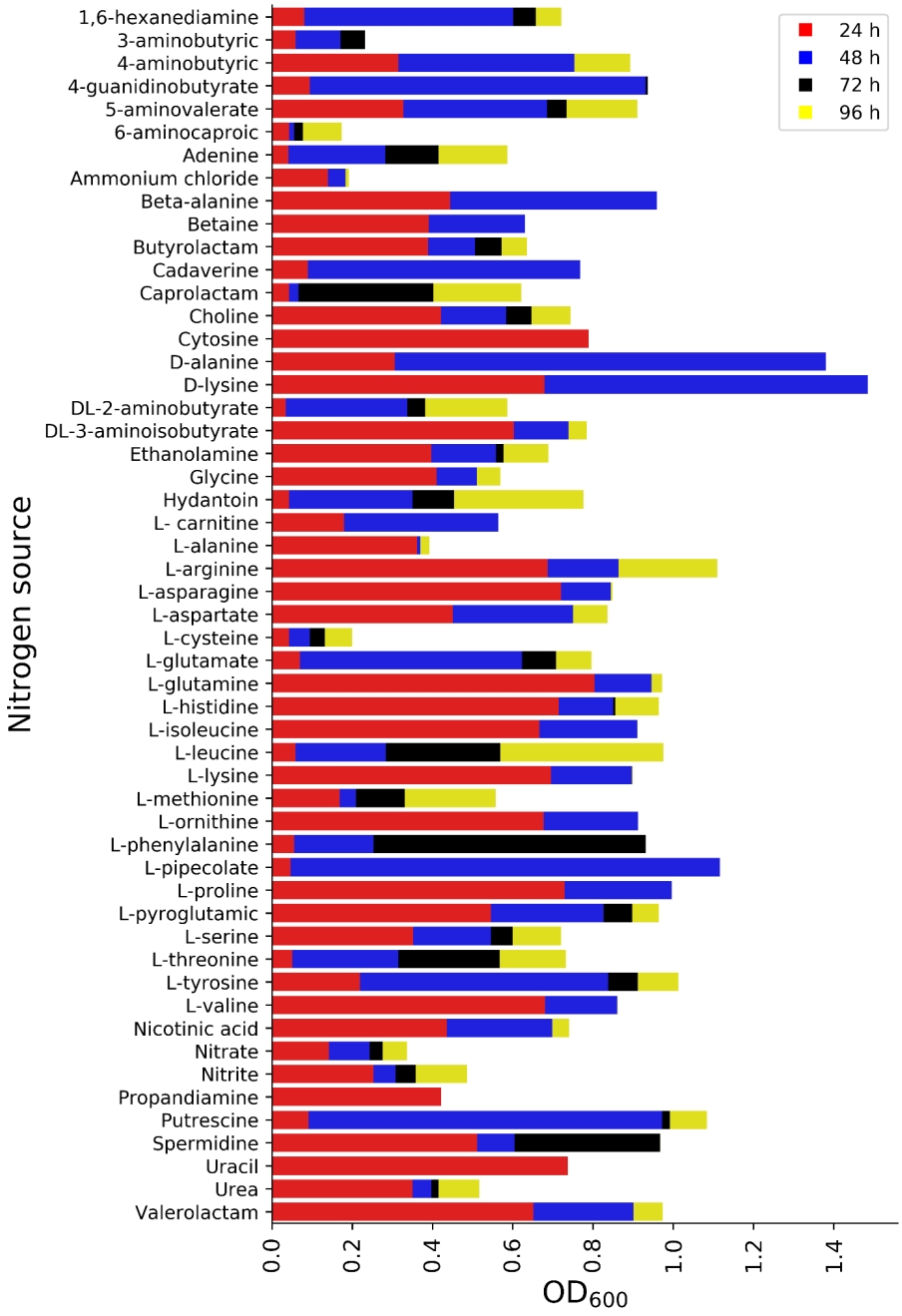


Figure 3.2. Plate-based growth assay showing the maximal OD for all the tested nitrogen sources.

Cells were grown as triplicates in a 24-deep well plate and OD was taken after 24, 48, 72 and 96 hours.

Figure 3.3. Global analysis of the *P. putida* KT2440 BarSeq data.

A) Significant genes ($|\text{fitness}| > 1$ and $|t| > 5$) from all 71 tested conditions sorted by their Cluster of Orthologous Groups (COGs) based on the eggNOG database [102–104]. 'Multiple COGs' indicates that there was more than one COG assigned. B) Image of the interactive t-SNE visualization showing the legend, t-SNE clustering (left) and cluster centroids (right). By clicking on a substrate in the legend, the corresponding cluster (left) and centroid (right) is highlighted, opening a list of cluster members and additional information. By clicking the highlighted centroid (right), the user is redirected to the Fitness Browser (https://fit.genomics.lbl.gov) [41], where the fitness data for all significant genes in the condition cluster is shown. More information about the

interactive figures can be found in the description of Figure 3.4. The Interactive figure is available at https://ppnitrogentsne.lbl.gov

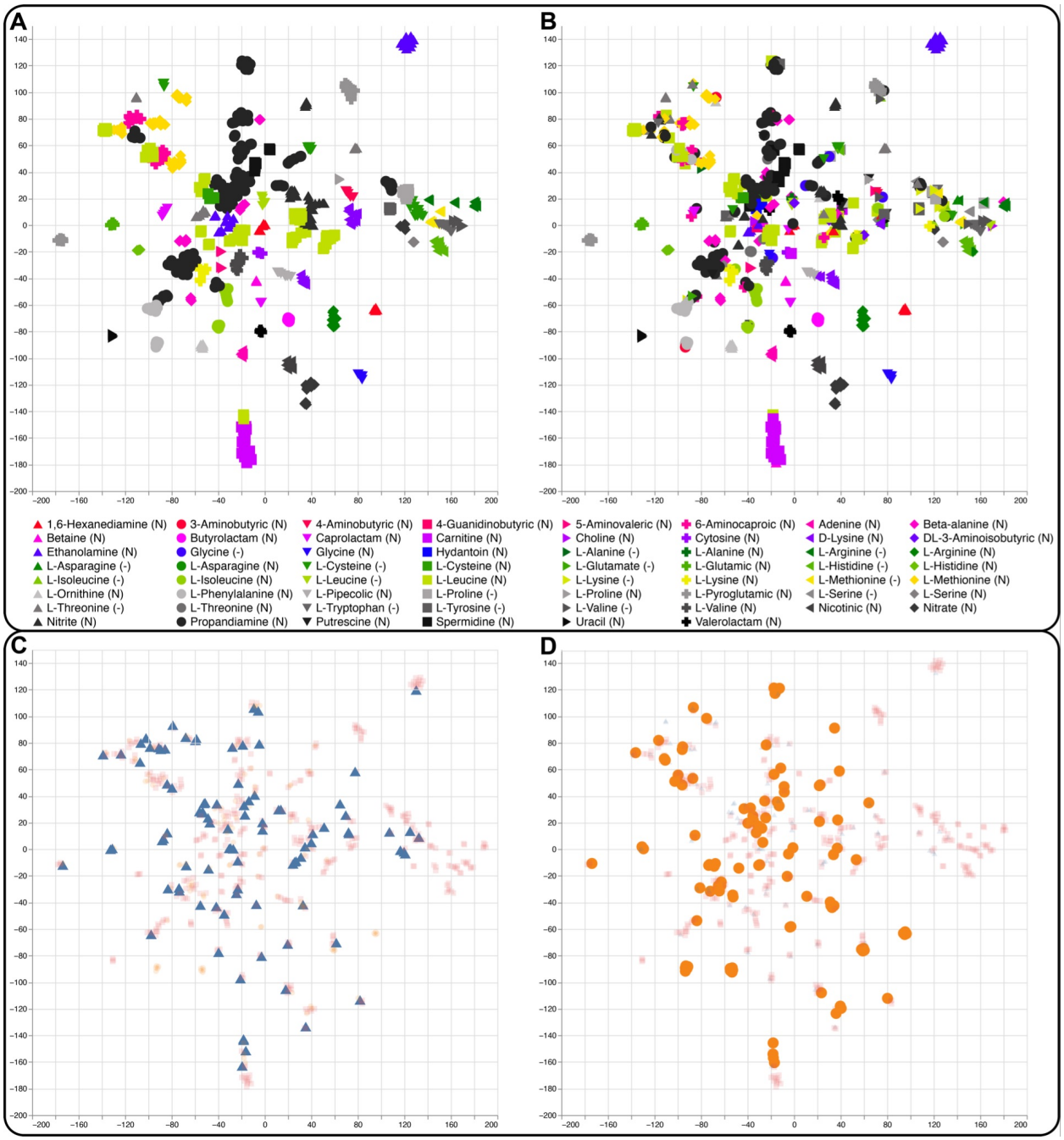


Figure 3.4. Additional t-SNE visualizations

A) t-SNE visualization of the nitrogen source and amino acid dropout BarSeq assays. Individual points represent genes with significant fitness changes in the dataset. The x and y axes are non-dimensional and distances between points within clusters are representative of the Euclidean distance between the genes in the dataset. Clusters were

defined using Ward clustering of the t-SNE results. B) Depiction of t-SNE results without Ward cluster definitions. C) Depiction of cluster locations for transcriptional regulators (blue triangles identified from MiSTdb 3.0. D) Depiction of cluster locations for genes predicted to be involved in transport reactions (orange circles) from TransportDB 2.0. Interactive versions of these plots are available in the supplementary materials (Figure 3.I.1). (Figure 3.I.2): An interactive version of the t-SNE analysis shown in Figure 3.4. with an extra panel (top right) excluding Ward clustering. Clicking in the top left chart highlights a cluster and generates a list of cluster components in the bottom right. Clicking in the top right panel highlights a specific gene. Clicking a point in the bottom left panel will open a comparison of the clustered genes in the Fitness Browser (https://fit.genomics.lbl.gov). (Figure 3.I.3): In this interactive figure, the shapes and colors are determined by the COG identifier for each individual gene. Clicking in the legend highlights all genes that are associated with a particular COG, and clicking in the scatterplot generates a list co-clustered genes

3.4.2. Global effectors of nitrogen metabolism

Mutants affected in the global regulators NtrB (PP_5047) and NtrC (PP_5048) have diverse phenotypes in the conditions tested. NtrBC is a two-component system that regulates the expression of numerous nitrogen assimilatory genes in *P. putida* [92]. When the cell is nitrogen starved, the uridylyltransferase protein GlnD (PP_1589) modulates the activity of GlnK (PP_5234, protein PII). GlnK then activates NtrB, which affects the phosphorylation state of NtrC. Subsequently, NtrC modulates expression of its target regulon [92]. While the library used for these assays has no insertions in *glnK*, the three other members of the signalling cascade are represented in the dataset and occupy the same cluster determined by t-SNE (Figure 3.I.1). It is unclear why certain nitrogen sources trigger a strong NtrC response (Figure 3.5A). The general role of NtrC is to counteract nitrogen starvation by activating the majority of nitrogen assimilatory genes, such as transporters [92]. By setting an arbitrary fitness score cutoff at > -1.5, we identified at least 8 out of 52 nitrogen sources that might not rely on a functional copy of *ntrC* (Table 3.2). The utilization pathways for those compounds seem to be less dependent on NtrC activation, suggesting specific regulation systems or the presence of constitutively expressed transporter and degradation pathways for these nitrogen sources. The fitness profile of mutants in the extracytoplasmic function sigma factor SigX (PP_2088) also seems to be correlated with that of *ntrBC* mutants. Previously, we found *sigX* to be partially essential for growth on D-lysine as a carbon source [18]. While no clear pattern in the fitness data could be determined for NtrC, another global factor involved in nitrogen metabolism, GltBD, illustrates a clear relationship between the conditions tested and their downstream metabolites.

Conditions on which NtrC fitness is >-1.5	
Condition	**NtrC fitness**
L-glutamine	-0.27
putrescine	-0.46
L-asparagine	-0.57
ammonium chloride	-0.66
4-aminobutyric acid	-0.86
4-guanidinobutyric acid	-1.30
L-histidine	-1.31
adenine	-1.43

Table 3.1. NtrC independent phenotypes

Nitrogen source conditions in which regulation may be less dependent on NtrC activation, indicated by *ntrC* fitness > -1.5 and illustrated by the dark blue points in Figure 3.5A.

Conditions on which GltBD fitness is >-1.5	
Condition	**GltBD fitness**
L-isoleucine	0.20
butyrolactam	0.12
L-pyroglutamic acid	0.10
L-valine	0.08
L-phenylalanine	0.04
L-aspartate	-0.07
L-glutamic acid	-0.15
L-proline	-0.16
L-glutamine	-0.32
4-aminobutyric acid	-0.45

Table 3.2. GltBD independent phenotypes

Nitrogen source conditions in which *gltBD* fitness > -1.5, illustrated by the scatter plot in Figure 3.5A.

GltB (PP_5076) and GltD (PP_5075) comprise the glutamate synthase (GOGAT) of *P. putida*, which plays an important role in the regulation of nitrogen assimilation. Interestingly, *gltBD* has diverse phenotypes across the conditions we tested (Figure 3.5). However, the second component of the central GS/GOGAT cycle, glutamine synthetase (*glnA*, PP_5046), has no insertions in our BarSeq library, indicating that it might have been essential during library construction due to glutamine auxotrophy of the mutants. We were able to observe particularly strong fitness defects for *gltBD* in nitrogen conditions that either produce free ammonium or do not yield L-glutamate (Figure 3.5A) [105]. For example, *gltBD* exhibits a very strong fitness defect in the nitrate (-6.17) and L-serine (-5.95) conditions but not in the L-phenylalanine condition (Figure 3.5A,

Table 3.2). Another example is in the comparison of 5-AVA (5-aminovalerate), 4-ABA (4-aminobutyrate), β-alanine, and 3-AIBA (3-aminoisobutyrate). 3-AIBA utilization likely uses pyruvate as an amino-acceptor (via PP_0596), 4-ABA and 5-AVA require 2-oxoglutarate (via PP_0214/DavT), and β-alanine requires both aminotransferases (Figure 3.5B, Figure 3.6). An exemption to this rule is the nitrogen source ammonium chloride. The fitness defects for *gltBD* in this condition were much lower ($|fitness| < 2$), but still significant ($|t| > 5$). It has been shown that at high concentrations of ammonium (>10 mM), its assimilation is achieved by both the glutamate dehydrogenase (GdhA) PP_0675 and GltBD [92,106,107].

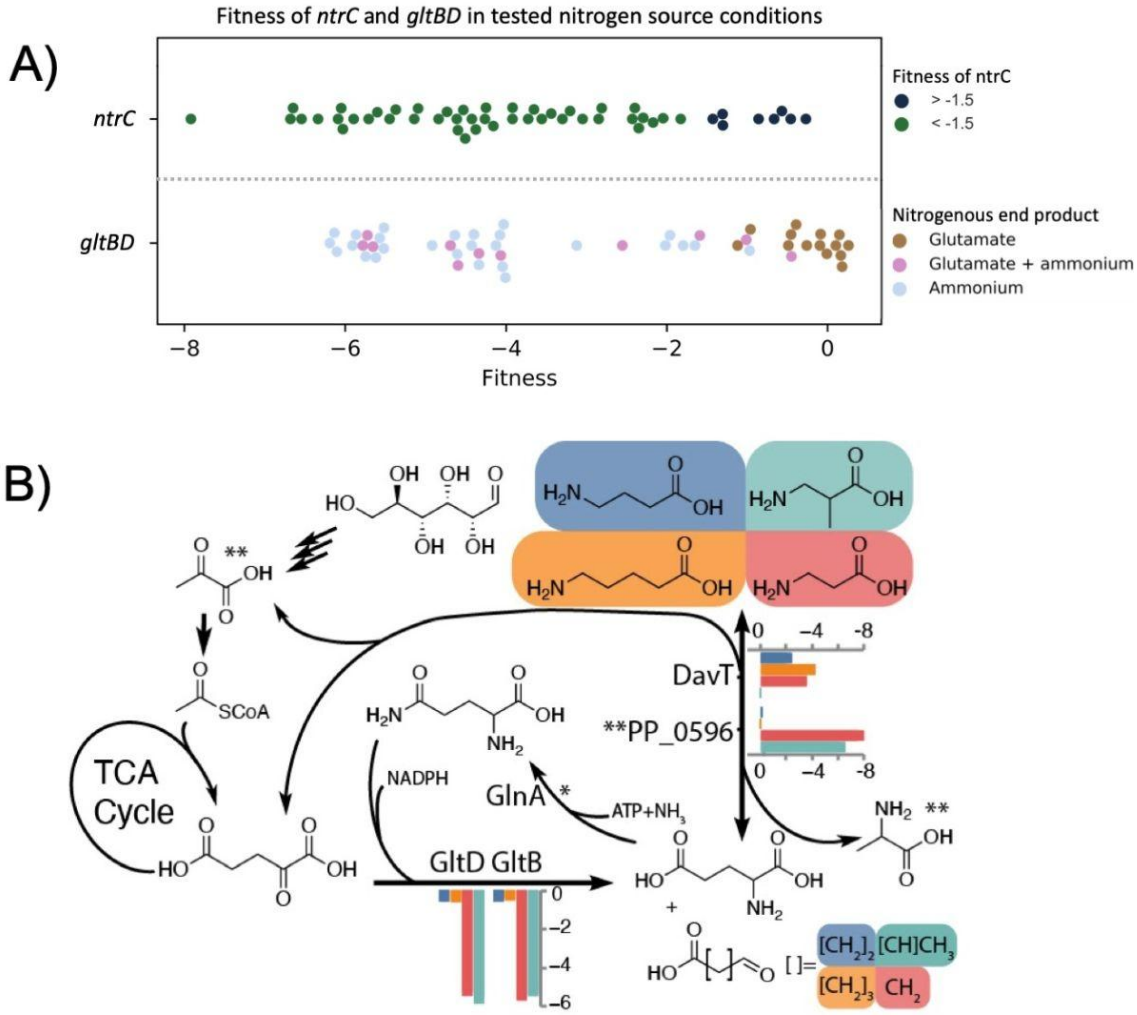


Figure 3.5. NtrC and GltBD phenotypes

A) Scatter plot of the average fitness values (n=2) for *ntrC* and *gltBD* in all the tested nitrogen conditions. For *ntrC*, nitrogen conditions are grouped based on whether the fitness phenotype of *ntrC* is < -1.5 (blue) or > -1.5 (green). Conditions where ntrC fitness is >-1.5 (green) may be less dependent on *ntrC* activation and are shown in Table 2. *gltBD* phenotypes are sorted based on putative glutamate (brown), glutamate+ammonium (pink), or ammonium (grey) release during nitrogen source

utilization. Fitness values for conditions resulting in *gltBD* fitness >-1.5 are shown in Table 3. B) The role of the GS/GOGAT cycle in the beta-alanine (red), 3-aminoisobutyrate (green), 4-aminobutyrate (blue) and 5-aminovalerate (orange) nitrogen source conditions. Average fitness values (n=2) are shown for *gltBD*, *davT*, and the pyruvate dependent transaminase PP_0596. *GlnA lacks fitness data because the library has no insertions. **Pyruvate and alanine are the specific nitrogen acceptor and product of PP_0596

The regulators *gacS* (PP_1650) and *gacA* (PP_4099) show strong fitness phenotypes in our dataset. This two-component system is homologous to the well-studied *barA*/*uvrY* system of *E. coli*, and it has been identified as a global regulator of cellular physiology in diverse organisms [58,108]. The conditions that elicited the strongest negative fitness phenotypes for *gacSA* were L-alanine (-1.3), β-alanine (-1.8), spermidine (-2.0), and propanediamine (-7.4). The direct targets of GacSA regulation are the *rsm* noncoding RNAs which modulate the activity of translational repressors, resulting in global changes in gene expression [59,109,110]. Through this mechanism, GacSA may be required for regulation of some portions of L-alanine, β-alanine, propanediamine, and spermidine metabolism.

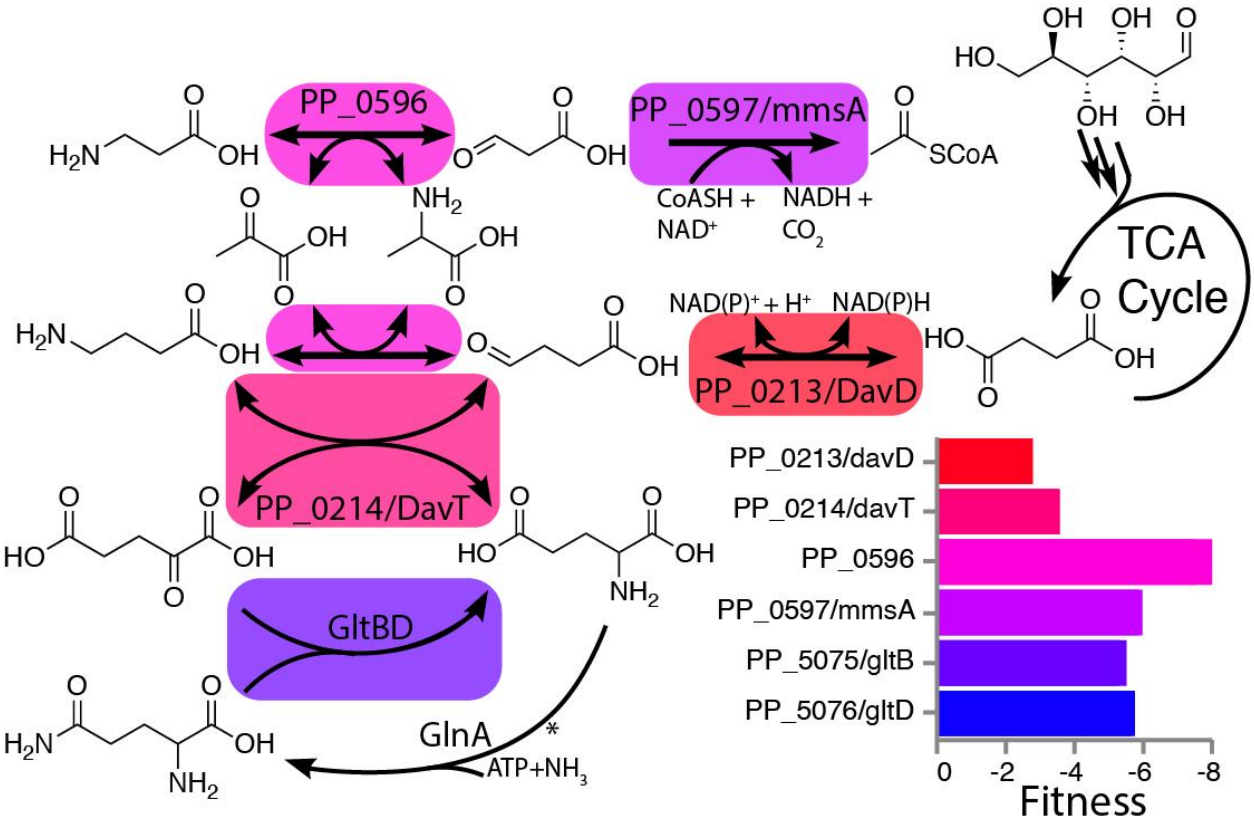


Figure 3.6 Beta-alanine metabolism in *P. putida* requires two amino transferases.

Proposed pathway for beta-alanine metabolism in *P. putida* that satisfies the requirements for *davD*, *davT*, *gltBD*, *mmsA*, and PP_0596. Inset bar chart shows the fitness values for the genes depicted in the pathway. While oxidation of 3-

oxopropionate (malonyl-semialdehyde) to malonate is also hypothetically possible via DavD, *P. putida* cannot use malonate as a sole carbon source indicating it may be a dead-end metabolite (unpublished data). Deeper biochemical or genetic analyses of this pathway are necessary to validate this mechanism of alanine and glutamate balancing during beta-alanine metabolism.

3.4.3. Inorganic Nitrogen Sources and Urea

P. putida KT2440 is a known obligate aerobe, and cannot use alternative terminal electron acceptors during oxidative phosphorylation [111]. Because of this, oxidized nitrogen species can only be used as nitrogen sources. Although the preferred inorganic nitrogen source for many bacteria is ammonium [112], other inorganic nitrogen species such as nitrate and nitrite can be utilized via the assimilatory nitrate reduction system, often organized in a single gene cluster [113,114]. The enzymes associated with the initial steps of the assimilatory nitrogen system in *P. putida* are the reductases NarB and NirDB, for which we see significant growth phenotypes (Figure 3.7). In Pseudomonads, the two-component system NasS/T is a common regulator of this operon [114–116]. In the absence of oxidized nitrogen sources, NasS and NasT form a complex that represses production of nitrate and nitrite reductases. When nitrate or nitrite are present, NasS dissociates from the NasS/T complex and the free RNA-binding antiterminator, NasT, enables full translation of the nitrate reduction operon [117] (Figure 3.7).

Our competitive growth assay using BarSeq indicated that the same system may also operate in *P. putida* KT2440. In the nitrate condition, the specific growth phenotypes for *nasT* (PP_2093) and *nasS* (PP_2094) are represented by the significant fitness values of -3.2 and -1.4, respectively. We were also able to identify a distinct nitrate transporter PP_2092 (-2.6). However, no significant phenotype could be detected for this gene in the nitrite condition, suggesting that it may be specific to nitrate. Since there were no mutants in predicted transporters with strong fitness phenotypes in the nitrite condition, it is also possible that PP_2092 or other unidentified transporters are responsible for nitrite transport. At pHs less than 7.2 and concentrations greater than 100 μM, nitrite is also known to be passively transported into the cell via diffusion of nitrous acid, likely explaining the lack of fitness data for this transport reaction [118].

Overall, the nitrite condition demonstrates a pattern of fitness phenotypes that resembles a stress response. In its free acid form, nitrite has been previously shown to have antimicrobial properties and was hypothesized to have wide-ranging mechanisms of action, including DNA damage, collapse of the proton motive force, and deleterious nitrosylation of cofactors and proteins [119]. Our fitness data indicate that disruption of the recently identified RES-Xre toxin-antitoxin (TA) module (PP_2433-4) is detrimental to growth in the nitrite condition. In this system, the toxin (PP_2434) rapidly degrades NAD^+ to halt bacterial growth and give the organism time to adapt its survival strategies, while the antitoxin (PP_2433) inhibits the toxin to allow restoration of NAD^+ levels [120,121]. The significant fitness defect of mutants in the antitoxin PP_2433 (-6.4) might indicate that either the RES-Xre system is part of *P. putida*'s stress response to nitrite, or that nitrite disturbs the redox homeostasis of the cell and further disruption of redox balance by the PP_2434 toxin is lethal in strains lacking a functional PP_2433 antitoxin. We also observed significant phenotypes for several pathways that oxidize NAD(P)H, which supports the idea that NAD^+ is depleted in the nitrite condition. One example is *P. putida*'s altered glucose utilization strategy. Fitness data indicates that glucose oxidation to 2-ketogluconate is preferred over the gluconate phosphorylation or direct glucose uptake pathways [122–124]. Indicators for the accumulation of 2-

ketogluconate in the nitrate condition include the specific fitness phenotypes for the transporter *kguT* (PP_3377; -1.5) and the NAD(P)H-dependent dehydrogenase *kguD* (PP_3376; -2.65). Compared to the other two pathways, glucose utilization via 2-ketogluconate produces more $NAD(P)^+$. Other conditions that seemed to lead to similar phenotypes for *kguT* were 1,6-hexanediamine (-1.3), caprolactam (-1.55), 2-aminobutyric acid (-1.5) and uracil (-1.2).

Urea, the simplest organic nitrogen source we tested, only elicited significant fitness phenotypes in *gltBD* (-5.45) and *ntrC* (-1.8). Transcriptome analysis has previously revealed that the expression of a putative urease operon in *P. putida* (PP_2842-9) is controlled by NtrC [92]. However, we observed no significant phenotypes in this operon. The presence of a second urea degradation pathway via urea carboxylase is unlikely, because it has been shown that organisms that possess the carboxylase pathway typically lack an urease [125–128]. More research is necessary to further characterize urea metabolism in *P. putida*.

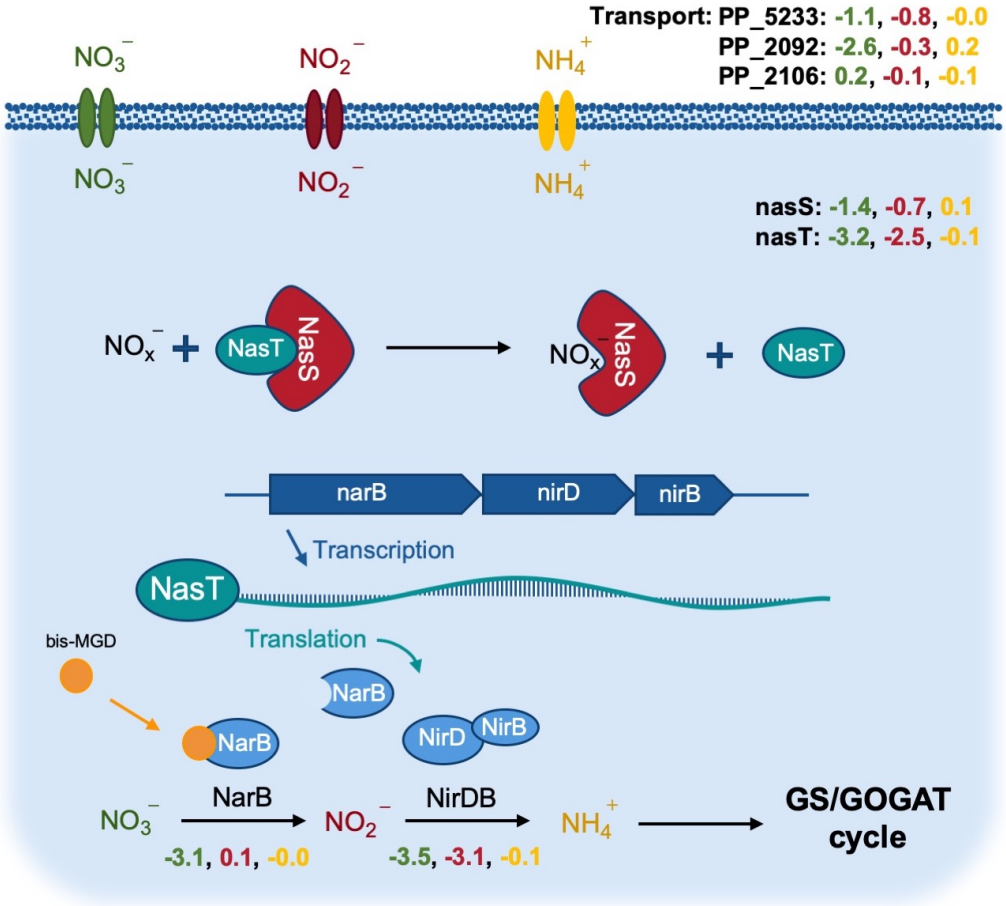


Figure 3.7 The assimilatory nitrate reduction system in *P. putida* KT2440.

Average fitness values (n=2) exhibited in the nitrate (green), nitrite (red) and ammonium (yellow) sole nitrogen source experiments. Shown are putative transporters, action of the NasST regulatory system, the assimilatory pathway, and the role of the bis-molybdopterin guanine dinucleotide (bis-MGD) cofactor. bis-MGD is the required

cofactor for NarB [129]. The fitness phenotypes for bis-MGD biosynthesis cluster together with the nitrate phenotypes and can be found in the interactive t-SNE visualization (Figure I1).

3.4.4. Quaternary amines and ethanolamine

Choline is a trimethylated, positively-charged amine and a common component of ionic liquids used in the depolymerization of lignin [130]. In nature, choline is metabolized by most bacteria to serve as a precursor for betaine, which is an important osmoprotectant [131]. While biosynthetic pathways for those compounds are rarely present in bacteria, transporters and catabolic routes for quaternary amines are ubiquitous in these organisms [132,133].

It is predicted that choline, betaine (trimethylglycine), and carnitine are metabolized through convergent pathways, with carnitine and choline entering betaine metabolism following thiolase cleavage and alcohol oxidation, respectively [2] (Figure 3.8A). In the subsequent reactions, betaine is demethylated to produce glycine. Even though the glycine cleavage system (PP_0986, PP_0988 and PP_0989) has a significant phenotype when glycine is used as the sole source of nitrogen (< -4), it is not essential for growth in the quaternary amine conditions (> -0.3). Instead, conversion into serine is more efficient due to the generation of the required methyl-group donor 5,10-methylenetetrahydrofolate during demethylase activity of PP_0310-1 [134]. As indicated by the strong fitness phenotypes elicited in the carnitine, choline, and betaine conditions, this reaction is likely catalyzed by PP_0322 (a predicted glycine/serine hydroxymethyltransferase) (< -2). Although PP_3144 (L-serine dehydratase) appears essential for growth on L-serine as a sole nitrogen source (-2.3), it is non-essential for growth on the tested quaternary amines. This suggests that other serine dehydratases (PP_0297 or PP_0987) or deaminating enzymes may be expressed in these conditions to release ammonium from serine.

Based on the results of the BarSeq assay and sequence homology, the initial steps in carnitine metabolism were also identified. We propose that the beta-oxidation of carnitine to betainyl-CoA proceeds through the genes PP_0301 (a putative thioesterase), PP_0303 (a dehydrocarnitine cleavage enzyme), and PP_0302 (L-carnitine dehydratase) (Figure 3.8A). The carnitine metabolism of *P. aeruginosa* was originally believed to proceed via CoA activation of oxidized carnitine [135]. Enzymatic characterization of a PP_0303 homolog has since suggested that this protein acts upon oxidized carnitine and acetyl-CoA, releasing acetoacetate and betainyl-CoA [136].

We also observed fitness phenotypes for the transporters and regulators involved in quaternary amine metabolism. Although all three quaternary amines cause a strong NtrC response, which tends to mask distinct transporters, PP_0294-6 (choline/betaine/carnitine ABC transporter) exhibits fitness defects for choline (-1.13) and carnitine (-2.55). The metabolism of betaine and its precursors is regulated by the repressor BetI (PP_5719) and the activators GbdR (PP_0298) and CdhR (PP_0305) [132,137]. While *gbdR* is essential in all three conditions (< -1.0), *cdhR* is specific to carnitine (-2.5) and *betI* is specific to choline (+1.15). The negative and positive fitness defects correlate with the predicted mode of action for these regulators [137]. The reason for the significant fitness phenotypes of PP_0308-9 in these conditions and their role in this operon are still unknown. Transcriptomic analysis suggested that they might be involved in the formation of filamentous biofilms [138].

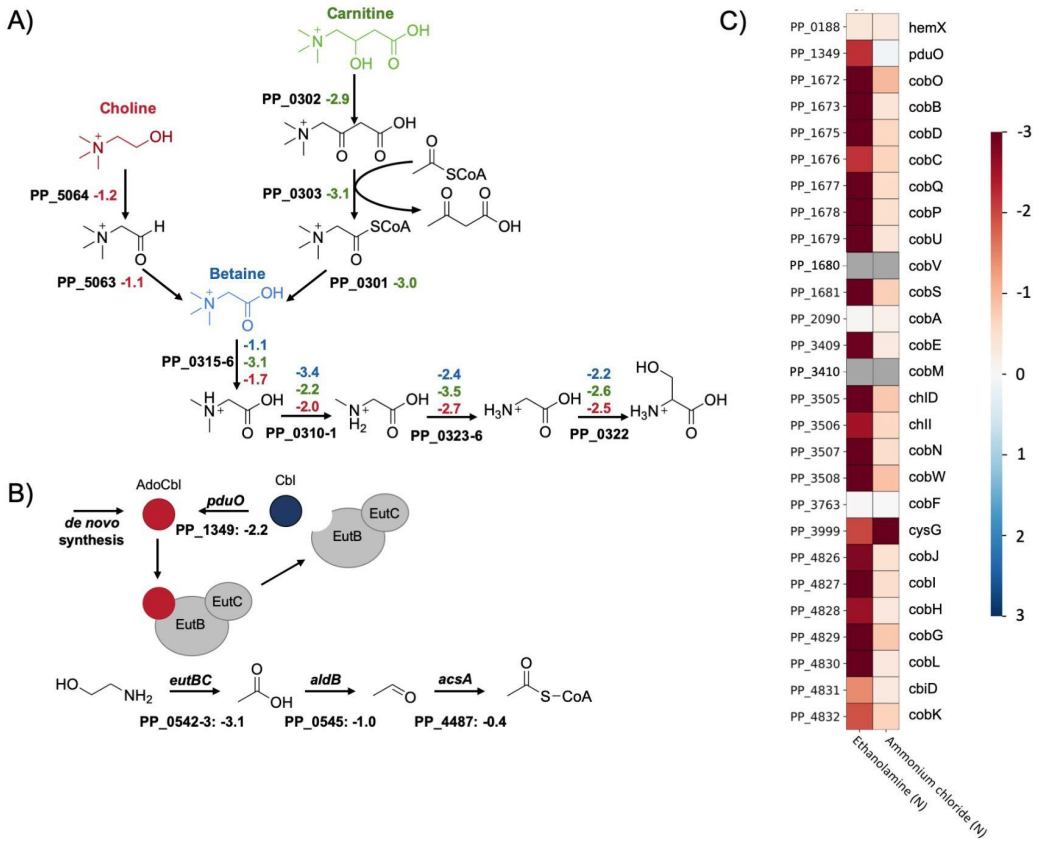


Figure 3.8. Quaternary amine and ethanolamine degradation in *P. putida*.

A) Putative routes for the quaternary amine catabolism in *P. putida* KT2440. The figure shows the degradation of choline (red), carnitine (green) and betaine (blue). The corresponding average fitness scores (n=2) are shown next to each gene. B) Ethanolamine degradation pathway, shown with fitness values (n=2) and regeneration of the AdoCbl cofactor. C) Heatmap with average fitness scores (n=2) of genes that are putatively involved in *P. putida*'s adenosylcobalamin biosynthesis. No fitness scores (grey) could be obtained for the genes PP_1680 (cobV) and PP_3410 (cobM).

Ethanolamine is a common molecule in nature and is involved in the choline and serine metabolism of plants [139,140]. Therefore, it is not surprising that several Pseudomonads are able to use it as a source for carbon and nitrogen [141]. Bacterial ethanolamine degradation can be divided into two routes: (i) via acetyl-CoA or (ii) via ethanol [141,142]. However, the first step in both pathways is catalyzed by the adenosylcobalamin-dependent ethanolamine ammonia-lyase (EAL) EutBC [142–144]. In *P. putida*, we observed a strong fitness phenotype for *eutBC* (PP_0542-3; -3.1) in the ethanolamine condition (Figure 3.8B). Furthermore, we were also able to support the requirement for its cofactor adenosylcobalamin (AdoCbl). Since there is no exogenous AdoCbl in our minimal media and no specific phenotypes for the corrinoid-specific transport system PP_0524-5, *P. putida* is likely capable of the *de novo* synthesis of AdoCbl [75,145–147]. This is also supported by the strong fitness phenotypes we observed for the putative *cob* genes in this condition (Figure 4C). The additional requirement of the adenosyltransferase PduO (PP_1349; -2.2) is most likely caused by the release of Cbl

during EAL activity [148–151]. Re-adenylation of Cbl by PduO is likely less metabolically demanding than the *de novo* synthesis of AdoCbl.

3.5. Conclusion

In conclusion, our study of *Pseudomonas putida* KT2440 has expanded the understanding of its diverse and robust metabolisms by identifying genes and proteins involved in assimilating 52 nitrogen-containing compounds and examining 19 amino acid drop-out conditions. We observed significant fitness phenotypes in 672 genes, classified these findings into 6 groups, and proposed assimilatory pathways based on our genetic data. Additionally, we characterized the substrate range of three aminotransferases, examined the specificity of five transcriptional regulators for potential synthetic biology applications, and developed an interactive visualization tool using manifold learning to enhance the accessibility of our BarSeq data for other researchers. Our findings contribute to the ongoing characterization of *P. putida* KT2440 and its potential applications in biotechnology.

3.6. Methods

3.6.1. Plate-based growth assays

Growth studies of bacterial strains were conducted using microplate reader kinetic assays as described previously [82]. Overnight cultures were washed three times with nitrogen-free MOPS minimal medium and used to inoculate 48-well plates at a ratio of 1:100 (Falcon, 353072). Each well contained 500 µL of MOPS medium with 10 mM of the tested nitrogen source. Plates were sealed with a gas-permeable microplate adhesive film (VWR, USA), and then optical density and fluorescence were monitored for 24-72 hours in an Biotek Synergy H1M plate reader (BioTek, USA) at 30 °C with fast continuous shaking. Optical density was measured at 600 nm. RFP was measured with an excitation wavelength of 535 nm, an emission of 620 nm, and a gain of 100.

3.6.2. BarSeq assays

BarSeq experiments utilized the *P. putida* library JBEI-1 and were completed as previously described [18]. Aliquots (2 mL) of libraries of JBEI-1 were thawed on ice, added to 25 mL of LB medium with kanamycin, and then grown at 30 °C to an OD600 of 0.5. Then, three 1-mL aliquots were removed, pelleted, and stored at -80 °C as zero timepoints. The libraries were washed three times in MOPS minimal medium with no nitrogen source, and then used to inoculate each experiment at a ratio of 1:100. Experiments were conducted in 24-well plates; each well contained 2 mL of nitrogen-free MOPS minimal medium with 10 mM of each tested nitrogen source. Plates were grown at 30 °C with shaking at 200 rpm, and 1 mL samples were collected after 24-72 hours, depending on when cultures appeared sufficiently turbid for DNA extraction. Samples were pelleted and stored at -80 °C until DNA extraction, which was done with a DNeasy UltraClean Microbial kit (Qiagen, Germany). BarSeq analysis was performed as previously described (19, 40). Strain fitness is defined as the normalized log2 ratio of the barcode reads in the experimental sample to the barcode reads in the time zero sample. The fitness of a gene is defined as the weighted average of the strain fitness for insertions in the central 10% to 90% of the gene. The gene fitness values are normalized such that the typical gene has a fitness of zero. The primary statistic t value represents the form of fitness divided by the estimated variance across different mutants of the

same gene. Statistic t values of >|4| were considered significant. A more detailed explanation of fitness score calculations can be found in Wetmore et al. [40]. All experiments described here passed the quality testing metrics described previously. Experiments were conducted in biological duplicates, and the fitness data are publically available at http://fit.genomics.lbl.gov.

3.6.3. Protein Production and Purification

The protocol for the ω-amino acid aminotransferases purification is a modified version of the previously described procedure by Yuzawa et al. [152]. Briefly, *E. coli* BL21(DE3) harboring the pET28a vectors for PP_0596, PP_2180 or PP_5182 were cultured in LB medium with kanamycin at 37 °C until OD600 reached 0.4. Protein expression was induced with 250 μM IPTG and cultures were grown at 18 °C for 24 h. Cells were harvested by centrifugation for 20 min at 5000*g* and the pellet was resuspended in 30 mL wash buffer (50 mM sodium phosphate pH 7.6, 300 mM NaCl, 10 mM imidazole, 4 °C). The cells were disrupted by sonication (8 x 30 sec) and cell debris was removed by three subsequent centrifugations (15 min, 8000*g*, 4 °C). The soluble fraction was mixed with 4 mL Nickel-NTA agarose beads (Thermo Fisher Scientific) for 1 h at 4 °C. The mixture was applied to a Nickel-NTA column and washed three times with wash buffer. The protein was eluted with 12 mL elution buffer (150 mM sodium phosphate buffer pH 7.6, 50 mM NaCl, 150 mM imidazole, 4 °C) and buffer was exchanged with stock buffer (100 mM sodium phosphate pH 7.5, 0.5 mM DTT, 10 % glycerol, 4 °C) by dialysis (SnakeSkin Dialysis Tubing, 10k MWCO, Thermo Fisher Scientific). Protein was concentrated in 30k MWCO Amicon Ultra-15 centrifugal filters (MilliporeSigma) and stored at -80 °C.

3.6.4. *In vitro* pyruvate transamination and product quantification

Substrate specificity of the purified aminotransferases was determined by *in vitro* L-alanine production. The reaction was carried out in 100 mM sodium phosphate buffer pH 9 with 500 μM pyruvate, 1 mM PLP, 5 μM enzyme and 5 mM substrate. After addition of the substrate, the reaction mixture was immediately incubated at 30 °C for 30 min. The reaction was stopped by boiling the mixture for 10 min at 100 °C. Alanine concentrations were determined enzymatically using the Alanine Assay Kit (Cell Biolabs, Inc.) following the manufacturer's instructions.

3.6.5. t-Stochastic Neighbor Embedding of BarSeq results

The data used for the t-SNE visualization was filtered by t-scores and fitness scores ($|t| > 5$ and $|\text{fitness}| > 1$). This resulted in a file containing fitness scores for 615 significantly affected genes across 129 sole-nitrogen source growth assays in 51 different conditions, and 19 amino acid dropout conditions. 2-ABA was excluded from the analysis due to the large number of significantly affected genes in this condition. For each gene, the "significant condition" was defined as the condition that caused the largest change in fitness score. The fitness values were then used as the input into the t-SNE module supplied in the python package Scikit-Learn [153,154]. Clusters (n=100) were defined using the Ward-clustering module in Scikit-Learn. The most frequently occurring "significant condition" was used to name each cluster. The Jupyter notebook used in these analyses can be found at:
https://github.com/mschmidt75/ppnitrogentsne.git

3.6.6. Bioinformatic Analyses

All statistical analyses were carried out using either the Python Scipy or Numpy libraries [85,155]. To identify the number of aminotransferases, the HMMER version 3.3.2 (http://hmmer.org/) was used to scan the *P. putida* KT2440 genome against a profile database file consisting out of the Pfam HMM files of PF00155, PF00202, PF01063 and PF00266. The E-value cutoff was set to 1e-20, and significant hits annotated as transcriptional regulators were excluded from the final dataset. HMMER was also used to identify possible genes of the γ-glutamyl cycle, with the Pfam HMM files of PF00120, PF01266, PF00171, and PF07722 corresponding to the glutamyl-polyamine synthetase, the glutamyl-polyamine oxidase, the aldehyde dehydrogenase, and the amide hydrolase.

The databases MiST 3.0 and TransportDB 2.0 were used to extract transcription factors and transport associated proteins from the BarSeq dataset [99,100]. EggNOGmapper was used to generate the Clusters of Orthologous Groups (COGs) for *P. putida* KT2440 [102,156]. Additionally, manual analysis of the data and proposals of metabolic pathways relied heavily on BioCyc and PaperBlast [24,62]

3.7. Acknowledgements

We would like to thank Morgan Price and Dr. Megan Garber for their assistance in analyzing BarSeq data, Alberto Nava for his advice on manifold learning methods, Dr. Elias Englund for his help on protein purification and Dr. Namil Lee for his advice on gene annotations. Mitchell Thompson is a Simons Foundation Awardee of the Life Sciences Research Foundation. The laboratory of LMB is partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy within the Cluster of Excellence FSC 2186 'The Fuel Science Center'. This work was part of the DOE Joint BioEnergy Institute (https://www.jbei.org) supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, supported by the U.S. Department of Energy, Energy Efficiency and Renewable Energy, Bioenergy Technologies Office, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy. The views and opinions of the authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).

3.8. Contributions

Conceptualization, M.S., A.N.P., M.R.I., M.G.T.; Methodology, M.S., A.N.P., M.R.I., M.G.T.; Investigation, M.S., A.N.P, M.R.I., E.E.K.B. M.G.T, R.K.; Writing – Original

Draft, M.S., A.N.P., M.R.I.; Writing – Review and Editing, All authors.; Resources and supervision, M.G.T., L.M.B., A.M.D., P.M.S, J.D.K.
M.S. and A.N.P. contributed equally to this work. Author order was determined by number of sesamoid bones.

3.9. Competing interests
J.D.K. has financial interests in Amyris, Ansa Biotechnologies, Apertor Pharma, Berkeley Yeast, Demetrix, Lygos, Napigen, ResVita Bio, and Zero Acre Farms.

Chapter 4. Creation of *Pseudomonas redtida* and *yellotida*

Including material from published work:
Matthew R. Incha, Mitchell G. Thompson, Jacquelyn M. Blake-Hedges, Yuzhong Liu, Allison N. Pearson, Matthias Schmidt, Jennifer W. Gin, Christopher J. Petzold, Adam M. Deutschbauer, Jay D. Keasling, "Leveraging host metabolism for bisdemethoxycurcumin production in *Pseudomonas putida*" MEC 2019

4.0. Chapter preface

With the strength of the functional genomics method, randomly barcoded transposon sequencing (RB-TnSeq) we were able to elucidate the metabolisms of aromatic hydrocarbons. This process took under a month and validated the findings from 50 years of biochemical and genomics research. With these data we sought to engineer *Pseudomonas putida* KT2440 to metabolize these aromatics and produce valuable compounds. This chapter explores the use of RB-TnSeq and the use of the data to inform metabolic pathway development in *P. putida* KT2440.

4.1. Abstract

Pseudomonas putida is a saprophytic bacterium with robust metabolisms and strong solvent tolerance making it an attractive host for metabolic engineering and bioremediation. Due to its diverse carbon metabolisms, its genome encodes an array of proteins and enzymes that can be readily applied to produce valuable products. In this work we sought to identify design principles and bottlenecks in the production of type III polyketide synthase (T3PKS)-derived compounds in *P. putida*. T3PKS products are widely used as nutraceuticals and medicines and often require aromatic starter units, such as coumaroyl-CoA, which is also an intermediate in the native coumarate catabolic pathway of *P. putida*. Using a randomly barcoded transposon mutant (RB-TnSeq) library, we assayed gene functions for a large portion of aromatic catabolism, confirmed known pathways, and proposed new annotations for two aromatic transporters. The 1,3,6,8-tetrahydroxynapthalene synthase of *Streptomyces coelicolor* (RppA), a microbial T3PKS, was then used to rapidly assay growth conditions for increased T3PKS product accumulation. The feruloyl/coumaroyl CoA synthetase (Fcs) of *P. putida* was used to supply coumaroyl-CoA for the curcuminoid synthase (CUS) of *Oryza sativa*, a plant T3PKS. We identified that accumulation of coumaroyl-CoA in this pathway results in extended growth lag times in *P. putida*. Deletion of the second step in coumarate catabolism, the enoyl-CoA hydratase-lyase (Ech), resulted in increased production of the type III polyketide bisdemethoxycurcumin.

4.2. Introduction

Secondary metabolites of fungi, plants, and bacteria have long been used as medicines and supplements [157]. Compounds such as naringenin, raspberry ketone, resveratrol, and curcumin are widely used nutraceuticals and are biosynthesized through similar pathways [158–161]. Commercially, these chemicals are either extracted directly from plants or produced synthetically, as in the case of raspberry ketone [160]. Renewable microbial production of these compounds will decrease reliance on agriculture and fossil fuel-derived chemical synthesis. The biosynthesis of these compounds (naringenin, raspberry ketone, resveratrol, and curcumin) relies on a class of enzymes called type III polyketide synthases (T3PKSs). T3PKSs carry out iterative

Claisen condensation reactions typically with coenzyme A (CoA)-based starter and extender units [162]. In the case of the 1,3,6,8-tetrahydroxynapthalene synthase of *Streptomyces coelicolor* (RppA) the starter and extender units are simply malonyl-CoA, while in many plant T3PKSs the starter unit is a phenylpropanoyl-CoA thioester, usually derived from ferulate, coumarate, or cinnamate [163,164].

Coumarate and ferulate are components of lignin found in lignocellulosic hydrolysate (LH), which has been proposed for use as a renewable feedstock for biocatalysis [165,166]. Characteristics such as high solvent tolerance and diverse carbon metabolisms are essential for microbes to be used in LH valorization. However, these robust traits are lacking in commonly used model organisms, such as *Escherichia coli*, hindering progress toward making LH a viable feedstock [167]. Rather than developing these characteristics in a microorganism *de novo*, a clear alternative is to source a microbe with the desirable traits already available to it. The saprophytic bacterium *Pseudomonas putida* has long been studied for its ability to catabolize aromatic compounds and withstand solvents, making it an attractive host for LH upcycling [168–170].

In this work we sought to leverage the native catabolism of *P. putida* KT2440 for use in the biosynthesis of a plant T3PKS product from coumarate. RppA of *S. coelicolor* was first expressed to determine production conditions conducive to T3PKS product accumulation. Using a randomly barcoded transposon mutant (RB-TnSeq) library of *P. putida*, we assayed for genes involved in the catabolism of coumarate and seven other related aromatic compounds. We then found that accumulation of coumaroyl-CoA by the activity of Fcs was toxic to *P. putida*. The native feruloyl/coumaroyl-CoA synthetase (Fcs) of *P. putida* was then used to produce the coumaroyl-CoA starter unit to the curcuminoid synthase (CUS) of *Oryza sativa*. Finally, relying on the native expression levels of *fcs* from the chromosome and plasmid-based expression of *CUS* (Figure 4.1), production of bisdemethoxycurcumin (BDC) from coumarate was achieved.

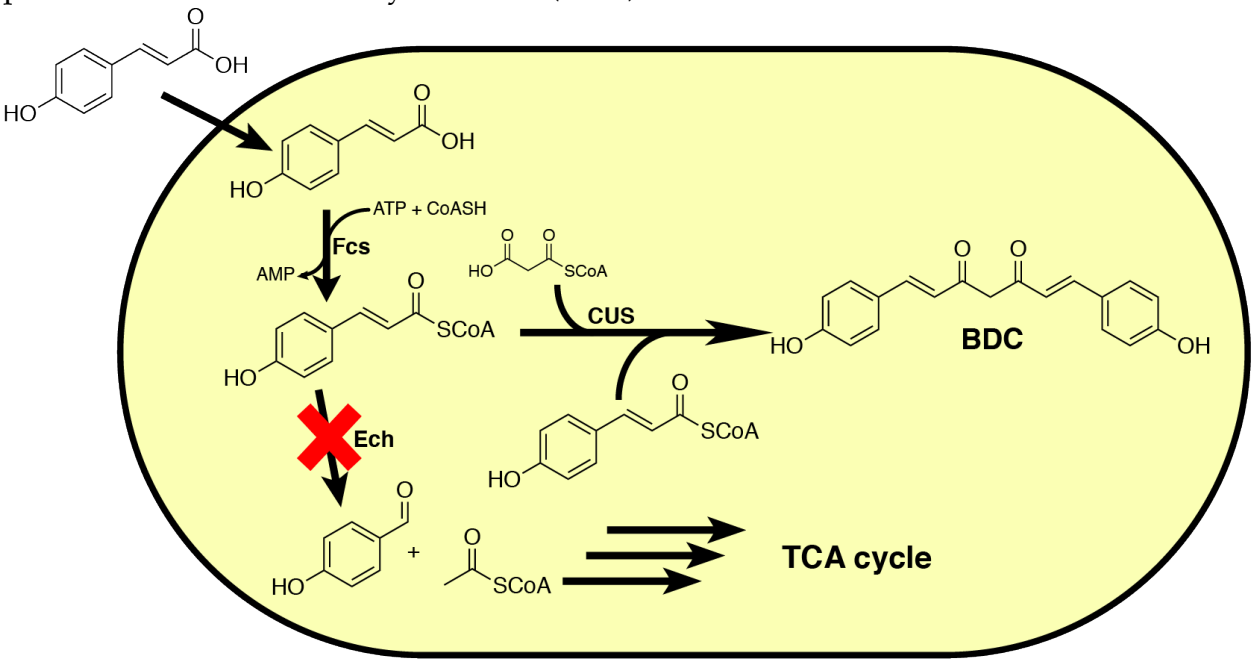


Figure 4.1. Diagram of the engineered *P. putida* strain used to produce bisdemethoxycurcumin.

Fcs is the feruloyl/coumaroyl-CoA synthetase native to *P. putida* KT2440. Ech is the native enoyl-CoA hydratase-lyase which carries out the second step in coumarate catabolism and was knocked out in our production host. CUS is the curcuminoid synthase from *O. sativa.* BDC is the final curcuminoid product, bisdemethoxycurcumin.

4.3. Results

4.3.1. RppA as a screen for optimal T3PKS production conditions

The tetrahydroxynaphthalene synthase of *Streptomyces coelicolor* was recently applied as a biosensor for malonyl-CoA concentrations in several hosts including *P. putida* [171]. We sought to express this protein from the broad host range pBADT vector to assay culture conditions for increased product accumulation in *P. putida*. RppA was initially codon optimized for expression in *P. putida* and cloned into the construct pBADT-*rppA*-OW. We then constructed two more plasmids, one with the complete *rppA* cloned from the *S. coelicolor* genome (pBADT-*rppA*-NW) and another with a 3' truncation of 75 base pairs (pBADT-*rppA*-NT) as this had been described previously to increase enzymatic activity [163].

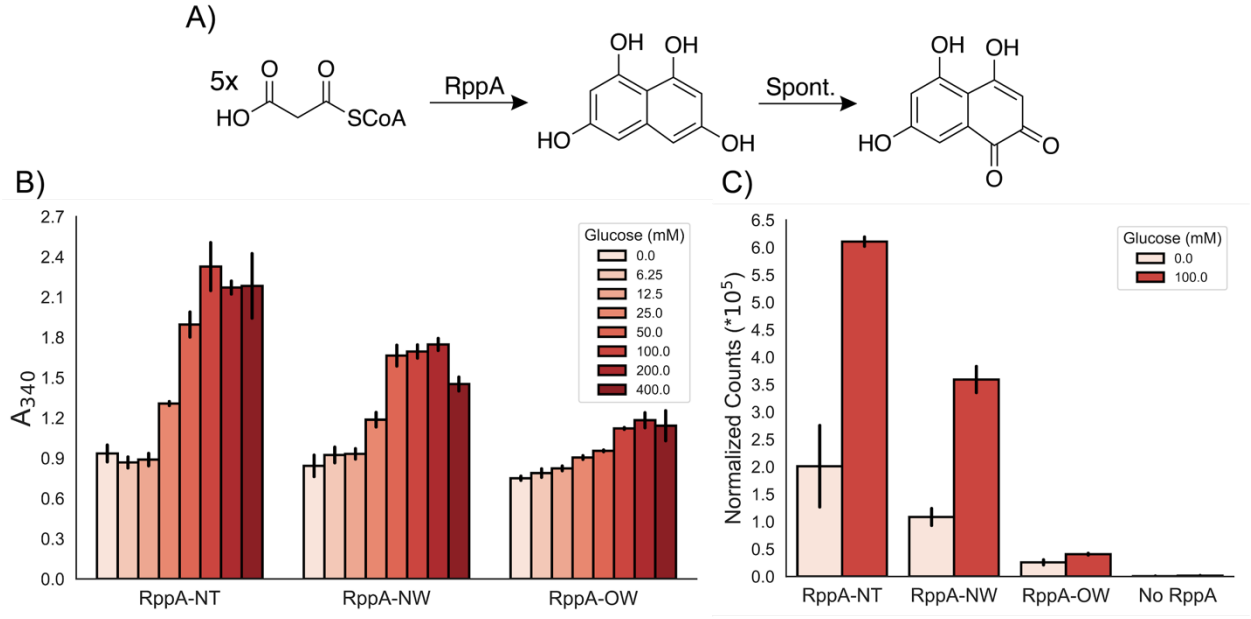


Figure 4.2. Flaviolin production as a reference for T3PKS activity.

(A) Diagram of the pathway to the red pigment, flaviolin, from RppA and subsequent spontaneous oxidation. (B) Absorbance measurements at 340 nm of supernatants from cells expressing RppA variants in LB containing different concentrations of glucose. Error bars indicate the standard deviation of three trials. RppA-NW contains natural *S. coelicolor* codons, RppA-NT is the same as RppA-NW but with a C-terminal truncation of 25 amino acids, and RppA-OW is translated from a transcript that was codon optimized for *P. putida.* (C) RppA expression profile as determined by mass spectrometry from lysates of cultures producing the RppA variants. The No RppA

sample was extracted from *P. putida* KT2440 carrying pBADT-RFP. Error bars indicate the standard deviation of three trials.

Since the product of RppA, 1,3,6,8-tetrahydroxynaphthalene, spontaneously oxidizes to the red pigment flaviolin, we used a colorimetric assay to determine how glucose concentrations affect the production of T3PKS products (Figure 2A). Concentrations of glucose ranging from 0 - 400 mM supplemented into LB medium were tested, and the production of flaviolin was measured by absorbance at 340 nm, as previously reported [171]. Unexpectedly, the *rppA* expression vectors using the native *S. coelicolor* codons produced more flaviolin than the *P. putida* codon optimized variant (Figure 2B). In all three constructs tested, there was an increase in the accumulation of flaviolin in cultures containing greater than 25 mM glucose and accumulation of flaviolin appeared to plateau at ~100 mM glucose (Figures 2B and S4). To validate the observed red compound was flaviolin, we extracted, purified, and obtained an ^{1}H-NMR spectrum for the red product using established methods (Figure 4.10) [172,173]. The final yield from 100 mL of culture was ~6.5 mg.

Following this result, we sought to identify if increased protein abundance was the cause for increased flaviolin production. When RppA was quantified using LC-MS, we observed an increase in its relative abundance for all variants when 100 mM glucose was supplemented to the medium. The abundance of RppA also was highest in the strains expressing the native genomic sequence of *rppA* (Figure 2C).

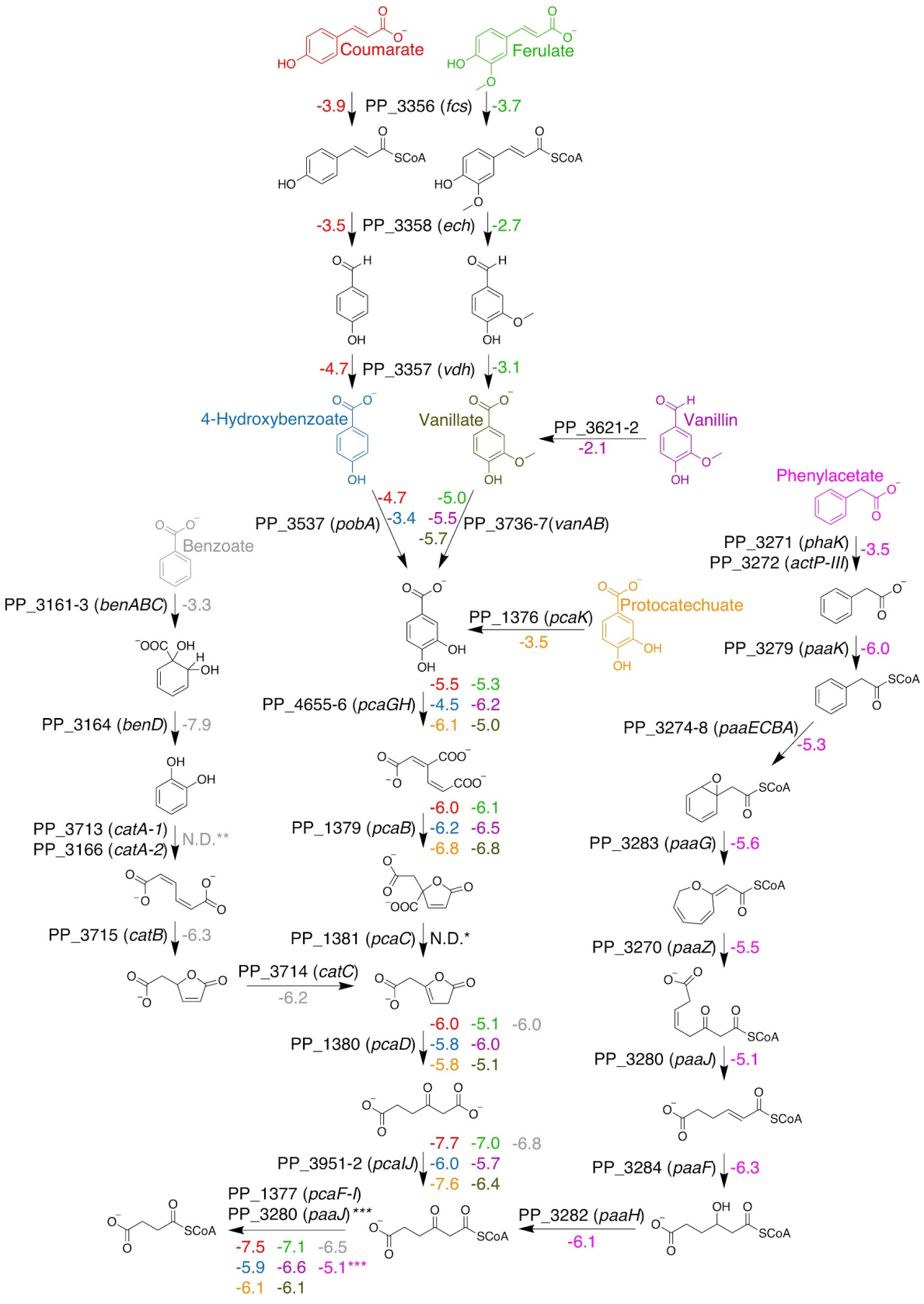

Coumarate
Ferulate
-3.9
PP_3356 (fcs)
-3.7
SCoA
-3.5
PP_3358 (ech)
-2.7
-4.7
PP_3357 (vdh)
-3.1
4-Hydroxybenzoate
Vanillate
PP_3621-2
-2.1
Vanillin
-4.7
-3.4
-5.0
-5.5
-5.7
PP_3537 (pobA)
PP_3736-7(vanAB)
Phenylacetate
PP_3271 (phaK)
PP_3272 (actP-III)
-3.5
Benzoate
PP_3161-3 (benABC)
-3.3
PP_1376 (pcaK)
-3.5
Protocatechuate
PP_3279 (paaK)
-6.0
-5.5
-5.3
PP_4655-6 (pcaGH)
-4.5
-6.2
-6.1
-5.0
PP_3274-8 (paaECBA)
-5.3
PP_3164 (benD)
-7.9
-6.0
-6.1
PP_1379 (pcaB)
-6.2
-6.5
-6.8
-6.8
PP_3283 (paaG)
-5.6
PP_3713 (catA-1)
PP_3166 (catA-2)
N.D.**
PP_3715 (catB)
-6.3
PP_1381 (pcaC)
N.D.*
PP_3270 (paaZ)
-5.5
PP_3714 (catC)
-6.2
-6.0
-5.1
-6.0
PP_1380 (pcaD)
-5.8
-6.0
-5.8
-5.1
PP_3280 (paaJ)
-5.1
-7.7
-7.0
-6.8
PP_3951-2 (pcaIJ)
-6.0
-5.7
-7.6
-6.4
PP_3284 (paaF)
-6.3
PP_1377 (pcaF-I)
PP_3280 (paaJ)***
-7.5
-7.1
-6.5
-5.9
-6.6
-5.1***
-6.1
-6.1
PP_3282 (paaH)
-6.1

Figure 4.3. Overview of aromatic catabolism in *P. putida* KT2440.

The colored and labelled compounds depicted were fed as sole carbon sources to the barcoded transposon library. Fitness values for each gene are the average of 2 replicate RB-TnSeq assays, and are colored corresponding to the carbon source (gray: benzoate, red: coumarate, blue: 4-hydroxybenzoate, light green: ferulate, dark green: vanillate, purple: vanillin, yellow: protocatechuate, pink: phenylacetate). For reactions where multiple genes are necessary, i.e. enzyme complexes, the fitness values for each gene involved in the reaction were averaged. *The t-score for *pcaC* was insignificant ($|t_{score}| < 4.0$) and was excluded from our analysis. **Fitness values for these genes were mild (fitness > -2.0) and excluded. ***Fitness value corresponds to *paaJ*.

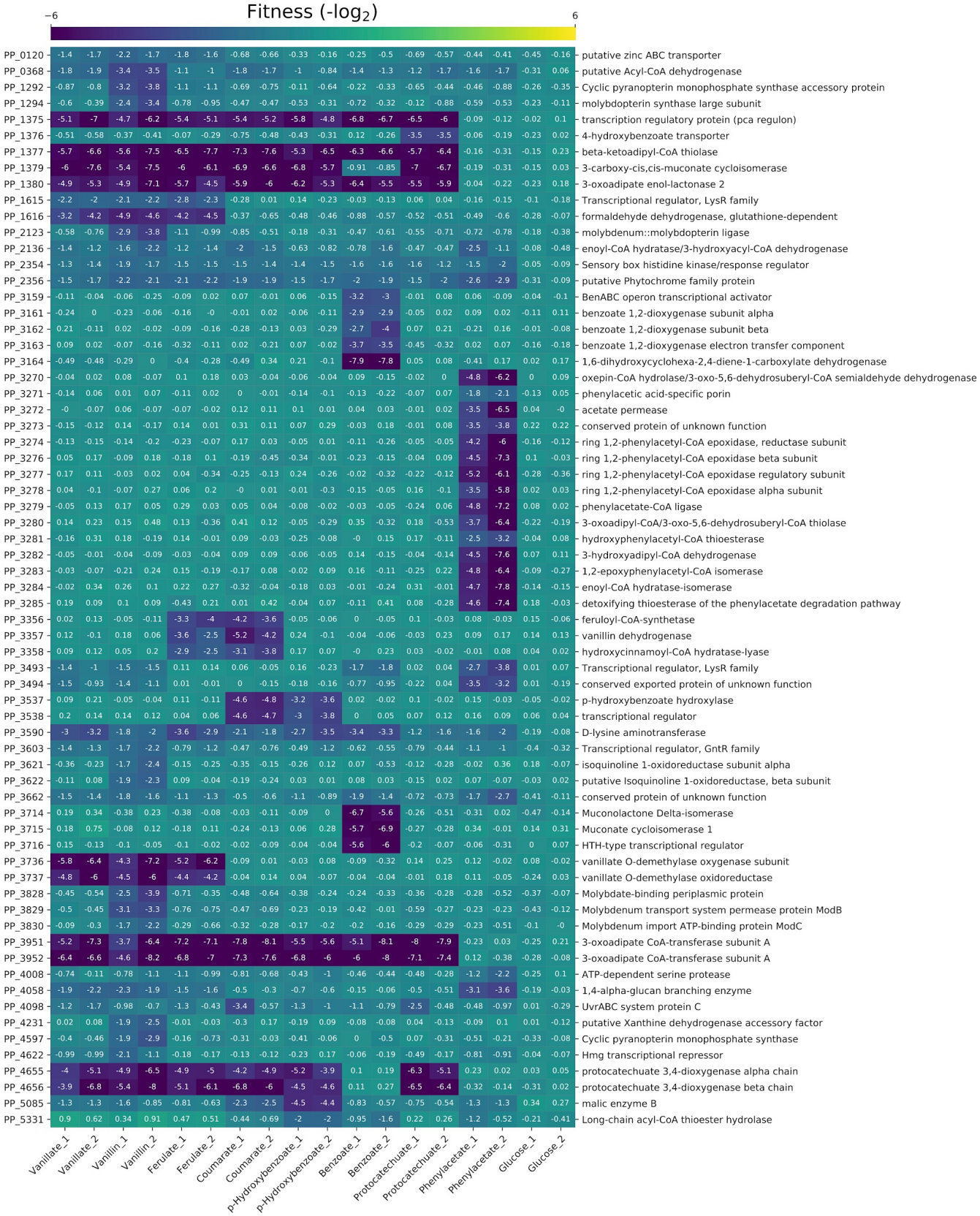

Fitness (-log2)
−6
6
PP_0120 -1.4 -1.7 -2.2 -1.7 -1.8 -1.6 -0.68 -0.66 -0.33 -0.16 -0.25 -0.5 -0.69 -0.57 -0.44 -0.41 -0.45 -0.16 putative zinc ABC transporter
PP_0368 -1.8 -1.9 -3.4 -3.5 -1.1 -1 -1.8 -1.7 -1 -0.84 -1.4 -1.3 -1.2 -1.7 -1.6 -1.7 -0.31 0.06 putative Acyl-CoA dehydrogenase
PP_1292 -0.87 -0.8 -3.2 -3.8 -1.1 -1.1 -0.69 -0.75 -0.11 -0.64 -0.22 -0.33 -0.65 -0.44 -0.46 -0.88 -0.26 -0.35 Cyclic pyranopterin monophosphate synthase accessory protein
PP_1294 -0.6 -0.39 -2.4 -3.4 -0.78 -0.95 -0.47 -0.47 -0.53 -0.31 -0.72 -0.32 -0.12 -0.88 -0.59 -0.53 -0.23 -0.11 molybdopterin synthase large subunit
PP_1375 -5.1 -7 -4.7 -6.2 -5.4 -5.1 -5.4 -5.2 -5.8 -4.8 -6.8 -6.7 -6.5 -6 -0.09 -0.12 -0.02 0.1 transcription regulatory protein (pca regulon)
PP_1376 -0.51 -0.58 -0.37 -0.41 -0.07 -0.29 -0.75 -0.48 -0.43 -0.31 0.12 -0.26 -3.5 -3.5 -0.06 -0.19 -0.23 0.02 4-hydroxybenzoate transporter
PP_1377 -5.7 -6.6 -5.6 -7.5 -6.5 -7.7 -7.3 -7.6 -5.3 -6.5 -6.3 -6.6 -5.7 -6.4 -0.16 -0.31 -0.15 0.23 beta-ketoadipyl-CoA thiolase
PP_1379 -6 -7.6 -5.4 -7.5 -6 -6.1 -6.9 -6.6 -6.8 -5.7 -0.91 -0.85 -7 -6.7 -0.19 -0.31 -0.15 0.03 3-carboxy-cis,cis-muconate cycloisomerase
PP_1380 -4.9 -5.3 -4.9 -7.1 -5.7 -4.5 -5.9 -6 -6.2 -5.3 -6.4 -5.5 -5.5 -5.9 -0.04 -0.22 -0.23 0.18 3-oxoadipate enol-lactonase 2
PP_1615 -2.2 -2 -2.1 -2.2 -2.8 -2.3 -0.28 0.01 0.14 -0.23 -0.03 -0.13 0.06 0.04 -0.16 -0.15 -0.1 -0.18 Transcriptional regulator, LysR family
PP_1616 -3.2 -4.2 -4.9 -4.6 -4.2 -4.5 -0.37 -0.65 -0.48 -0.46 -0.88 -0.57 -0.52 -0.51 -0.49 -0.6 -0.28 -0.07 formaldehyde dehydrogenase, glutathione-dependent
PP_2123 -0.58 -0.76 -2.9 -3.8 -1.1 -0.99 -0.85 -0.51 -0.18 -0.31 -0.47 -0.61 -0.55 -0.71 -0.72 -0.78 -0.18 -0.38 molybdenum::molybdopterin ligase
PP_2136 -1.4 -1.2 -1.6 -2.2 -1.2 -1.4 -2 -1.5 -0.63 -0.82 -0.78 -1.6 -0.47 -0.47 -2.5 -1.1 -0.08 -0.48 enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase
PP_2354 -1.3 -1.4 -1.9 -1.7 -1.5 -1.5 -1.5 -1.4 -1.4 -1.5 -1.6 -1.6 -1.6 -1.2 -1.5 -2 -0.05 -0.09 Sensory box histidine kinase/response regulator
PP_2356 -1.5 -1.7 -2.2 -2.1 -2.1 -2.2 -1.9 -1.9 -1.5 -1.7 -2 -1.9 -1.5 -2 -2.6 -2.9 -0.31 -0.09 putative Phytochrome family protein
PP_3159 -0.11 -0.04 -0.06 -0.25 -0.09 0.02 0.07 -0.01 0.06 -0.15 -3.2 -3 -0.01 0.08 0.06 -0.09 -0.04 -0.1 BenABC operon transcriptional activator
PP_3161 -0.24 0 -0.23 -0.06 -0.16 -0 -0.01 0.02 -0.06 -0.11 -2.9 -2.9 -0.05 0.02 0.09 0.02 -0.11 0.11 benzoate 1,2-dioxygenase subunit alpha
PP_3162 0.21 -0.11 0.02 -0.02 -0.09 -0.16 -0.28 -0.13 0.03 -0.29 -2.7 -4 0.07 0.21 -0.21 0.16 -0.01 -0.08 benzoate 1,2-dioxygenase subunit beta
PP_3163 0.09 0.02 -0.07 -0.16 -0.32 -0.11 0.02 -0.21 0.07 -0.02 -3.7 -3.5 -0.45 -0.32 0.02 0.07 -0.16 -0.18 benzoate 1,2-dioxygenase electron transfer component
PP_3164 -0.49 -0.48 -0.29 0 -0.4 -0.28 -0.49 0.34 0.21 -0.1 -7.9 -7.8 0.05 0.08 -0.41 0.17 0.02 0.17 1,6-dihydroxycyclohexa-2,4-diene-1-carboxylate dehydrogenase
PP_3270 -0.04 0.02 0.08 -0.07 0.1 0.18 0.03 -0.04 -0.06 -0.04 0.09 -0.15 -0.02 0 -4.8 -6.2 0 0.09 oxepin-CoA hydrolase/3-oxo-5,6-dehydrosuberyl-CoA semialdehyde dehydrogenase
PP_3271 -0.14 0.06 0.01 0.07 -0.11 0.02 0 -0.01 -0.14 -0.1 -0.13 -0.22 -0.07 0.07 -1.8 -2.1 -0.13 0.05 phenylacetic acid-specific porin
PP_3272 -0 -0.07 0.06 -0.07 -0.07 -0.02 0.12 0.11 0.1 0.01 0.04 0.03 -0.01 0.02 -3.5 -6.5 0.04 -0 acetate permease
PP_3273 -0.15 -0.12 0.14 -0.17 0.14 0.01 0.31 0.11 0.07 0.29 -0.03 0.18 -0.01 0.08 -3.5 -3.8 0.22 0.22 conserved protein of unknown function
PP_3274 -0.13 -0.15 -0.14 -0.2 -0.23 -0.07 0.17 0.03 -0.05 0.01 -0.11 -0.26 -0.05 -0.05 -4.2 -6 -0.16 -0.12 ring 1,2-phenylacetyl-CoA epoxidase, reductase subunit
PP_3276 0.05 0.17 -0.09 0.18 -0.18 0.1 -0.19 -0.45 -0.34 -0.01 -0.23 -0.15 -0.04 0.09 -4.5 -7.3 0.1 -0.03 ring 1,2-phenylacetyl-CoA epoxidase beta subunit
PP_3277 0.17 0.11 -0.03 0.02 0.04 -0.34 -0.25 -0.13 0.24 -0.26 -0.02 -0.32 -0.22 -0.12 -5.2 -6.1 -0.28 -0.36 ring 1,2-phenylacetyl-CoA epoxidase regulatory subunit
PP_3278 0.04 -0.1 -0.07 0.27 0.06 0.2 -0 0.01 -0.01 -0.3 -0.15 -0.05 0.16 -0.1 -3.5 -5.8 0.02 0.03 ring 1,2-phenylacetyl-CoA epoxidase alpha subunit
PP_3279 -0.05 0.13 0.17 0.05 0.29 0.03 0.05 0.04 -0.08 -0.02 -0.03 -0.05 -0.24 0.06 -4.8 -7.2 0.08 0.02 phenylacetate-CoA ligase
PP_3280 0.14 0.23 0.15 0.48 0.13 -0.36 0.41 0.12 -0.05 -0.29 0.35 -0.32 0.18 -0.53 -3.7 -6.4 -0.22 -0.19 3-oxoadipyl-CoA/3-oxo-5,6-dehydrosuberyl-CoA thiolase
PP_3281 -0.16 0.31 0.18 -0.19 0.14 -0.01 0.09 -0.03 -0.25 -0.08 -0 0.15 0.17 -0.11 -2.5 -3.2 -0.04 0.08 hydroxyphenylacetyl-CoA thioesterase
PP_3282 -0.05 -0.01 -0.04 -0.09 -0.03 -0.04 0.09 0.09 -0.06 -0.05 0.14 -0.15 -0.04 -0.14 -4.5 -7.6 0.07 0.11 3-hydroxyadipyl-CoA dehydrogenase
PP_3283 -0.03 -0.07 -0.21 0.24 0.15 -0.19 -0.17 0.08 -0.02 0.09 0.16 -0.11 -0.25 0.22 -4.8 -6.4 -0.09 -0.27 1,2-epoxyphenylacetyl-CoA isomerase
PP_3284 -0.02 0.34 0.26 0.1 0.22 0.27 -0.32 -0.04 -0.18 0.03 -0.01 -0.24 0.31 -0.01 -4.7 -7.8 -0.14 -0.15 enoyl-CoA hydratase-isomerase
PP_3285 0.19 0.09 0.1 0.09 -0.43 0.21 0.01 0.42 -0.04 0.07 -0.11 0.41 0.08 -0.28 -4.6 -7.4 0.18 -0.03 detoxifying thioesterase of the phenylacetate degradation pathway
PP_3356 0.02 0.13 -0.05 -0.11 -3.3 -4 -4.2 -3.6 -0.05 -0.06 0 -0.05 0.1 -0.03 0.08 -0.03 0.15 -0.06 feruloyl-CoA-synthetase
PP_3357 0.12 -0.1 0.18 0.06 -3.6 -2.5 -5.2 -4.2 0.24 -0.1 -0.04 -0.06 -0.03 0.23 0.09 0.17 0.14 0.13 vanillin dehydrogenase
PP_3358 0.09 0.12 0.05 0.2 -2.9 -2.5 -3.1 -3.8 0.17 0.07 -0 0.23 0.03 -0.02 -0.01 0.08 0.04 0.02 hydroxycinnamoyl-CoA hydratase-lyase
PP_3493 -1.4 -1 -1.5 -1.5 0.11 0.14 0.06 -0.05 0.16 -0.23 -1.7 -1.8 0.02 0.04 -2.7 -3.8 0.01 0.07 Transcriptional regulator, LysR family
PP_3494 -1.5 -0.93 -1.4 -1.1 0.01 -0.01 0 -0.15 -0.18 -0.16 -0.77 -0.95 -0.22 0.04 -3.5 -3.2 0.01 -0.19 conserved exported protein of unknown function
PP_3537 0.09 0.21 -0.05 -0.04 0.11 -0.11 -4.6 -4.8 -3.2 -3.6 0.02 -0.02 0.1 -0.02 0.15 -0.03 -0.05 -0.02 p-hydroxybenzoate hydroxylase
PP_3538 0.2 0.14 0.14 0.12 0.04 0.06 -4.6 -4.7 -3 -3.8 0 0.05 0.07 0.12 0.16 0.09 0.06 0.04 transcriptional regulator
PP_3590 -3 -3.2 -1.8 -2 -3.6 -2.9 -2.1 -1.8 -2.7 -3.5 -3.4 -3.3 -1.2 -1.6 -1.6 -2 -0.19 -0.08 D-lysine aminotransferase
PP_3603 -1.4 -1.3 -1.7 -2.2 -0.79 -1.2 -0.47 -0.76 -0.49 -1.2 -0.62 -0.55 -0.79 -0.44 -1.1 -1 -0.4 -0.32 Transcriptional regulator, GntR family
PP_3621 -0.36 -0.23 -1.7 -2.4 -0.15 -0.25 -0.35 -0.15 -0.26 0.12 0.07 -0.53 -0.12 -0.28 -0.02 0.36 0.18 -0.07 isoquinoline 1-oxidoreductase subunit alpha
PP_3622 -0.11 0.08 -1.9 -2.3 0.09 -0.04 -0.19 -0.24 0.03 0.01 0.08 0.03 -0.15 0.02 0.07 -0.07 -0.03 0.02 putative Isoquinoline 1-oxidoreductase, beta subunit
PP_3662 -1.5 -1.4 -1.8 -1.6 -1.1 -1.3 -0.5 -0.6 -1.1 -0.89 -1.9 -1.4 -0.72 -0.73 -1.7 -2.7 -0.41 -0.11 conserved protein of unknown function
PP_3714 0.19 0.34 -0.38 0.23 -0.38 -0.08 -0.03 -0.11 -0.09 0 -6.7 -5.6 -0.26 -0.51 -0.31 0.02 -0.47 -0.14 Muconolactone Delta-isomerase
PP_3715 0.18 0.75 -0.08 0.12 -0.18 0.11 -0.24 -0.13 0.06 0.28 -5.7 -6.9 -0.27 -0.28 0.34 -0.01 0.14 0.31 Muconate cycloisomerase 1
PP_3716 0.15 -0.13 -0.1 -0.05 -0.1 -0.02 -0.02 -0.09 -0.04 -0.04 -5.6 -6 -0.2 -0.07 -0.06 -0.31 0 0.07 HTH-type transcriptional regulator
PP_3736 -5.8 -6.4 -4.3 -7.2 -5.2 -6.2 -0.09 0.01 -0.03 0.08 -0.09 -0.32 0.14 0.25 0.12 -0.02 0.08 -0.02 vanillate O-demethylase oxygenase subunit
PP_3737 -4.8 -6 -4.5 -6 -4.4 -4.2 -0.04 0.14 0.04 -0.07 -0.04 -0.04 -0.01 0.18 0.11 -0.05 -0.24 0.03 vanillate O-demethylase oxidoreductase
PP_3828 -0.45 -0.54 -2.5 -3.9 -0.71 -0.35 -0.48 -0.64 -0.38 -0.24 -0.24 -0.33 -0.36 -0.28 -0.28 -0.52 -0.37 -0.07 Molybdate-binding periplasmic protein
PP_3829 -0.5 -0.45 -3.1 -3.3 -0.76 -0.75 -0.47 -0.69 -0.23 -0.19 -0.42 -0.01 -0.59 -0.27 -0.23 -0.23 -0.43 -0.12 Molybdenum transport system permease protein ModB
PP_3830 -0.09 -0.3 -1.7 -2.2 -0.29 -0.66 -0.32 -0.28 -0.17 -0.2 -0.2 -0.16 -0.29 -0.29 -0.23 -0.51 -0.1 -0 Molybdenum import ATP-binding protein ModC
PP_3951 -5.2 -7.3 -3.7 -6.4 -7.2 -7.1 -7.8 -8.1 -5.5 -5.6 -5.1 -8.1 -8 -7.9 -0.23 0.03 -0.25 0.21 3-oxoadipate CoA-transferase subunit A
PP_3952 -6.4 -6.6 -4.6 -8.2 -6.8 -7 -7.3 -7.6 -6.8 -6 -6 -8 -7.1 -7.4 0.12 -0.38 -0.28 -0.08 3-oxoadipate CoA-transferase subunit A
PP_4008 -0.74 -0.11 -0.78 -1.1 -1.1 -0.99 -0.81 -0.68 -0.43 -1 -0.46 -0.44 -0.48 -0.28 -1.2 -2.2 -0.25 0.1 ATP-dependent serine protease
PP_4058 -1.9 -2.2 -2.3 -1.9 -1.5 -1.6 -0.5 -0.3 -0.7 -0.6 -0.15 -0.06 -0.5 -0.51 -3.1 -3.6 -0.19 -0.03 1,4-alpha-glucan branching enzyme
PP_4098 -1.2 -1.7 -0.98 -0.7 -1.3 -0.43 -3.4 -0.57 -1.3 -1 -1.1 -0.79 -2.5 -0.48 -0.48 -0.97 0.01 -0.29 UvrABC system protein C
PP_4231 0.02 0.08 -1.9 -2.5 -0.01 -0.03 -0.3 0.17 -0.19 0.09 -0.08 -0.08 0.04 -0.13 -0.09 0.1 0.01 -0.12 putative Xanthine dehydrogenase accessory factor
PP_4597 -0.4 -0.46 -1.9 -2.9 -0.16 -0.73 -0.31 -0.03 -0.41 -0.06 0 -0.5 0.07 -0.31 -0.51 -0.21 -0.33 -0.08 Cyclic pyranopterin monophosphate synthase
PP_4622 -0.99 -0.99 -2.1 -1.1 -0.18 -0.17 -0.13 -0.12 -0.23 0.17 -0.06 -0.19 -0.49 -0.17 -0.81 -0.91 -0.04 -0.07 Hmg transcriptional repressor
PP_4655 -4 -5.1 -4.9 -6.5 -4.9 -5 -4.2 -4.9 -5.2 -3.9 0.1 0.19 -6.3 -5.1 0.23 0.02 0.03 0.05 protocatechuate 3,4-dioxygenase alpha chain
PP_4656 -3.9 -6.8 -5.4 -8 -5.1 -6.1 -6.8 -6 -4.5 -4.6 0.11 0.27 -6.5 -6.4 -0.32 -0.14 -0.31 0.02 protocatechuate 3,4-dioxygenase beta chain
PP_5085 -1.3 -1.3 -1.6 -0.85 -0.81 -0.63 -2.3 -2.5 -4.5 -4.4 -0.83 -0.57 -0.75 -0.54 -1.3 -1.3 0.34 0.27 malic enzyme B
PP_5331 0.9 0.62 0.34 0.91 0.47 0.51 -0.44 -0.69 -2 -2 -0.95 -1.6 0.22 0.26 -1.2 -0.52 -0.21 -0.41 Long-chain acyl-CoA thioester hydrolase
Vanillate_1
Vanillate_2
Vanillin_1
Vanillin_2
Ferulate_1
Ferulate_2
Coumarate_1
Coumarate_2
p-Hydroxybenzoate_1
p-Hydroxybenzoate_2
Benzoate_1
Benzoate_2
Protocatechuate_1
Protocatechuate_2
Phenylacetate_1
Phenylacetate_2
Glucose_1
Glucose_2

Figure 4.4. Heatmap of RB-TnSeq mutant fitness data.

Genes shown have fitness defects ($-\log_2 < -2$) and with statistical significance ($|t_{score}| > 4$) in at least one tested condition which show no significant fitness defects when grown on glucose.

4.3.2. Functional genomics to validate aromatic catabolisms of *P. putida*

While aromatic catabolism has been extensively studied in *P. putida*, essential genes implicated in these pathways have been described as recently as 2019 [20]. The genes involved in the first steps of coumarate catabolism reside in an operon with a putative acyl-CoA dehydrogenase (*PP_3354*) and a putative beta-ketothiolase (*PP_3355*), which have been proposed to be involved in an alternative catabolic pathway [174]. As functional redundancy in coumarate catabolism could result in loss of the type III polyketide precursor, coumaroyl-CoA, we sought to identify any pathways that could potentially impact product titers. To assay for genes involved in coumarate and related aromatic metabolisms, we grew a randomly barcoded transposon mutant (RB-TnSeq) library of *P. putida* KT2440 in minimal medium with a variety of different aromatic compounds often found in LH (p-coumarate, ferulate, benzoate, p-hydroxybenzoate, protocatechuate, vanillin, vanillate, phenylacetate) and glucose as sole carbon sources. The fitness of each gene was calculated by comparing the abundance of barcodes before versus after growth selection, using barcode sequencing (BarSeq) [19,40]. Negative values indicate that the gene was important for growth in that condition.

The results of the RB-TnSeq assay validated that the primary route for ferulate and coumarate catabolism is through the feruloyl/coumaroyl-CoA synthetase (Fcs) and the enoyl-CoA hydratase lyase (Ech) (Figure 3 and Figure 4). The genes in the proposed secondary pathway of coumarate and ferulate catabolism, *PP_3354* and *PP_3355*, had no significant fitness phenotype, indicating that these genes are likely not necessary for coumarate or ferulate catabolism [175].

Of all the known reactions depicted in the map of aromatic catabolism (Figure 3), we observed significant negative fitness values for all but three of their corresponding genes: *catA-1*, *catA-2*, and *pcaC*. Even though the fitness score was higher than our cutoff at -1.4, *catA-1* had a significant |t-score| of 6.2 when grown on benzoate. This reflects previous work demonstrating CatA-I is the preferred catechol 1,2-dioxygenase, while CatA-II acts as a "safety valve" to handle high intracellular concentrations of catechol. It was also demonstrated that both catA-I and catA-II need to be deleted to abolish growth on benzoate, and this likely explains why the fitness value for catA-I was above our cut-off [176]. In the case of *pcaC*, we noticed that there was a strong phenotype (fitness < -2.0), but the significance fell below our cutoff in all conditions tested ($|t_{score}| < 4.0$). This is likely due to the low frequency of transposon insertions into this gene in the library (n=4).

While our results heavily support the current models of aromatic metabolism in *P. putida*, our data also indicated that some gene annotations should be revised. *PP_3272* is currently annotated as encoding an acetate permease [177–179]. We observed that *PP_3272* had a significant phenotype when grown on phenylacetate (Figure 3). Given these data and previously described homology to other systems [180], *PP_3272* should be reannotated as the phenylacetate transporter (phaJ). The *PP_1376* gene is annotated as encoding a 4-hydroxybenzoate transporter [177–179]; however, we only observed a fitness

detriment for this gene with protocatechuate as the sole carbon source (Figure 3). Because of this, *PP_1376* should be reannotated as a protocatechuate transporter.

4.3.3. Accumulation of coumaroyl-CoA is toxic to *P. putida*

The first step in the biosynthetic pathway for bisdemethoxycurcumin is the activation of coumarate with coenzyme A (CoASH) (Figure 4.1). Because *Pseudomonas putida* KT2440 natively produces coumaroyl-CoA during coumarate catabolism, we knocked out the subsequent gene in the native catabolic pathway, *ech*, to prevent *P. putida* from consuming this necessary precursor (Figure 4.1) [180]. Initial production experiments in Δ*ech* strains overexpressing *fcs* and *CUS* from a synthetic operon resulted in an extended lag phase (data not shown). To determine the cause, we overexpressed *fcs* alone under control of the arabinose-inducible *araBAD* promoter (P_{BAD}) in the presence and absence of coumarate. Increasing the inducer concentration resulted in increased lag times only when coumarate was present in the medium (Figures 4A and S6). This suggested that the coumaroyl-CoA intermediate is toxic to *P. putida*.

To determine if coumaroyl-CoA concentrations are elevated in cultures expressing *fcs*, cultures carrying pBADT-*fcs* were induced with L-arabinose. After a 7 hour growth period, 5 mM coumarate was supplemented to the medium. Following supplementation, growth was stunted and remained relatively constant over the course of 11 hours. The concentration of coumaroyl-CoA was highest in samples one hour after coumarate was supplemented into the medium; this was observed regardless of the *ech* genotype (Figure 4B). We therefore sought to minimize the burden of this intermediate by relying on the native chromosomal expression of *fcs* in future experiments.

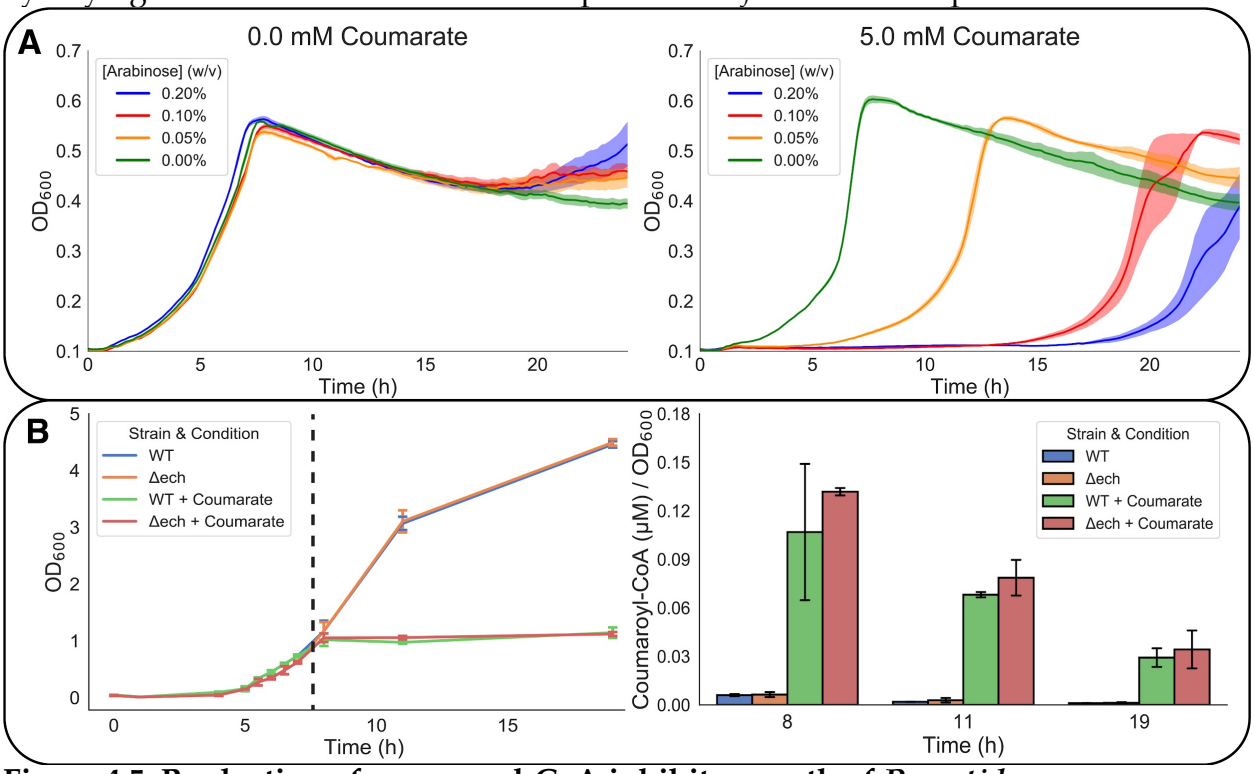


Figure 4.5. Production of coumaroyl-CoA inhibits growth of *P. putida*.

(A) 24 hour growth curves of KT2440 Δ*ech* carrying pBADT-*fcs* in LB induced with varying concentrations of arabinose and no added coumarate (left), or 5 mM coumarate

(right). Error bars are +/- the standard deviation of three replicates. (B) Growth curve (left) and normalized abundance of coumaroyl-CoA (right) in KT2440 (WT) and KT2440 Δ*ech* (Δ*ech*). Cultures expressing *fcs* were supplemented with 0 or 5 mM coumarate at hour 7 (marked with dashed line). Coumaroyl-CoA concentrations were normalized by OD_{600}. Error bars are +/- the standard deviation of three replicates.

4.3.4. Production of bisdemethoxycurcumin

In order to produce bisdemethoxycurcumin in *P. putida*, we constructed the biochemical pathway outlined in Figure 4.1. Exogenously added coumarate is activated by the native feruloyl/coumaroyl-CoA synthetase of *P. putida* (Fcs), then two resultant coumaroyl-CoA molecules are condensed with malonyl-CoA by the curcuminoid synthase of *O. sativa* (CUS) to yield bisdemethoxycurcumin [181]. The CoA synthetase, *fcs*, was expressed from its native chromosomal locus, while *CUS* was expressed from the pBADT plasmid and induced with L-arabinose [182].

Our initial production strategy was to induce *CUS* until the cultures reached stationary phase. Then the cultures were pelleted and resuspended in fresh LB medium supplemented with 5 mM coumarate. These samples were then incubated for another 72 hours. A similar approach had been used successfully in *E. coli* [183]; however, our titers were less than 0.5 mg/L (Figure 4.6). Given that bisdemethoxycurcumin is insoluble in water [184], a 10 % v/v oleyl alcohol overlay was used to extract the product as the fermentation progressed. Production levels were low (approximately 0.1 mg/L) in *P. putida* Δ*ech* when the medium was supplemented with 10 mM coumarate, likely due to the toxicity of coumaroyl-CoA (Figure 5). Supplementation with 5 mM coumarate resulted in a ~5-fold increase in bisdemethoxycurcumin titers in the *P. putida* Δ*ech* strain relative to wild-type (Figure 5).

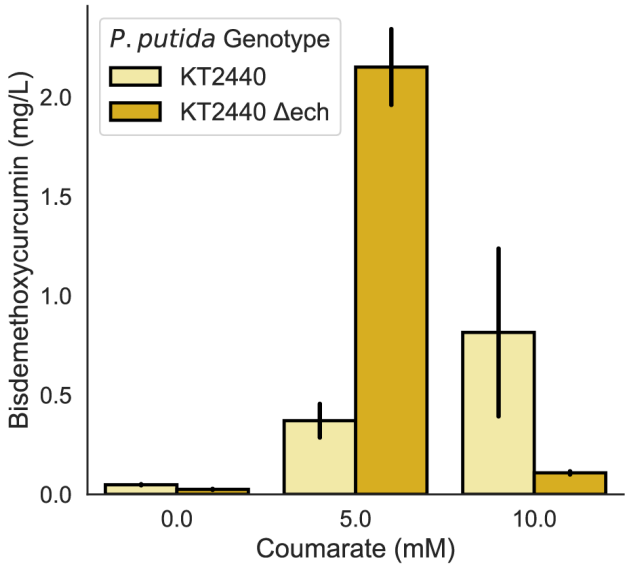


Figure 4.6: Bisdemethoxycurcumin titers in *P. putida* KT2440 Δ*ech* and in wildtype KT2440 harboring pBADT-CUS.

CUS was induced with 2 % w/v L-arabinose. Error bars are +/- the standard deviation of three biological replicates.

3. DISCUSSION

Pseudomonas putida is among the most well studied saprophytic bacteria. Its diverse metabolisms enable it to catabolize a wide variety of complex carbon sources, including lignocellulosic hydrolysate [185]. The robust catabolic pathways of *P. putida*, while useful for producing valuable molecules from diverse carbon sources, can also serve as an obstacle to achieving high product titers as it can often metabolize the desired products [10]. Given recent advances in gene editing techniques [4–6,186,187] and our ability to rapidly assay for gene function with transposon site sequencing [19,40,41], engineering non-model hosts like *P. putida* for industrial applications has become less challenging.

Using RB-TnSeq mutant fitness assays we were able to rapidly confirm entire pathways of aromatic catabolism (Figure 3 and S1). The genes downstream of *fcs*, *PP_3354* and *PP_3355* previously described as a possible alternative route for coumarate/ferulate catabolism [174,180,188], showed no significant fitness detriment on any of the carbon sources tested. These genes may be structural remnants of a β-oxidation pathway that eventually evolved into the coumarate/ferulate pathway requiring *fcs*, *ech*, and *vdh* [174,180,188]. We then proposed revised annotations of two genes required for transport of the aromatic compounds, phenylacetate (*PP_3272*) and protocatechuate (*PP_1376*), an important plant hormone and lignin metabolite respectively [189]. There is a large amount of information in these data about regulatory and structural genetic elements that could be useful to engineers and biologists.

Heterologous expression of bacterial T3PKSs, including the tetrahydroxynapthalene synthase of *S. coelicolor* (RppA), has previously been demonstrated in *P. putida* KT2440 [171,172]. Using variants of *rppA*, we were able to rapidly screen for optimal T3 polyketide production conditions. Expressing the codon

optimized variant we created in this study, *rppA*-OW, resulted in less flaviolin production and less protein production than the native codon variants *rppA*-NW and *rppA*-NT. It is possible that there are some factors affecting heterologous protein expression that are not sufficiently accounted for in current codon optimization algorithms [190]. However, we demonstrated in all our constructs that increasing the glucose concentration had a considerable effect on the production of flaviolin (Figure 2 and S4).

Higher flaviolin accumulation with increased glucose concentrations could be due to increased flux to malonyl-CoA, a known limiting reactant to the biosynthesis of plant T3PKS products [161,191], or due to increased functional protein expression. As the relative abundance of RppA increased when the culture was grown in the presence of 100 mM glucose, we hypothesize this likely contributed to the increased flaviolin titer observed in this condition (Figure 2B and 2C). The exact cause for the increase in protein abundance requires more detailed investigation. These T3PKS "sensors" have broad utility in both rapidly assaying culture conditions, as described here, and as high-throughput screens of genetic libraries for increased malonyl-CoA accumulation [171]. Future work could employ these sensors to screen for increased intracellular malonyl-CoA concentrations from a complex growth medium like LH.

To provide the coumaroyl-CoA substrate for the bisdemethoxycurucumin T3PKS, CUS, we sought to use the native CoA synthetase (Fcs) of *P. putida*. Plasmid-based induction of *fcs* expression in the presence of coumarate, however, resulted in an increase in lag times due to the build up of the toxic coumaroyl-CoA intermediate (Figures 4A, 4B, and S6). Neither the accumulation of coumaroyl-CoA nor the resultant growth inhibition seemed to be significantly affected by the genetic presence of *ech* (Figure 4B). The reduction in intracellular coumaroyl-CoA concentrations following the addition of coumarate was also not dependent on the presence of *ech* (Figure 4B). This may be the result of nonspecific thioesterase activity, or spontaneous hydrolysis of coumaroyl-CoA.

This toxicity has been observed in *Acinetobacter baylyi* following disruptions in the gene encoding its enoyl-CoA hydratase lyase, and it was observed in *E. coli* expressing the *A. baylyi fcs* homolog in the presence of coumarate [192]. This toxicity could have been at play in other systems using bacterial coumaroyl-CoA synthetases, but the defective growth phenotype may not have been observed due to differences in experimental design [193,194]. The exact cause for coumaroyl-CoA toxicity remains unclear and will be the subject of future investigations.

To engineer *P. putida* for bisdemethoxycurcumin production, we deleted the native enoyl-CoA hydratase lyase (*ech*) responsible for the conversion of coumaroyl-CoA to acetyl-CoA and p-hydroxybenzaldehyde. In order to relieve some coumaroyl-CoA toxicity, we relied on the native genomic copy of *fcs* instead of a plasmid-based system. We demonstrated that native expression of *fcs* generates sufficient coumaroyl-CoA for curcuminoid synthase (CUS). Extraction of the product during growth using an oleyl alcohol overlay also significantly enhanced titers (Figure 4.7 and 5). In the final *P. putida* production strain, we achieved production of bisdemethoxycurcumin at titers of 2.15 mg/L (Figure 5).

This work is a significant first step towards the production of plant T3PKS-derived compounds in *P. putida*, but several issues remain to be addressed. In particular, our final bisdemthoxycurcumin titer of 2.15 mg/L is far lower than the 91.3 mg/1 titer first described in *E. coli* [161]. Several methods have been proposed to enhance the production of curcuminoids and similar T3PKS derived compounds in microbial

hosts. In one such approach, the metabolic burden of product, protein, and plasmid DNA synthesis was spread across several microbial strains in a synthetic co-culture [195]. Since phenylpropanoid overproducing strains of *P. putida* have been described [196,197], a co-culture method could be explored for the biosynthesis of plant T3PKS products in this host. It is also quite clear that more extensive pathway balancing will be required to mitigate the intermediate toxicity we have observed. As we continue to understand more about the metabolism of *P. putida* and learn to engineer more effectively,

4.4. Methods

4.4.1. Media, chemicals, and culture conditions

General *E. coli* cultures were grown in Luria-Bertani (LB) Miller medium (BD Biosciences, USA) at 37 °C, while *P. putida* was grown at 30 °C. MOPS minimal medium was used where indicated and comprised of the following: 32.5 μM $CaCl_2$, 0.29 mM K_2SO_4, 1.32 mM K_2HPO_4, 8 μM $FeCl_2$, 40 mM MOPS, 4 mM tricine, 0.01 mM $FeSO_4$, 9.52 mM NH_4Cl, 0.52 mM $MgCl_2$, 50 mM NaCl, 0.03 μM $(NH_4)_6Mo_7O_{24}$, 4 μM H_3BO_3, 0.3 μM $CoCl_2$, 0.1 μM $CuSO_4$, 0.8 μM $MnCl_2$, and 0.1 μM $ZnSO_4$ [75]. Cultures were supplemented with kanamycin (50 mg/L, Sigma Aldrich, USA) when indicated. Technical grade oleyl alcohol was acquired from Alfa Aesar (Alfa Aesar, Thermo Fisher Scientific). Coenzyme A (lithium salt) was purchased from CoALA Biosciences (CoALA Biosciences, USA). All other compounds were purchased through Sigma Aldrich (Sigma Aldrich, USA). Construction of *P. putida* deletion mutants was performed as described previously [18].

4.4.2. Strains and plasmids

Bacterial strains and plasmids used in this work are listed in Table 1. All strains and plasmids created in this work are available through the public instance of the JBEI registry https://public-registry.jbei.org. All plasmids were designed using Device Editor and Vector Editor software, while all primers used for the construction of plasmids were designed using j5 software [76–78]. Plasmids were assembled via Gibson Assembly using standard protocols [79], or Golden Gate Assembly using standard protocols [80]. Plasmids were routinely isolated using the Qiaprep Spin Miniprep kit (Qiagen, USA), and all primers were purchased from Integrated DNA Technologies (IDT, Coralville, IA).

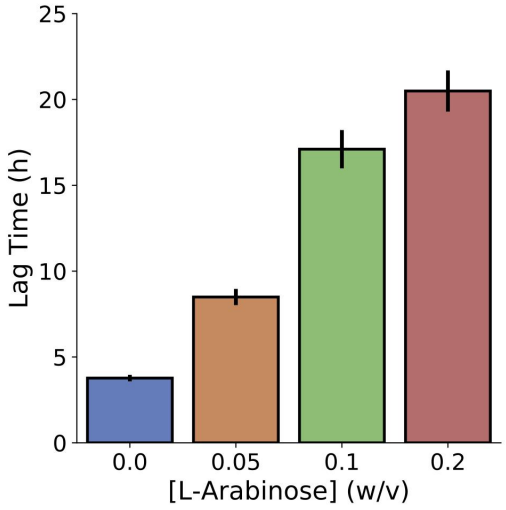
Table 4.1. Strains and plasmids used in this study

Strain	**JBEI Part ID**	**Reference**
E. coli DH10B		[198]
E. coli BL21(DE3)		Novagen
P. putida KT2440		ATCC 47054
P. putida KT2440 Δ*ech*		
Plasmids		
pET28		Novagen
pBADT		[182]
pBADT-*CUS*	JBx_134311	This work
pBADT-*rppA*-NT	JBx_134312	This work
pBADT-*rppA*-NW	JBx_134313	This work
pBADT-*rppA*-OW	JBx_134314	This work
pBADT-*fcs*	JBx_134315	This work
pBADT-*fcs*-*CUS*	JBx_134316	This work
pMQ30		[81]
pMQ30-*PP_3358*		[199]

4.4.3. Plate based growth assays

Growth studies of bacterial strains were conducted using microplate reader kinetic assays. Overnight cultures were inoculated into 10 mL of LB medium from single colonies, and grown at 30 °C. These cultures were then diluted 1:100 into 500 μL

of LB medium with appropriate concentrations of arabinose and p-coumarate in 48-well plates (Falcon, 353072). Plates were sealed with a gas-permeable microplate adhesive film (VWR, USA), and then optical density and fluorescence were monitored for 48 hours in an Biotek Synergy 4 plate reader (BioTek, USA) at 30 °C with fast continuous shaking. Optical density was measured at 600 nm. The amount of time necessary for the culture to reach an OD_{600} of 0.16 was defined as the lag time.



[L-arabinose] (w/v)	Lag Time +/− s.d. (hrs)
0%	3.77 +/− 0.10
0.05%	8.50 +/− 0.44
0.10%	17.10 +/− 1.23
0.20%	20.49 +/− 1.33

Figure 4.7: Calculated lagtimes for cultures expressing *fcs* in the presence of 5 mM coumarate.

The *fcs* gene was expressed with 0, 0.05, 0.1, or 0.2% w/v L-arabinose. The lagtimes and their standard deviations are included in the table (right). Lag time is defined here as the amount of time required for the culture to reach OD_{600} = 0.16. Standard deviation was calculated from 3 replicates

4.4.4. HPLC detection of bisdemethoxycurcumin

HPLC analysis was performed on an Agilent Technologies 1200 series liquid chromatography instrument coupled to a Diode Array Detector (Agilent Technologies, USA). Compounds were separated at a constant flow rate of 0.4 mL/min over a Kinetex C18 column (2.6 μm diameter, 100Å particle size, dimensions 100 x 3.00 mm, Phenomenex, USA) held at 50 °C. The mobile phase consisted of H_2O + 0.1% trifluoroacetic acid (A) and acetonitrile + 0.1% trifluoroacetic acid (B). Separation was performed using the following gradient method: 0-3 minutes 95% A, 3-15 minutes 95-5% A, 15-17 minutes 5% A, 17-17.5 minutes 5-95% A, 17.5-20 minutes 95% A. The presence of bisdemethoxycurcumin was monitored and quantified at 440 nm.

4.4.5. RB-TnSeq experiments and analysis

BarSeq-based experiments utilized the *P. putida* RB-TnSeq library, JBEI-1, which has been described previously [18]. An aliquot of JBEI-1 was thawed on ice, diluted into 25 mL of LB medium supplemented with kanamycin and grown to an OD_{600} of 0.5 at 30 °C. Three 1 mL aliquots of the library were pelleted and stored at -80 °C to later serve as the

t_0 of gene abundance. Libraries were then washed in MOPS minimal medium, and diluted 1:50 in MOPS minimal medium with 10 mM p-coumarate, ferulate, benzoate, p-hydroxybenzoate, protocatechuate, vanillin, vanillate, phenylacetate, or D-glucose. Cells were grown in 600 µL of medium in 96-well deep well plates (VWR). Plates were sealed with a gas-permeable microplate adhesive film (VWR, USA), and then grown at 30 °C in an INFORS HT Multitron (Infors USA Inc.), with shaking at 700 rpm. Two 600-µL samples were combined, pelleted, and stored at -80 °C until analysis by BarSeq , which was performed as previously described [19,40]. All fitness data are publically available at http://fit.genomics.lbl.gov.

4.4.6. Curcuminoid production

For production of bisdemthoxycurcumin without an overlay, an overnight culture of *P. putida* KT2440 Δ*ech* + pBADT-CUS was diluted 1:100 into 5 mL of LB supplemented with 50 mg/L kanamycin, 1 % w/v L-arabinose, and 100 mM glucose. The culture was grown to stationary phase over 12 hours then pelleted in a centrifuge at 5000 xg for 5 minutes. The cell pellets were resuspended in 2 mL of fresh LB with 50 mg/L kanamycin, 0.5 % L-arabinose, 100 mM glucose, and 5 mM coumarate. The culture was allowed to proceed for 72 hours.

Overnights of *P. putida* harboring pBADT-CUS were diluted 1:100 into 25 mL of fresh LB supplemented with 50 mg/L kanamycin, and 100 mM glucose. Arabinose and coumarate were added at the beginning of the fermentation at concentrations indicated. A 2.5 mL overlay of oleyl alcohol was added to extract the bisdemethoxycurcumin during growth. The fermentation was allowed to proceed for 72 hours.

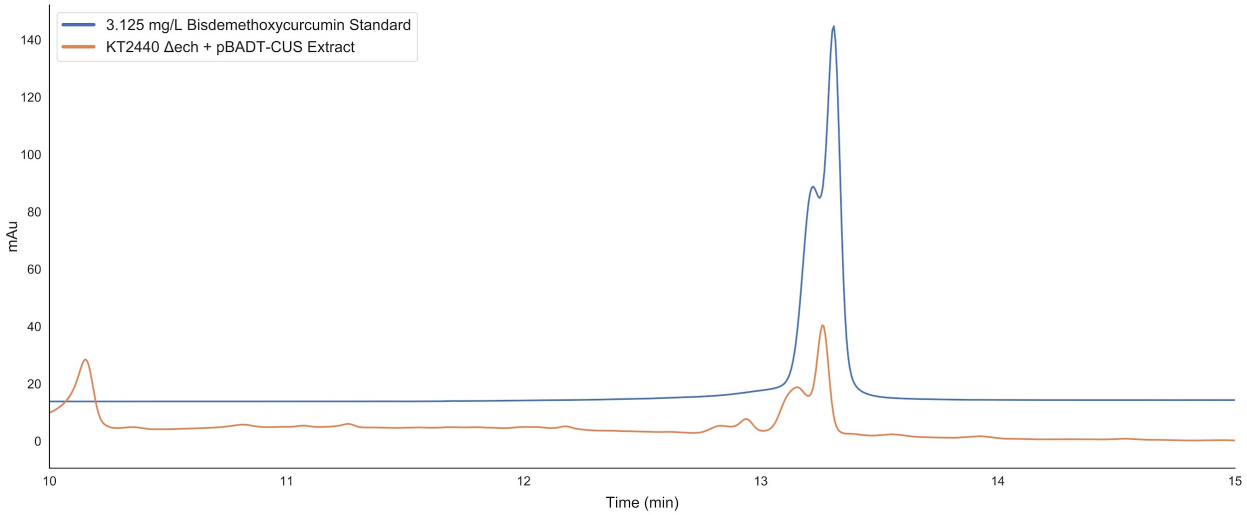


Figure 4.8. HPLC trace of curcumin extracted from *P. putida* KT2440 Δ*ech* + pBADT-CUS.

Overlaid with a 3.125 mg/L standard

4.4.7. Curcuminoid extraction

For cultures lacking an overlay, 0.5 mL of culture was acidified to pH 3 with 3 N HCl. Bisdemethoxycurcumin was then extracted with an equal volume of ethyl acetate. 250 µL of the ethyl acetate layer was removed and the solvent was allowed to evaporate

overnight. The dried samples were then resuspended in 50 µL of acetonitrile for analysis with HPLC-DAD.

For cultures with an overlay, the cultures were acidified to pH 3 with 3 N HCl. Acidified cultures were then pelleted in a centrifuge and the oleyl alcohol overlays were removed. To quantify bisdemethoxycurcumin, 100 µL of the extracted overlays were added to a black, clear bottom 96-well plate and absorbance was measured at 425 nm in a Biotek Synergy 4 plate reader (BioTek, USA). A standard curve was made with bisdemethoxycurcumin standards dissolved in oleyl alcohol (Figure 4.7)

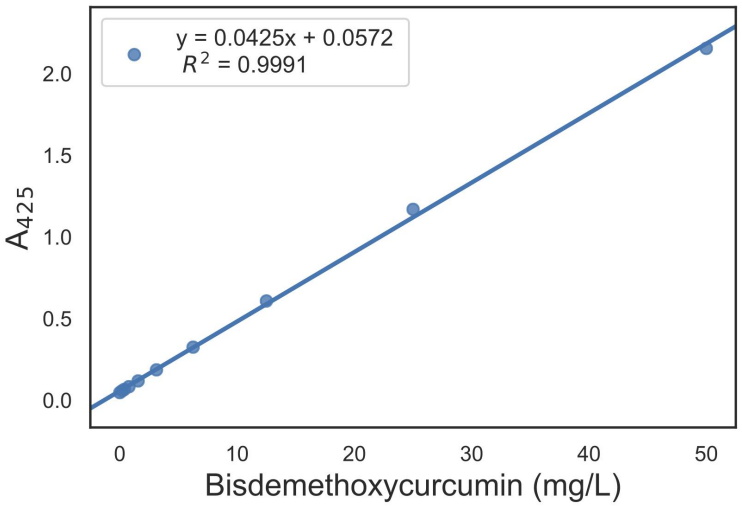


Figure 4.9. Standard curve of bisdemethoxycurcumin.

Absorbance measurements at 425 nm from 50 to 0 mg/L of bisdemethoxycurcumin dissolved in oleyl alcohol.

4.4.8. Flaviolin production and targeted proteomic analysis

Colonies of *P. putida* KT2440 strains carrying pBADT-*rppA*-NW, pBADT-*rppA*-NT, and pBADT-*rppA*-OW were used to inoculate LB or MOPS minimal medium with 50 mg/L kanamycin and cultured overnight. The overnight culture was then diluted 1/100 into fresh LB with 0.2 % w/v L-arabinose, 50 mg/L kanamycin, and 100, 50, 25, 12.5, 6.25, 3.125, 1.5625 or 0 mM glucose. Cultures were conducted in 24-well deep-well plates, and allowed to proceed for 48 hours. Cultures were pelleted in a centrifuge at 5000 xg for 5 minutes. Supernatants were removed, aliquoted into a 96-well black clear bottom plate, and the absorbance was measured at 340 nm in a Biotek Synergy 4 plate reader (BioTek, USA). For cultures expressing *rppA* variants in LB, the supernatants were too opaque with red product and needed to be diluted 1:5 in fresh LB for accurate absorbance measurements. Supernatants of strains cultured in minimal medium were not diluted before absorbance measurements.

Samples for intracellular targeted proteomic analysis were grown for 48 hours in 5 mL LB supplemented with 0 mM or 100 mM glucose. Then, 2 mL of the culture were then pelleted by centrifugation, the supernatant was decanted, and the pellets were stored at -80 °C. Proteins were extracted and analyzed using a variation of a previously-described workflow [200].

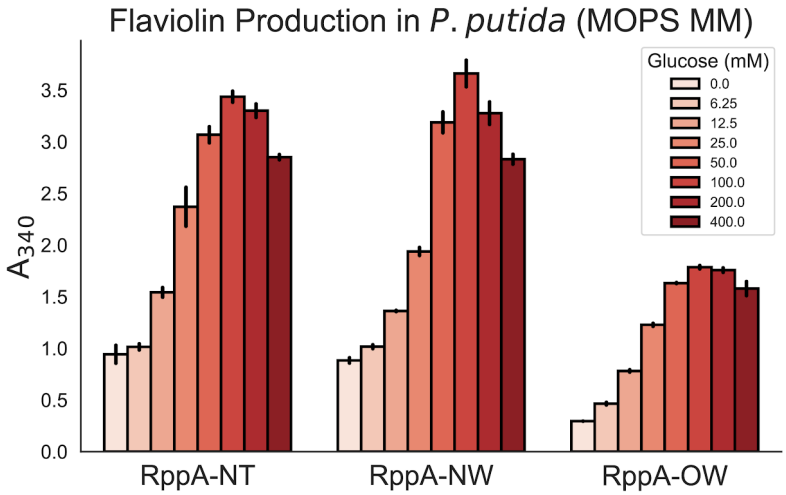


Figure 4.10. Production of flaviolin in *P. putida* KT2440 measured with absorbance at 340 nm.

Glucose concentrations were varied from 0 mM to 400 mM in MOPS minimal medium.

4.4.9. Flaviolin purification

An overnight culture of *P. putida* KT2440 carrying pBADT-rppA-NT was diluted 1:100 into 100 mL of LB supplemented with 25 mg/L kanamycin, 0.2% *w*/*v* *L*-arabinose, and 100 mM glucose. The culture was shaken for 48 hours at 30 °C in a 500 mL baffled Erlenmeyer flask. Following fermentation, the red supernatant was acidified with 10 mL 10 *N* HCl and extracted twice with equal volumes of ethyl acetate. The combined organic phase was evaporated leaving a black solid, which was redissolved in ethyl acetate and methanol (7:3). Flaviolin was purified using preparative thin layer chromatography (TLC, silica gel). The red band was isolated from the TLC plate, extracted with methanol, and filtered. The methanol was evaporated *in vacuo* leaving dark red solid flaviolin with a yield of ~6.5 mg. The identity of the flaviolin was validated with ^{1}H-NMR. ^{1}H-NMR (400 MHz, CD_3OD) δ 6.99 (s, 1 H), 6.47 (s, 1 H), 5.58 (s, 1 H) (Figure 4.9).

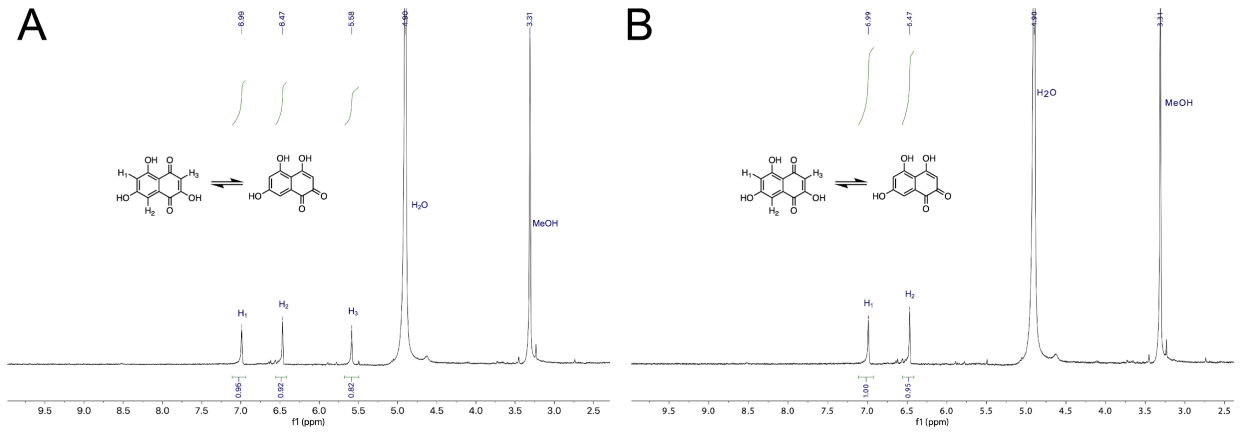


Figure 4.11. ^{1}H-NMR spectrum of flaviolin isolated from *P. putida* KT2440 expressing *rppA*-NT from the plasmid pBADT-*rppA*-NT.

A) ^{1}H-NMR spectrum from flaviolin freshly resuspended in CD_3OD. B) ^{1}H-NMR spectrum from flaviolin resuspended in CD_3OD and incubated at room temperature overnight. The signal at 5.58 (s, 1 H) is absent in the overnight treatment due to proton exchange with CD_3OD [201].

4.4.10. Synthesis of coumaroyl-CoA standard

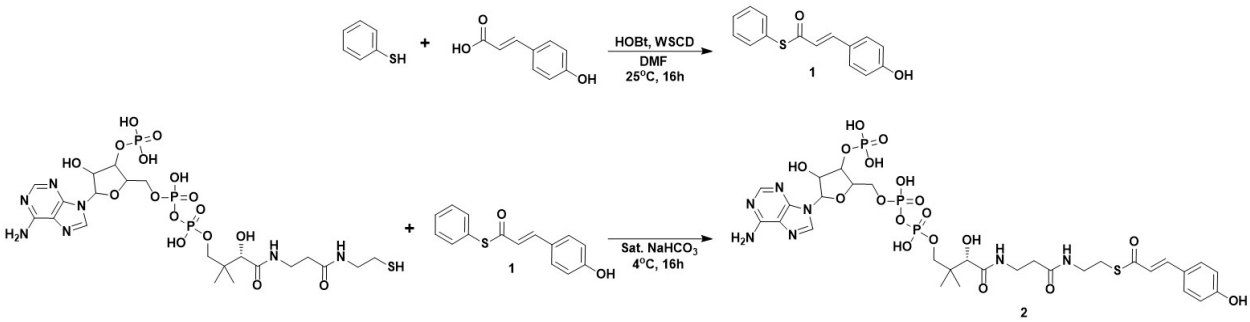


Figure 4.12. Synthesis of coumaroyl-CoA

Coumaroyl-Coenzyme A (**2**) was synthesized in two steps according to Scheme X. *S*-phenyl (*E*)-3-(4-hydroxyphenyl)prop-2-enethioate (**1**) was synthesized according to previous literature [202]. 1-(3-Dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (WSCD·HCl, 1.48 g, 7.3 mmol), 1-hydroxybenzotriazole (HOBt, 1.12 g, 7.3 mmol), and 4-hydroxycinnamic acid (1 g, 6.1 mmol) were dissolved in DMF (100 mL) and stirred for 1 h at room temperature. Thiophenol (0.610 mL, 6.1 mmol) was subsequently added to the mixture and stirred overnight at room temperature. After the solvent was removed *in vacuo*, water and dichloromethane (DCM) were added to the mixture. The organic layers were combined, dried over Na_2SO_4, and evaporated. The product was purified with silica-gel column chromatography and eluted with DCM to yield *S*-phenyl (*E*)-3-(4-hydroxyphenyl)prop-2-enethioate as a white solid (**1**, 468 mg, 30%).

To synthesize coumaroyl-CoA (**2**), coenzyme A lithium salt (7.7 mg, 10 μmol) was dissolved in a solution of saturated sodium bicarbonate (750 μL) which had been pre-chilled to 4 °C. *S*-phenyl (*E*)-3-(4-hydroxyphenyl)prop-2-enethioate (30 mg, 117 μmol) was added to the solution. The solution was mixed for 16 hours at 4 °C then quenched via the addition of 6M HCl to bring the pH of the solution to ~2. The mixture was extracted with ethyl acetate (2 x 2 mL), then coumaroyl-CoA was purified from the remaining aqueous phase via preparatory HPLC.

Coumaroyl-CoA was purified using an Agilent 1260 series preparatory HPLC system equipped with a 900 μL sample loop and a UV detector coupled to a fraction collector. Compounds were separated over an Agilent Zorbax SB-C18 preparatory column (9.4 x 50 mm, 5 μm particle size) using a mobile phase composed of 10 mM ammonium formate, pH 4.5 (solvent A) and methanol (solvent B) using the following method:

Time (min)	**%A**	**%B**	**Flow rate (mL/min)**
0	95	5	5
1	95	5	5
8	5	95	5
10	5	95	5
11	95	5	5
14.5	95	5	5

Table 4.2. HPLC protocol for coumaroyl-CoA separation

The fraction collector was programmed to collect peaks absorbing at a wavelength of 260 nm and that eluted between 3.6 and 4.6 minutes. The fractions containing coumaroyl-CoA were identified by direct infusion onto an Applied Biosystems 4000 QTRAP mass spectrometer system operating with the following parameters: Q1 MS mode, scan range 100-1000 m/z, declustering potential: 70, entrance potential: 10, curtain gas: 10, IonSpray Voltage: 4800, Temperature: 300, Ion Source Gases: 40. Fractions containing coumaroyl-CoA were combined and lyophilized, yielding coumaroyl-CoA as a white powder (6 mg)

4.10 LC-MS/MS analysis of intracellular Coenzyme A esters

P. putida KT2440 and *P. putida* Δ*ech* cells carrying pBADT-*fcs* were cultured overnight. Cultures were inoculated 1:100 into 25 mL LB supplemented with 100 mM glucose and 0.2% w/v L-arabinose. Cells were cultured at 30 °C for 7 hours then coumarate was added to a final concentration of 5 mM. 10 mL of culture were taken at hours 8, 11, and 19 and pelleted in a centrifuge. Cell pellets were resuspended in 1 mL ice cold LC grade methanol and stored at -80 °C until workup.

Coenzyme A esters were extracted as follows: to the 1 mL methanol resuspensions, 0.5 mL of ice cold LC grade water and 0.25 mL ice cold LC chloroform. Samples were vortexed to mix after each addition. Samples were then centrifuged at maximum speed for 2 minutes to promote phase separation, after which 0.5 mL of the top aqueous phase was removed and transferred to an Amicon 3 kDa molecular weight cutoff centrifuge filter (Sigma Aldrich, USA). Extracts were filtered by spinning in a benchtop centrifuge at 4 °C for 90 minutes at 13000g. The flow through was lyophilized for 24 hours to concentrate the sample. Lyophilized fractions were resuspended in 0.1 mL of ice cold 60:40 (v/v) LC grade acetonitrile: LC grade water, and 35 μL of the aqueous layer was removed for LC-MS/MS analysis.

CoA extracts were analyzed using an Applied Biosystems 4000 QTRAP mass spectrometer coupled to an Aglient 1100-1200 series HPLC. Compounds were separated over a Phenomenex Kinetex XB-C18 column (100 x 3 mm, 100 Å, 2.6 μm particle size) held at 50 °C. The mobile phase was composed of 10 mM ammonium formate, pH 4.5 (solvent A) and acetonitrile (solvent B) the following LC method was used:

Time (min)	**%A**	**%B**	**Flow rate (μL/min)**
0	98	2	400
2	98	2	400
5.5	2	98	400
8.5	2	98	400
9.5	98	2	400
12.75	98	2	400

Table 4.3 HPLC protocol for validation of synthesized coumaroyl-CoA standard via mass spectrometry

The mass spectrometer was operated in positive ion mode using the following parameters: Curtain gas: 20 L/min, Collision gas: medium, Ionspray voltage: 4500 V, Temperature: 250 °C, Ion source gas 1: 20 L/min, Ion source gas 2: 10 L/min,

Declustering potential: 0, Entrance potential: 10, Collision cell exit potential: 10. Coumaroyl-CoA was detected in multiple reaction monitoring (MRM) mode by monitoring the following transition: 914.16 m/z → 407.16 m/z using a collision energy of 47.8 V. Data was imported into Skyline targeted mass spectrometry environment (v. 3.7.0 11317) and peaks were integrated using Skyline's embedded integration function. Coumaroyl-CoA was quantified relative to a standard curve that was generated by calculating the linear regression of 3 individual injections of each concentration within the standard curve ($R^2 = 0.9982$).

4.5. Acknowledgements

We would like to thank Professor Mattheos Koffas for providing us with the plasmid pEMT6-4CL-CUS. We thank Jesus Barajas for his careful reading of this manuscript and helpful suggestions during preparation. We also thank Morgan Price for assistance in analyzing RB-TnSeq data.

This work was part of the DOE Joint BioEnergy Institute (https://www.jbei.org) supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, supported by the U.S. Department of Energy, Energy Efficiency and Renewable Energy, Bioenergy Technologies Office, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy. The views and opinions of the authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan)

Contributions

Conceptualization, M.R.I., M.G.T., J.M.B; Methodology, M.R.I., M.G.T., J.M.B.,Y.L.,C.J.P.; Investigation, M.R.I., M.G.T., J.M.B., Y.L., M.S., A.N.P., J.W.G.; Writing – Original Draft, M.R.I.; Writing – Review and Editing, All authors.; Resources and supervision, C.J.P, A.M.D., J.D.K.

Competing Interests

J.D.K. has financial interests in Amyris, Lygos, Demetrix, Napigen and Maple Bio

Chapter 5. Conclusion

This thesis has thoroughly investigated the biotechnological potential of Pseudomonas alloputida (formerly P. putida KT2440), focusing on its ability to catabolize a wide range of carbon sources and its genetic tractability. The study aimed to improve the fundamental understanding of the genes, pathways, and proteins involved in P. putida KT2440's metabolism of inexpensive feedstocks, such as lignin monomers, and their conversion into valuable biomolecules like alcohols and fatty acids.

The research identified the limitations of general gene annotations due to the presence of paralogous genes and the resulting impact on the predictive power of metabolic models. To overcome these challenges, the study employed a high-throughput functional genomics method, barcode abundance sequencing (BarSeq or RB-TnSeq), to reveal conditionally essential genes involved in the metabolism of feedstocks and desired bioproducts.

Utilizing the data obtained from BarSeq, an engineered pathway was developed to produce a curcuminoid by employing some native genes. Furthermore, the study elucidated genes involved in the metabolism of various fatty acids, fatty alcohols, and nitrogen sources, resulting in a comprehensive dataset containing over 1000 genes with significant phenotypes.

To enhance the visualization and interpretation of this extensive dataset, we employed, t-stochastic neighbor embedding, which facilitated a deeper understanding of the functional roles of these genes in *P. putida* KT2440. The findings of this thesis contribute to the knowledge base necessary for optimizing the biotechnological applications of *P. putida* KT2440 and pave the way for future metabolic engineering efforts in this promising microorganism.

Chapter 6. References

Bibliography

1. Park, M. *et al.* Response of *Pseudomonas putida* to Complex, Aromatic-Rich Fractions from Biomass. *ChemSusChem* **13**, 1–14 (2020).
2. Belda, E. *et al.* The revisited genome of Pseudomonas putida KT2440 enlightens its value as a robust metabolic chassis. *Environ. Microbiol.* **18**, 3403–3424 (2016).
3. Tiso, T. *et al.* Bio-upcycling of polyethylene terephthalate. *BioRxiv* (2020) doi:10.1101/2020.03.16.993592.
4. Aparicio, T., de Lorenzo, V. & Martínez-García, E. CRISPR/Cas9-enhanced ssDNA recombineering for Pseudomonas putida. *Microb. Biotechnol.* **12**, 1076–1089 (2019).
5. Aparicio, T., de Lorenzo, V. & Martínez-García, E. CRISPR/Cas9-Based Counterselection Boosts Recombineering Efficiency in Pseudomonas putida. *Biotechnol. J.* **13**, e1700161 (2018).
6. Cook, T. B. *et al.* Genetic tools for reliable gene expression and recombineering in Pseudomonas putida. *J. Ind. Microbiol. Biotechnol.* **45**, 517–527 (2018).
7. Banerjee, D. *et al.* Genome-scale metabolic rewiring to achieve predictable titers rates and yield of a non-native product at scale. *BioRxiv* (2020) doi:10.1101/2020.02.21.954792.
8. Blank, L. M., Ionidis, G., Ebert, B. E., Bühler, B. & Schmid, A. Metabolic response of Pseudomonas putida during redox biocatalysis in the presence of a second octanol phase. *FEBS J.* **275**, 5173–5190 (2008).
9. Ebert, B. E., Kurth, F., Grund, M., Blank, L. M. & Schmid, A. Response of Pseudomonas putida KT2440 to increased NADH and ATP demand. *Appl. Environ. Microbiol.* **77**, 6597–6605 (2011).
10. Thompson, M. G. *et al.* Omics-driven identification and elimination of valerolactam catabolism in Pseudomonas putida KT2440 for increased product titer. *Metab. Eng. Commun.* **9**, e00098 (2019).
11. Incha, M. R. *et al.* Leveraging host metabolism for bisdemethoxycurcumin production in Pseudomonas putida. *Metab. Eng. Commun.* **10**, e00119 (2020).
12. Zhang, M. *et al.* Increased glutarate production by blocking the glutaryl-CoA dehydrogenation pathway and a catabolic pathway involving L-2-hydroxyglutarate. *Nat. Commun.* **9**, 2114 (2018).
13. Dong, J. *et al.* Methyl ketone production by Pseudomonas putida is enhanced by plant-derived amino acids. *Biotechnol. Bioeng.* **116**, 1909–1922 (2019).
14. Tiso, T. *et al.* Designer rhamnolipids by reduction of congener diversity: production and characterization. *Microb. Cell Fact.* **16**, 225 (2017).
15. Kohlstedt, M. *et al.* From lignin to nylon: Cascaded chemical and biochemical conversion using metabolically engineered Pseudomonas putida. *Metab. Eng.* **47**, 279–293 (2018).
16. Loeschcke, A. & Thies, S. Engineering of natural product biosynthesis in Pseudomonas putida. *Curr. Opin. Biotechnol.* **65**, 213–224 (2020).
17. Nogales, J. *et al.* High-quality genome-scale metabolic modelling of *Pseudomonas putida* highlights its broad metabolic capabilities. *Environ. Microbiol.* **22**, 255–269 (2020).
18. Thompson, M. G. *et al.* Massively parallel fitness profiling reveals multiple novel enzymes in Pseudomonas putida lysine metabolism. *MBio* **10**, (2019).

19. Rand, J. M. *et al.* A metabolic pathway for catabolizing levulinic acid in bacteria. *Nat. Microbiol.* **2**, 1624–1634 (2017).
20. Price, M. N. *et al.* Oxidative pathways of deoxyribose and deoxyribonate catabolism. *mSystems* **4**, (2019).
21. Zhu, Z. *et al.* Multidimensional engineering of Saccharomyces cerevisiae for efficient synthesis of medium-chain fatty acids. *Nat. Catal.* **3**, 64–74 (2020).
22. Luo, X. *et al.* Complete biosynthesis of cannabinoids and their unnatural analogues in yeast. *Nature* **567**, 123–126 (2019).
23. Mehrer, C. R., Incha, M. R., Politz, M. C. & Pfleger, B. F. Anaerobic production of medium-chain fatty alcohols via a β-reduction pathway. *Metab. Eng.* **48**, 63–71 (2018).
24. Karp, P. D. *et al.* The BioCyc collection of microbial genomes and metabolic pathways. *Brief. Bioinformatics* **20**, 1085–1093 (2019).
25. Wehrmann, M., Billard, P., Martin-Meriadec, A., Zegeye, A. & Klebensberger, J. Functional Role of Lanthanides in Enzymatic Activity and Transcriptional Regulation of Pyrroloquinoline Quinone-Dependent Alcohol Dehydrogenases in Pseudomonas putida KT2440. *MBio* **8**, (2017).
26. Wehrmann, M., Toussaint, M., Pfannstiel, J., Billard, P. & Klebensberger, J. The Cellular Response to Lanthanum Is Substrate Specific and Reveals a Novel Route for Glycerol Metabolism in Pseudomonas putida KT2440. *MBio* **11**, (2020).
27. Nikel, P. I. & de Lorenzo, V. Robustness of Pseudomonas putida KT2440 as a host for ethanol biosynthesis. *N. Biotechnol.* **31**, 562–571 (2014).
28. Cuenca, M. del S. *et al.* Understanding butanol tolerance and assimilation in Pseudomonas putida BIRD-1: an integrated omics approach. *Microb. Biotechnol.* **9**, 100–115 (2016).
29. Simon, O. *et al.* Analysis of the molecular response of Pseudomonas putida KT2440 to the next-generation biofuel n-butanol. *J. Proteomics* **122**, 11–25 (2015).
30. Li, W.-J. *et al.* Unraveling 1,4-Butanediol Metabolism in Pseudomonas putida KT2440. *Front. Microbiol.* **11**, 382 (2020).
31. Basler, G., Thompson, M., Tullman-Ercek, D. & Keasling, J. A Pseudomonas putida efflux pump acts on short-chain alcohols. *Biotechnol. Biofuels* **11**, 136 (2018).
32. Nikel, P. I., Chavarría, M., Danchin, A. & de Lorenzo, V. From dirt to industrial applications: Pseudomonas putida as a Synthetic Biology chassis for hosting harsh biochemical reactions. *Curr. Opin. Chem. Biol.* **34**, 20–29 (2016).
33. Burlage, R. S., Hooper, S. W. & Sayler, G. S. The TOL (pWW0) catabolic plasmid. *Appl. Environ. Microbiol.* **55**, 1323–1328 (1989).
34. Nozzi, N. E., Desai, S. H., Case, A. E. & Atsumi, S. Metabolic engineering for higher alcohol production. *Metab. Eng.* **25**, 174–182 (2014).
35. Yan, Q. & Pfleger, B. F. Revisiting metabolic engineering strategies for microbial synthesis of oleochemicals. *Metab. Eng.* **58**, 35–46 (2020).
36. Pfleger, B. F., Gossing, M. & Nielsen, J. Metabolic engineering strategies for microbial synthesis of oleochemicals. *Metab. Eng.* **29**, 1–11 (2015).
37. Kukurugya, M. A. *et al.* Multi-omics analysis unravels a segregated metabolic flux network that tunes co-utilization of sugar and aromatic carbons in Pseudomonas putida. *J. Biol. Chem.* **294**, 8464–8479 (2019).
38. Nikel, P. I. *et al.* Redox stress reshapes carbon fluxes of Pseudomonas putida for cytosolic glucose oxidation and NADPH generation. *BioRxiv* (2020) doi:10.1101/2020.06.13.149542.
39. Nikel, P. I., Chavarría, M., Fuhrer, T., Sauer, U. & de Lorenzo, V. *Pseudomonas*

putida KT2440 Strain Metabolizes Glucose through a Cycle Formed by Enzymes of the Entner-Doudoroff, Embden-Meyerhof-Parnas, and Pentose Phosphate Pathways. *J. Biol. Chem.* **290**, 25920–25932 (2015).
40. Wetmore, K. M. *et al.* Rapid quantification of mutant fitness in diverse bacteria by sequencing randomly bar-coded transposons. *MBio* **6**, e00306-15 (2015).
41. Price, M. N. *et al.* Mutant phenotypes for thousands of bacterial genes of unknown function. *Nature* **557**, 503–509 (2018).
42. Guzik, M. W. *et al.* Identification and characterization of an acyl-CoA dehydrogenase from Pseudomonas putida KT2440 that shows preference towards medium to long chain length fatty acids. *Microbiology (Reading, Engl)* **160**, 1760–1771 (2014).
43. Hume, A. R., Nikodinovic-Runic, J. & O'Connor, K. E. FadD from Pseudomonas putida CA-3 is a true long-chain fatty acyl coenzyme A synthetase that activates phenylalkanoic and alkanoic acids. *J. Bacteriol.* **191**, 7554–7565 (2009).
44. McMahon, B. & Mayhew, S. G. Identification and properties of an inducible phenylacyl-CoA dehydrogenase in Pseudomonas putida KT2440. *FEMS Microbiol. Lett.* **273**, 50–57 (2007).
45. Leščić Ašler, I. *et al.* Probing enzyme promiscuity of SGNH hydrolases. *Chembiochem* **11**, 2158–2167 (2010).
46. McMahon, B., Gallagher, M. E. & Mayhew, S. G. The protein coded by the PP2216 gene of Pseudomonas putida KT2440 is an acyl-CoA dehydrogenase that oxidises only short-chain aliphatic substrates. *FEMS Microbiol. Lett.* **250**, 121–127 (2005).
47. del Peso-Santos, T. *et al.* Coregulation by phenylacetyl-coenzyme A-responsive PaaX integrates control of the upper and lower pathways for catabolism of styrene by Pseudomonas sp. strain Y2. *J. Bacteriol.* **188**, 4812–4821 (2006).
48. Ferrández, A., García, J. L. & Díaz, E. Transcriptional regulation of the divergent paa catabolic operons for phenylacetic acid degradation in Escherichia coli. *J. Biol. Chem.* **275**, 12214–12222 (2000).
49. Görisch, H. The ethanol oxidation system and its regulation in Pseudomonas aeruginosa. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1647**, 98–102 (2003).
50. Hempel, N., Görisch, H. & Mern, D. S. Gene ercA, encoding a putative iron-containing alcohol dehydrogenase, is involved in regulation of ethanol utilization in Pseudomonas aeruginosa. *J. Bacteriol.* **195**, 3925–3932 (2013).
51. Bator, I., Karmainski, T., Tiso, T. & Blank, L. M. Killing Two Birds With One Stone - Strain Engineering Facilitates the Development of a Unique Rhamnolipid Production Process. *Front. Bioeng. Biotechnol.* **8**, 899 (2020).
52. Wehrmann, M., Berthelot, C., Billard, P. & Klebensberger, J. The PedS2/PedR2 Two-Component System Is Crucial for the Rare Earth Element Switch in Pseudomonas putida KT2440. *mSphere* **3**, (2018).
53. Arias, S., Olivera, E. R., Arcos, M., Naharro, G. & Luengo, J. M. Genetic analyses and molecular characterization of the pathways involved in the conversion of 2-phenylethylamine and 2-phenylethanol into phenylacetic acid in Pseudomonas putida U. *Environ. Microbiol.* **10**, 413–432 (2008).
54. Fernández, M. *et al.* Mechanisms of resistance to chloramphenicol in Pseudomonas putida KT2440. *Antimicrob. Agents Chemother.* **56**, 1001–1009 (2012).
55. Gliese, N., Khodaverdi, V. & Görisch, H. The PQQ biosynthetic operons and their transcriptional regulation in Pseudomonas aeruginosa. *Arch. Microbiol.* **192**, 1–14 (2010).

56. García-Hidalgo, J. *et al.* Vanillin Production in Pseudomonas: Whole-Genome Sequencing of Pseudomonas sp. Strain 9.1 and Reannotation of Pseudomonas putida CalA as a Vanillin Reductase. *Appl. Environ. Microbiol.* **86**, (2020).
57. Heeb, S. & Haas, D. Regulatory roles of the GacS/GacA two-component system in plant-associated and other gram-negative bacteria. *Mol. Plant Microbe Interact.* **14**, 1351–1363 (2001).
58. Venturi, V. Control of rpoS transcription in Escherichia coli and Pseudomonas: why so different? *Mol. Microbiol.* **49**, 1–9 (2003).
59. Bentley, G. J. *et al.* Engineering glucose metabolism for enhanced muconic acid production in Pseudomonas putida KT2440. *Metab. Eng.* **59**, 64–75 (2020).
60. Ryan, W. J., O'Leary, N. D., O'Mahony, M. & Dobson, A. D. W. GacS-dependent regulation of polyhydroxyalkanoate synthesis in Pseudomonas putida CA-3. *Appl. Environ. Microbiol.* **79**, 1795–1802 (2013).
61. Jacob, K. *et al.* Regulation of acetyl-CoA synthetase transcription by the CrbS/R two-component system is conserved in genetically diverse environmental pathogens. *PLoS ONE* **12**, e0177825 (2017).
62. Price, M. N. & Arkin, A. P. PaperBLAST: Text Mining Papers for Information about Homologs. *mSystems* **2**, (2017).
63. Monteagudo-Cascales, E., García-Mauriño, S. M., Santero, E. & Canosa, I. Unraveling the role of the CbrA histidine kinase in the signal transduction of the CbrAB two-component system in Pseudomonas putida. *Sci. Rep.* **9**, 9110 (2019).
64. Valentini, M. *et al.* Hierarchical management of carbon sources is regulated similarly by the CbrA/B systems in *Pseudomonas aeruginosa* and *Pseudomonas putida*. *Microbiology (Reading, Engl)* **160**, 2243–2252 (2014).
65. Werle, P. *et al.* Alcohols, Polyhydric. in *Ullmann's encyclopedia of industrial chemistry* (ed. Wiley-VCH Verlag GmbH & Co. KGaA) (Wiley-VCH Verlag GmbH & Co. KGaA, 2000). doi:10.1002/14356007.a01_305.pub2.
66. Wang, Y. *et al.* Reconstruction of lactate utilization system in Pseudomonas putida KT2440: a novel biocatalyst for l-2-hydroxy-carboxylate production. *Sci. Rep.* **4**, 6939 (2014).
67. Jiang, T. *et al.* A Bacterial Multidomain NAD-Independent d-Lactate Dehydrogenase Utilizes Flavin Adenine Dinucleotide and Fe-S Clusters as Cofactors and Quinone as an Electron Acceptor for d-Lactate Oxidization. *J. Bacteriol.* **199**, (2017).
68. Feller, C., Günther, R., Hofmann, H.-J. & Grunow, M. Molecular basis of substrate recognition in D-3-hydroxybutyrate dehydrogenase from Pseudomonas putida. *Chembiochem* **7**, 1410–1418 (2006).
69. Atsumi, S., Hanai, T. & Liao, J. C. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **451**, 86–89 (2008).
70. Conrad, R. S., Massey, L. K. & Sokatch, J. R. D- and L-isoleucine metabolism and regulation of their pathways in Pseudomonas putida. *J. Bacteriol.* **118**, 103–111 (1974).
71. Roberts, C. M., Conrad, R. S. & Sokatch, J. R. The role of enoyl-coa hydratase in the metabolism of isoleucine by Pseudomonas putida. *Arch. Microbiol.* **117**, 99–108 (1978).
72. Kang, A. *et al.* Isopentenyl diphosphate (IPP)-bypass mevalonate pathways for isopentenol production. *Metab. Eng.* **34**, 25–35 (2016).
73. Sasaki, Y. *et al.* Engineering Corynebacterium glutamicum to produce the biogasoline isopentenol from plant biomass hydrolysates. *Biotechnol. Biofuels* **12**,

41 (2019).
74. Hammer, S. K., Zhang, Y. & Avalos, J. L. Mitochondrial Compartmentalization Confers Specificity to the 2-Ketoacid Recursive Pathway: Increasing Isopentanol Production in Saccharomyces cerevisiae. *ACS Synth. Biol.* **9**, 546–555 (2020).
75. LaBauve, A. E. & Wargo, M. J. Growth and laboratory maintenance of Pseudomonas aeruginosa. *Curr. Protoc. Microbiol.* **Chapter 6**, Unit 6E.1. (2012).
76. Ham, T. S. *et al.* Design, implementation and practice of JBEI-ICE: an open source biological part registry platform and tools. *Nucleic Acids Res.* **40**, e141 (2012).
77. Chen, J., Densmore, D., Ham, T. S., Keasling, J. D. & Hillson, N. J. DeviceEditor visual biological CAD canvas. *J. Biol. Eng.* **6**, 1 (2012).
78. Hillson, N. J., Rosengarten, R. D. & Keasling, J. D. j5 DNA assembly design automation software. *ACS Synth. Biol.* **1**, 14–21 (2012).
79. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
80. Engler, C., Kandzia, R. & Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. *PLoS ONE* **3**, e3647 (2008).
81. Shanks, R. M. Q., Kadouri, D. E., MacEachran, D. P. & O'Toole, G. A. New yeast recombineering tools for bacteria. *Plasmid* **62**, 88–97 (2009).
82. Thompson, M. G. *et al.* Identification, Characterization, and Application of a Highly Sensitive Lactam Biosensor from Pseudomonas putida. *ACS Synth. Biol.* **9**, 53–62 (2020).
83. George, K. W. *et al.* Metabolic engineering for the high-yield production of isoprenoid-based C_5 alcohols in E. coli. *Sci. Rep.* **5**, 11128 (2015).
84. van der Walt, S., Colbert, S. C. & Varoquaux, G. The NumPy Array: A Structure for Efficient Numerical Computation. *Comput. Sci. Eng.* **13**, 22–30 (2011).
85. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
86. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermiin, L. S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
87. Cruz-Morales, P. *et al.* Actinobacteria phylogenomics, selective isolation from an iron oligotrophic environment and siderophore functional characterization, unveil new desferrioxamine traits. *FEMS Microbiol. Ecol.* **93**, (2017).
88. López-Farfán, D., Reyes-Darias, J. A., Matilla, M. A. & Krell, T. Concentration Dependent Effect of Plant Root Exudates on the Chemosensory Systems of Pseudomonas putida KT2440. *Front. Microbiol.* **10**, 78 (2019).
89. Fernández, M., Morel, B., Corral-Lugo, A. & Krell, T. Identification of a chemoreceptor that specifically mediates chemotaxis toward metabolizable purine derivatives. *Mol. Microbiol.* **99**, 34–42 (2016).
90. Neal, A. L., Ahmad, S., Gordon-Weeks, R. & Ton, J. Benzoxazinoids in root exudates of maize attract Pseudomonas putida to the rhizosphere. *PLoS ONE* **7**, e35498 (2012).
91. Neal, A. L. & Ton, J. Systemic defense priming by Pseudomonas putida KT2440 in maize depends on benzoxazinoid exudation from the roots. *Plant Signal. Behav.* **8**, e22655 (2013).
92. Hervás, A. B., Canosa, I. & Santero, E. Transcriptome analysis of *Pseudomonas putida* in response to nitrogen availability. *J. Bacteriol.* **190**, 416–420 (2008).
93. Poblete-Castro, I. *et al.* The metabolic response of P. putida KT2442 producing high levels of polyhydroxyalkanoate under single- and multiple-nutrient-limited

growth: highlights from a multi-level omics approach. *Microb. Cell Fact.* **11**, 34 (2012).
94. Mozejko-Ciesielska, J., Dabrowska, D., Szalewska-Palasz, A. & Ciesielski, S. Medium-chain-length polyhydroxyalkanoates synthesis by Pseudomonas putida KT2440 relA/spoT mutant: bioprocess characterization and transcriptome analysis. *AMB Express* **7**, 92 (2017).
95. Galman, J. L. *et al.* Biocatalytic transamination with near-stoichiometric inexpensive amine donors mediated by bifunctional mono- and di-amine transaminases. *Green Chem.* **19**, 361–366 (2017).
96. Han, T., Kim, G. B. & Lee, S. Y. Glutaric acid production by systems metabolic engineering of an l-lysine-overproducing Corynebacterium glutamicum. *Proc Natl Acad Sci USA* **117**, 30328–30334 (2020).
97. Banerjee, D. *et al.* Genome-scale metabolic rewiring improves titers rates and yields of the non-native product indigoidine at scale. *Nat. Commun.* **11**, 5385 (2020).
98. Thompson, M. G. *et al.* Fatty acid and alcohol metabolism in Pseudomonas putida: functional analysis using random barcode transposon sequencing. *Appl. Environ. Microbiol.* **86**, (2020).
99. Gumerov, V. M., Ortega, D. R., Adebali, O., Ulrich, L. E. & Zhulin, I. B. MiST 3.0: an updated microbial signal transduction database with an emphasis on chemosensory systems. *Nucleic Acids Res.* **48**, D459–D464 (2020).
100. Elbourne, L. D. H., Tetu, S. G., Hassan, K. A. & Paulsen, I. T. TransportDB 2.0: a database for exploring membrane transporters in sequenced genomes from all domains of life. *Nucleic Acids Res.* **45**, D320–D324 (2017).
101. Unable to find information for 11446352.
102. Huerta-Cepas, J. *et al.* eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* **47**, D309–D314 (2019).
103. Tatusov, R. L., Koonin, E. V. & Lipman, D. J. A genomic perspective on protein families. *Science* **278**, 631–637 (1997).
104. Galperin, M. Y. *et al.* COG database update: focus on microbial diversity, model organisms, and widespread pathogens. *Nucleic Acids Res.* **49**, D274–D281 (2021).
105. Sonawane, A. M. & Röhm, K. H. A functional gltB gene is essential for utilization of acidic amino acids and expression of periplasmic glutaminase/asparaginase (PGA) by Pseudomonas putida KT2440. *Mol. Genet. Genomics* **271**, 33–39 (2004).
106. Eberl, L. *et al.* Inactivation of gltB abolishes expression of the assimilatory nitrate reductase gene (nasB) in Pseudomonas putida KT2442. *J. Bacteriol.* **182**, 3368–3376 (2000).
107. Reitzer, L. Nitrogen assimilation and global regulation in Escherichia coli. *Annu. Rev. Microbiol.* **57**, 155–176 (2003).
108. Lapouge, K., Schubert, M., Allain, F. H.-T. & Haas, D. Gac/Rsm signal transduction pathway of gamma-proteobacteria: from RNA recognition to regulation of social behaviour. *Mol. Microbiol.* **67**, 241–253 (2008).
109. Martínez-Gil, M., Ramos-González, M. I. & Espinosa-Urgel, M. Roles of cyclic Di-GMP and the Gac system in transcriptional control of the genes coding for the Pseudomonas putida adhesins LapA and LapF. *J. Bacteriol.* **196**, 1484–1495 (2014).
110. Eng, T. *et al.* Engineering Pseudomonas putida for efficient aromatic conversion to bioproduct using high throughput screening in a bioreactor. *Metab. Eng.* **66**, 229–238 (2021).

111. Demling, P. *et al.* Pseudomonas putida KT2440 endures temporary oxygen limitations. *Biotechnol. Bioeng.* **118**, 4735–4750 (2021).
112. Merrick, M. J. & Edwards, R. A. Nitrogen control in bacteria. *Microbiol. Rev.* **59**, 604–622 (1995).
113. Pino, C. *et al.* The assimilatory nitrate reduction system of the phototrophic bacterium Rhodobacter capsulatus E1F1. *Biochem. Soc. Trans.* **34**, 127–129 (2006).
114. Luque-Almagro, V. M. *et al.* Bacterial nitrate assimilation: gene distribution and regulation. *Biochem. Soc. Trans.* **39**, 1838–1843 (2011).
115. Caballero, A., Esteve-Núñez, A., Zylstra, G. J. & Ramos, J. L. Assimilation of nitrogen from nitrite and trinitrotoluene in Pseudomonas putida JLR11. *J. Bacteriol.* **187**, 396–399 (2005).
116. Romeo, A. *et al.* Transcriptional regulation of nitrate assimilation in Pseudomonas aeruginosa occurs via transcriptional antitermination within the nirBD-PA1779-cobA operon. *Microbiology (Reading, Engl)* **158**, 1543–1552 (2012).
117. Wang, B., Rensing, C., Pierson, L. S., Zhao, H. & Kennedy, C. Translational coupling of nasST expression in Azotobacter vinelandii prevents overexpression of the nasT gene. *FEMS Microbiol. Lett.* **361**, 123–130 (2014).
118. Maeda, S., Okamura, M., Kobayashi, M. & Omata, T. Nitrite-specific active transport system of the cyanobacterium Synechococcus sp. strain PCC 7942. *J. Bacteriol.* **180**, 6761–6763 (1998).
119. Gao, S.-H., Fan, L., Yuan, Z. & Bond, P. L. The concentration-determined and population-specific antimicrobial effects of free nitrous acid on Pseudomonas aeruginosa PAO1. *Appl. Microbiol. Biotechnol.* **99**, 2305–2312 (2015).
120. Kusumawardhani, H., van Dijk, D., Hosseini, R. & de Winde, J. H. Novel Toxin-Antitoxin Module SlvT-SlvA Regulates Megaplasmid Stability and Incites Solvent Tolerance in Pseudomonas putida S12. *Appl. Environ. Microbiol.* **86**, (2020).
121. Skjerning, R. B., Senissar, M., Winther, K. S., Gerdes, K. & Brodersen, D. E. The RES domain toxins of RES-Xre toxin-antitoxin modules induce cell stasis by degrading NAD+. *Mol. Microbiol.* **111**, 221–236 (2019).
122. Fuhrer, T., Fischer, E. & Sauer, U. Experimental identification and quantification of glucose metabolism in seven bacterial species. *J. Bacteriol.* **187**, 1581–1590 (2005).
123. Nikel, P. I. *et al.* Reconfiguration of metabolic fluxes in Pseudomonas putida as a response to sub-lethal oxidative stress. *ISME J.* **15**, 1751–1766 (2021).
124. del Castillo, T. *et al.* Convergent peripheral pathways catalyze initial glucose catabolism in Pseudomonas putida: genomic and flux analysis. *J. Bacteriol.* **189**, 5142–5152 (2007).
125. Roon, R. J. & Levenberg, B. An adenosine triphosphate-dependent, avidin-sensitive enzymatic cleavage of urea in yeast and green algae. *J. Biol. Chem.* **243**, 5213–5215 (1968).
126. Roon, R. J. & Levenberg, B. Urea amidolyase. I. Properties of the enzyme from Candida utilis. *J. Biol. Chem.* **247**, 4107–4113 (1972).
127. Mackay, E. M. & Pateman, J. A. The regulation of urease activity in Aspergillus nidulans. *Biochem. Genet.* **20**, 763–776 (1982).
128. Kanamori, T., Kanou, N., Atomi, H. & Imanaka, T. Enzymatic characterization of a prokaryotic urea carboxylase. *J. Bacteriol.* **186**, 2532–2539 (2004).
129. Unable to find information for 11844658.
130. Gadilohar, B. L. & Shankarling, G. S. Choline based ionic liquids and their applications in organic transformation. *J. Mol. Liq.* **227**, 234–261 (2017).
131. Wood, J. M. Bacterial osmoregulation: a paradigm for the study of cellular

homeostasis. *Annu. Rev. Microbiol.* **65**, 215–238 (2011).
132. Wargo, M. J. Homeostasis and catabolism of choline and glycine betaine: lessons from Pseudomonas aeruginosa. *Appl. Environ. Microbiol.* **79**, 2112–2120 (2013).
133. Geiger, O., López-Lara, I. M. & Sohlenkamp, C. Phosphatidylcholine biosynthesis and function in bacteria. *Biochim. Biophys. Acta* **1831**, 503–513 (2013).
134. Leys, D., Basran, J. & Scrutton, N. S. Channelling and formation of "active" formaldehyde in dimethylglycine oxidase. *EMBO J.* **22**, 4038–4048 (2003).
135. Wargo, M. J. & Hogan, D. A. Identification of genes required for Pseudomonas aeruginosa carnitine catabolism. *Microbiology (Reading, Engl)* **155**, 2411–2419 (2009).
136. Bastard, K. *et al.* Revealing the hidden functional diversity of an enzyme family. *Nat. Chem. Biol.* **10**, 42–49 (2014).
137. Meadows, J. A. & Wargo, M. J. Transcriptional Regulation of Carnitine Catabolism in Pseudomonas aeruginosa by CdhR. *mSphere* **3**, (2018).
138. Lee, S. *et al.* Effects of carbazole-degradative plasmid pCAR1 on biofilm morphology in Pseudomonas putida KT2440. *Environ. Microbiol. Rep.* **8**, 261–271 (2016).
139. Nuccio, M. L. *et al.* The endogenous choline supply limits glycine betaine synthesis in transgenic tobacco expressing choline monooxygenase. *Plant J.* **16**, 487–496 (1998).
140. Rontein, D. *et al.* Plants synthesize ethanolamine by direct decarboxylation of serine using a pyridoxal phosphate enzyme. *J. Biol. Chem.* **276**, 35523–35529 (2001).
141. Lundgren, B. R., Sarwar, Z., Pinto, A., Ganley, J. G. & Nomura, C. T. Ethanolamine Catabolism in *Pseudomonas aeruginosa* PAO1 Is Regulated by the Enhancer-Binding Protein EatR (PA4021) and the Alternative Sigma Factor RpoN. *J. Bacteriol.* **198**, 2318–2329 (2016).
142. Roof, D. M. & Roth, J. R. Ethanolamine utilization in Salmonella typhimurium. *J. Bacteriol.* **170**, 3855–3863 (1988).
143. Chang, G. W. & Chang, J. T. Evidence for the B12-dependent enzyme ethanolamine deaminase in Salmonella. *Nature* **254**, 150–151 (1975).
144. Scarlett, F. A. & Turner, J. M. Microbial metabolism of amino alcohols. Ethanolamine catabolism mediated by coenzyme B12-dependent ethanolamine ammonia-lyase in Escherichia coli and Klebsiella aerogenes. *J. Gen. Microbiol.* **95**, 173–176 (1976).
145. Cameron, B., Briggs, K., Pridmore, S., Brefort, G. & Crouzet, J. Cloning and analysis of genes involved in coenzyme B12 biosynthesis in Pseudomonas denitrificans. *J. Bacteriol.* **171**, 547–557 (1989).
146. Heller, K., Mann, B. J. & Kadner, R. J. Cloning and expression of the gene for the vitamin B12 receptor protein in the outer membrane of Escherichia coli. *J. Bacteriol.* **161**, 896–903 (1985).
147. Cadieux, N. *et al.* Identification of the periplasmic cobalamin-binding protein BtuF of Escherichia coli. *J. Bacteriol.* **184**, 706–717 (2002).
148. Warren, M. J., Raux, E., Schubert, H. L. & Escalante-Semerena, J. C. The biosynthesis of adenosylcobalamin (vitamin B12). *Nat. Prod. Rep.* **19**, 390–412 (2002).
149. Martens, J. H., Barg, H., Warren, M. J. & Jahn, D. Microbial production of vitamin B12. *Appl. Microbiol. Biotechnol.* **58**, 275–285 (2002).
150. Zayas, C. L. & Escalante-Semerena, J. C. Reassessment of the late steps of coenzyme B12 synthesis in Salmonella enterica: evidence that dephosphorylation

of adenosylcobalamin-5′-phosphate by the CobC phosphatase is the last step of the pathway. *J. Bacteriol.* **189**, 2210–2218 (2007).
151. Mori, K., Bando, R., Hieda, N. & Toraya, T. Identification of a reactivating factor for adenosylcobalamin-dependent ethanolamine ammonia lyase. *J. Bacteriol.* **186**, 6845–6854 (2004).
152. Yuzawa, S., Eng, C. H., Katz, L. & Keasling, J. D. Broad substrate specificity of the loading didomain of the lipomycin polyketide synthase. *Biochemistry* **52**, 3791–3793 (2013).
153. van der Maaten, L. & Hinton, G. Visualizing Data using t-SNE. *Journal of Machine Learning Research* (2008).
154. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* (2011).
155. Harris, C. R. *et al.* Array programming with NumPy. *Nature* **585**, 357–362 (2020).
156. Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P. & Huerta-Cepas, J. eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. *Mol. Biol. Evol.* **38**, 5825–5829 (2021).
157. Hewlings, S. J. & Kalman, D. S. Curcumin: A review of its effects on human health. *Foods* **6**, (2017).
158. Schanz, S., Schröder, G. & Schröder, J. Stilbene synthase from Scots pine (*Pinus sylvestris*). *FEBS Lett.* **313**, 71–74 (1992).
159. Heller, W. & Hahlbrock, K. Highly purified "flavanone synthase" from parsley catalyzes the formation of naringenin chalcone. *Arch. Biochem. Biophys.* **200**, 617–619 (1980).
160. Smith, L. R. Rheosmin ("Raspberry Ketone") and Zingerone, and Their Preparation by Crossed Aldol-Catalytic Hydrogenation Sequences. *Chem. Educator* **1**, 1–18 (1996).
161. Katsuyama, Y., Matsuzawa, M., Funa, N. & Horinouchi, S. Production of curcuminoids by Escherichia coli carrying an artificial biosynthesis pathway. *Microbiology (Reading, Engl)* **154**, 2620–2628 (2008).
162. Yu, D., Xu, F., Zeng, J. & Zhan, J. Type III polyketide synthases in natural product biosynthesis. *IUBMB Life* **64**, 285–295 (2012).
163. Izumikawa, M. *et al.* Expression and characterization of the type III polyketide synthase 1,3,6,8-tetrahydroxynaphthalene synthase from Streptomyces coelicolor A3(2). *J. Ind. Microbiol. Biotechnol.* **30**, 510–515 (2003).
164. Wakimoto, T., Morita, H. & Abe, I. Engineering of plant type III polyketide synthases. *Meth. Enzymol.* **515**, 337–358 (2012).
165. Lawther, J. M. & Sun, R. The fractional characterisation of polysaccharides and lignin components in alkaline treated and atmospheric refined wheat straw. *Industrial Crops and Products* **5**, 87–95 (1996).
166. Jönsson, L. J., Alriksson, B. & Nilvebrant, N.-O. Bioconversion of lignocellulose: inhibitors and detoxification. *Biotechnol. Biofuels* **6**, 16 (2013).
167. Mills, T. Y., Sandoval, N. R. & Gill, R. T. Cellulosic hydrolysate toxicity and tolerance mechanisms in Escherichia coli. *Biotechnol. Biofuels* **2**, 26 (2009).
168. Guarnieri, M. T., Ann Franden, M., Johnson, C. W. & Beckham, G. T. Conversion and assimilation of furfural and 5-(hydroxymethyl)furfural by Pseudomonas putida KT2440. *Metab. Eng. Commun.* **4**, 22–28 (2017).
169. Linger, J. G. *et al.* Lignin valorization through integrated biological funneling and chemical catalysis. *Proc Natl Acad Sci USA* **111**, 12013–12018 (2014).
170. Nozaki, M., Kagamiyama, H. & Hayaishi, O. Crystallization and some properties

of metapyrocatechase. *Biochem. Biophys. Res. Commun.* **11**, 65–69 (1963).
171. Yang, D. *et al.* Repurposing type III polyketide synthase as a malonyl-CoA biosensor for metabolic engineering in bacteria. *Proc Natl Acad Sci USA* **115**, 9835–9844 (2018).
172. Gross, F. *et al.* Bacterial type III polyketide synthases: phylogenetic analysis and potential for the production of novel secondary metabolites by heterologous expression in pseudomonads. *Arch. Microbiol.* **185**, 28–38 (2006).
173. Soga, O. Stimulative Production of Flaviolin by *Phoma wasabiae*. *Agric. Biol. Chem.* **46**, 1061–1063 (1982).
174. Overhage, J., Priefert, H., Rabenhorst, J. & Steinbüchel, A. Biotransformation of eugenol to vanillin by a mutant of Pseudomonas sp. strain HR199 constructed by disruption of the vanillin dehydrogenase (vdh) gene. *Appl. Microbiol. Biotechnol.* **52**, 820–828 (1999).
175. Plaggenborg, R., Overhage, J., Steinbüchel, A. & Priefert, H. Functional analyses of genes involved in the metabolism of ferulic acid in Pseudomonas putida KT2440. *Appl. Microbiol. Biotechnol.* **61**, 528–535 (2003).
176. Jiménez, J. I., Pérez-Pantoja, D., Chavarría, M., Díaz, E. & de Lorenzo, V. A second chromosomal copy of the catA gene endows Pseudomonas putida mt-2 with an enzymatic safety valve for excess of catechol. *Environ. Microbiol.* **16**, 1767–1778 (2014).
177. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
178. Kanehisa, M., Sato, Y., Furumichi, M., Morishima, K. & Tanabe, M. New approach for understanding genome variations in KEGG. *Nucleic Acids Res.* **47**, D590–D595 (2019).
179. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **28**, 1947–1951 (2019).
180. Jiménez, J. I., Miñambres, B., García, J. L. & Díaz, E. Genomic analysis of the aromatic catabolic pathways from Pseudomonas putida KT2440. *Environ. Microbiol.* **4**, 824–841 (2002).
181. Katsuyama, Y., Matsuzawa, M., Funa, N. & Horinouchi, S. In vitro synthesis of curcuminoids by type III polyketide synthase from Oryza sativa. *J. Biol. Chem.* **282**, 37702–37709 (2007).
182. Bi, C. *et al.* Development of a broad-host synthetic biology toolbox for Ralstonia eutropha and its application to engineering hydrocarbon biofuel production. *Microb. Cell Fact.* **12**, 107 (2013).
183. Katsuyama, Y., Funa, N., Miyahisa, I. & Horinouchi, S. Synthesis of unnatural flavonoids and stilbenes by exploiting the plant biosynthetic pathway in Escherichia coli. *Chem. Biol.* **14**, 613–621 (2007).
184. Javeri, I. & Chand, N. Curcumin. in *Nutraceuticals* 435–445 (Elsevier, 2016). doi:10.1016/B978-0-12-802147-7.00031-0.
185. Wang, Y. *et al.* Growth of engineered *Pseudomonas putida* KT2440 on glucose, xylose, and arabinose: Hemicellulose hydrolysates and their major sugars as sustainable carbon sources. *Glob. Change Biol. Bioenergy* **11**, 249–259 (2019).
186. Martínez-García, E. & de Lorenzo, V. Engineering multiple genomic deletions in Gram-negative bacteria: analysis of the multi-resistant antibiotic profile of Pseudomonas putida KT2440. *Environ. Microbiol.* **13**, 2702–2716 (2011).
187. Wirth, N. T., Kozaeva, E. & Nikel, P. I. Accelerated genome engineering of Pseudomonas putida by I-SceI-mediated recombination and CRISPR-Cas9

counterselection. *Microb. Biotechnol.* **13**, 233–249 (2020).
188. Priefert, H., Rabenhorst, J. & Steinbüchel, A. Biotechnological production of vanillin. *Appl. Microbiol. Biotechnol.* **56**, 296–314 (2001).
189. Wightman, F. & Lighty, D. L. Identification of phenylacetic acid as a natural auxin in the shoots of higher plants. *Physiol. Plant.* **55**, 17–24 (1982).
190. Cambray, G., Guimaraes, J. C. & Arkin, A. P. Evaluation of 244,000 synthetic sequences reveals design principles to optimize translation in Escherichia coli. *Nat. Biotechnol.* **36**, 1005–1015 (2018).
191. Zang, Y. *et al.* In Vitro Naringenin Biosynthesis from p-Coumaric Acid Using Recombinant Enzymes. *J. Agric. Food Chem.* **67**, 13430–13436 (2019).
192. Parke, D. & Ornston, L. N. Toxicity caused by hydroxycinnamoyl-coenzyme A thioester accumulation in mutants of Acinetobacter sp. strain ADP1. *Appl. Environ. Microbiol.* **70**, 2974–2983 (2004).
193. Santos, C. N. S., Koffas, M. & Stephanopoulos, G. Optimization of a heterologous pathway for the production of flavonoids from glucose. *Metab. Eng.* **13**, 392–400 (2011).
194. Park, S. R. *et al.* Engineering of plant-specific phenylpropanoids biosynthesis in Streptomyces venezuelae. *J. Biotechnol.* **141**, 181–188 (2009).
195. Fang, Z., Jones, J. A., Zhou, J. & Koffas, M. A. G. Engineering Escherichia coli Co-Cultures for Production of Curcuminoids From Glucose. *Biotechnol. J.* **13**, e1700576 (2018).
196. Nijkamp, K., Westerhof, R. G. M., Ballerstedt, H., de Bont, J. A. M. & Wery, J. Optimization of the solvent-tolerant Pseudomonas putida S12 as host for the production of p-coumarate from glucose. *Appl. Microbiol. Biotechnol.* **74**, 617–624 (2007).
197. Nijkamp, K., van Luijk, N., de Bont, J. A. M. & Wery, J. The solvent-tolerant Pseudomonas putida S12 as host for the production of cinnamic acid from glucose. *Appl. Microbiol. Biotechnol.* **69**, 170–177 (2005).
198. Grant, S. G., Jessee, J., Bloom, F. R. & Hanahan, D. Differential plasmid rescue from transgenic mouse DNAs into Escherichia coli methylation-restriction mutants. *Proc Natl Acad Sci USA* **87**, 4645–4649 (1990).
199. Park, M.-R. *et al.* The response of *Pseudomonas putida* to a complex lignolysate. *BioRxiv* (2019) doi:10.1101/764399.
200. Chen, Y. *et al.* A rapid methods development workflow for high-throughput quantitative proteomic applications. *PLoS ONE* **14**, e0211582 (2019).
201. Snyder, S. A., Tang, Z.-Y. & Gupta, R. Enantioselective total synthesis of (-)-napyradiomycin A1 via asymmetric chlorination of an isolated olefin. *J. Am. Chem. Soc.* **131**, 5744–5745 (2009).
202. Hori, Y., Nakaki, K., Sato, M., Mizukami, S. & Kikuchi, K. Development of protein-labeling probes with a redesigned fluorogenic switch based on intramolecular association for no-wash live-cell imaging. *Angew. Chem. Int. Ed* **51**, 5611–5614 (2012).