

# Comparative transcriptomics elucidates adaptive phenol tolerance and utilization in lipid-accumulating *Rhodococcus opacus* PD630

Aki Yoneda<sup>1,2,†</sup>, William R. Henson<sup>3,†</sup>, Nicholas K. Goldner<sup>2</sup>, Kun Joo Park<sup>3</sup>, Kevin J. Forsberg<sup>2</sup>, Soo Ji Kim<sup>3</sup>, Mitchell W. Pesesky<sup>2</sup>, Marcus Foston<sup>3</sup>, Gautam Dantas<sup>1,2,4,5,\*</sup> and Tae Seok Moon<sup>3,\*</sup>

<sup>1</sup>Department of Pathology & Immunology, Washington University in St. Louis School of Medicine, St. Louis, MO 63108, USA, <sup>2</sup>Center for Genome Sciences & Systems Biology, Washington University in St. Louis School of Medicine, St. Louis, MO 63108, USA, <sup>3</sup>Department of Energy, Environmental & Chemical Engineering, Washington University in St. Louis, St. Louis, MO 63130, USA, <sup>4</sup>Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, MO 63130, USA and <sup>5</sup>Department of Molecular Microbiology, Washington University in St. Louis School of Medicine, St. Louis, MO 63108, USA

Received September 22, 2015; Revised December 17, 2015; Accepted January 20, 2016

## ABSTRACT

Lignin-derived (e.g. phenolic) compounds can compromise the bioconversion of lignocellulosic biomass to fuels and chemicals due to their toxicity and recalcitrance. The lipid-accumulating bacterium *Rhodococcus opacus* PD630 has recently emerged as a promising microbial host for lignocellulose conversion to value-added products due to its natural ability to tolerate and utilize phenolics. To gain a better understanding of its phenolic tolerance and utilization mechanisms, we adaptively evolved *R. opacus* over 40 passages using phenol as its sole carbon source (up to 373% growth improvement over wild-type), and extensively characterized two strains from passages 33 and 40. The two adapted strains showed higher phenol consumption rates (~20 mg/l/h) and ~2-fold higher lipid production from phenol than the wild-type strain. Whole-genome sequencing and comparative transcriptomics identified highly-upregulated degradation pathways and putative transporters for phenol in both adapted strains, highlighting the important linkage between mechanisms of regulated phenol uptake, utilization, and evolved tolerance. Our study shows that the *R. opacus* mutants are likely to use their transporters to import phenol rather than export them, suggesting a

new aromatic tolerance mechanism. The identified tolerance genes and pathways are promising candidates for future metabolic engineering in *R. opacus* for improved lignin conversion to lipid-based products.

## INTRODUCTION

Lignocellulosic biomass, comprised of cellulose, hemicellulose and lignin (1,2), remains an underutilized substrate in sustainable microbial production of fuels and chemicals (3–6). One main challenge is that current biorefinery pretreatment approaches release diverse toxic degradation compounds from lignin during conversion of lignocellulosic biomass to fermentable sugars (7). These lignin degradation compounds include a wide array of phenolics that can severely inhibit microbial production of fuels or chemicals, leading to lower yields (8). Currently, unconverted lignin is typically burned to provide thermal energy onsite, but the amount of waste lignin is predicted to escalate as lignocellulose-based biorefinery output increases (4,9,10).

Lignin, the second most abundant terrestrial polymer, constitutes ~15–30% of lignocellulose (11) and is more energy dense than cellulose and hemicellulose due to its higher carbon-to-oxygen ratio. Unfortunately, lignin is much more difficult to depolymerize due to its complex molecular structure. Structural heterogeneity also leads to a broad spectrum of breakdown products, substantially compromising

\*To whom correspondence should be addressed. Tel: +1 314 362 7238; Fax: +1 314 362 2156; Email: dantas@wustl.edu  
Correspondence may also be addressed to Tae Seok Moon. Tel: +1 314 935 5026; Fax: +1 314 935 7211; Email: tsmoon@wustl.edu

†These authors contributed equally to the paper as first authors.

Present addresses:

Kevin J. Forsberg, Basic Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA.  
Mitchell W. Pesesky, Department of Chemical Engineering, University of Washington, Seattle, WA 98195, USA.

the efficiency of chemical catalysis approaches for product synthesis and purification. Some bacteria and fungi can consume lignin breakdown products and utilize them as carbon sources (12), potentiating fuel and chemical production via lignin consolidated bioprocessing (13–15). One such bacterium, *Rhodococcus opacus*, is a promising microbial host for converting lignocellulose to useful products due to its naturally robust lipid production and ability to both tolerate and metabolize diverse phenolic compounds (14,16–18).

*Rhodococcus* strains are found in diverse environments (19,20) and can tolerate environmental stresses such as desiccation and high salinity (21,22) as well as chemical stresses such as high concentrations of butanol (23,24). Often isolated from polluted or contaminated environmental samples (25,26), *R. opacus* strains have a strong innate tolerance to benzene, toluene and lignocellulosic hydrolysates from different sources (27,28) and can metabolize aromatic compounds (14,16). *Rhodococcus* species can convert aromatics to acetyl-CoA and succinyl-CoA (12,29), which are important precursors for converting phenolics to bioproducts (30). Originally isolated from soil at a gas works plant, *R. opacus* PD630 (hereafter *R. opacus* unless specified) is known to accumulate large amounts of the biodiesel precursors triacylglycerols (TAGs, up to 76% of cell dry weight) when using sugars as a carbon source. Thus, *R. opacus* has been a target strain for commercial-scale lipid production using sugars derived from lignocellulose (18,31–34).

Growth inhibition by toxic compounds (either end-product or feedstock) is a major limiting factor for commercialization of biochemical processes (35,36). Developing production hosts with natural tolerance to toxic inhibitors may significantly reduce time and efforts for host optimization. The tolerance capabilities of *R. opacus* are hypothesized to come from its highly hydrophobic cell wall (22) and/or its ability to consume a diverse array of compounds (20), but few studies have directly investigated phenolic tolerance mechanisms in this organism. While we have recently explored the central metabolism in wild-type (WT) *R. opacus* and found simultaneous utilization of glucose and phenol (i.e. no catabolite repression) using <sup>13</sup>C-metabolite fingerprinting (37), more work is necessary to thoroughly characterize the catabolic pathways of aromatic compounds and global metabolism to elucidate tolerance mechanisms in *R. opacus*.

In this study, we adaptively evolved increased phenol tolerance of *R. opacus* by sequentially sub-culturing in phenol as a sole carbon source and screening fast-growing mutants over 40 passages. We selected phenol as a model lignin degradation product to avoid the confounding effects of many compounds present in heterogeneous lignin degradation product streams. Phenol has a shared substructure to many of the compounds that can be derived from lignin (38), and it also has a similar level of toxicity to that of other compounds derived from lignin (39,40).

Even though previous studies have demonstrated bioconversion of lignocellulose-derived compounds into lipids by *R. opacus* (14,27,40–44), the connection between the tolerance phenotype and specific cellular mechanisms remains elusive. Here, we present a combined adaptive evolution/omics approach leveraging multiple phenol-

adapted strains to identify possible mechanisms for phenolic tolerance and utilization in *R. opacus* (Supplementary Figure S1). This approach builds on previous studies by examining multiple paths to phenolic tolerance from the same strain background and resolving these tolerance strategies by comparing the transcriptomic response in different growth conditions. We performed adaptive evolution of *R. opacus* on increasing concentrations of phenol to select for increasingly tolerant strains and identified two high-performing strains through in-depth phenotyping. Next, we performed whole-genome sequencing of the selected strains to identify genomic alterations during adaptive evolution and employed comparative transcriptomics to identify transcriptional changes between strains in different growth conditions (i.e. in glucose and different concentrations of phenol). This approach proved to be effective in gaining insights into tolerance mechanisms and identifying promising gene candidates that can facilitate future metabolic engineering efforts in *Rhodococcus* to produce fuels and chemicals directly from lignin breakdown products or from lignocellulosic hydrolysates rich in toxic lignin degradation compounds.

## MATERIALS AND METHODS

### Adaptive evolution

To increase the strain's tolerance to phenol, *R. opacus* PD630 was adapted to grow in increasing concentrations of phenol for 40 successive subcultures. To start the experiment, the WT strain was grown in 2 ml of 0.3 g/l phenol for 2 days, and then subcultured at 0.3 g/l phenol with an initial OD<sub>600</sub> of 0.03 during the first 10 subcultures. Once the culture reached an OD<sub>600</sub> of 0.3 within 24–36 h, the phenol concentration was increased in increments of 0.25 g/l up to a final concentration of 1.5 g/l phenol. Frozen stocks of each subculture were saved for later testing. To pick phenol-tolerant strains, frozen stocks from intermediate subcultures were streaked onto minimal media plates supplemented with 1.5% agar containing 0.3 g/l phenol. From three of the intermediate subcultures (17, 26 and 33) and the final subculture (40), six colonies were picked for growth screening and lipid assays in 96 well plates. The two fastest growing strains with best lipid productivities (named evol33 and evol40) were chosen from this screening for further analysis. For detailed growth conditions for other experiments and assays, see Supplementary Materials and Methods.

### Lipid assay

The two best performing strains were chosen for lipid assays using Nile red. Growth experiments were started with single colonies from LB plates, and cells were cultured in 2 ml of 0.3 g/l phenol in a 50 ml glass tube for 36–48 h. Next, 2 ml of cells was subcultured in 10 ml of 0.3 g/l phenol for 24 h in a 50 ml glass tube. Next, 10 ml of cells was added to 140 ml of 0.3 g/l phenol in a 250 ml flask and grown for 24 h. Cells were centrifuged at 3000 x g for 10 min, resuspended in carbon- and nitrogen-free media, and then each condition (1 g/l glucose, 0.75 g/l phenol or 1.5 g/l phenol) was tested for each strain with an initial OD<sub>600</sub> = 0.3 in low nitrogen

conditions (0.05 g/l  $(\text{NH}_4)_2\text{SO}_4$ ) at a 10 ml volume in a 50 ml glass culture tube. Cells were stained by Nile red according to the reported method (45) with some modifications. Briefly, 200  $\mu\text{l}$  of cell culture was stored in 20% DMSO at  $-20^\circ\text{C}$  until staining. For staining, cells were centrifuged at 1000  $\times$  g and room temperature for 3 min and resuspended in 200  $\mu\text{l}$  of 1x PBS. 5  $\mu\text{l}$  of 1 mM Nile red in DMSO was added to each sample in a 96-well plate, which was subsequently incubated in the dark at room temperature for 30 min. The plate was wrapped in foil, centrifuged at 1000  $\times$  g for 3 min and washed twice with 0.9% NaCl solution. Flow cytometry analysis was performed using a Millipore Guava EasyCyte High Throughput flow cytometer with a 488 nm excitation laser and a 575 nm emission filter. For all data, at least 5000 events were collected and gated by forward and side scatter. Samples were stained and analyzed on the same day to decrease variability. FlowJo (TreeStar Inc.) was used to obtain the arithmetic mean of the Nile red fluorescence distributions of stained cells, and the autofluorescence of unstained cells was subtracted from each measurement value. Because of differences in lag phases between strains, the lipid productivity was calculated as the average change in Nile red fluorescence (per cell per time) between at least four time points taken from the exponential to the stationary phase.

### Illumina library preparation

A protocol adapted from Bowman *et al.* and Fisher *et al.* (46,47) was used to generate Illumina sequencing libraries. Double-stranded cDNA (for RNA-Seq) or genomic DNA (for whole genome sequencing) in 50  $\mu\text{l}$  elution buffer (Qiagen) was fragmented to 200–400 bp using Covaris E-220 and microTUBEs at a recommended setting by Covaris.  $\sim 50$   $\mu\text{l}$  of fragmented DNA in Buffer EB (Qiagen) was recovered from Covaris tubes and transferred to a semi-skirted 96-well polymerase chain reaction (PCR) plate. 20  $\mu\text{l}$  of end repair mix (7  $\mu\text{l}$  10x T4 ligase buffer, 2.8  $\mu\text{l}$  10 mM dNTPs, 3.5  $\mu\text{l}$  T4 polymerase, 3.5  $\mu\text{l}$  T4 polynucleotide kinase, 0.7  $\mu\text{l}$  *Escherichia coli* DNA polymerase I and 2.5  $\mu\text{l}$  nuclease-free water; all enzymes were purchased from NEB) was added to the fragmented DNA to blunt and phosphorylate the ends at room temperature for 30 min. 130  $\mu\text{l}$  of AMPure XP beads (Agencourt) were added to each well to remove reagents from the previous step. The 96-well plate was placed on a magnetic plate after pipetting thoroughly to mix, and supernatant was discarded when it became clear. The beads with DNA on the walls of 96 wells were washed with 150  $\mu\text{l}$  80% ethanol once, before eluting the DNA with 32  $\mu\text{l}$  of water. 18  $\mu\text{l}$  of dA-tailing mix (5  $\mu\text{l}$  10x NEBuffer 2, 1 mM dATP and 3  $\mu\text{l}$  Klenow fragment (3'-5' exo-)) was added to the beads with eluted DNA to add a dA-tail at  $37^\circ\text{C}$  for 30 min. After dA-tailing, 110  $\mu\text{l}$  of 20% polyethylene glycol (PEG) 8000 in 2.5 M NaCl was used to precipitate DNA onto the magnetic beads. The plate was placed on a magnet again to collect DNA and washed with 80% ethanol. Barcoded adapters were added to the eluted DNA (20.5  $\mu\text{l}$  water added to DNA with beads) to a 10:1 molar ratio, and 6.5  $\mu\text{l}$  of ligation mix was added to ligate adapters at room temperature for 20 min. 21  $\mu\text{l}$  of 20% PEG in 2.5 M NaCl was added to precipitate DNA onto the beads,

and two barcoded samples were pooled after washing for elution. Adapter-ligated DNA was then PCR-amplified using KAPA HiFi Hot start Readymix (KAPA Biosystems) and gel purified using a Zymoclean Gel DNA Recovery Kit (Zymo Research). The resulting purified DNA samples were quantified using a Qubit DNA HS kit, combined and diluted with Buffer EB to yield a 10 nM pool. 20  $\mu\text{l}$  of the pool was submitted to HiSeq-2500 sequencing at the Center for Genome Sciences & Systems Biology, Washington University in St. Louis School of Medicine. For other details regarding DNA and RNA preparations, see Supplementary Materials and Methods.

### Genome sequencing and SNP analysis

Raw Illumina reads were demultiplexed by barcodes and trimmed using an in-house Python script, and mapped to an indexed genome of *R. opacus* (both annotated supercontigs maintained by Broad Institute (Genbank#ABRH01000000 (20)) and the finished genome by Chinese Academy of Science (ASM59954v1 (48)) using Bowtie2 (49). The mapped SAM files were converted to sorted BAM files with BAM index files using SAMtools (50,51). SNPs and possible indels were analyzed using Pilon (52). SNPs that were also found in the WT strain were filtered out using an in-house Python script (see Supplementary Computational Methods). Additionally, MIRA (53) was used to confirm SNPs and indels predicted by Pilon. In addition to utilizing MIRA's SNP detection function, we performed *de novo*-assembly of a genome for each strain using MIRA. Assembled genomes were then used to further confirm presence or absence of SNPs in each strain. This deep sequencing confirmation eliminated SNPs that were inconsistently found due to Illumina sequencing errors or biases during Illumina preparations. Table 1A shows all SNPs found in evolved strains and not in WT.

### RNA-Seq data analysis

After demultiplexing and trimming, raw Illumina reads were mapped to an indexed genome of *R. opacus* (both annotated supercontigs (20) and the finished genome (48)) using Bowtie2 (49). The mapped SAM files were converted to sorted BAM files with BAM index files using SAMtools (50,51). FeatureCounts (54) was used to count the number of reads per coding sequence, and the raw result was used for differential expression analysis using DESeq2 (55) as implemented in the R software package.

## RESULTS

### Adaptive evolution of *R. opacus* using phenol as a sole carbon source

In the adaptive evolution approach, we increased phenol concentrations in a minimal salts medium (MSM) over time to select for the fastest growing strains using phenol as a sole carbon source (Supplementary Figure S1). Growth of the WT strain on phenol was characterized by a lag phase that increases with higher concentrations of phenol (Figure 1). To prevent long lag phases in adapted strains,

**Table 1.** Single nucleotide polymorphisms (SNPs) found in the adapted strains. (A) All SNPs found in the adapted strains but not in the WT strain are summarized. The mutations were called by both Pilon and MIRA. (B) Expression levels of the genes listed in A, shown in DESeq2-normalized counts, using plotCounts function. (C) Expression levels of the same genes listed in A, in Reads per Kilobase of transcript per Million mapped reads (rpkm). Rpk values are calculated as explained in the Materials and Method section, and the average of triplicate samples is shown. Glc, 1 g/l glucose; LowP, 0.75 g/l phenol; HighP, 1.5 g/l phenol.

A. SNP summary					
Both Evol33 and Evol40					
CAS gene ID	MIT gene ID	Annotation	WT	SNP	Amino acid change
Between LPD04711 and LPD04712	OPAG_05248	Intergenic (CAS)/ acyl-CoA thioesterase II	ttc	ttt	F120 to F (silent)
LPD03112	OPAG_03447	cytochrome aa3 quinol oxidase, subunit I dicarboxylate carrier protein/citrate transporter/MFS transporter	ttc	tcc	F335 to S
LPD03747	OPAG_04053	protein/citrate transporter/MFS transporter	atc	ttc	I296 to F
Evol33 only					
LPD00118	OPAG_07911	4-hydroxyphenylacetate permease	gga	gta	G162 to V

### B. Expression levels using DESeq2 calculation (average of triplicates)

Gene ID	WT Glc	WT LowP	evol33 Glc	evol33 LowP	evol33 HighP	evol40 Glc	evol40 LowP	evol40 HighP
LPD04711	3	6	2	10	13	3	12	4
LPD04712	11	14	13	9	5	11	8	2
LPD03112	2580	18633	2908	9718	12462	2373	13083	13086
LPD03747	152	29	116	136	165	119	59	75
LPD00118	376	64	289	312	343	312	163	165

### C. Expression levels using rpkm calculation (average of triplicates)

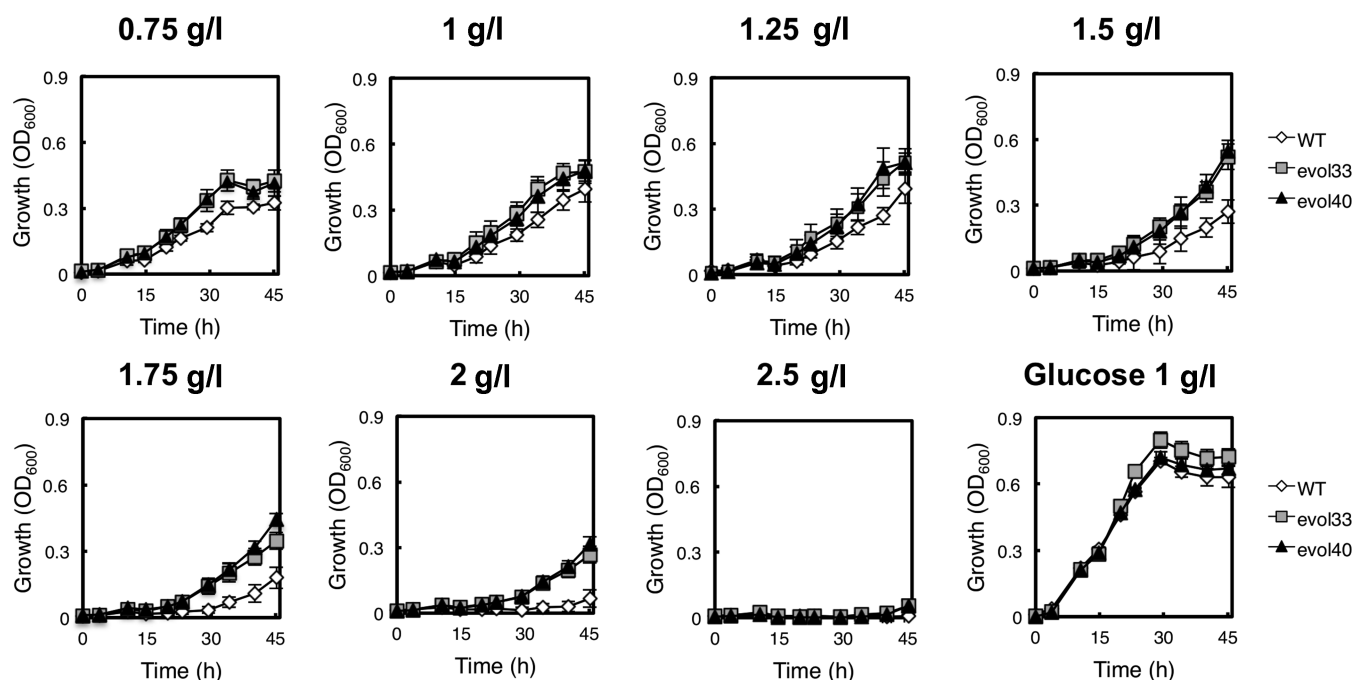
Gene ID	WT Glc	WT LowP	evol33 Glc	evol33 LowP	evol33 HighP	evol40 Glc	evol40 LowP	evol40 HighP
LPD04711	2	3	1	5	6	2	7	2
LPD04712	6	7	8	5	2	7	4	1
LPD03112	152	900	189	498	591	153	753	697
LPD03747	11	2	9	8	9	9	4	5
LPD00118	27	4	23	20	20	25	12	11

cells were grown in non-inhibitory concentrations of phenol for the initial rounds of selection before increasing to inhibitory phenol concentrations. Serial transfers were performed when exponential phase was reached (optical density at 600 nm (OD<sub>600</sub>) of ~0.3). For the total adaptive evolution experiment (40 serial passages), the estimated cumulative cell divisions (56) were  $\sim 6 \times 10^9$ , assuming four doublings per subculture and an initial cell concentration of  $\sim 1 \times 10^7$  CFU/ml (colony forming unit per ml; initial OD<sub>600</sub> of ~0.03 was experimentally determined to be  $\sim 1 \times 10^7$  CFU/ml). Over the course of the 40 serial transfers (~200 generations), six colonies each from passages 17, 26, 33 and 40 (total of 24 colonies) were randomly selected for high-throughput phenotypic screening, along

with WT (Supplementary Figure S1). We chose the two best-performing phenol-adapted strains based on improved growth rate, biomass yields and lipid yields: one derived from passage 33 (named evol33) and the other from passage 40 (evol40). These two adapted strains were subjected to further phenotypic characterization (i.e. growth profiles and lipid accumulation assay), deep whole-genome sequencing and comparative transcriptomic profiling by RNA-Seq.

### Phenotypic characterization of phenol-adapted strains

To characterize improvements in tolerance from adaptive evolution, both adapted and WT strains were cultivated in a range of concentrations of phenol (Figure 1). We charac-

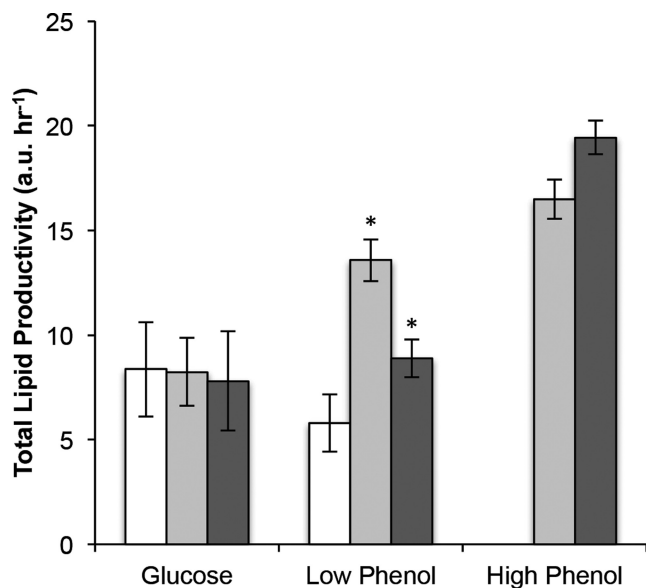


**Figure 1.** Growth comparison between the two phenol-adapted strains and the WT strain in diverse growth conditions. The adapted strains and the WT strain were grown in multiple concentrations of phenol as their sole carbon source to characterize improvements in tolerance after adaptive evolution (phenol concentrations as indicated; 1 g/l ammonium sulfate). The adapted strains grew to higher optical densities than the WT strain in all concentrations of phenol (Supplementary Table S1), had shorter lag phases in higher concentrations of phenol (Supplementary Table S2) and demonstrated growth even in 2 g/l phenol. White diamonds, gray squares and black triangles represent WT, evol33 and evol40, respectively. Data show the average of six biological replicates grown in 96 well plates (with 200  $\mu$ l culture each). The error bars represent one standard deviation.

terized phenol tolerance by comparing growth rates, final biomass accumulation ( $OD_{600}$ ), lag phases and estimated  $IC_{50}$  (half maximal inhibitory concentration; Supplementary Figure S2). The adapted strains evol33 and evol40 had 20% and 24% higher  $IC_{50}$  values than WT, respectively, after 45 h of culture ( $P < 0.001$  for both evol33 and evol40, one mean, two-tailed Student's  $t$ -test). Both adapted strains also reached 22–373% higher cell densities ( $P < 0.05$ ) than WT in all concentrations of phenol tested where growth was observed (0.75–2 g/l), with evol33 and evol40 having a 92% and 102% increase in  $OD_{600}$  at 1.5 g/l ( $P = 0.0005$  and  $P = 0.0002$ , respectively; Supplementary Table S1). The lag phase of the adapted strains was 17–54% shorter compared to that of the WT strain, with evol33 showing a 46% shorter lag phase at 1.5 g/l ( $P = 0.029$ ) and evol40 showing a 34% shorter lag phase compared to WT ( $P = 0.054$ ; Supplementary Table S2). The adapted strains also demonstrated significant growth rate increases at intermediate concentrations of phenol, with 30% and 38% increases in growth rates at 1.5 g/l for evol33 ( $P = 0.03$ ) and evol40 ( $P = 0.04$ ), respectively (Supplementary Table S3). When cells were grown in 1.5 g/l phenol, we also observed significantly higher phenol consumption rates in the adapted strains compared to the WT strain ( $P = 0.002$  and  $P = 0.003$  for evol33 and evol40, respectively), suggesting that phenol uptake, consumption and growth are related (Supplementary Figure S3).

To determine whether adaptive evolution affected lipid accumulation, we measured lipid amounts under nitrogen-limited conditions using either glucose or phenol as a sole carbon source (Figure 2 and Supplementary Table S4). WT

*R. opacus* accumulates intracellular lipids in the form of TAGs under nitrogen limitation in the stationary phase (18). Strains were grown in 1 g/l glucose ('glucose'), 0.75 g/l phenol ('low phenol') or 1.5 g/l phenol ('high phenol') as sole carbon sources. Relative lipid amounts at different time points were estimated by staining cells with Nile red and measuring single cell fluorescence, which has been used to characterize total lipid contents in many oleaginous organisms (15,45) and correlates well with total lipid content (Supplementary Figure S4). Peak lipid accumulation occurs in the mid-stationary phase, followed by a decrease in lipid contents due to lipid consumption caused by extended carbon starvation. Because cells would be harvested at peak lipid accumulation for an industrial-scale process, we used the maximum Nile red fluorescence in the stationary phase to compare the lipid titer (Supplementary Table S4) between strains. The adapted strains showed similar maximum Nile red fluorescence to that of the WT strain when grown in glucose ( $P = 0.1$  and  $P = 0.09$  for evol33 and evol40, respectively). In contrast, when grown in low phenol, evol33 and evol40 showed a 2.5-fold ( $P = 5 \times 10^{-5}$ ) and 1.9-fold ( $P = 0.001$ ) increase in maximum Nile red fluorescence compared to WT, respectively (Supplementary Table S4). The adapted strains also showed high lipid accumulation in high phenol (Supplementary Table S4 and Supplementary Figure S5). In addition to titer, productivity is also an important factor for industrial scale processes, and we compared lipid productivity between strains by examining the average increase in Nile red fluorescence over time in exponential to stationary phase (Figure 2, Supplementary



**Figure 2.** Characterization of lipid accumulation. Comparison of total lipid (Nile red fluorescence) productivity between WT and the adapted strains. For all experiments, the nitrogen source was 0.05 g/l ammonium sulfate and an initial OD<sub>600</sub> was 0.3 with a culture volume of 10 ml. White bar = WT, gray bar = evol33 and black bar = evol40. Glucose = 1 g/l, Low Phenol = 0.75 g/l and High Phenol = 1.5 g/l. The data are the average of the arithmetic means of the Nile red fluorescence distribution obtained from three biological replicates. To take into account differences in the lag phase for each condition, time points were taken from the exponential to the stationary phase to determine the average increase in Nile red fluorescence over time for each condition. The asterisk indicates significant increases in lipid productivity compared to that of WT in the same condition ( $P < 0.05$ ; one mean, two-tailed Student's *t*-test). No growth was observed for the WT strain in 1.5 g/l phenol (<1 cell doubling in 10 ml cultures). Error bars represent one standard deviation with all staining and flow cytometry measurements performed on the same day.

Table S5, and Supplementary Figure S5). Here, we define lipid productivity as the lipid amount per cell per unit time. In glucose, all three strains showed similar lipid productivities ranging from 7.8 to 8.4 arbitrary unit (a.u.) h<sup>-1</sup>. While the WT strain demonstrated a lower lipid productivity in low phenol (5.8 a.u. h<sup>-1</sup>), the adapted strains showed similar or higher lipid productivities in low phenol compared to glucose conditions (13.6 and 8.9 a.u. h<sup>-1</sup> for evol33 and evol40, respectively). Compared to the WT strain, evol33 and evol40 has a 2.3- and 1.5-fold increase in lipid productivity in low phenol. The adapted strains also had higher lipid productivity in high phenol compared to glucose conditions (16.5 a.u. h<sup>-1</sup>,  $P = 0.001$  for evol33; and 19.4 a.u. h<sup>-1</sup>,  $P = 0.0001$  for evol40). These results prompted us to further characterize these strains using whole-genome sequencing and RNA-Seq.

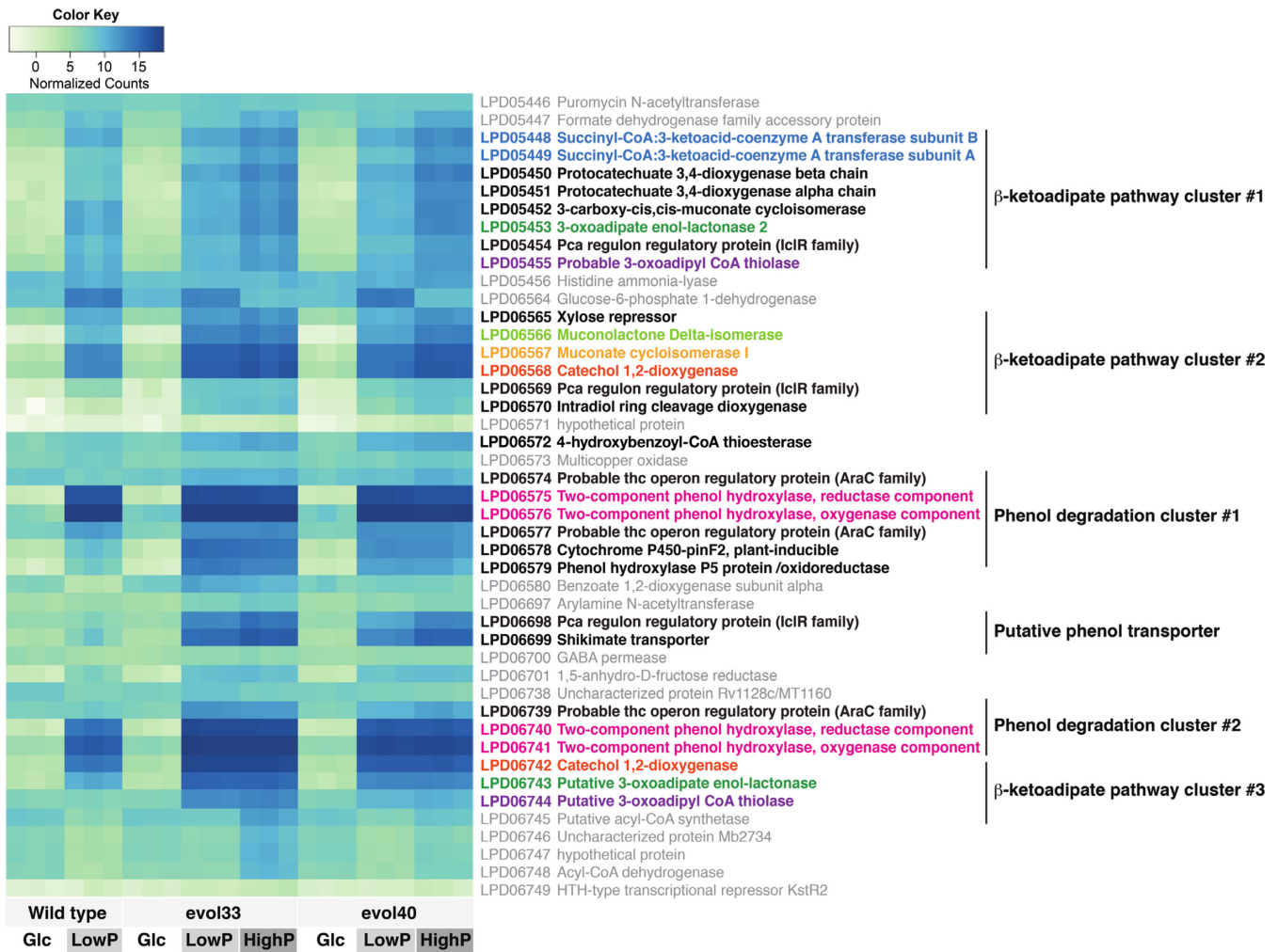
### Deep whole-genome sequencing of phenol-adapted strains

We sequenced the genomes of phenol-adapted *R. opacus* strains evol33 and evol40 along with the original wild-type strain to > 90X-coverage. For a reference genome, we used both annotated supercontigs maintained by the Broad Institute (Genbank # ABRH01000000 (20); OPAG\_#s) and the finished genome by the Chinese Academy of Science

(ASM59954v1 (48); LPD#s). Contrary to our initial expectations, there were no single nucleotide polymorphisms (SNPs) or insertion/deletions (indels) in promoter regions, transcriptional regulators, or sensors. In fact, the numbers of mutations were low, and when excluding mutations that were found in WT as well, there were only three consistent mutations (which were not due to sequencing errors) in evol40 and four mutations in evol33. While we explored the possibility of large duplications as genomes of *Rhodococcus* species are known to contain duplications (57–59)), we could not obtain consistent large duplication prediction based on short read sequencing only. The three common point mutations shared in the genomes of both evol33 and evol40 (Table 1A) are the following: LPD03112, annotated as ‘cytochrome c oxidase/cytochrome aa3 quinol oxidase’, contained a Phe to Ser mutation at amino acid position 335. LPD03747, annotated as ‘dicarboxylate carrier protein/citrate transporter/MFS transporter’, contained an Ile296 to Phe mutation. The third shared mutation occurred in OPAG\_05248, which is annotated as ‘acyl-CoA thioesterase II’, and had a silent mutation of G to A (Phe120) when mapped to the supercontigs (20). When mapped to the complete genome (48), however, this mutation occurred in an intergenic region between LPD04711 and LPD04712 (Table 1A). Regardless of the annotations, overall reads mapped to this region were very low (reads per kilobase of transcript per million reads (rpkm) 0–10) and not phenol-responsive (Table 1B,C), suggesting caution when considering the biological implication of this mutation. evol33 additionally had a gene LPD00118 with a SNP; this gene is annotated as ‘4-hydroxyphenylacetate permease/MFS transporter’ (Gly162 to Val). Given two transporter/permease genes containing consistently found SNPs, these results suggest that adapted strains have altered transport of phenol or related compounds.

### Transcriptomic analysis of phenol-adapted strains

*Stress adaptation in evol33 and evol40 compared to WT.* To investigate tolerance mechanisms further, we employed a comparative transcriptomic approach by sequencing ribosomal RNA-depleted total RNA from WT, evol33 and evol40 in the three medium conditions described in Materials and Methods. High phenol (1.5 g/l) was used for adapted strains only, since the wild-type strain did not reach adequate cell densities at this culture condition. High phenol condition data were used to assess whether the up- or downregulation of genes was phenol concentration-dependent. Three biological replicates of each strain and condition showed consistent gene expression results (Figure 3 and Supplementary Figure S7; in Figure 3, each square on a heat map represents a biological replicate). In addition to coping with the low nutrient-culture conditions (cells were grown in minimal media with low nitrogen to promote lipid accumulation), cells came in contact with phenol, which is generally toxic to cell membranes and cellular macromolecules (60,61). As expected, multiple putative stress genes were upregulated due to nitrogen-limited growth conditions and the toxicity of phenol. Some of the annotated stress genes that were upregulated in all strains in phenol growth conditions were ‘DNA protection



**Figure 3.** Transcriptomic and genomic information of genes involved in phenol degradation and utilization. Differential expression is shown for the phenol-responsive degradation operons, β-ketoadipate pathway gene clusters, a putative phenol transporter gene and transcriptional regulator genes. Raw counts were normalized using variance stabilizing transformation in DESeq2 to fit in the range of 20 across all genes. Darker colors indicate higher normalized counts (as shown in Color Key). Glc, 1 g/l glucose; LowP, 0.75 g/l phenol; and HighP, 1.5 g/l phenol. The color scheme is the same throughout Figures 3 and 4 for each step of the pathway. Each square on the heat map represents a biological replicate.

during starvation protein' LPD04106, 'heat shock protein 20 homolog' LPD07748 and 'molecular chaperone DnaK' LPD02079 (Table 2 and Supplementary Table S6, Sheet 13). LPD04106 was highly expressed in glucose, but even more upregulated in low phenol in WT and evol40 (Table 2 and Supplementary Table S6). evol33 showed modest upregulation of LPD04106 (2.64-fold) in low phenol. In adapted strains in high phenol, expression levels of LPD04106 were similar to that of glucose (no change in evol33 and 1.3-fold downregulation in evol40). LPD07748 was highly upregulated in low phenol in all strains, but to a less extent in adapted strains in high phenol compared to low phenol (Table 2 and Supplementary Table S6). LPD02079 was 16-fold upregulated in WT in phenol, while it was 3.1- to 6.3- fold upregulated in adapted strains in phenol (Table 2 and Supplementary Table S6). These results suggest that adapted strains show less stress response to phenol-containing media than WT.

*Phenol degradation and utilization genes were highly up-regulated in adapted strains.* Genes that were more than 1024-fold upregulated in phenol compared to glucose across all strains were LPD06575 and LPD06576, which are annotated as 'flavin-dependent monooxygenase, reductase subunit/phenol monooxygenase reductase' and '4-hydroxyphenylacetate 3-monooxygenase', respectively. In addition to these two genes, LPD06740, LPD06741 and LPD06742 were also more than ~1000-fold upregulated in phenol in evol33 and evol40 (Figure 3, Table 2, and Supplementary Figure S8). LPD06740 and LPD06741 have high identity to LPD06575 and LPD06576, and are in the same orientation in tandem (Figure 4B). LPD06742 is annotated as catechol 1,2-dioxygenase (Figures 3 and 4). WT did upregulate the LPD06740 operon in phenol, but the expression levels were 3- to 18-fold lower than that of the adapted strains (Figure 3, Table 2, and Supplementary Figure S8). LPD06575/LPD06740 and LPD06576/LPD06741 are likely to be small and large components of two-component phe-

**Table 2.** Summarized log<sub>2</sub> fold changes of selected genes (over glucose condition). WT LowP, WT in low phenol versus glucose; evol33 LowP, evol33 in low phenol versus glucose; evol33 HighP, evol33 in high phenol versus glucose; evol40 LowP, evol40 in low phenol versus glucose; evol40 HighP, evol40 in high phenol versus glucose. For DESeq2-normalized count data, counts are shown in each condition specified, rather than comparing to glucose (Glc, 1 g/l glucose; LowP, 0.75 g/l phenol; HighP, 1.5 g/l phenol). Count numbers were obtained using plotCounts function of DESeq2 and averaging triplicate samples. More information of this chart can be found in Supplementary Table S6.

Gene ID	Annotation / Strain:conditions	WT LowP	evol33 LowP	evol33 HighP	evol40 LowP	evol40 HighP		
<b>Genes with SNPs (* evol33 only)</b>								
LPD03112	Probable cytochrome c oxidase subunit 1-alpha	2.83	1.75	2.09	2.46	2.45		
LPD03747	Uncharacterized protein YfdE/MFS transporter	0.32	-0.40	0.62	0.62	1.59		
Between LPD04711 and 04712 (OPAG_05248)	Acyl-CoA thioesterase II	0.56	-0.12	-0.55	0.55	-1.03		
LPD00118*	Putative tartrate transporter/4-hydroxyphenylacetate permease	-2.54	0.13	0.25	-1.25	-0.92		
<b>Differentially expressed phenol degradation and utilization genes</b>								
LPD05449	Probable succinyl-CoA:3-ketoacid-coenzyme A transferase subunit A	4.12	4.45	6.00	4.09	6.41		
LPD05450	Protocatechuate 3,4-dioxygenase beta chain	4.78	6.39	8.07	5.73	8.24		
LPD05451	Protocatechuate 3,4-dioxygenase alpha chain	4.84	6.08	7.72	5.92	8.15		
LPD05452	3-carboxy-cis,cis-muconate cycloisomerase	6.14	5.75	7.55	5.26	7.81		
LPD05453	3-oxoadipate enol-lactonase 2	5.99	5.41	7.19	5.07	7.30		
LPD05454	Pca regulon regulatory protein	4.63	4.12	5.60	4.09	5.99		
LPD05455	Probable acetyl-CoA acyltransferase	4.10	3.69	5.08	3.73	5.23		
LPD06565	Xylose repressor	4.14	4.62	5.09	3.68	5.19		
LPD06566	Muconolactone Delta-isomerase	7.44	9.88	10.34	8.85	11.01		
LPD06567	Muconate cycloisomerase 1	7.19	9.73	10.12	8.67	10.46		
LPD06568	Catechol 1,2-dioxygenase	6.33	9.27	9.81	8.03	9.99		
LPD06570	Catechol 1,2-dioxygenase	2.77	6.66	7.13	5.08	6.42		
LPD06575	Two-component phenol hydroxylase, reductase component	12.57	12.51	12.11	12.88	12.21		
LPD06576	Two-component phenol hydroxylase, oxygenase component	12.18	11.48	11.24	10.98	10.99		
LPD06740	Two-component phenol hydroxylase, reductase component	9.36	12.53	12.52	11.34	11.37		
LPD06741	Two-component phenol hydroxylase, oxygenase component	9.17	12.25	11.87	10.43	10.45		
LPD06742	Catechol 1,2-dioxygenase	8.36	11.97	11.53	9.95	9.94		
LPD06743	Putative 3-oxoadipate enol-lactonase	5.24	9.74	9.48	7.45	7.49		
LPD06744	Putative beta-ketoadipyl CoA thiolase	0.13	4.68	4.82	2.01	2.32		
<b>Putative phenol transporters (&gt; 2 fold up in HighP compared to Glc)</b>								
LPD06699	Shikimate transporter	2.34	8.46	9.87	6.68	9.16		
LPD07505	Shikimate transporter	-1.68	2.82	3.09	0.61	1.53		
<b>Stress response genes</b>								
LPD04106	DNA protection during starvation protein	3.69	1.40	0.04	2.03	-0.44		
LPD07748	18 kDa antigen 2/heat shock protein 20	5.33	3.69	3.32	4.31	3.40		
LPD02079	Chaperone protein DnaK	4.00	1.63	2.48	2.66	2.24		
<b>Highly expressed genes (count data below)</b>								
LPD00472	Hypothetical protein/ porin protein	-1.59	-0.72	-1.94	-0.84	-1.98		
LPD04147	Linear gramicidin synthase subunit C	-1.15	-0.79	-1.81	-1.55	-2.51		
LPD03031	Nitrite extrusion protein 1	-0.32	-0.37	1.19	-0.15	1.49		
LPD03032	Nitrite reductase NAD(P)H large subunit	-0.57	-1.37	0.94	-1.05	1.24		
LPD06131	DNA-directed RNA polymerase subunit beta'	0.41	-0.64	0.22	-0.40	-0.04		
LPD06132	DNA-directed RNA polymerase subunit beta	0.16	-0.25	0.17	-0.09	-0.03		
LPD05487	ATP-dependent Clp protease ATP-binding subunit ClpX	-0.18	-0.28	-0.05	-0.07	0.45		
LPD06091	Elongation factor Tu	-0.29	-0.91	0.36	-0.74	0.15		
LPD04933	Proteasome-associated ATPase	0.01	-0.42	-0.86	0.03	-0.67		
LPD04928	Proteasome subunit beta	0.12	-0.26	-0.53	-0.13	-0.64		
<b>Expression levels of differentially expressed phenol degradation/utilization genes and highly expressed genes using plotCounts function (average counts from triplicate samples)</b>								
	WT Glc	WT LowP	evol33 Glc	evol33 LowP	evol33 HighP	evol40 Glc	evol40 LowP	evol40 HighP
LPD05449	34	592	47	995	3036	46	784	4008
LPD05450	25	702	30	2446	8167	35	1862	10944
LPD05451	23	609	21	1472	4716	20	1236	5917
LPD05452	27	2085	31	1626	5905	32	1233	7366
LPD05453	32	2108	42	1803	6392	46	1511	7274
LPD05454	44	1140	61	1065	3031	51	862	3286
LPD05455	87	1525	101	1270	3411	100	1315	3793
LPD06565	110	1962	145	3527	5008	128	1632	4750
LPD06566	6	991	8	8152	11264	5	2407	10971
LPD06567	48	7001	59	48979	67851	43	17726	62444
LPD06568	91	7284	101	61673	94392	80	21091	83313
LPD06569	12	193	11	654	696	10	430	759
LPD06570	4	32	6	640	862	7	237	605
LPD06574	642	425	549	1601	1548	611	850	1618
LPD06575	20	138793	30	190664	144683	25	186689	120411



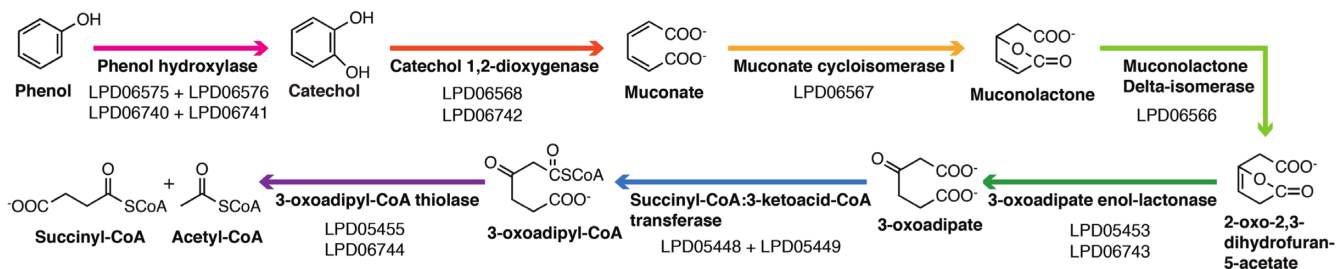
Table 2. Continued

LPD06576	226	1115301	534	1546914	1339798	495	1147992	1052801
LPD06577	372	3206	288	6635	7135	312	5360	4835
LPD06698	144	229	134	7329	18985	141	2107	8299
LPD06699	77	419	74	26215	70678	77	8062	45006
LPD06739	356	485	269	5067	4041	278	1392	2480
LPD06740	27	18588	29	224065	186503	26	67421	70570
LPD06741	119	69820	200	1005004	779070	166	229458	238395
LPD06742	50	16975	74	312714	232725	65	63793	65619
LPD06743	50	1944	48	41968	35435	44	7767	8113
LPD06744	255	282	265	6691	7520	297	1187	1496
LPD05449	34	592	47	995	3036	46	784	4008
LPD05450	25	702	30	2446	8167	35	1862	10944
LPD05451	23	609	21	1472	4716	20	1236	5917
LPD00472	391881	130913	280688	168207	72748	307073	171411	78144
LPD04147	276825	125049	254571	145768	71941	321709	109833	56532
LPD03031	176283	141517	213166	163206	485834	161971	145702	460250
LPD03032	199408	134490	234314	89615	450441	165498	79799	394788
LPD06131	34784	46649	37322	23675	43500	35730	27179	35080
LPD06132	44905	50714	40698	33754	45703	43683	40984	43259
LPD05487	15674	13971	13293	10789	12796	13163	12571	18072
LPD06091	20925	17195	26436	13873	29313	24111	14338	31098
LPD04933	104368	105771	108367	79769	59627	95300	97451	60274
LPD04928	47296	52024	41704	34480	28722	43987	40396	28430
<b>Most downregulated in phenol</b>								
<b>Gene ID</b>	<b>Annotation / Strain:conditions</b>		<b>WT LowP</b>	<b>evol33 LowP</b>	<b>evol33 HighP</b>	<b>evol40 LowP</b>	<b>evol40 HighP</b>	
LPD02697	NDMA-dependent methanol dehydrogenase		-12.68	-7.24	-6.78	-7.01	-10.83	
LPD02728	Inner membrane metabolite transport protein yjhE		-10.00	-6.56	-5.51	-8.71	-8.58	
LPD04550	Methane monooxygenase component A alpha chain		-9.02	-8.26	-6.55	-7.57	-9.00	
LPD04551	Methane monooxygenase component C		-8.40	-8.35	-6.50	-8.64	-8.44	
LPD04552	Methane monooxygenase component A beta chain		-7.38	-7.07	-6.56	-7.57	-6.93	
LPD04555	Uncharacterized protein MJ1129		-7.41	-7.35	-7.09	-7.31	-7.59	
LPD04556	NAD-dependent alcohol dehydrogenase		-7.91	-7.24	-6.78	-7.01	-6.98	
LPD04557	60 kDa chaperonin 3		-7.64	-7.03	-6.67	-6.43	-6.56	
<b>Lipid biosynthesis/metabolism</b>								
LPD06283	Heparin-binding hemagglutinin (TadA)		4.67	4.76	3.77	4.58	3.14	
LPD05549	Putative sterigmatocystin biosynthesis fatty acid synthase subunit alpha		6.01	5.04	5.84	5.04	5.82	
LPD00131	Stearoyl-CoA 9-desaturase		11.80	5.10	6.66	6.52	6.24	
LPD00130	Long-chain specific acyl-CoA dehydrogenase, mitochondrial		8.06	2.65	5.14	3.79	3.79	
LPD03000	Stearoyl-CoA 9-desaturase		7.24	5.15	5.67	5.48	5.85	
LPD05528	Stearoyl-CoA 9-desaturase		6.18	4.52	7.55	5.03	6.31	
LPD05529	Stearoyl-CoA 9-desaturase electron transfer partner		5.80	4.23	6.93	4.79	5.85	
LPD01195	Long-chain-fatty-acid--CoA ligase FadD19		-5.92	-2.13	-2.98	-4.41	-5.03	
LPD06540	Long-chain specific acyl-CoA dehydrogenase, mitochondrial		-3.82	-2.20	-1.27	-2.50	-2.72	
LPD04435	Long-chain-fatty-acid--CoA ligase		-2.34	-2.71	-3.93	-2.66	-3.51	
LPD03408	Long-chain specific acyl-CoA dehydrogenase, mitochondrial		-3.76	-2.88	-2.76	-3.12	-3.06	
LPD07841	Long-chain-fatty-acid--CoA ligase		-4.16	-3.66	-3.73	-3.96	-2.94	
LPD03067	Short/branched chain specific acyl-CoA dehydrogenase, mitochondrial		-3.08	-3.71	-2.96	-3.10	-2.72	
LPD03066	Short-chain specific acyl-CoA dehydrogenase, mitochondrial		-3.36	-3.78	-3.19	-3.35	-3.26	
<b>Transcriptional regulators</b>								
LPD06577	Probable thc operon regulatory protein		3.09	4.52	4.61	4.09	3.93	
LPD06739	Probable thc operon regulatory protein		0.44	4.23	3.89	2.31	3.14	
LPD06574	Probable thc operon regulatory protein		-0.59	1.55	1.49	0.47	1.39	
LPD03777	Probable thc operon regulatory protein		-1.78	-0.78	1.09	-2.19	0.81	
LPD03770	Probable serine/threonine-protein kinase pknK		0.38	0.08	1.62	-0.30	1.62	

nol hydroxylase, where LPD06575/LPD06740 are flavin reductases and LPD6576/LPD06741 are NADH-dependent phenol hydroxylases, as there are reports of biochemically proven two-component phenol hydroxylases in other *Rhodococcus* species (62,63). Gene duplication of this region is also documented, and *R. opacus* strain ICP is thought to have a third copy of the genes (59). Based on our transcriptional data, *R. opacus* appears to utilize both copies of the two-component phenol hydroxylase genes, but

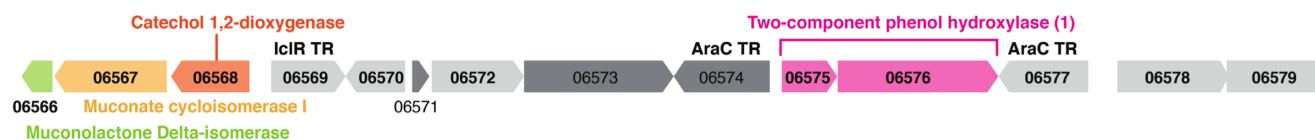
the adapted strains have the ability to upregulate the second copy (phenol degradation cluster #2; Figures 3 and 4B) more than the WT strain. Notably, the second copy of the phenol hydroxylase genes is directly upstream of a catechol 1,2-dioxygenase gene (Figure 4B), which is the first gene in the catechol branch of the  $\beta$ -ketoacid pathway (Figure 4A; 12). Both copies were tightly regulated by the presence of phenol, and AraC-family transcriptional regulators are found in the same operons (Figures 3 and 4B). Moreover,

A

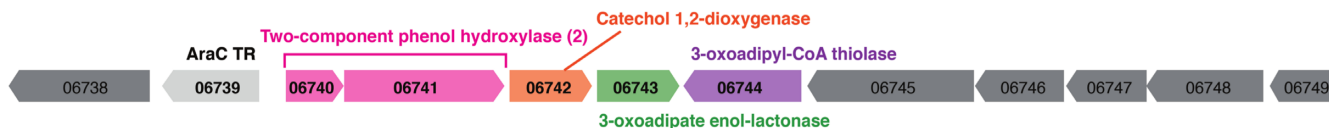


B

### $\beta$ -ketoadipate pathway cluster #2 + Phenol degradation cluster #1



### Phenol degradation cluster #2 + $\beta$ -ketoadipate pathway cluster #3



**Figure 4.** Pathway and genome maps showing genes involved in phenol degradation and utilization. (A) Ortho-cleavage branch of the  $\beta$ -ketoadipate pathway in *R. opacus*. An expanded pathway map including other phenol degradation pathways can be found in Supplementary Figure S10. The color scheme is the same throughout Figures 3 and 4 for each step of the pathway. (B) Genomic organization of the two phenol degradation operons. Light gray genes are highly upregulated in phenol. Darker gray genes are flanking genes that were not upregulated in phenol.

other downstream genes in the  $\beta$ -ketoadipate pathway, such as muconolactone delta isomerase and muconate cycloisomerase, are upregulated more in the adapted strains, especially in high phenol ( $\sim 2000$ -fold; Figures 3 and 4A), and they are located in the same operon as IclR-family regulators (LPD05454, LPD06569 and LPD06698; Figures 3 and 4A). We searched for genomic mutations near differentially regulated operons such as the phenol degradation cluster #2, but the closest SNP observed was over 2 Mbp upstream of differentially regulated genes in adapted strains, suggesting that the differential gene expression could be affected by an indirect cause, such as intracellular concentrations of regulatory molecules or metabolites.

RNA-Seq also revealed that *R. opacus* upregulates the gentisate pathway of aromatic degradation and utilization when phenol concentration was elevated to 1.5 g/l, while predicted meta-cleavage aromatic degradation pathway genes and anaerobic aromatic degradation genes did not seem utilized (Supplementary Figures S9 and S10). These results strongly suggest that adapted strains efficiently detoxify phenol mainly via the ortho-cleavage pathway and utilize it as a carbon source through the  $\beta$ -ketoadipate pathway (12).

There are over 380 genes annotated as transcriptional regulators or regulatory proteins in *R. opacus*. Of those,

ten genes are annotated as ‘probable the operon regulatory protein’ (AraC-family transcriptional regulator). Of the ten genes, LPD06577, LPD06739 and potentially LPD06574 are putative regulators of the highly phenol-responsive phenol degradation genes based on the proximity of the genes to phenol degradation genes and expression profiles (Figure 3). A gene annotated as ‘xylose repressor’ (LPD06565) may be involved in regulating the  $\beta$ -ketoadipate pathway cluster #2 because it is co-regulated with the genes in the cluster.  $\beta$ -ketoadipate pathway operons seem to be associated with genes annotated as Pca regulon regulatory protein (IclR family transcriptional regulator; LPD05454, LPD06569 and LPD06698). Although the gentisate pathway operon was adjacent to an AraC-family transcriptional regulator (LPD03777), their expression patterns were not similar (Supplementary Figure S9). However, their expression patterns were similar to an annotated serine/threonine-protein kinase gene (LPD03770) downstream of the gentisate pathway genes (Supplementary Figure S9).

Genes most downregulated by phenol in WT were LPD02697 (NDMA-dependent methanol dehydrogenase/lactaldehyde reductase;  $>6500$ -fold downregulated) and two hypothetical protein genes downstream of LPD02697 (Table 2 and Supplementary Table S6, Sheet 14). Adapted strains did not show dramatic downregulation

of this gene in low phenol, but did show strong downregulation in high phenol (Table 2 and Supplementary Table S6, Sheets 15-18). A homolog of LPD02697 is implicated in oligotrophic growth of *Rhodococcus erythropolis* N9T-4 (64). In all strains, methane monooxygenase component genes were among the most downregulated in phenol conditions.

When comparing the low phenol and glucose conditions in terms of global regulatory changes, evol33 had higher average levels of upregulation from fewer genes than WT and evol40, and both adapted strains had fewer downregulated genes than WT, which were also generally downregulated to a lesser extent than WT (Supplementary Tables S6 and S7 and Supplementary Figure S11).

*Upregulation of transporter genes in phenol.* While genomic data suggested that phenol transport could be altered in the adapted strains, it did not give a complete picture about how the adapted strains could tolerate and utilize higher concentrations of phenol compared to WT. Significant expression differences were not observed between WT and adapted strains in transporter genes identified to have SNPs by whole-genome sequencing. However, a few other putative transporter genes were observed to have altered expression levels (Table 2 and Supplementary Table S6). We identified a highly upregulated gene in both adapted strains (401-fold and 102-fold upregulated in evol33 and evol40, respectively, compared to 5-fold in WT in low phenol, and upregulated to higher degrees in high phenol in adapted strains), which is annotated as a shikimate transporter gene (LPD06699). Shikimate is a cyclohexene compound that is a precursor for aromatic amino acids, and it contains functional groups similar to many phenolic compounds derived from lignin (65). LPD07505, also annotated as a shikimate transporter, was 3.2-fold downregulated in WT in low phenol, but was upregulated in both low and high phenol in adapted strains, especially in evol33 (8.5-fold compared to glucose in high phenol; Table 2 and Supplementary Table S6). No other genes annotated as shikimate transporters showed similar upregulation patterns, and shikimate utilization genes were not upregulated in phenol (Supplementary Table S6, Sheet 12). Moreover, phenol was more quickly consumed by the adapted strains than WT (Supplementary Figure S3), suggesting that increased phenol uptake may play a role in phenol tolerance and utilization in *R. opacus*.

*Lipid biosynthesis and metabolism genes were differentially regulated in adapted strains.* *R. opacus* TadA is a heparin-binding hemagglutinin-like protein required for normal lipid droplet morphology, which is thought to colocalize with lipid droplets (48,66). Interestingly, the gene encoding TadA (LPD06283 (48,66)) was upregulated in phenol in all strains (Supplementary Table S6, Sheet 13; >24-fold in WT and >27-fold in evol33). This finding prompted us to speculate that lipid droplet morphology may be different in *R. opacus* strains depending on the carbon source. However, there was no marked change in size or number of lipid droplets when comparing cells in glucose to cells in phenol (Supplementary Figure S12). LPD05549, annotated as a fatty acid synthesis gene (FASI), was upregulated 33- to

64-fold in all strains in phenol compared to glucose. 8 out of the 11 genes annotated as stearoyl-CoA 9-desaturase or stearoyl-CoA 9-desaturase electron transfer partner genes were at least 4-fold upregulated in WT in phenol, whereas 24 of the 45 genes annotated as fatty acid-CoA ligase were greater than 2-fold downregulated in phenol in WT (19/45 genes in evol33 and 22/45 genes in evol40 were >2-fold upregulated), although the degree of fold change differed significantly between WT and adapted strains (representative genes shown in Table 2 and Supplementary Table S6, Sheet 13).

*Effects of nitrogen limitation on R. opacus transcriptome.* In glucose, strain-to-strain variation in global expression was low (Supplementary Figures S7 and S13), and the most highly expressed genes assessed by DESeq2 were LPD00472 (porin protein), LPD04147 (linear gramicidin synthase subunit C/non-ribosomal peptide synthase), LPD03031 (nitrite extrusion protein) and LPD03032 (nitrite reductase; Supplementary Table S6). LPD00472 was a highly abundant transcript and was slightly downregulated (up to 4-fold) in phenol growth conditions, probably reflecting the hydrophobic nature of phenol compared to hydrophilic glucose. Normalized counts using plotCounts function of DESeq2 comparing these genes, phenol degradation genes and constitutively highly expressed housekeeping genes (such as DNA-dependent RNA polymerase subunits (LPD06131/6132), ClpX (LPD05487), and elongation factor Tu (LPD06091)) are shown in Table 2. Because we used a minimal salts medium with a reduced nitrogen concentration, we expected that genes involved in nitrogen limitation would be highly expressed. This speculation was confirmed by the extremely high expression of LPD03031 and LPD03032 (Table 2). In fact, some of the most highly expressed genes across all strains in all medium conditions included proteasome genes LPD04933 (Mpa homolog/ATPase) and LPD04928 (proteasome endopeptidase complex beta subunit). As shown in *R. erythropolis* (67), *R. opacus* is likely equipped with a complete set of components to constitute active proteasomes, where proteins are presumably recycled and utilized in nitrogen-limited conditions. Changes in carbon source (glucose or phenol) did not seem to affect proteasome gene expression (Table 2; less than 2-fold up- or downregulated).

## DISCUSSION

In this work, we analyzed genomic and transcriptomic changes between adaptively evolved *R. opacus* mutants and the WT strain grown in phenol. Laboratory adaptive evolution has been used for many different bacteria to improve growth on diverse compounds (68). We used phenol as a sole carbon source in our adaptive evolution process and identified strains with higher tolerance to the toxic chemical by measuring their total biomass accumulation, lag phase and growth rate (Figure 1 and Supplementary Tables S1-S3). After ~200 generations, adapted strains contained 3-4 SNPs in the genomes, which was more rapid accumulation of SNPs than reported by Lenski et al. in their long-term evolution experiment and other labs ( $8.9 \times 10^{-11}$  per base-pair per generation; 69). Strong selective pressure by the

toxic compound being the sole carbon source probably contributed to the faster rate of SNP accumulation compared to the *E. coli* experiment, although the 'normal' rate of SNP accumulation during adaptive evolution in *R. opacus* has not been determined. Unfortunately, we did not identify 'obvious' SNPs that contributed to the striking transcriptomic profile difference between WT and evolved strains. The two adapted strains had improved growth profiles with increased biomass accumulation and shorter lag phases that could be due to altered transport of phenol in the mutants, leading to changes in intracellular concentrations of phenol and thus expression of phenol-responsive genes for phenol detoxification and degradation.

Differences were also observed between WT and adapted strains in terms of lipid accumulation. From RNA-Seq expression data, phenol appears to be mainly degraded into acetyl-CoA and succinyl-CoA via the  $\beta$ -ketoacid pathway (12). Acetyl-CoA can be integrated directly into fatty acid biosynthesis pathways, and succinyl-CoA can enter the TCA cycle to generate reducing equivalents and ATP for cell growth and lipid synthesis (70). Compared to WT, the adapted strains maintained the ability to synthesize and accumulate lipids when grown in glucose, but they both accumulated significantly more lipids than WT when grown in phenol (Figure 2, Supplementary Tables S4 and S5).

The adapted strains in this study demonstrated high phenol consumption capabilities that match or surpass other phenol-degrading bacterial strains and mixed cultures. *Pseudomonas putida* MTCC1194, which was adapted for growth in phenol as a sole carbon source, had maximum phenol degradation rates of  $\sim 12$  mg phenol/l/h during growth at its maximum phenol concentration of 1 g/l (71). Another approach used mixed cultures to increase phenol degradation capacity, with tolerances up to 0.8 g/l phenol and degradation rates of 15.4 mg/l/h (72). Another approach is to isolate strains directly from environments such as contaminated wastewater. *Bacillus brevis*, isolated from an industrial wastewater, showed the highest phenol tolerance and utilization at concentrations up to 1.75 g/l phenol with degradation rates of  $\sim 20$  mg/l/h (73). Our adapted strains showed higher tolerance than *B. brevis* with growth in 2 g/l phenol, and experiments at 1.5 g/l showed phenol degradation rates of 22 and 21 mg/l/h for evol33 and evol40, respectively (Supplementary Figure S3). In summary, evol33 and evol40 demonstrated comparative phenol degradation rates and tolerances to the best-performing environmental isolates.

Our transcriptomic data suggested that phenol-adapted strains express lower levels of stress-response genes in the conditions tested (low nitrogen, minimal media and phenol as a sole carbon source). Phenol is known to cause oxidative stress response in bacteria (74). In our study, *R. opacus* upregulated 'DNA protection during starvation protein' LPD04106, or Dps, which is thought to be involved in oxidative stress resistance (75,76). It seems likely that low nutrient condition can induce the gene, but the adapted strain, especially evol33, showed significantly lower upregulation of this gene compared to WT in phenol, suggesting that evolved strains sensed less oxidative stress in our experiment. A similar trend was observed with the major housekeeping chaperone genes. DnaK is a major bacterial

Hsp70 (70 kDa heat shock protein) and functions with co-chaperones GrpE and DnaJ to prevent aggregation of denatured proteins (77,78). *R. opacus* expressed 2 copies of the dnaK operons (LPD2079–02081 and LPD03845–03847) in the conditions we tested, and the major operon appeared to be LPD02079 and the following 3 genes. All DnaK and co-chaperone genes were upregulated more in WT than in the phenol-adapted strains in low phenol, indicating that the low-nutrient, phenol medium condition encountered by WT may cause proteins to misfold, and that the phenol-adapted strains encounter lower level of cellular stress than WT at the same concentration of phenol in the media.

We observed clear patterns of altered gene regulation in phenol-adapted strains compared to WT *R. opacus* in response to phenol. In WT, one of the two copies of phenol hydroxylase genes was dominantly upregulated, but in adapted strains the second copy (LPD06740 and LPD06741) was also upregulated to the same or even higher level (Figure 3, Table 2 and Supplementary Figure S8). Both SNP analysis and RNA-Seq results suggest that transport of phenol or compounds related to its metabolism is altered in the adapted strains. In fact, phenol was consumed at faster rates by the adapted strains than the WT strain (Supplementary Figure S3). evol33, which additionally had a putative transporter gene with a single point mutation compared to evol40 (Table 1), upregulated phenol degradation genes more strongly than evol40 (Figure 3). It is well documented that many identified transcriptional regulators for aromatic degradation and catabolism are AraC/XylS family (79–81). AraC/XylS-family transcriptional regulators bind to ligands (e.g. arabinose), and upregulation of effector genes is tunable by adjusting the concentration of ligands (82–84). Therefore, it is possible that adapted strains allow more phenol molecules to be transported into the cell, which in turn activate the second phenol degradation operons to the degree observed (6307-fold upregulation in evol33, 2624-fold upregulation in evol40 and 657-fold upregulation in WT; Figure 3 and Supplementary Table S6). Higher upregulation of  $\beta$ -ketoacid pathway genes in adapted strains could also be influenced by higher intracellular concentrations of phenol or its downstream metabolite, catechol, which has been observed with aromatic degradation pathways in other organisms (79,80).

There are 46 genes annotated as shikimate transporters in the complete genome of *R. opacus*. Of the 46, LPD06699 was uniquely co-upregulated with phenol degradation genes, at 353- and 103-fold in evol33 and evol40 comparing low phenol to glucose, and 937- and 572-fold in evol33 and evol40 comparing high phenol to glucose, respectively. In WT, LPD06699 was upregulated 5-fold in low phenol compared to glucose. Except for LPD07505, which was upregulated in phenol only in adapted strains, none of the other genes annotated as a shikimate transporter were regulated in the same manner (Supplementary Table S6). Furthermore, predicted shikimate utilization genes LPD05461, LPD06023, LPD17011 and LPD05485 were not highly transcribed or upregulated in phenol (Supplementary Table S6), suggesting that LPD06699 and potentially LPD07505 may be involved in phenol transport.

Another interesting finding was the upregulation of a putative YRNA and Rsr (85) in adapted strains in phe-

nol. YRNAs were recently found to exist in a wide variety of bacterial species, including *R. opacus* (86). *R. opacus* Rsr homolog, LPD06269, caught our attention because of its phenol-responsive upregulation in evolved strains, even though it was annotated as a hypothetical protein. LPD06269 was constitutively expressed at a medium-low level in all strains in glucose, but upregulated 52-fold in low phenol compared to glucose and 115-fold in high phenol compared to glucose in evol33 (8.3-fold and 17-fold in evol40), and upregulation of the predicted YRNA region showed the same pattern (Supplementary Figure S14). In WT, transcript levels of LPD06269 slightly decreased in low phenol (0.87-fold), and no upregulation of YRNA region was observed, suggesting that this response was phenol-dependent and specific to the adapted strains. We are in a process of focusing on small noncoding RNAs in *R. opacus*, since we removed the majority of small RNAs from the total RNA samples in order to obtain maximal amount of mRNA in this study. Nonetheless, our results indicate that small RNA-mediated global response may also be involved in adaptive evolution of *R. opacus* against phenol as a sole carbon source, and it may explain the ability of *R. opacus* to rapidly adapt to diverse environments without accumulating substantial genomic changes.

This work suggests that characterizing funneling pathways (which convert lignin-derived compounds into pathway metabolites such as catechol; 87) and transporters for phenolics is important for conversion of phenolics into fuels and chemicals. Genes converting phenol to catechol were some of the most highly upregulated genes when strains were grown in phenol compared to glucose, suggesting that phenol-to-catechol conversion may be a limiting step for growth on phenol. Under the condition that we tested (low nitrogen, minimal medium and aerobic condition), phenol seems to be degraded mainly by ortho-cleavage via the catechol branch of the  $\beta$ -ketoacid pathway into succinyl-CoA and acetyl-CoA, consistent with our previous result obtained from WT cultures grown in 0.5 g/l phenol supplemented with  $^{13}\text{C}$  glucose (37). Other phenolic compounds would require other funneling enzymes to convert them into precursors such as catechol and protocatechuic acid for degradation (87). This study also identified putative transporters for phenol, and transport of many different phenolic compounds would require either promiscuity for transporters or different transporters for each compound. Note that while previous studies on transporters of toxic chemicals focused on efflux pumps to minimize the intracellular concentrations of these compounds (88–90), *R. opacus* is likely to use its transporters to import toxic phenolics, suggesting a new mechanism of bacterial aromatic tolerance. Increased phenolic compound flux into the cell requires balancing between transport and degradation to prevent accumulation of toxic compounds within the cell.

A recent study demonstrated that adaptive evolution can increase *R. opacus* tolerance to lignocellulose-derived inhibitors (40). While they used media supplemented with glucose, our evolution process is different in that phenol was used as a sole carbon source to obtain highly tolerant mutants to phenol. Consequently, we were able to decouple the effects of other potentially available carbon sources in lignocellulose-derived feedstock on changes in the genome

and transcriptome. We also demonstrated that phenol as a sole carbon source can support *R. opacus* for both growth and accumulation of lipids (biodiesel precursors) and that adaptive evolution can increase lipid productivity, an important factor for current lignin conversion processes (15). Combining adaptive evolution and omics analyses, our approach provides insights into the tolerance mechanism of the promising production host, facilitating future lignocellulose, specifically lignin, valorization efforts.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

The authors thank Merima Beganovic and Yu Gao for scientific discussions, Dr Wandy Beatty for performing transmission electron microscopy and assisting with fluorescence microscopy and Jessica Hoisington-López for her Illumina sequencing expertise.

## FUNDING

U.S. Department of Energy [DE-SC0012705 to G.D., T.S.M. and M.F.]. Funding for open access charge: U.S. Department of Energy [DE-SC0012705].

*Conflict of interest statement.* None declared.

## REFERENCES

- Chundawat, S.P., Beckham, G.T., Himmel, M.E. and Dale, B.E. (2011) Deconstruction of lignocellulosic biomass to fuels and chemicals. *Annu. Rev. Chem. Biomol. Eng.*, **2**, 121–145.
- Sun, Y. and Cheng, J. (2002) Hydrolysis of lignocellulosic materials for ethanol production: a review. *Bioresour. Technol.*, **83**, 1–11.
- Peralta-Yahya, P.P., Zhang, F., del Cardayre, S.B. and Keasling, J.D. (2012) Microbial engineering for the production of advanced biofuels. *Nature*, **488**, 320–328.
- Ragauskas, A.J., Beckham, G.T., Biddy, M.J., Chandra, R., Chen, F., Davis, M.F., Davison, B.H., Dixon, R.A., Gilna, P., Keller, M. *et al.* (2014) Lignin valorization: improving lignin processing in the biorefinery. *Science*, **344**, 1246843.
- Steen, E.J., Kang, Y., Bokinsky, G., Hu, Z., Schirmer, A., McClure, A., del Cardayre, S.B. and Keasling, J.D. (2010) Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature*, **463**, 559–562.
- Rubin, E.M. (2008) Genomics of cellulosic biofuels. *Nature*, **454**, 841–845.
- Keating, D.H., Zhang, Y., Ong, I.M., McIlwain, S., Morales, E.H., Grass, J.A., Tremaine, M., Bothfeld, W., Higbee, A., Ulbrich, A. *et al.* (2014) Aromatic inhibitors derived from ammonia-pretreated lignocellulose hinder bacterial ethanologenesis by activating regulatory circuits controlling inhibitor efflux and detoxification. *Front. Microbiol.*, **5**, 402.
- Palmqvist, E. and Hahn-Hagerdal, B. (2000) Fermentation of lignocellulosic hydrolysates. II: inhibitors and mechanisms of inhibition. *Bioresour. Technol.*, **74**, 25–33.
- Strassberger, Z., Tanase, S. and Rothenberg, G. (2014) The pros and cons of lignin valorisation in an integrated biorefinery. *RSC Adv.*, **4**, 25310–25318.
- Sannigrahi, P. and Ragauskas, A.J. (2011) Characterization of Fermentation Residues from the Production of Bio-Ethanol from Lignocellulosic Feedstocks. *J. Biobased Mater. Bioenergy*, **5**, 514–519.
- Carrier, M., Loppinet-Serani, A., Denux, D., Lasnier, J.-M., Ham-Pichavant, F., Cansell, F. and Aymonier, C. (2011) Thermogravimetric analysis as a new method to determine the lignocellulosic composition of biomass. *Biomass Bioenergy*, **35**, 298–307.

12. Harwood, C.S. and Parales, R.E. (1996) The beta-ketoadipate pathway and the biology of self-identity. *Annu. Rev. Microbiol.*, **50**, 553–590.
13. Johnson, C.W. and Beckham, G.T. (2015) Aromatic catabolic pathway selection for optimal production of pyruvate and lactate from lignin. *Metab. Eng.*, **28**, 240–247.
14. Kosa, M. and Ragauskas, A.J. (2012) Bioconversion of lignin model compounds with oleaginous *Rhodococci*. *Appl. Microbiol. Biotechnol.*, **93**, 891–900.
15. Salvachua, D., Karp, E.M., Nimlos, C.T., Vardon, D.R. and Beckham, G.T. (2015) Towards lignin consolidated bioprocessing: simultaneous lignin depolymerization and product generation by bacteria. *Green Chem.*, **17**, 4951–4967.
16. Zaitsev, G.M., Uotila, J.S., Tsitko, I.V., Lobanok, A.G. and Salkinoja-Salonen, M.S. (1995) Utilization of Halogenated Benzenes, Phenols, and Benzoates by *Rhodococcus opacus* GM-14. *Appl. Environ. Microbiol.*, **61**, 4191–4201.
17. Di Gennaro, P., Terreni, P., Masi, G., Botti, S., De Ferra, F. and Bestetti, G. (2010) Identification and characterization of genes involved in naphthalene degradation in *Rhodococcus opacus* R7. *Appl. Microbiol. Biotechnol.*, **87**, 297–308.
18. Alvarez, H.M., Mayer, F., Fabritius, D. and Steinbuechel, A. (1996) Formation of intracytoplasmic lipid inclusions by *Rhodococcus opacus* strain PD630. *Arch. Microbiol.*, **165**, 377–386.
19. Taylor, C.R., Hardiman, E.M., Ahmad, M., Sainsbury, P.D., Norris, P.R. and Bugg, T.D. (2012) Isolation of bacterial strains able to metabolize lignin from screening of environmental samples. *J. Appl. Microbiol.*, **113**, 521–530.
20. Holder, J.W., Ulrich, J.C., DeBono, A.C., Godfrey, P.A., Desjardins, C.A., Zucker, J., Zeng, Q., Leach, A.L., Ghiviriga, I., Dancel, C. *et al.* (2011) Comparative and functional genomics of *Rhodococcus opacus* PD630 for biofuels development. *PLoS Genet.*, **7**, e1002219.
21. LeBlanc, J.C., Goncalves, E.R. and Mohn, W.W. (2008) Global response to desiccation stress in the soil actinomycete *Rhodococcus jostii* RHA1. *Appl. Environ. Microbiol.*, **74**, 2627–2636.
22. de Carvalho, C.C., Marques, M.P., Hachicho, N. and Heipieper, H.J. (2014) Rapid adaptation of *Rhodococcus erythropolis* cells to salt stress by synthesizing polyunsaturated fatty acids. *Appl. Microbiol. Biotechnol.*, **98**, 5599–5606.
23. Fischer, C.R., Klein-Marcuschamer, D. and Stephanopoulos, G. (2008) Selection and optimization of microbial hosts for biofuels production. *Metab. Eng.*, **10**, 295–304.
24. van der Geize, R. and Dijkhuizen, L. (2004) Harnessing the catabolic diversity of rhodococci for environmental and biotechnological applications. *Curr. Opin. Microbiol.*, **7**, 255–261.
25. Uz, I., Duan, Y.P. and Ogram, A. (2000) Characterization of the naphthalene-degrading bacterium, *Rhodococcus opacus* M213. *FEMS Microbiol. Lett.*, **185**, 231–238.
26. Taki, H., Syutsubo, K., Mattison, R.G. and Harayama, S. (2007) Identification and characterization of o-xylene-degrading *Rhodococcus* spp. which were dominant species in the remediation of o-xylene-contaminated soils. *Biodegradation*, **18**, 17–26.
27. Wang, B., Rezenom, Y.H., Cho, K.C., Tran, J.L., Lee do, G., Russell, D.H., Gill, J.J., Young, R. and Chu, K.H. (2014) Cultivation of lipid-producing bacteria with lignocellulosic biomass: effects of inhibitory compounds of lignocellulosic hydrolysates. *Bioresour. Technol.*, **161**, 162–170.
28. Na, K.S., Kuroda, A., Takiguchi, N., Ikeda, T., Ohtake, H. and Kato, J. (2005) Isolation and characterization of benzene-tolerant *Rhodococcus opacus* strains. *J. Biosci. Bioeng.*, **99**, 378–382.
29. Fuchs, G., Boll, M. and Heider, J. (2011) Microbial degradation of aromatic compounds - from one strategy to four. *Nat. Rev. Microbiol.*, **9**, 803–816.
30. Sainsbury, P.D., Hardiman, E.M., Ahmad, M., Otani, H., Seghezzi, N., Eltis, L.D. and Bugg, T.D. (2013) Breaking down lignin to high-value chemicals: the conversion of lignocellulose to vanillin in a gene deletion mutant of *Rhodococcus jostii* RHA1. *ACS Chem. Biol.*, **8**, 2151–2156.
31. Kurosawa, K., Boccazzi, P., de Almeida, N.M. and Sinskey, A.J. (2010) High-cell-density batch fermentation of *Rhodococcus opacus* PD630 using a high glucose concentration for triacylglycerol production. *J. Biotechnol.*, **147**, 212–218.
32. Xiong, X., Wang, X. and Chen, S. (2012) Engineering of a xylose metabolic pathway in *Rhodococcus* strains. *Appl. Environ. Microbiol.*, **78**, 5483–5491.
33. Kurosawa, K., Wewetzer, S.J. and Sinskey, A.J. (2013) Engineering xylose metabolism in triacylglycerol-producing *Rhodococcus opacus* for lignocellulosic fuel production. *Biotechnol. Biofuels*, **6**, 134.
34. Waltermann, M., Luftmann, H., Baumeister, D., Kalscheuer, R. and Steinbuechel, A. (2000) *Rhodococcus opacus* strain PD630 as a new source of high-value single-cell oil? Isolation and characterization of triacylglycerols and other storage lipids. *Microbiology*, **146**, 1143–1149.
35. Nicolaou, S.A., Gaida, S.M. and Papoutsakis, E.T. (2010) A comparative view of metabolite and substrate stress and tolerance in microbial bioprocessing: From biofuels and chemicals, to biocatalysis and bioremediation. *Metab. Eng.*, **12**, 307–331.
36. Nielsen, D.R. and Moon, T.S. (2013) From promise to practice. The role of synthetic biology in green chemistry. *EMBO Rep.*, **14**, 1034–1038.
37. Hollinshead, W.D., Henson, W.R., Abernathy, M., Moon, T.S. and Tang, Y.J. (2016) Rapid metabolic analysis of *Rhodococcus opacus* PD630 via parallel <sup>13</sup>C-metabolite fingerprinting. *Biotechnol. Bioeng.*, **113**, 91–100.
38. van der Pol, E.C., Bakker, R.R., Baets, P. and Eggink, G. (2014) By-products resulting from lignocellulose pretreatment and their inhibitory effect on fermentations for (bio)chemicals and fuels. *Appl. Microbiol. Biotechnol.*, **98**, 9579–9593.
39. Shumkova, E.S., Solianikova, I.P., Plotnikova, E.G. and Golovleva, L.A. (2009) Phenol degradation by *Rhodococcus opacus* strain 1G. *Prikl. Biokhim. Mikrobiol.*, **45**, 51–57.
40. Kurosawa, K., Laser, J. and Sinskey, A.J. (2015) Tolerance and adaptive evolution of triacylglycerol-producing *Rhodococcus opacus* to lignocellulose-derived inhibitors. *Biotechnol. Biofuels*, **8**, 76.
41. Kosa, M. and Ragauskas, A.J. (2013) Lignin to lipid bioconversion by oleaginous *Rhodococci*. *Green Chem.*, **15**, 2070–2074.
42. Wei, Z., Zeng, G., Huang, F., Kosa, M., Sun, Q., Meng, X., Huang, D. and Ragauskas, A.J. (2015) Microbial lipid production by oleaginous *Rhodococci* cultured in lignocellulosic autohydrolysates. *Appl. Microbiol. Biotechnol.*, **99**, 7369–7377.
43. Wei, Z., Zeng, G., Kosa, M., Huang, D. and Ragauskas, A.J. (2015) Pyrolysis oil-based lipid production as biodiesel feedstock by *Rhodococcus opacus*. *Appl. Biochem. Biotechnol.*, **175**, 1234–1246.
44. Kurosawa, K., Anthony Debono, C. and Sinskey, A.J. (2015) Lignocellulose-derived inhibitors improve lipid extraction from wet *Rhodococcus opacus* cells. *Bioresour. Technol.*, **193**, 206–212.
45. Blazeck, J., Hill, A., Liu, L., Knight, R., Miller, J., Pan, A., Otoupal, P. and Alper, H.S. (2014) Harnessing *Yarrowia lipolytica* lipogenesis to create a platform for lipid and biofuel production. *Nat. Commun.*, **5**, 3131.
46. Fisher, S., Barry, A., Abreu, J., Minie, B., Nolan, J., Delorey, T.M., Young, G., Fennell, T.J., Allen, A., Ambrogio, L. *et al.* (2011) A scalable, fully automated process for construction of sequence-ready human exome targeted capture libraries. *Genome Biol.*, **12**, R1.
47. Bowman, S.K., Simon, M.D., Deaton, A.M., Tolstorukov, M., Borowsky, M.L. and Kingston, R.E. (2013) Multiplexed Illumina sequencing libraries from picogram quantities of DNA. *BMC Genomics*, **14**, 466.
48. Chen, Y., Ding, Y., Yang, L., Yu, J., Liu, G., Wang, X., Zhang, S., Yu, D., Song, L., Zhang, H. *et al.* (2014) Integrated omics study delineates the dynamics of lipid droplets in *Rhodococcus opacus* PD630. *Nucleic Acids Res.*, **42**, 1052–1064.
49. Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–359.
50. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
51. Li, H. (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, **27**, 2987–2993.
52. Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C.A., Zeng, Q., Wortman, J., Young, S.K. *et al.* (2014) Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One*, **9**, e112963.

53. Chevreur, B., Wetter, T. and Suhai, S. (1999) Computer Science and Biology. *Proceedings of the German Conference on Bioinformatics (GCB)*. Hannover, Vol. **99**, pp. 45–56.
54. Liao, Y., Smyth, G.K. and Shi, W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, **30**, 923–930.
55. Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
56. Lee, D.H., Feist, A.M., Barrett, C.L. and Palsson, B.O. (2011) Cumulative number of cell divisions as a meaningful timescale for adaptive laboratory evolution of *Escherichia coli*. *PLoS One*, **6**, e26172.
57. Leahy, J.G., Batchelor, P.J. and Morcomb, S.M. (2003) Evolution of the soluble diiron monooxygenases. *FEMS Microbiol. Rev.*, **27**, 449–479.
58. Choi, K.Y., Kim, D., Chae, J.C., Zylstra, G.J. and Kim, E. (2007) Requirement of duplicated operons for maximal metabolism of phthalate by *Rhodococcus* sp. strain DK17. *Biochem. Biophys. Res. Commun.*, **357**, 766–771.
59. Groning, J.A., Eulberg, D., Tischler, D., Kaschabek, S.R. and Schlomann, M. (2014) Gene redundancy of two-component (chloro)phenol hydroxylases in *Rhodococcus opacus* ICP. *FEMS Microbiol. Lett.*, **361**, 68–75.
60. Sikkema, J., de Bont, J.A. and Poolman, B. (1995) Mechanisms of membrane toxicity of hydrocarbons. *Microbiol. Rev.*, **59**, 201–222.
61. Putrins, M., Ilves, H., Lilje, L., Kivisaar, M. and Horak, R. (2010) The impact of ColRS two-component system and TtgABC efflux pump on phenol tolerance of *Pseudomonas putida* becomes evident only in growing bacteria. *BMC Microbiol.*, **10**, 110.
62. Saa, L., Jaureguibeitia, A., Largo, E., Llama, M.J. and Serra, J.L. (2010) Cloning, purification and characterization of two components of phenol hydroxylase from *Rhodococcus erythropolis* UPV-1. *Appl. Microbiol. Biotechnol.*, **86**, 201–211.
63. Szokol, J., Rucka, L., Simcikova, M., Halada, P., Nesvera, J. and Patek, M. (2014) Induction and carbon catabolite repression of phenol degradation genes in *Rhodococcus erythropolis* and *Rhodococcus jostii*. *Appl. Microbiol. Biotechnol.*, **98**, 8267–8279.
64. Ohhata, N., Yoshida, N., Egami, H., Katsuragi, T., Tani, Y. and Takagi, H. (2007) An extremely oligotrophic bacterium, *Rhodococcus erythropolis* N9T-4, isolated from crude oil. *J. Bacteriol.*, **189**, 6824–6831.
65. Maeda, H. and Dudareva, N. (2012) The Shikimate Pathway and Aromatic Amino Acid Biosynthesis in Plants. *Annu. Rev. Plant Biol.*, **63**, 73–105.
66. MacEachran, D.P., Prophete, M.E. and Sinskey, A.J. (2010) The *Rhodococcus opacus* PD630 heparin-binding hemagglutinin homolog TadA mediates lipid body formation. *Appl. Environ. Microbiol.*, **76**, 7217–7225.
67. Yun, H.Y., Tamura, N. and Tamura, T. (2012) *Rhodococcus* prokaryotic ubiquitin-like protein (Pup) is degraded by deaminase of pup (Dop). *Biosci. Biotechnol. Biochem.*, **76**, 1959–1966.
68. Dragosits, M. and Mattanovich, D. (2013) Adaptive laboratory evolution—principles and applications for biotechnology. *Microb. Cell Fact.*, **12**, 64.
69. Wielgoss, S., Barrick, J.E., Tenaillon, O., Cruveiller, S., Chane-Woon-Ming, B., Medigue, C., Lenski, R.E. and Schneider, D. (2011) Mutation Rate Inferred From Synonymous Substitutions in a Long-Term Evolution Experiment With *Escherichia coli*. *G3: Genes, Genomes, Genet.*, **1**, 183–186.
70. Wells, T. Jr and Ragauskas, A.J. (2012) Biotechnological opportunities with the beta-ketoadipate pathway. *Trends Biotechnol.*, **30**, 627–637.
71. Kumar, A., Kumar, S. and Kumar, S. (2005) Biodegradation kinetics of phenol and catechol using *Pseudomonas putida* MTCC 1194. *Biochem. Eng. J.*, **22**, 151–159.
72. Saravanan, P., Pakshirajan, K. and Saha, P. (2008) Growth kinetics of an indigenous mixed microbial consortium during phenol degradation in a batch reactor. *Bioresour. Technol.*, **99**, 205–209.
73. Arutchelvan, V., Kanakasabai, V., Elangovan, R., Nagarajan, S. and Muralikrishnan, V. (2006) Kinetics of high strength phenol degradation using *Bacillus brevis*. *J. Hazard. Mater.*, **129**, 216–222.
74. Santos, P.M., Benndorf, D. and Sa-Correia, I. (2004) Insights into *Pseudomonas putida* KT2440 response to phenol-induced stress by quantitative proteomics. *Proteomics*, **4**, 2640–2652.
75. Almiron, M., Link, A.J., Furlong, D. and Kolter, R. (1992) A novel DNA-binding protein with regulatory and protective roles in starved *Escherichia coli*. *Genes Dev.*, **6**, 2646–2654.
76. Wei, X., Mingjia, H., Xiufeng, L., Yang, G. and Qingyu, W. (2007) Identification and biochemical properties of Dps (starvation-induced DNA binding protein) from cyanobacterium *Anabaena* sp. PCC 7120. *IUBMB Life*, **59**, 675–681.
77. Harrison, C. (2003) GrpE, a nucleotide exchange factor for DnaK. *Cell Stress Chaperones*, **8**, 218–224.
78. Calloni, G., Chen, T., Schermann, S.M., Chang, H.C., Genevaux, P., Agostini, F., Tartaglia, G.G., Hayer-Hartl, M. and Hartl, F.U. (2012) DnaK functions as a central hub in the *E. coli* chaperone network. *Cell Rep.*, **1**, 251–264.
79. Gallegos, M.T., Williams, P.A. and Ramos, J.L. (1997) Transcriptional control of the multiple catabolic pathways encoded on the TOL plasmid pWW53 of *Pseudomonas putida* MT53. *J. Bacteriol.*, **179**, 5024–5029.
80. Arengi, F.L., Pinti, M., Galli, E. and Barbieri, P. (1999) Identification of the *Pseudomonas stutzeri* OX1 toluene-o-xylene monooxygenase regulatory gene (touR) and of its cognate promoter. *Appl. Environ. Microbiol.*, **65**, 4057–4063.
81. Tropel, D. and van der Meer, J.R. (2004) Bacterial transcriptional regulators for degradation pathways of aromatic compounds. *Microbiol. Mol. Biol. Rev.*, **68**, 474–500.
82. Gallegos, M.T., Schleif, R., Bairoch, A., Hofmann, K. and Ramos, J.L. (1997) AraC/XylS family of transcriptional regulators. *Microbiol. Mol. Biol. Rev.*, **61**, 393–410.
83. Schleif, R. (2010) AraC protein, regulation of the l-arabinose operon in *Escherichia coli*, and the light switch mechanism of AraC action. *FEMS Microbiol. Rev.*, **34**, 779–796.
84. Khlebnikov, A., Risa, O., Skaug, T., Carrier, T.A. and Keasling, J.D. (2000) Regulatable arabinose-inducible gene expression system with consistent control in all cells of a culture. *J. Bacteriol.*, **182**, 7029–7034.
85. Wolin, S.L., Belair, C., Boccitto, M., Chen, X., Sim, S., Taylor, D.W. and Wang, H.W. (2013) Non-coding Y RNAs as tethers and gates: Insights from bacteria. *RNA Biol.*, **10**, 1602–1608.
86. Chen, X., Sim, S., Wurtmann, E.J., Feke, A. and Wolin, S.L. (2014) Bacterial noncoding Y RNAs are widespread and mimic tRNAs. *RNA*, **20**, 1715–1724.
87. Linger, J.G., Vardon, D.R., Guarnieri, M.T., Karp, E.M., Hunsinger, G.B., Franden, M.A., Johnson, C.W., Chupka, G., Strathmann, T.J., Pienkos, P.T. et al. (2014) Lignin valorization through integrated biological funneling and chemical catalysis. *Proc. Natl. Acad. Sci. U. S. A.*, **111**, 12013–12018.
88. Dunlop, M.J., Dossani, Z.Y., Szmids, H.L., Chu, H.C., Lee, T.S., Keasling, J.D., Hadi, M.Z. and Mukhopadhyay, A. (2011) Engineering microbial biofuel tolerance and export using efflux pumps. *Mol. Syst. Biol.*, **7**, 487.
89. Foo, J.L. and Leong, S.S. (2013) Directed evolution of an *E. coli* inner membrane transporter for improved efflux of biofuel molecules. *Biotechnol. Biofuels*, **6**, 81.
90. Fisher, M.A., Boyarskiy, S., Yamada, M.R., Kong, N., Bauer, S. and Tullman-Ercek, D. (2014) Enhancing tolerance to short-chain alcohols by engineering the *Escherichia coli* AcrB efflux pump to secrete the non-native substrate n-butanol. *ACS Synth. Biol.*, **3**, 30–40.