Adaptive laboratory evolution of Pseudomonas putida KT2440 improves p-coumaric and ferulic acid catabolism and tolerance

Mohamed, Elsayed Tharwat Tolba; Werner, Allison Z.; Salvachúa, Davinia; Singer, Christine A.; Szostkiewicz, Kiki; Rafael Jiménez-Diaz, Manuel; Eng, Thomas; Radi, Mohammad S.; Simmons, Blake A.; Mukhopadhyay, Aindrila

Total number of authors: 14

Published in: Metabolic Engineering Communications

Link to article, DOI: 10.1016/j.mec.2020.e00143

Publication date: 2020

Publisher's Version, also known as Version of record

Link back to DTU Orbit

Adaptive laboratory evolution of *Pseudomonas putida* KT2440 improves *p*-coumaric and ferulic acid catabolism and tolerance

Elsayed T. Mohamed, Allison Z. Werner, Davinia Salvachú, Christine A. Singer, Kiki Szostkiewicz, Manuel Rafael Jiménez-Díaz, Thomas Eng, Mohammad S. Radia, Blake A. Simmons, Aindrila Mukhopadhyay, Markus J. Herrgård, Steven W. Singer, Gregg T. Beckham, Adam M. Feist

Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Lyngby, Denmark
Renewable Resources and Enabling Sciences Center, National Renewable Energy Laboratory, Golden, CO, USA
Center for Bioenergy Innovation, Oak Ridge, TN, USA
Joint BioEnergy Institute, Emeryville, CA, USA
Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
Department of Bioengineering, University of California, San Diego, CA, USA

**ARTICLE INFO**

**Keywords:**
- Microbial lignin conversion
- Adaptive laboratory evolution
- Hydroxycinnamic acids

**ABSTRACT**

*Pseudomonas putida* KT2440 is a promising bacterial chassis for the conversion of lignin-derived aromatic compound mixtures to biofuels and bioproducts. Despite the inherent robustness of this strain, further improvements to aromatic catabolism and toxicity tolerance of *P. putida* will be required to achieve industrial relevance. Here, tolerance adaptive laboratory evolution (TALE) was employed with increasing concentrations of the hydroxycinnamic acids *p*-coumaric acid (*p*CA) and ferulic acid (FA) individually and in combination (*p*CA + FA). The TALE experiments led to evolved *P. putida* strains with increased tolerance to the targeted acids as compared to wild type. Specifically, a 37 h decrease in lag phase in 20 g/L *p*CA and a 2.4-fold increase in growth rate in 30 g/L FA was observed. Whole genome sequencing of intermediate and endpoint evolved *P. putida* populations revealed several expected and non-intuitive genetic targets underlying these aromatic catabolic and toxicity tolerance enhancements. *PP_3350* and *ttgB* were among the most frequently mutated genes, and the beneficial contributions of these mutations were verified via gene knockouts. Deletion of *PP_3350*, encoding a hypothetical protein, recapitulated improved toxicity tolerance to high concentrations of *p*CA, but not an improved growth rate in high concentrations of FA. Deletion of *ttgB*, part of the TtgABC efflux pump, severely inhibited growth in *p*CA + FA TALE-derived strains but did not affect growth in *p*CA + FA in a wild type background, suggesting epistatic interactions. Genes involved in flagellar movement and transcriptional regulation were often mutated in the TALE experiments on multiple substrates, reinforcing ideas of a minimal and deregulated cell as optimal for domesticated growth. Overall, this work demonstrates increased tolerance towards and growth rate at the expense of hydroxycinnamic acids and presents new targets for improving *P. putida* for microbial lignin valorization.

1. Introduction

Lignin is an abundant aromatic macromolecule found in plant cell walls where it plays key roles in defense, structure, and nutrient transport (Boerjan et al., 2003). Despite over a century of research, the industrial-scale conversion of lignin to high-value products is not profitable, with the exception of a few niche products (Ragauskas et al., 2014; Tuck et al., 2012; Zakzeski et al., 2010). For the burgeoning global industry utilizing lignocellulosic biomass for the production of renewable fuels and chemicals, where most of the emphasis to date has been on...
polysaccharide conversion, techno-economic analysis and life-cycle assessments have shown that lignin valorization to co-products can lead to lower carbohydrate-derived fuel prices and improve the overall sustainability of a lignocellulosic biorefinery (Baral et al., 2019; Corona et al., 2018; Davis et al., 2013; van Duuren et al., 2011).

The main technical challenge for valorization of most natural and processed lignin results from the inherent heterogeneity of the macromolecule, stemming from its biosynthesis, and results in a mixture of C-O and C-C bonds in the polymer (Boerjan et al., 2003). Correspondingly, most catalytic and thermal depolymerization processes for lignin produce heterogeneous mixtures of aromatic products that pose a challenge to standard separations unit operations (Li et al., 2015; Rinaldi et al., 2016; Schutyser et al., 2018; Sun et al., 2018; Zakzeski et al., 2010). However, the heterogeneous nature of lignin-derived aromatic compound mixtures can be overcome by aromatic catalytic microbials, which commonly exhibit a battery of enzymes to funnel aromatic compounds into single intermediates (Abdelaziz et al., 2016; Becker et al., 2018; Bekham et al., 2016; Bugg and Rampanour, 2015; Fuchs et al., 2011; Hartwood and Parales, 1996; Linger et al., 2014). In aerobic pathways, ring-opening dioxygenases oxidatively cleave the aromatic rings of these central intermediates and assimilate the substrates into central carbon metabolism. The application of metabolic engineering and synthetic biology to microbial lignin valorization has demonstrated multiple products, including polyhydroxyalkanoates (Lin et al., 2016; Linger et al., 2014; Salvachúa et al., 2018; Sonoki et al., 2018; Vardon et al., 2015), aromatic catalytic intermediates (Barton et al., 2018; Becker and Wittmann, 2019; Johnson et al., 2019, 2017; Kohlstedt et al., 2018; Okamura-Abe et al., 2016; Salvachúa et al., 2018; Sonoki et al., 2015; Vardon et al., 2015), fatty acids (Liu et al., 2018; Salvachúa et al., 2015; Zhao et al., 2016), and aromatic compounds (Mycroft et al., 2015; Sainsbury et al., 2013), among others (Johnson and Beckham, 2015).

Among microbial hosts under development for biological lignin valorization, Pseudomonas putida KT2440 (hereafter P. putida) is a particularly promising bacterial chassis. P. putida exhibits broad metabolic versatility and genetic flexibility (Clarke, 1982; von Graevenitz, 1976), and harbors a substantial number of aromatic catalytic pathways relevant to lignin (Jimenez et al., 2002). These advantages have made P. putida attractive for biotechnological applications including production of a broad range of chemicals (Belda et al., 2016; Nikel et al., 2016, 2014; Poblete-Castro et al., 2012; Schmid et al., 2001; Timmis et al., 1994), and bioremediation of pollutants in the environment (Dejonghe et al., 2001; Timmis et al., 1994).

Much of the research to date on aromatic monomer conversion with P. putida focuses on the hydroxycinnamic acids, p-coumaric acid (pCA) and ferulic acid (FA), as substrates (Beckham et al., 2016; Calero et al., 2018; Johnson et al., 2017; Linger et al., 2014; Ravi et al., 2017; Salvachúa et al., 2020, 2018; Vardon et al., 2015). These compounds are both found as ester-linked pendant groups on lignin, and FA can be incorporated into both hemicellulose and lignin (Ralph, 2010; Ralph et al., 1994). Alkaline treatments of grasses, where hydroxycinnamic acids are abundant, can yield substantial amounts of these monomers as “clip-off” compounds via saponification reactions (Karlen et al., 2020; Karp et al., 2016; Munson et al., 2016). P. putida exhibits higher tolerance to pCA (and likely FA) compared to Escherichia coli, potentially due to membrane integrity (Calero et al., 2018). However, cellular growth of P. putida strains engineered convert pCA to value-added compounds is inhibited at >10 g/L pCA (Salvachúa et al., 2018), indicating that further improvement of pCA and FA catabolic capacity and toxicity tolerance is necessary.

Here, we employ tolerance adaptive laboratory evolution (TALE) with increasing levels of pCA, FA, or an equal mass ratio of both acids (hereafter referred to as pCA + FA) to generate P. putida strains capable of growing at elevated hydroxycinnamic acid concentrations. Endpoint populations have improved growth rates in high concentrations of pCA and FA, and mutations in intermediate and endpoint populations are described. Reverse engineering demonstrates that Pp3350, encoding a hypothetical protein, and tgtB, encoding the efflux pump membrane protein TgT, both contribute to the improved phenotype. Additionally, static adaptive laboratory evolution (ALE) in glucose identified mutations related to flagellar transport and transcriptional regulation which may be generally beneficial under the employed cultivation conditions. Overall, the evolved strains and the genes identified in this work are promising targets for future improvement of P. putida aromatic catalytic capacity and toxicity tolerance, ultimately enhancing lignin bioconversion processes.

2. Materials and methods

2.1. Strains and media

P. putida KT2440 provided by Dr. Pablo I. Nikel was cultivated in modified M9 minimal media (1X M9 salts, 2 mM MgSO4, 100 μM CaCl2, 1X trace elements) (Merck, Germany). M9 salts (10X) consisted of 68 g/L Na2HPO4 anhydrous, 30 g/L KH2PO4, 5 g/L NaCl, and 20 g/L (NH4)2SO4. Trace elements (200X) consisted of 3 g/L FeSO4 7H2O, 4.5 g/L ZnSO4 7H2O, 0.3 g/L CoCl2 6H2O, 0.4 g/L Na2MoO4 2H2O, 4.5 g/L CaCl2 2H2O, 0.2 g/L CuSO4 5H2O, 1 g/L H3BO3, 15 g/L EDTA, 0.1 g/L KI, and 0.7 g/L MnCl2 4H2O adjusted to pH 4 with HCl. pCA and FA (TCI, Belgium) stocks were prepared fresh prior to each experiment by dissolving each acid in H2O and adjusting to pH 7 with NaOH. Carbon sources were added to the media in the specified concentration and the medium was sterilized by 0.22-μm filtration.

2.2. TALE and ALE experiments

The TALE and ALE experiments were conducted using an automated liquid handler platform as previously described (LaCroix et al., 2015; Mohamed et al., 2019, 2017; Sandberg et al., 2017). To prepare pre-cultures, single isolates of P. putida on LB agar plates were inoculated at 1% (v/v) in 15 mL of M9 medium with 10 g/L glucose, 8 g/L pCA, 10 g/L FA, or an equal mass mixture of 3 g/L pCA and 3 g/L FA in 30 mL test tubes with four biological replicates for the glucose ALE and six biological replicates for each aromatic TALE (Table S1). Cultures were incubated at 30 °C in a heat block, mixed at 1100 rpm with full aeration, and OD600nm was measured periodically on Sunrise plate reader (Tecan Group Ltd., Switzerland). For aromatic TALEs, 900 μL of cells were transferred to new media with an 18–20% (v/v) increase in substrate concentration when the OD600nm reached approximately 1.3 (i.e., cells were passaged during the exponential phase of growth to maintain a constant selection pressure). TALE experiments were terminated when the first of two criteria were met: either reaching substrate solubility limit (40 g/L for FA or pCA) or reaching the allowed run timeframe (67 days for pCA and pCA + FA). A glucose ALE experiment (wherein the glucose concentration does not change and therefore is referred to as ALE, as opposed to TALE) was also performed to identify media- or cultivation-specific adaptation mutations, as well as the added benefit of potentially identifying mutations to improve fitness on glucose. For the glucose ALE experiment, 150 μL of cells were transferred to fresh media with 10 g/L glucose when the OD600nm reached approximately 1.3. Glucose ALE experiments were terminated when the parallel replicates showed no change in growth rate. During the exponential phase, the slope of the natural logarithm of OD600nm as a function of time of at least three OD600nm measurements from each flask was used to compute the growth rate. The cumulative number of cell divisions (CCDs) was calculated from the summation of all cell division events, from the initial cell number and the total number of generations per flask using the equation:

$$CCD = \sum_{i=1}^{n} N_0 (2^w - 1),$$

where m is the total number of achieved flasks and the number of generations (n) were calculated from the initial ($N_0$) and the final number of cells ($N$) measured in the working volume using equation.
\[ n = \log(N / N_0) / \log(2). \]

CCD has been demonstrated to serve as a meaningful timescale for ALE experiments (Lee et al., 2011). At the defined flask numbers, aliquots of cultures were sampled, mixed with 50% glycerol at a ratio of (1:1), and frozen at \(-80^\circ C\) before preparation of samples for genomic DNA sequencing. Individual isolates were isolated at a later date by streaking population samples on agar plates.

2.3. Whole genome sequencing and analysis

For each aromatic TALE (pCA, FA, and pCA + FA) and the glucose ALE, 1–2 intermediate isolates, the endpoint population, and a single endpoint isolate were selected and prepared for whole genome sequencing as follows. Isolates were chosen from agar plates with a media and the corresponding carbon source, inoculated into LB medium and cultivated overnight, genomic DNA was extracted with PureLink® Genomic DNA Extraction kit (Invitrogen, CA, USA), quality was assessed by evaluating Abs260nm/Ab280nm using a Nanodrop (Thermo Fisher scientific, USA), DNA concentration was measured using a Qubit (Thermo Fisher scientific, USA) broad range assay, and paired-end sequencing libraries were generated using the Illumina 300 cycle (150 bp x 2) kit (San Diego, CA, USA). Sequencing was performed on an Illumina NextSeq system (Illumina, USA). The average coverage for each sample was over 60X. The sequencing files were analyzed using a previously described in-house script (Phaneuf et al., 2019) based on bowtie2 (Detherage and Barrick, 2014) and the NCBI NC_002947 version 4 reference genome for \( P. putida \) KT2440, (www.ncbi.nlm.nih.gov/nucore/NC_002947.4/) was used for annotation of genes. For population samples, a filter was applied to exclude mutations with a frequency of less than 0.50, unless the same mutation was found in an isolate. This value was selected to focus on clearly causal mutations.

2.4. Evaluation of TALE endpoint populations

Cells were revived from glycerol stocks on M9 minimal medium plates supplemented with 5 g/L of the carbon compound(s) from which the populations were evolved to maintain selection pressure for the acquired mutations, inoculated in 30 mL tubes into 15 mL of M9 minimal medium containing 5 g/L of the same aromatic compound used in the plates, and incubated at 30 °C, 1100 rpm until the OD\(_{500}\) reached 3.0. This provided adequate cell density for inoculation without subjecting the cells to an extended stationary phase. Wild type \( P. putida \) cells were revived in the same manner as the cells subject to comparison. Cells were inoculated into 15 mL of fresh M9 minimal medium supplemented with the carbon source specified in each experiment at a 1:100 dilution in 30 mL test tubes with biological duplicates and incubated at 30 °C with aeration provided by a magnetic stir bar, which was found to provide the best aeration and mixing in this cultivation format. OD\(_{500}\) was measured using the Sunrise plate reader (Tecan Group Ltd., Switzerland).

2.5. Plasmid and strain construction

Deletion of PP_3350 was performed using the antibiotic/sacB method of gene replacement, as described previously (Blomfield et al., 1991; Johnson and Beckham, 2015). The plasmid pC222 for deleting PP_3350 was constructed by amplifying 1 kilobase (kb) PP_3350 homology regions with Q5® High-Fidelity DNA Polymerase (New England Biolabs, USA) from \( P. putida \) genomic DNA (Table S2; Table S3), assembly into the plK18mobsacB plasmid digested with EcoRI and HindIII using NEBbuild® HiFi DNA Assembly Master Mix (New England Biolabs, USA), transformation into \( NEB \) E. coli DH5-alpha F’ cells, and the correct sequence was verified using Sanger sequencing (GENEWIZ, Germany). pC222 was transformed into \( P. putida \) as previously described (Choi et al., 2006) and sucrose selection and diagnostic colony PCR with MyTaq™ DNA Polymerase (Bioline, USA) were followed to identify a clone with proper deletion mutation. The resulting strain was named CJ782 (Table S4). Deletion of tggB was also performed using the anti-biotic/sacB method. The plasmid pTE289 was constructed using the isothermal DNA assembly method (Gibson et al., 2009) using ~1 kb homology arms tggBPP_1385 flanking the 5’ and 3’ upstream and downstream sequences flanking the tggB (PP_1385) open reading frame. (Table S2-3). A 389 bp spacer sequence derived from Saccharomyces cerevisiae SMCI containing a unique, randomly generated DNA barcode (‘S’-TACTGGCACC-3’) was inserted in the deleted region to aid identification of the strain using next generation sequencing methods. Homolog arms and spacer sequences were synthesized by IDT (Coraville, IA) or amplified from \( P. putida \) genomic DNA. DNA fragments were inserted into plK18mobsacB following plasmid linearization with the Sall endonuclease (Table S2-3). Following verification of plasmid assembly with Sanger sequencing, the plasmid was transformed into wild type \( P. putida \) via conjugation with E. coli S-17 as a donor strain or TALE endpoint isolates using electroporation and sucrose selection on LB agar plates with 25% (w/v) sucrose and diagnostic colony PCR was followed to identify a clone with the tggB deletion (strain MJ11, Table S4). For colony PCR, colonies directly from a selective agar plate were picked with a sterile toothpick into 50 μL 20 mM NaOH and boiled for 45 min at 95 °C. 2 μL of the reaction was used to inoculate a 25 μL PCR reaction with Q5® High-Fidelity DNA Polymerase (New England Biolabs, USA). When needed, antibiotics were used at the following concentrations: 30 μg/mL chloramphenicol, 50 μg/mL kanamycin, or 10 μg/mL gentamicin.

2.6. Evaluation of reverse-engineered \( P. putida \) ΔPP_3350 and endpoint TALE populations

Cells were revived from glycerol stocks in 50 mL M9 minimal medium supplemented with 3 g/L of the aromatic compound from which the populations were evolved or 10 g/L glucose in baffled flasks and incubated at 30 °C and 225 rpm until the OD\(_{500}\) reached 3.0. Cells were inoculated in fresh medium to an OD\(_{500}\) of 0.15–0.20 in 300 μL and incubated in 100-well honeycomb plates in a BioscreenC™ (Growth Curves USA, USA) at 30 °C, maximum continuous shaking, and abs\(_{420-580}\) measurements were taken every 15 min. Lag phase was calculated as the time that the cells required to reach an OD\(_{500}\) of 0.50.

2.7. Evaluation of reverse-engineered \( P. putida \) ΔttgB and endpoint TALE populations

Cells were revived from glycerol stocks by two sequential overnight incubations at 30 °C in M9 minimal media supplemented with 5 g/L of the carbon compound(s) from which the populations were evolved, diluted to an OD\(_{500}\) of 0.01–0.02 into 800 μL of growth medium, and incubated in 48-well flow plates in a BioLector® at 30 °C and 1500 rpm (m2p-labs GmbH, Germany). Light scattering measurements were collected every 15 min intervals and converted to optical density OD\(_{500}\) using standard curve calibration.

3. Results

TALE experiments with aromatic compounds were performed by cultivating six independent parallel replicates of \( P. putida \) in minimal media supplemented with pCA, FA, or an equal mass mixture of pCA and FA (“pCA + FA”). When the growth rate reached 0.15 h\(^{-1}\), the cells were propagated into media with a 18–20% increase in substrate concentration (Fig. S1). The gradual concentration increases, starting with concentrations that permitted growth, were designed to evolve cells with improved tolerance to pCA and FA. Here, growth rate and lag phase were used as measures to assess toxicity tolerance with increases in growth rate and decreases in lag time deemed superior. Additionally, a constant condition ALE experiment was performed with \( P. putida \) propagated in minimal media supplemented with a static concentration of glucose to identify general adaptation to media components and/or the cultivation
The aromatic TALEs were passaged for 350 to 921 generations, enabling evolution over 23–35 \times 10^{11} \text{ cumulative cell divisions (CCDs, Fig. 1a, Fig. S1). By the final passage, which occurred when TALEs reached each termination criteria (see Methods Section 2.2), evolved \textit{P. putida} populations developed the capacity to grow in 3.3-, 5.0-, and 2.5-fold increased concentrations of \textit{pCA}, \textit{FA}, and \textit{pCA} + \textit{FA}, respectively (Fig. 1b). The four independent glucose ALEs progressed much faster than the aromatic TALEs, achieving 64 \times 10^{11} \text{ CCDs (Fig. 1a, Table S1). Growth rate increased by 1.69 \pm 0.04 fold in glucose ALEs, primarily in the first third of the experiment (Fig. S2). The aromatic TALEs displayed oscillatory growth rate increases/decreases (Fig. S1), as expected given the provided changes in aromatic acid concentration. Interestingly, the final growth rate of the FA populations increased from the start of the TALE experiments while the concentration increase by a fold-change (FC) of 4 (Fig. 1), likely due to the significant amount of time it was evolved at a constant solubility-limit concentration (Fig. S1; Table S1). In \textit{pCA} TALEs, no large changes in the population growth rates were observed during concentration increases (Fig. S1). In \textit{pCA} + FA TALEs, the population growth rates generally decreased with increases in acid concentrations (Fig. S1).

We next evaluated the growth rate of evolved populations to determine if the improved growth rate could be recapitulated. Population screens were performed on evolved lineages to determine the fold change in growth rate at different hydroxycinnamic acid concentrations as compared to the starting wild type strain. Two of the four to six replicate endpoint populations were selected per TALE condition and cultivated in test tubes alongside the parent wild type \textit{P. putida}. The growth rates were compared to estimate catabolic capacity; notably, these values are not directly comparable to the values recorded during the evolution experiments on the automated platform due to differences in cultivation set-up. Growth rate improvements on \textit{pCA} were modest, ranging from 0.93- to 1.37-FC with TALE#31 (evolved in \textit{pCA} + \textit{FA}, 1.37-FC) performing the best on 10 g/L \textit{pCA} and TALE#19 (evolved in \textit{FA}, 1.72-FC) performing best on 20 g/L \textit{pCA} (Fig. 2a). Growth rate improvements were observed on 30 g/L \textit{FA} from \textit{pCA}, \textit{FA}, and \textit{pCA} + \textit{FA} TALEs with 2.5, 3.4-, 2.0-average FCs observed, respectively (Fig. 2b). Glucose-evolved strains

- **Fig. 1.** Experimental parameters for \textit{p}-coumarate (\textit{pCA}) tolerance adaptive laboratory evolution (TALE), ferulate (\textit{FA}) TALE, equal mass mixture \textit{pCA} and \textit{FA} (\textit{pCA} + \textit{FA}) TALE, and glucose adaptive laboratory evolution (ALE). (a) Duration of each experiment in number of generations and cumulative cell divisions (CCD). (b) Substrate concentration (g/L) at the start (C_i) and end (C_f) of each experiment, and the fold-change (FC) increase from t_0 to t_f. For the \textit{pCA} + \textit{FA} conditions, substrate concentration is presented as a sum where each individual compound is present in a 1:1 mass ratio. (a-b) Error bars represent the standard deviation of independent replicate evolution experiments (n = 6 for \textit{pCA}, \textit{FA}, and \textit{pCA} + \textit{FA}, n = 4 for glucose).

- **Fig. 2.** Growth rate screens of wild type \textit{P. putida} and evolved populations from \textit{p}-coumarate (\textit{pCA}, TALE#3 and TALE#9), ferulate (\textit{FA}, TALE#13 and TALE#19), equal mass mixture of \textit{p}-coumarate and ferulate (\textit{pCA} + \textit{FA}, TALE#31 and TALE#37) TALEs and glucose (ALE#37 and ALE#38) taken at the endpoint. Cells were cultivated in test tubes in M9 minimal media supplemented with (a) 10 g/L \textit{pCA} or 20 g/L \textit{pCA}, (b) 10 g/L \textit{FA} or 30 g/L \textit{FA}, (c) 10 g/L \textit{pCA} + \textit{FA} (5 g/L each) or 20 g/L \textit{pCA} + \textit{FA} (10 g/L each), and (d) 10 g/L glucose. Dashed grey lines are provided to aid in the visual comparison of mutant values to wild type values. Error bars represent the absolute difference between biological duplicates.
were found in all three stress conditions for (Fig. 3, Table 1, Supplementary Data File 1). Many of the converged accD, kefB-III, and the intergenic region between acid TALEs but not the glucose ALE, including replicates on the same carbon source or in a different carbon source.

23 mutations were identified in clonal isolates or populations from each TALE and ALE experiment. Starting strain-specific mutations and misaligned reads were filtered (see Section 2.3, for more detail on analysis of the whole genome sequencing data). In total, 62, 139, 76, and 23 mutations were identified in clonal isolates or populations from pCA TALE, FA TALE, pCA + FA TALE, and glucose ALE experiments, respectively, (Supplementary Data File 1). In each TALE and ALE experiment, over 70% of the mutations were single nucleotide polymorphisms (SNPs) with insertion and deletion prevalence varying by experiment (Fig. S3).

Converged mutations were identified by comparing all clonal and population samples and selecting genes or genetic regions which presented a mutation in ≥2 independent TALE or ALE experiments, either replicates on the same carbon source or in a different carbon source. Several genes were mutated in more than one of the hydroxycinnamic acid TALEs but not the glucose ALE, including sucA, rho, PP_3350, tggB, accD, kefB-III, and the intergenic region between PP_5245 and kefB-III (Fig. 3, Table 1, Supplementary Data File 1). Many of the converged mutations from glucose ALEs were shared with at least one of the aromatic TALEs (Fig. 3, Table 1, Supplementary Data File 1) implying that these adaptations are related to the cultivation conditions and/or M9 medium rather than the carbon source. Several of the frequently mutated genes have known functions, and in some cases have demonstrated involvement in tolerance or increased aromatic catabolic capacity.

To identify mutations which may contribute to the improved phenotypes in the evolved populations, whole genome sequencing was performed on three intermediate isolates and the endpoint populations in each TALE and ALE species (Anand et al., 2020), see also Supplementary Text 1 for additional cases.

We then analyzed whether distinct mutations arose early (when substrate concentrations were low) or late (when substrate concentrations were high) in the TALE time-course, which could suggest distinct catabolism enhancement or toxicity tolerance adaptations, respectively. To do this, we divided the experiment into three conditions, i.e. periods at specific aromatic acids concentration: “low stress” condition (15, 20, and 16 g/L pCA, FA, and pCA + FA, respectively), “moderate stress” condition (25, 40, and 32 g/L pCA, FA, and pCA + FA, respectively), and “high stress” condition (40 g/L (70 asks), 26 g/L (62 asks), and 40 g/L (100 asks) for FA, pCA, and pCA + FA, respectively). *Mutations in PP_3350 were found in all three stress conditions for pCA TALEs, but only moderate and high stress conditions for FA and pCA + FA TALEs.

Fig. 3. Overview of mutations. (a) A landscape of the converged mutations identified in P. putida from pCA TALEs (purple), FA TALEs (green), pCA + FA TALEs (orange), or glucose ALEs (grey). Genes/genetic loci are listed on the left-hand side. Line width is proportional to the number of unique occurrences in which a given gene or genetic region was mutated. See Supplementary Data File 1 for a full list of mutations and Table 1 for a list of converged mutations. (b) Venn Diagram of converged mutations identified in P. putida from pCA, FA, and pCA + FA TALEs at different evolutionary time points of the experiments from clones and populations at the approximate achieved number of passages (flasks). TALEs were divided into a “low stress” condition (20 g/L (28 flasks), 15 g/L (27 flasks), and 16 g/L (31 flasks) for FA, pCA, and pCA + FA, respectively), a “moderate stress” condition (40 g/L (70 flasks), 25 g/L (48 flasks), and 32 g/L (65 flasks) for FA, pCA, and pCA + FA, respectively), and a “high stress” condition (40 g/L (100 flasks), 26 g/L (62 flasks), and 40 g/L (100 flasks) for FA, pCA, and pCA + FA, respectively). *Mutations in PP_3350 were found in all three stress conditions for pCA TALEs, but only moderate and high stress conditions for FA and pCA + FA TALEs.
as well as groups common to both moderate and high stress levels (Fig. 3b). Mutations to accD, sucA, rho, and PP_3350 were shared between the moderate and high stress condition; a notable exception was that PP_3350 was identified in the designated low stress pCA condition. Mutations to wbpl were only found in the pCA low stress condition, suggesting that this mutation serves to enhance pCA catabolism. Conversely, mutations to the intergenic region between the moderate and high stress levels which harbor mutations in the corresponding gene.

We next sought to understand if individual mutations that underlie the improved phenotypes could be identified. Two of the most frequently mutated genes in pCA, FA, and pCA + FA TALEs were PP_3350, encoding a hypothetical protein, and ttgB, a subunit of a membrane efflux pump (Table 1, Supplementary Data File 1). PP_3350 is a hypothetical protein with low homology (~30% identity) to known alginate porins. We identified SNPs, insertions, and deletions in PP_3350 in multiple independent pCA, FA, and pCA + FA TALEs. Conversely, we identified only SNPs in ttgB in multiple independent and parallel pCA and pCA + FA TALEs. Accordingly, we hypothesized that deletion of PP_3350 would improve growth in pCA and/or FA given the prevalence of predicted loss of function mutations, whereas deletion of ttgB would diminish growth in pCA and/or FA given the role of TtgB as an efflux pump membrane protein (Duque et al., 2001; Teran et al., 2003) with the potential role to export aromatic compounds. To test these hypotheses, we constructed single-gene deletions in wild type P. putida of PP_3350 (strain CJ782) and ttgB (strain MJ1D1) and examined growth alongside endpoint populations which harbor mutations in the corresponding gene.

We first evaluated the effect of PP_3350 deletions on growth on hydroxycinnamic acids. Endpoint populations from pCA TALE#7, FA TALE#23, and pCA + FA TALE#25 were selected for comparison with wild type and CJ782 as they each harbor deletions in PP_3350 of varying length, in addition to other mutations (Fig. 4a, Table S5, Supplementary Data File 1). Growth was evaluated in “low” (10 g/L) and “high” (20 g/L for pCA and 30 g/L for FA) substrate concentrations corresponding to the near-starting and late substrate levels, respectively, in the TALEs (Fig. 1, Table S1). Growth rate was calculated to estimate catabolic capacity and lag phase was calculated to estimate substrate toxicity tolerance. P. putida ΔPP_3350 (strain CJ782), TALE#7, TALE#23, and TALE#25 endpoint populations exhibited 1.4-, 5.4-, 4.9-, and 4.4-fold decreases in lag phase and 1.8-, 2.1-, 1.7-, and 2.3-fold increases in growth rate, respectively, in 20 g/L pCA (Fig. 4b). In 30 g/L FA, the growth improvement was more specific to the conditions under which the strains were evolved, with FA-evolved TALE#23 displaying a 2.2-fold increase in growth rate and a 1.9-fold decrease in lag phase as compared to wild type P. putida (Fig. 4c). The deletion of PP_3350 does not fully recapitulate the observed phenotype in TALE#23, indicating that additional mutations – including both those directly affecting FA catabolic pathways as well as general changes to cellular physiology, regulation, and/or metabolism – present in this strain may underlie the improved phenotype. All evolved strains display a decrease in growth rate in 10 g/L pCA + FA, yet improved tolerance (as measured by decreased lag phase) and catabolic capacity in 20 g/L pCA + FA (Fig. 4d). In summary, a PP_3350 deletion strain recapitulates the improved pCA tolerance, but not the improved FA growth rate observed in the TALE endpoint populations. We next examined the effect of ttgB mutations on growth in aromatic-containing media. TtgB is part of the TtgABC resistance-nodulation-division (RND) efflux pump. TtgABC is involved in resistance to antibiotics in P. putida KT2440 (Fernandez et al., 2012; Godoy et al., 2010; Kim et al., 2017) with direct evidence for the efflux of short chain alcohols and tolune in P. putida DOT-T1 (Basler et al., 2018; Ramos et al., 1998). TtgB SNPs, the only mutation type associated with this ORF, occurred in the pCA + FA and pCA TALE experiments with 5 and 3 unique non-synonymous SNPs occurring across all replicates in each condition, respectively. Interestingly, all of the SNPs were detected in the last flask sequenced (Fig. 3b) for all experiments where the ttgB mutations were present, along with additional mutations, suggesting that perhaps the adaptive mechanism involving TtgB is synergistic or epistatic with other accumulated mutations. Thus, we evaluated the effect of ttgB deletions in clones from TALE#25, #27, and #31 populations. The pCA + FA condition was chosen for evaluation as this condition resulted in the most unique ttgB mutations identified from all conditions, suggesting its

Table 1
Converged mutations identified in P. putida from pCA TALEs, FA TALEs, pCA + FA TALEs, or glucose ALEs. See Supplementary Data File 1 for a list of all identified mutations.

<table>
<thead>
<tr>
<th>Gene name (Locus ID)</th>
<th>Mutation type (unique counts)</th>
<th>Gene annotation(s)</th>
<th>Experiment(s) in which mutations were identified</th>
</tr>
</thead>
<tbody>
<tr>
<td>gacS (PP_1650)</td>
<td>SNP (7), DEL (3)</td>
<td>sensor protein GacS</td>
<td>pCA, FA, pCA + FA, Glucose</td>
</tr>
<tr>
<td>oaprB (PP_1445)</td>
<td>SNP (2), DEL (1)</td>
<td>carbohydrate-selective porin</td>
<td>pCA + FA, Glucose</td>
</tr>
<tr>
<td>urk1</td>
<td>SNP (6), INS (1)</td>
<td>BarA/UvrX two-component system response regulator</td>
<td>pCA, FA, pCA + FA, Glucose</td>
</tr>
<tr>
<td>sucs (PP_4189)</td>
<td>SNP (7)</td>
<td>2-oxoglutarate dehydrogenase subunit E3</td>
<td>pCA, FA, pCA + FA</td>
</tr>
<tr>
<td>rho (PP_5214)</td>
<td>SNP (10)</td>
<td>transcription termination factor Rho</td>
<td>pCA, FA, pCA + FA</td>
</tr>
<tr>
<td>PP_3350</td>
<td>SNP (6), DEL (12), INS/DUP (3), INS (1)</td>
<td>hypothetical protein</td>
<td>pCA, FA, pCA + FA</td>
</tr>
<tr>
<td>tgb (PP_1385)</td>
<td>SNP (2)</td>
<td>efflux pump membrane protein Tgb</td>
<td>pCA, pCA + FA</td>
</tr>
<tr>
<td>acdC (PP_1996)</td>
<td>SNP (3)</td>
<td>acetyl-CoA carboxylase carboxyltransferase subunit beta</td>
<td>pCA, pCA + FA</td>
</tr>
<tr>
<td>accA (PP_1607)</td>
<td>SNP (2)</td>
<td>acetyl-CoA carboxylase carboxyltransferase subunit alpha</td>
<td>pCA</td>
</tr>
<tr>
<td>PP_5245, kefB-III</td>
<td>SNP (2)</td>
<td>ArgC family transcriptional regulator/glutathione-regulated</td>
<td>pCA, FA, pCA + FA</td>
</tr>
<tr>
<td>(PP_5246)</td>
<td>SNP (4)</td>
<td>potassium/H+ antiporter</td>
<td>pCA, pCA + FA</td>
</tr>
<tr>
<td>fBQ (PP_4173)</td>
<td>SNP (6), DEL (1)</td>
<td>transenprololar FliQ</td>
<td>FA, pCA + FA, Glucose</td>
</tr>
<tr>
<td>dnaK (PP_4272)</td>
<td>SNP (2)</td>
<td>chaperone protein DnaK</td>
<td>FA</td>
</tr>
<tr>
<td>wbpl (PP_1804)</td>
<td>SNP (1), INS (2)</td>
<td>glycolys transferase Wbpl</td>
<td>pCA</td>
</tr>
<tr>
<td>PP_1060, PP_1061</td>
<td>SNP (2)</td>
<td>glutamate synthas large subunit/ATP-dependent DNA helicase</td>
<td>pCA, pCA + FA</td>
</tr>
<tr>
<td>rpmB (PP_6052)</td>
<td>SNP (1), DEL (1)</td>
<td>RNA polymerase sigma-54 factor</td>
<td>pCA + FA</td>
</tr>
<tr>
<td>sdxB (PP_4191)</td>
<td>SNP (1)</td>
<td>succinate dehydrogenase flavoprotein subunit</td>
<td>pCA + FA</td>
</tr>
<tr>
<td>oxyB (PP_5309)</td>
<td>SNP (1), INS/DUP (1)</td>
<td>oxidative and nitrosative stress transcriptional dual regulator</td>
<td>pCA + FA, Glucose</td>
</tr>
</tbody>
</table>

* Locus identifiers (IDs) are provided from the NCBI Prokaryotic Genome Annotation Pipeline Version 4.10 (Oct. 2019).

b Mutation type abbreviations: SNP, single nucleotide polymorphism; INS, insertion; DEL, deletion; DUP: duplication. The number in parentheses indicates the unique number of counts this type of mutation was observed across the re-sequenced samples, in either a population or isolate.

Annotations assigned from NCBI Reference Genome NC_002947.4.

Experiment abbreviations: pCA, TALE in p-coumaric acid; FA, TALE in ferulic acid; pCA + FA, TALE in equal mass mixture of p-coumaric acid and ferulic acid; Glucose, ALE in glucose.
importance. The selected populations each contained a non-synonymous SNP in \texttt{ttgB}, as well as 6-9 additional mutations, and were screened along with wild type (Fig. 5a, Table S6) in low (5 g/L each) and high (10 g/L each) pCA.\footnote{P. \textit{putida} \texttt{ΔttgB} (strain MJD1) grows at 80\%, i.e., a 20\% reduction, and 104\%, i.e., a 4\% increase, of the growth rate as compared to wild type at 10 and 20 g/L pCA, respectively (Fig. 5b-c). In examining the evolved strains’ phenotypes, at 20 g/L pCA, TALE#27 grows at 209\% of the wild type rate, which is drastically reduced to 19\% upon deletion of \texttt{ttgB} (Fig. 5c). A similar reduction in growth rate is observed upon deletion of \texttt{ttgB} in TALE#25, #27, and #31 in both low and high pCA + FA (Fig. 5b-c). These data suggest that a mutated TtgB in TtgABC may play a role in efflux of aromatic intermediates or by-products, which is essential as the strains become specialized with the additional mutations contained in the evolved TALE clones.}

### 4. Discussion and conclusions

\textit{P. \textit{putida}} has long been considered exceptionally tolerant to toxic compounds such as wastewater, environmental pollutants, and aromatic compounds from biomass waste due in large part to its metabolic versatility (Belda et al., 2016; dos Santos et al., 2004; Regenhardt et al., 2002). Evolutionary (Calero et al., 2018) and rational engineering approaches (Franden et al., 2018; Jayakody et al., 2018) have been successful at further improving toxicity tolerance towards developing \textit{P. \textit{putida}} as a robust chassis for bioremediation and valorization of waste streams (Calero and Nikel, 2019; Dvorák et al., 2017; Martínez-García et al., 2015, 2014a; Nikel et al., 2016; Nikel and de Lorenzo, 2018). Here, we expand on this work with a focus on two plant-based hydroxycinnamic acids, pCA and FA (Ralph, 2010), with both intrinsic and indirect (i.e., metabolite intermediates and byproducts) toxicity via TALE. Accordingly, the major outcomes of this study are: i.) the generation of a set of optimized strains which have increased compound-specific and cross-tolerance phenotypes for use as platform strains for further engineering; ii.) the identification of specific mutational mechanisms associated with the enhanced phenotypes, with two specific frequently occurring mutations validated to be causal; and iii.) the demonstration that the already perceived tolerant \textit{P. \textit{putida}} can be improved via a TALE approach and the specific performance of the characterized strains establish a high mark of tolerance for the chosen substrate environments for \textit{P. \textit{putida}}. Notably, the generated strains also have implications for use as catabolically-enhanced platform strains, including glucose consumption, gives the nature of the experimental approach (i.e., starting with a relatively low initial concentration).
The experimental design of TALE provides a continually changing selection pressure that could be generalized as more selective toward fast growth rate early in the experiment (when substrate concentration is low) and more selective toward toxicity tolerance late in the experiment (when substrate concentration is high). We observed an increased growth rate in endpoint populations and isolates from FA TALEs and, to a lesser extent, pCA TALEs (Figs. 1 and 2) whereas evolved strains from pCA TALEs displayed substantially reduced lag phase (Fig. 5). Distinct mutations were identified in low, moderate, and high stress conditions, defined as early, mid, and late time-points in the TALE experiments (Fig. 3b). Interestingly, PP_3350 mutations were identified only in moderate and high stress conditions for FA and pCA + FA TALEs but all three conditions for pCA TALEs. Considering pCA is more toxic to P. putida than FA (Salvachúa et al., 2018), the pressure for toxicity tolerance is likely higher in pCA TALEs and thus may underlie the earlier occurrence of mutations which increase tolerance to pCA. Differences in solubility and therefore bioavailability between pCA and FA could also contribute to this observation.

The dual nature of aromatic compounds as both substrates and growth inhibitors raises interesting questions regarding the evolutionary pressures P. putida may face in natural environments. Likely, the concentration of pCA and FA in nature is lower than those utilized here, but the environmental complexity (e.g., additional metabolites) is higher, which would give rise to different evolutionary trajectories. Similar to previous work (Salvachúa et al., 2018), our data suggest that pCA is a more toxic substrate than FA, and that both substrates elicit a lag phase prior to growth that can be overcome, in part, by evolution. These results underscore the interplay between tolerance and utilization, and further suggest that strain design for valorization of lignin streams may be optimal in a manner specific to cultivation mode (e.g., batch mode relative to fed-batch mode) and that employing different strain engineering strategies may result in an optimal approach for a given application.

PP_3350 was frequently mutated with predicted loss of function mutations in the pCA, FA, and pCA + FA TALEs (Table 1). Recapitulation of a PP_3350 deletion in a wild type background decreased the lag phase in 20 g/L pCA by > 30 h, in line with the evolved isolates (Fig. 5). The PP_3350 protein sequence has high homology to the AlgE porin identified from multispecies Pseudomonas (accession: WP.010954226) with 100% identity using protein blast BLASTP 2.10.0+ (Altschul et al., 1997), suggesting PP_3350 is an outer membrane porin. While the function of PP_3350 has yet to be determined, deletion of a porin which enables import of pCA/FA may result in the decreased entry, and therefore decreased toxicity, effects due to pCA/FA in the periplasm and/or cytoplasm. However, it is unclear why ttgB, as opposed to other efflux pumps such as ttg2B, which has previously been described to play a role in pCA tolerance (Calero et al., 2018), was mutated in this study. We speculate that the SNPs observed in ttgB may enable enhanced efflux of pCA.
pCA- or FA-derived compounds and/or catabolic intermediates. Given the difference in phenotype between the wild type and TALE endpoints, which also harbor additional mutations (some containing PP_3350 mutations (Table S6)), it is possible that we are observing an epistatic effect whereby mutations are initially enhancing growth, but TgB is then essential for efflux of pCA or intermediates at higher concentrations. Future work will explore the effect of individual SNPs on TgB substrate selectivity/function in the efflux of pCA, FA, or derived intermediates.

Several gene targets identified may be important towards generally improving growth in the media and/or cultivation conditions. For example, gacS was mutated in all TALE and ALE experiments (Supplementary Data File 1). The GacS/GacA two-component regulatory system is implicated in biofilm formation (Martínez-García et al., 2014) and flagellar transport (Kim et al., 2014), suggesting alternate regulation of these functions may be generally beneficial in well-mixed planktonic cultures, as has been previously reported (Bentley et al., 2020). Genes related to flagellar transport (i.eq and fliQ (Blanco-Romero et al., 2018)) were also mutated in several TALE/ALE conditions (Fig. 2, Supplementary Data File 1), further strengthening the argument that deregulation of flagellar transport improves growth under the experimental conditions applied here due to conservation of resources as has been demonstrated previously (Martínez-García et al., 2014b; Utrilla et al., 2016). These results support the concept of a “minimal cell” for industrial biotechnology (Lara and Guillermo, 2020), both conceptualized and realized in P. putida by the de Lorenzo group (Lieder et al., 2015; Martínez-García et al., 2014a; Martínez-García and de Lorenzo, 2016).

The benefit of engineering some genetic targets may be substrate-specific whereas others may be broadly beneficial. Thus, it would be interesting to understand the effect of these mutations in toxicity tolerance and substrate utilization toward other lignin-derived aromatic compounds (e.g., vanillate, syringate, and 4-hydroxybenzoate) as well as non-polar aromatic compounds present in lignin-derived streams from thermal biomass deconstruction processes (e.g., guaiacol, syringol, and phenol) (Black et al., 2016; Jayakody et al., 2018). However, general adaptations observed in shake flasks may not always translate to improved performance in scale-up bioreactors or scaled-down plate-based culturing (e.g., differences were seen in the phenotypes for the strains evolved here depending on the culturing method utilized). Investigating the impact of selected key mutations on a broader range of lignin-derived aromatics and at-scale is warranted in future work.

Data availability

The DNA resequencing read files supporting the results of this article are available in the NCBI repository under Sequence Read Archive (SRA) BioProject ID: PRJNA624660. The data set supporting the results of this article is included in the article (and its Supplementary files). Mutation Tables and Breseq output analysis should be accessible on ALEdb (aledb.org) under project name: “P. putida Hydroxycinnamic TALE”.

Credit author contributions

Conceptualization: GTB, AMF; Data curation: ETM; Formal analysis: ETM, AZW, DS, KS, MJD, TE, MSR, AM, MJH, SWS, GTB, AMF; Funding acquisition: BAS, AM, MJH, SWS, GTB, AMF; Investigation/Experimen-tation: ETM, AZW, DS, CS, KS, MJD, TE, MSR; Supervision: DS, AM, SWS, GTB, AMF; Roles/Writing - Original Draft: ETM, AZW, GTB, AMF; Writing - Review & Editing: All authors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Funding was provided to ETM, MSR, MJH, and AMF by The Novo Nordisk Foundation Grant No. NNFI400C011269 and No. NNFI00C016517. This work was authored in part by Alliance for Sustainable Energy, LLC, the manager and operator of the National Renewable Energy Laboratory for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding was provided to DS, CAS, KS, and GTB by the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy BioEnergy Technologies Office. AZW and GTB were funded by The Center for Bioenergy Innovation, a US DOE Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science. MRJ-D, TE, BAS, SWS, AM and AMF are funded by JBEI. The DOE Joint BioEnergy Institute (http://www.jbei.org) is supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy. We thank Patrick Planef, Alexander Strimeyer, and Morten Formnitskog for their help in preparing the manuscript. We thank Pablo I. Nikel for providing us with P. putida KT2440. We want Christopher W. Johnson for his assistance in strain construction.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmecc.2020.e00143.

References


Salvachúa, D., Johnson, C.W., Franden, M.A., Johnson, C.W., Karp, E.M., Guarnieri, M.T., Linger, J.G., 2015. Understanding the flux pump is...