Laboratory evolution reveals the metabolic and regulatory basis of ethylene glycol metabolism by *Pseudomonas putida* KT2440

Wing-Jin Li[1], Lahiru N. Jayakody[2], Mary Ann Franden[2], Matthias Wehrmann[3], Tristan Daun[1], Bernhard Hauer[3], Lars M. Blank[1], Gregg T. Beckham[2*], Janosch Klebensberger[3*], Nick Wierckx[1,4*]

[1] Institute of Applied Microbiology-iAMB, Aachen Biology and Biotechnology-ABBt, RWTH Aachen University, Worringerweg 1, 52074 Aachen, Germany
[2] National Bioenergy Center, National Renewable Energy Laboratory, Golden, CO, 80401, USA
[3] University of Stuttgart, Institute of Biochemistry and Technical Biochemistry, Allmandring 31, 70569 Stuttgart, Germany
[4] Institute of Bio- and Geosciences IBG-1: Biotechnology, Forschungszentrum Jülich, 52425 Jülich, Germany

*Address correspondence to:
Nick Wierckx: n.wierckx@fz-juelich.de
Janosch Klebensberger: janosch.klebensberger@itb.uni-stuttgart.de
Gregg T. Beckham: gregg.beckham@nrel.gov

Author copy

Summary

Pollution from ethylene glycol, and plastics containing this monomer, represent a significant environmental problem. The investigation of its microbial metabolism therefore provides insights into the environmental fate of this pollutant, but also enables its utilization as a carbon source for microbial biotechnology. Here, we reveal the genomic and metabolic basis of ethylene glycol metabolism in *Pseudomonas putida* KT2440. Although this strain cannot grow on ethylene glycol as sole carbon source, it can be used to generate growth-enhancing reducing equivalents upon co-feeding with acetate. Mutants that utilize ethylene glycol as sole carbon source were isolated through adaptive laboratory evolution. Genomic analysis of these mutants revealed a central role of the transcriptional regulator GclR, which represses the glyoxylate carboligase pathway as part of a larger metabolic context of purine and allantoin metabolism. Secondary mutations in a transcriptional regulator encoded by PP_2046 and a porin encoded by PP_2662 further improved growth on ethylene glycol in evolved strains, likely by balancing fluxes through the initial oxidations of ethylene glycol to glyoxylate. With this knowledge we reverse engineered an ethylene glycol utilizing strain and thus, revealed the metabolic and regulatory basis that are essential for efficient ethylene glycol metabolism in *P. putida* KT2440.

Originality-Significance Statement

Ethylene glycol is a ubiquitous environmental pollutant in monomer form and in common plastics. This manuscript characterizes the metabolism of ethylene glycol by *Pseudomonas putida* KT2440, both as auxiliary energy source and as the sole carbon source. Adaptive laboratory evolution enabled efficient growth *of P. putida* on ethylene glycol. Genomic analysis of evolved strains followed by reverse engineering uncovered the underlying regulatory mechanisms and provided insights into the non-intuitive metabolic optimization strategies that emerged during the evolution.

Keywords

Pseudomonas putida, ethylene glycol, glyoxylate, adaptive laboratory evolution, genomics, reverse engineering

Introduction

Environmental pollution with plastics is a global problem with far-reaching implications (Geyer et al. 2017; Garcia and Robertson 2017; PlasticsEurope 2018). In 2016, approximately 335 million tonnes of plastic waste was produced, a large proportion of which is mismanaged leading to the dissemination of plastic particles into mainly aquatic ecosystems (Jambeck et al. 2015; Ogunola et al. 2018; Narancic and O'Connor 2017; PlasticsEurope 2018). Despite the fact that these plastics represent non-natural chemicals, several organisms capable of metabolizing these structures have been identified in recent years. These include bacteria in the gut of the wax worm *Plodia interpunctella* capable of degrading polyethylene (PE) (Yang et al. 2014; Bombelli et al. 2017) and the bacterium *Ideonella sakaiensis,* which can depolymerize poly(ethylene terephthalate) (PET) and grow on the resulting terephthalate component (Yoshida et al. 2016). The identification and engineering of plastic-degrading organisms and enzymes provides a compelling opportunity to increase plastic recycling and thereby reduce plastic pollution through the utilization of plastic waste as carbon source for microbial biotechnology (Wierckx et al. 2015; Cho et al. 2015; Narancic and O'Connor 2017; Wei and Zimmermann 2017a, 2017b; Austin et al. 2018).

Pseudomonas putida has been implicated as a promising host for a broad variety of biotechnological approaches including the metabolic use of plastics (Wierckx et al. 2015; Wilkes and Aristilde 2017). It is a model organism in bioremediation and also a work horse in biotechnology and synthetic biology (Nikel et al. 2014; Loh and Cao 2008) due to its metabolic versatility (Nikel et al. 2015), genetic tractability (Belda et al. 2016), and tolerance towards chemical and oxidative stresses (Martínez-García et al. 2014b). A proof of principle for biotechnological conversion of polystyrene, PE, and PET through pyrolysis and subsequent conversion to polyhydroxyalkanoate using different Pseudomonads has been demonstrated (Guzik et al. 2014; Ward et al. 2006; Kenny et al. 2008). In the case of PET, pyrolysis resulted in solid terephthalate, which was used as feedstock for *Pseudomonas putida* GO16 (Kenny et al. 2012).

The microbial metabolism of ethylene glycol, the second component of PET besides terephthalate, is at least equally important. The global production of ethylene glycol was about 20 million tonnes in 2010, with applications in a wide range of polyester resins and fibers (Yue et al. 2012). Ethylene glycol is also used as coolant or antifreeze agent, and as solvent or humectant (Dobson 2000). Its oxidation products glycolaldehyde, glyoxal, glycolate, and in particular glyoxylate, also represent chemicals of industrial relevance, due to their use as reactive building blocks in the production of agro-, aroma-, and polymer chemicals, or pharmaceuticals (Mattioda G. 2000; Sajtos A. 1991; Yue et al. 2012). Glycolaldehyde is also a significant component of aqueous thermochemical wastewater streams and lignocellulosic hydrolysates, which renders those substrates highly toxic, but also a potentially attractive feedstock for microbes if toxicity tolerance improvements can be achieved (Black et al. 2016; Czernik and Bridgwater 2004; Jayakody et al. 2017; Kumar and Gupta 2008; Lu et al. 2009; Vispute et al. 2010; Yu et al. 2008; Jayakody et al. 2018). In nature, glycolate is a significant overflow metabolite of phytoplankton during autotrophic photorespiration, making up 10-50 % of excreted dry organic matter in marine environments (Lau and Armbrust 2006).

Pseudomonas putida possesses the genetic inventory for several pathways potentially enabling ethylene glycol metabolism. Initially, the diol is converted into glyoxylate in a series of oxidation reactions catalyzed by a set of redundant dehydrogenases, with PP_0545, PedI, PedE, and PedH as predominant enzymes (**Figure 1A**) (Wehrmann et al. 2017; Mückschel et al. 2012). Further oxidation to oxalate also occurs in whole-cell biotransformations, but generally only relatively small traces of this dead-end product are observed (Mückschel et al. 2012). The complete conversion of ethylene glycol to glyoxylate yields three reducing equivalents, either in the form of NADH, $PQQH_2$, or in a direct coupling to the electron transport chain. Glyoxylate can be further metabolized by the AceA or GlcB enzymes involved in the glyoxylate shunt (Blank et al. 2008b). Paradoxically, although this shunt is usually a carbon conservation pathway for growth on C2 substrates that enter primary metabolism at the level of acetyl-CoA, the overall stoichiometry of either of the two pathways starting with these enzymes can only yield two molecules of CO_2 and two reducing equivalents (**Figure 1**), making it unsuitable for the utilization of glyoxylate as a sole carbon source. As an alternative to this energy yielding pathway, *P. putida* KT2440 also has the genetic inventory (PP_4297-PP_4301) for a route to metabolize glyoxylate through the glyoxylate carboligase (Gcl) enzyme, which converts two glyoxylate molecules into tartronate semialdehyde and CO_2. The former is converted to glycerate, either directly or via hydroxypyruvate, and subsequently to 2-phosphoglycerate (**Figure 1B**, hence, termed the 'Gcl pathway') (Franden et al. 2018). Theoretically, this pathway could enable the utilization of ethylene glycol as carbon source, however, it is not induced in *P. putida* KT2440 under these conditions (Mückschel et al. 2012).

In this work, we aimed to identify the underlying cause of the existing growth deficiency and to characterize the metabolism of ethylene glycol as an energy-yielding secondary substrate for redox biocatalytic processes. With adaptive laboratory evolution, we isolated strains that use ethylene glycol as a sole source of carbon. The subsequent identification of the corresponding mutations combined with a reverse engineering approach, finally unraveled the metabolic and regulatory systems that underlie efficient ethylene glycol catabolism.

Results and Discussion

Redox equivalent homeostasis in the utilization of ethylene glycol as co-substrate

Although wildtype *P. putida* KT2440 is unable to use ethylene glycol as a sole carbon source, non-growing cells can consume it as studies in detail by Mückschel et al. (2012). The catabolism of ethylene glycol involves three oxidation steps towards glyoxylate, which is subsequently converted to two molecules of CO_2 through reactions of the glyoxylate shunt (**Figure 1**). This overall conversion of ethylene glycol to CO_2 can supply 2.5 reducing equivalents per C-mole of substrate, making it a promising energy-yielding co-substrate. Such co-substrates can be useful in the application of *Pseudomonas* for redox biocatalysis, for example the NADH-dependent epoxidation of styrene to (*S*)-styrene epoxide (Park et al. 2007), or for maintaining energy-demanding solvent-tolerance mechanisms (Blank et al. 2008b). In comparison, typical redox energy co-substrates such as formate can only yield one reducing equivalent per C-mole of substrate (Zobel et al. 2017). Even the complete oxidation of glucose through primary metabolism only yields approximately 1.66 reducing equivalents per C-mole of substrate (Blank et al. 2008a). In addition, co-consumption is likely to occur in an environmental context, making it relevant for the degradation of ethylene glycol as pollutant through cometabolic bioremediation (Hazen 2018).

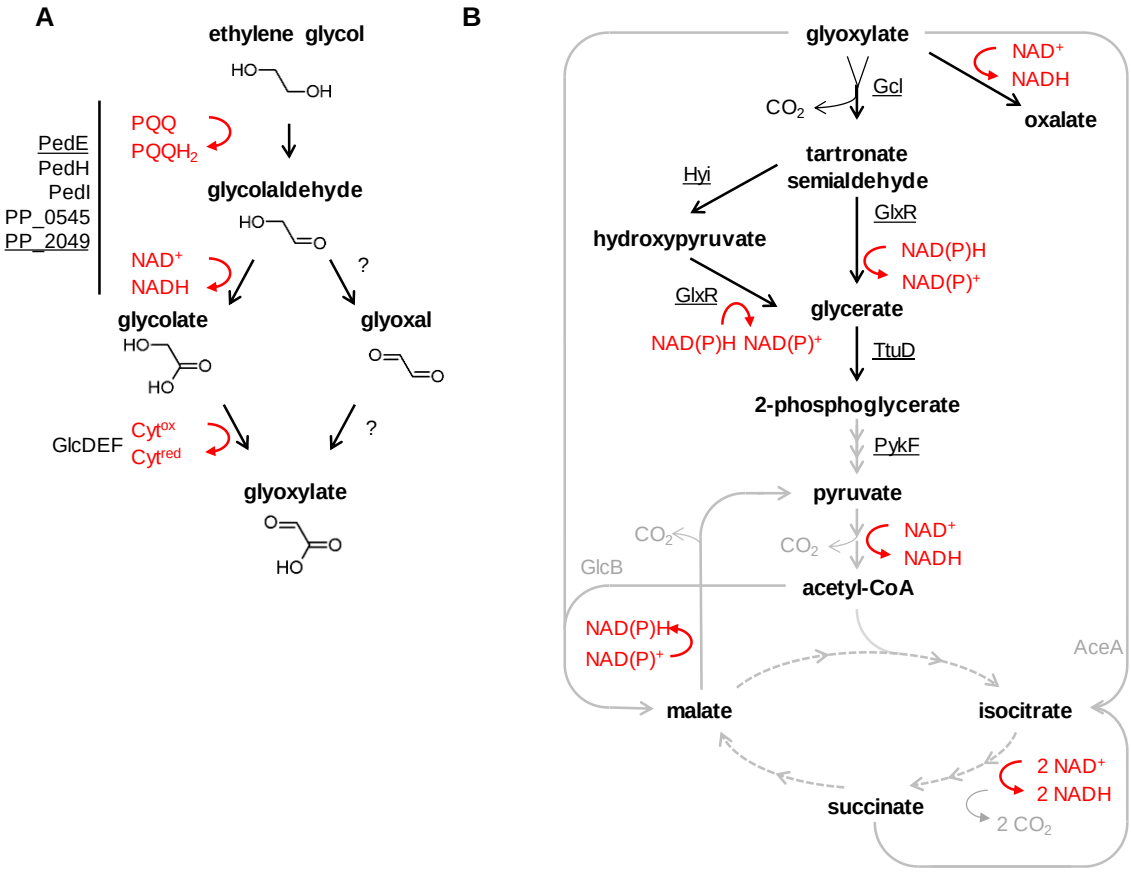


Figure 1 Reaction scheme of the ethylene glycol metabolism in *P. putida* KT2440. **A**: Oxidation steps of ethylene glycol to glyoxylate. **B**: Possible routes of glyoxylate metabolism. The pathway which enables the utilization of glyoxylate as a carbon source is shown in black. Routes which only generate redox equivalents and CO_2 are shown in grey. Redox equivalents are indicated in red. The forked arrow indicates that two glyoxylate are used by the glyoxylate carboligase. Enzymes investigated in are underlined.

To investigate the applicability of ethylene glycol as a co-substrate, *P. putida* KT2440 was cultured in carbon-limited chemostats with acetate as a carbon source, and either ethylene glycol or glyoxylate as the energy source. This setup was chosen as acetate is known to induce enzymes of the glyoxylate shunt (Ahn et al. 2016; Blank et al. 2008a) and because previous experiments showed diauxic utilization of acetate and ethylene glycol in batch cultivations, with acetate being metabolized first (**Figure S1**). Compared to the control with only acetate, a co-feed with ethylene glycol or glyoxylate significantly increased the biomass yield on acetate by 29.6 ± 1.1 % (p = 0.03) or 22.2 ± 8.2 % (p = 0.05), respectively (**Figure 2A**). However, at steady state, ethylene glycol was not completely metabolized to CO_2 and could be detected in the culture medium together with its corresponding oxidation products (**Figure 2B**). The increase in biomass can likely be attributed to the additional reducing equivalents generated through the co-substrate metabolism. Assuming that all glyoxylate is metabolized through the glyoxylate shunt, 27.3 ± 2.6 mmol g_{CDW}^{-1} h^{-1} of reducing equivalents were generated through the co-metabolism of ethylene glycol (**Figure 2C**). In contrast, under similar conditions Zobel et al. (2017) reached the maximal achievable biomass yield on glucose already with an NADH regeneration rate of 7.6 ± 0.9 mmol g_{CDW}^{-1} h^{-1} using a co-feed of formate. Notably, a limitation in the initial oxidation reactions can be excluded as the primary cause for this observation, since glyoxylate was also not completely metabolized under these conditions, yielding only 14.4 ± 0.5 mmol g_{CDW}^{-1} h^{-1} of reducing equivalents while enabling almost the same biomass yield increase as the ethylene glycol co-feed. In all, this experiment demonstrates the potential of ethylene glycol as an energy-yielding cometabolite, but also indicates limitations regarding the full consumption which should be addressed in further studies.

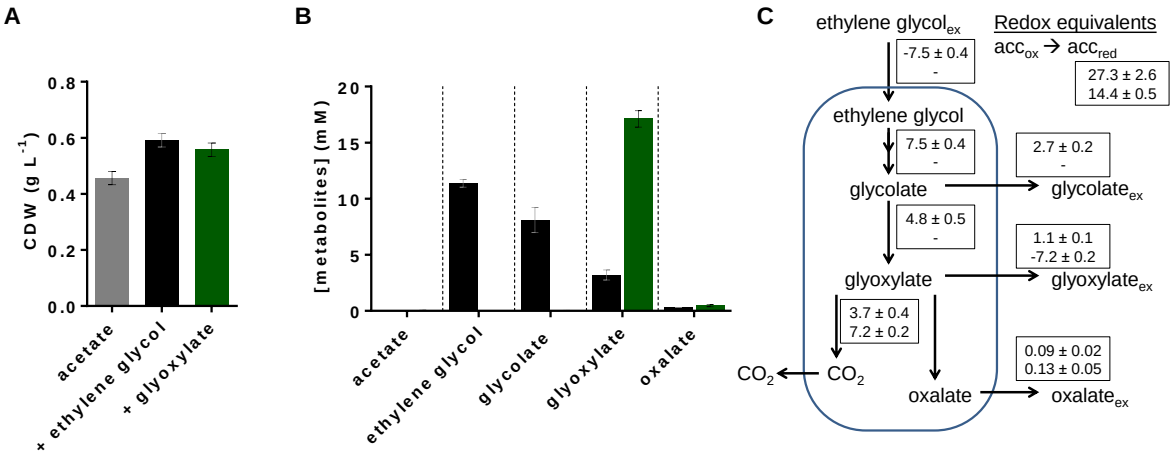


Figure 2 Co-feeding of *P. putida* KT2440 in C-limited chemostats on MSM with 30 mM acetate supplemented with 30 mM ethylene glycol (black), 30 mM glyoxylate (green), or no co-feed (grey) at a dilution rate of 0.2. **A**: Comparison of biomass at steady state. **B**: Extracellular metabolites at steady state. For the acetate control no metabolites were detected. **C**: *In vivo* flux during growth on ethylene glycol (upper value) or glyoxylate (lower value) in mmol g_{CDW}^{-1} h^{-1} obtained from the substrate utilization and metabolites secretion rates on the respective substrate with a growth rate of 0.2 h^{-1}. The flux analysis allowed the estimation of redox cofactor regeneration rates according to the stoichiometry of Figure 1. (acc = electron acceptor oxidized or reduced) Error bars indicate the deviation from the mean (n = 2).

Isolation of mutants able to utilize ethylene glycol as sole carbon source by adaptive laboratory evolution

Adaptive laboratory evolution (ALE) is a common method to adapt strains to specific environments (Dragosits and Mattanovich 2013). Given the inefficient utilization of ethylene glycol as co-substrate with acetate, and the fact that *P. putida* KT2440 possess the genetic inventory allowing growth on ethylene glycol as C-source, we speculated that ALE might select for ethylene glycol utilizing mutants. Therefore, we performed two independent ALE experiments, in different laboratories in Stuttgart and Aachen, each using their own laboratory wildtype strain of *P. putida* KT2440 and different MSM (mineral salt medium) supplemented with ethylene glycol as sole carbon source (see Experimental procedures for details). Adaptive mutants reproducibly emerged, leading to visible growth after a lag phase of 4-6 days (**Figure S2**). In Stuttgart, clones from three independently evolved cultures were isolated after initial growth and subcultured three times on LB-agar plates to obtain strains E1.1, E1.2 and E1.3. In Aachen, a series of three parallel ALE cultivations was performed, where batches were sequentially re-inoculated into fresh MSM with ethylene glycol after growth became apparent by visual inspection (OD_{600} > 0.5) (**Figure 3A**). After six serial transfers into fresh growth medium, 36 individual strains were isolated on LB-agar plates and assessed in liquid cultures in a Growth Profiler to select the best growing strains, finally obtaining strains E6.1 and E6.2 from two parallel ALE lines.

All five resulting strains showed a stable phenotype of growth on ethylene glycol. No major differences in growth and substrate oxidation and uptake could be observed within the E1 and E6 groups. Nevertheless, when comparing all strains collectively, E1 and E6 behaved differently when growing on higher ethylene glycol concentrations. With 120 mM ethylene glycol as sole carbon source, E6 strains showed 1.4 times faster growth than the E1 strains (E1: 0.083 ± 0.004 h^{-1}; E6: 0.118 ± 0.004 h^{-1}; p= 0.005) and reached a higher final biomass concentration (**Figure 3B, Figure S3**). When grown in MSM with 26.7 ± 0.4 mM ethylene glycol as the sole carbon source, all strains grew at approximately the same initial rate (0.19 ± 0.02 h^{-1}). However, the maximum biomass concentration (CDW = cell dry weight) of the E6 cultures (0.63 ± 0.02 g_{cdw} L^{-1}) was significantly higher (p = 0.011) than that of the E1 cultures (0.49 ± 0.07 g_{cdw} L^{-1}) (**Figure 3C, Figure S4**). The difference between E1 and E6 groups was also

reflected in the metabolism of ethylene glycol and the formation of intermediate oxidation products. Illustrated in **Figure 3D**, the ethylene glycol degradation rate of E6.1 was 1.8-fold higher than that of E1.1 and E6.1 transiently produced up to 7.9-fold more glycolate than E1.1 (p = 0.011). In general, these experiments show the strength of ALE for the selection of new and/or improved phenotypes. This approach may also be extended in order to select strains with even more efficient utilization of ethylene glycol, i.e. with higher growth rates, lower accumulation of intermediates, or better growth in high substrate concentrations. However, we chose a relatively short evolutionary selection to facilitate the downstream genomic identification of causal mutations, as long ALE experiments can easily lead to the accumulation of additional background mutations.

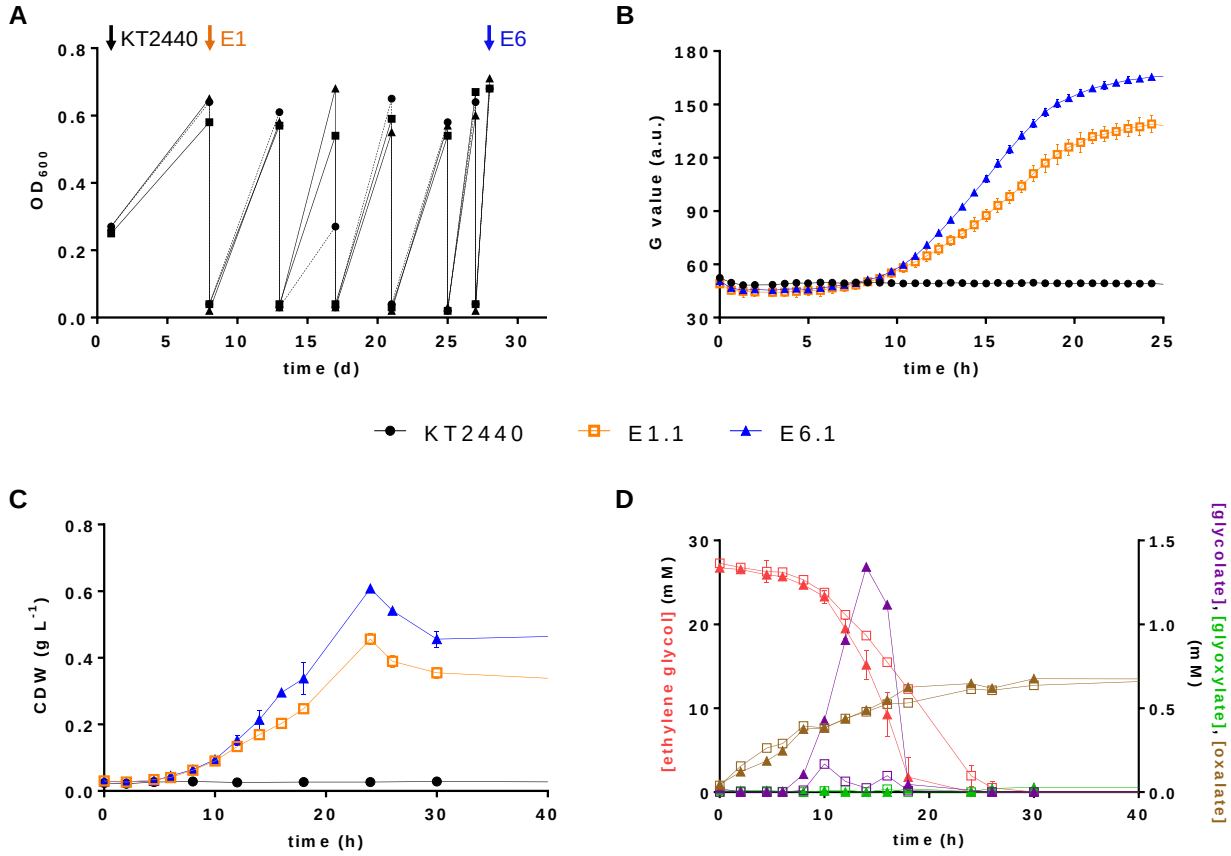


Figure 3 Adaptive laboratory evolution of *P. putida* KT2440 on ethylene glycol and characterization of adapted strains. **A:** Sequential batch cultivation on MSM with 15 mM ethylene glycol. Arrows indicate the time points where strains were isolated. **B:** Growth comparison of *P. putida* KT2440 and the adapted strains E1.1 and E6.1 in MSM containing 120 mM ethylene glycol. Growth was detected via a Growth Profiler in 96-square-well plates. **C** and **D:** Biomass growth of the isolated ALE strains E1.1 (empty square) and E6.1 (filled triangles) in comparison with KT2440, and extracellular metabolic products (ethylene glycol in red, glycolate in purple, glyoxylate in green, oxalate in brown) of the isolated ALE strains E1.1 and E6.1 growing on 26.7 ± 0.4 mM ethylene glycol in a shake flask cultivation on MSM. Error bars indicate the deviation from the mean (n = 2).

Genomic and metabolic context of adaptive mutations

To identify mutations underlying the stable phenotypic switch in the ALE strains, whole genome resequencing was performed on the E1 and E6 strains, as well as the two respective wildtypes from the different laboratories. With this approach, we uncovered 80-83 Single Nucleotide Polymorphisms (SNP) and Insertion-Deletion polymorphisms (InDel) (**Table S1**, **Supplemental data 1**), compared to the updated genome sequence of *P. putida* KT2440 (AE015451.2, (Belda et al. 2016)). However, the vast majority of these differences were also present in the two parental strains, and consist mainly of

mutations in non-coding regions, silent mutations, or errors due to low coverage and read quality. Notably, the two parental strains were very similar to each other, even though they share no known immediate common history. After subtracting the parental, silent, and intergenic mutations, three mutated regions remained. One was mutated in all evolved strains, while the other two were only mutated in the E6 strains (**Figure 4**).

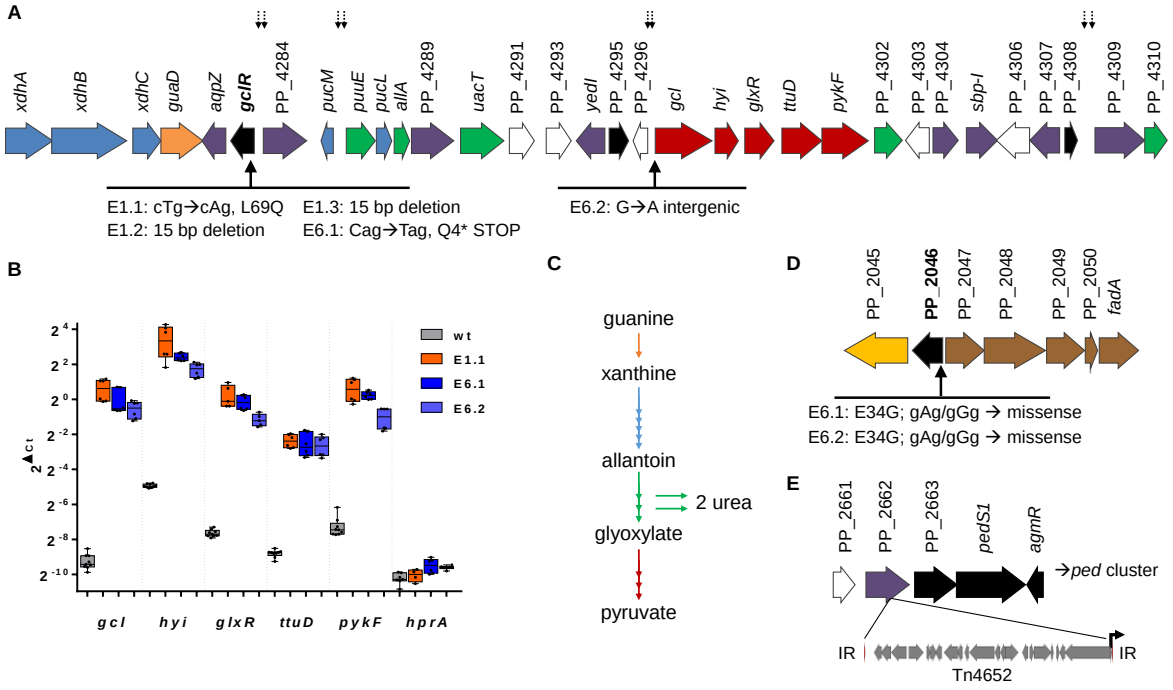


Figure 4 Schematic representation of genomic regions mutated after adaptive laboratory evolution on ethylene glycol. **A:** Gene organization from *xdhA* to PP_4310 (coordinates: 4866804-4902814, 36010 bp in total). Small dotted arrows indicate putative *gclR* binding sites. **B:** Box-and-whisker plot of relative expression levels of genes implicated in ethylene glycol metabolism in cells of *P. putida* KT2440 wildtype (wt) and cells of evolved mutants growing on 20 mM ethylene glycol and 40 mM acetate determined by qRT-PCR. The $2^{\Delta Ct}$ values were normalized to *rpoD*. Individual data points are plotted onto the graph, whiskers indicate minimum to maximum values. **C:** Simplified pathway scheme from guanine to pyruvate. **D:** Gene organization from PP_2045 to *fadA* (coordinates: 2325342-2334253, 8911 bp in total). PP_2045, coding for a metallo-beta-lactamase family protein is shown in a yellow arrow. Brown arrows show genes related to beta-oxidation. The mutated regulator PP_2046 is indicated in black. **E:** Gene organization and direction of transcription from PP_2667 to *agmR* (coordinates: 3047946-3053910, 5964 bp in total). The transposon Tn4652, flanked by inverted repeats (IR in red), is shown in grey arrows, with an angled black arrow indicating the direction of a putative promotor. The annotated sites of SNP and InDels found in the adapted strains are indicated with capital letters below the relevant genes. Colored arrows indicate putative gene functions related to the conversion of guanine to xanthine (orange), from xanthine to allantoin (blue), from allantoin to glyoxylate (green) and from glyoxylate to pyruvate (red). Purple arrows indicate genes coding for transporters or porins. Genes coding for regulators are shown in black. White arrows show genes which are predicted as hypothetical proteins.

In the first region, 4 out of 5 strains contained mutations in the gene with locus tag PP_4283, encoding a putative GntR-type transcriptional regulator with DNA binding function, hence called *gclR* (Donald et al. 2001). These mutations include a missense mutation in E1.1, two identical 15 bp deletions in E1.2 and E1.3 and one nonsense mutation in E6.1 giving rise to a stop codon in the 4th triplet (**Figure 4A**). All of these mutations are located in the first third of the gene. The latter mutation makes it likely that the gene function is disrupted. Strain E6.2 did not contain a mutation in *gclR*, instead bearing a SNP

approximately 12.5 kb downstream of this locus, in the promotor region of the *gcl* gene, which is the first gene in the PP_4297-PP_4301 cluster that encodes the enzymes of the Gcl pathway (Franden et al. 2018). According to RegPrecise (Novichkov et al. 2013), GclR is predicted to be a regulator of xanthine metabolism with, among others, two predicted binding sites upstream of *gcl* (**Figure 4A**). The mutation in strain E6.2 may disrupt one of these binding sites, while also leading to the emergence of a putative promoter (**Supporting Information 1-3** and **Table S2**)**.** Quantitative reverse transcription polymerase chain reaction (qRT-PCR) analysis of *P. putida* KT2440 and evolved strains cultured on 20 mM ethylene glycol and 40 mM acetate indicated that transcript levels of all five genes in the *gcl* cluster were very low in cells of the wildtype, while being significantly upregulated in cells of strains E1.1, E6.1 and E6.2 (**Figure 4B**, 71-fold to 842-fold upregulated compared to wildtype cells, $p < 0.01$). A similar effect can be assumed for strains E1.2 and E1.3. In contrast, the distantly located *hprA* gene (PP_0762), which encodes a second possible glyoxylate/hydroxypyruvate reductase (Cartwright and Hullin 1966), was not expressed under these conditions, independent of the strain used. From these data, we conclude that GclR is likely a repressor of the PP_4297-PP_4301 gene cluster. The disruption of *gclR*, or the disruption of the putative GclR binding site in the case of strain E6.2, alleviates this repression, enabling growth on ethylene glycol.

These results raise the question: why are ethylene glycol or glyoxylate not the effectors that bind GclR to relieve repression of the downstream metabolic genes? Clues to the answer can be found in the genomic context of the *gclR* gene in *P. putida* KT2440, which, similarly to other organisms such as *Escherichia coli*, *Streptomyces coelicolor*, and *Bacillus subtilis* (Hasegawa et al. 2008; Cusa et al. 1999; Rintoul et al. 2002; Navone et al. 2015; Schultz et al. 2001), encodes multiple genes known or predicted to be involved in the metabolism of purines via the intermediates allantoin and glyoxylate (**Figure 4AC**). In aerobically growing cells of *E. coli*, the genes encoding the Gcl pathway are repressed in the presence of allantoin through the action of the AllR regulator. This repression is alleviated by glyoxylate, which concomitantly induces an alternative allantoin metabolic pathway, ultimately yielding NH_3, CO_2, and ATP, predominantly active under anaerobic conditions (Hasegawa et al. 2008; Cusa et al. 1999; Rintoul et al. 2002). This alternative pathway seems to be absent in *P. putida* KT2440, befitting its obligate aerobic lifestyle. Although the situation in *P. putida* KT2440 is different from that in *E. coli*, we hypothesize that the genomic context of *gclR* suggests that the failure of *P. putida* KT2440 to activate the Gcl pathway in the presence of ethylene glycol or glyoxylate may be because it is part of a larger metabolic context, governed by effectors that lie upstream of their metabolism. Indeed, both the parental strain *P. putida* KT2440 and strain E6.1 are able to grow on allantoin as a sole carbon and nitrogen source (**Figure 5**), indicating that allantoin, and not glyoxylate, serves as inducer of the genes encoding the Gcl pathway. A co-feed of allantoin and ethylene glycol resulted in a higher biomass concentration compared to allantoin alone, likely through the activation of the Gcl pathway, since this is the only pathway known to enable growth on these substrates as sole C-source. Similar results were obtained with xanthine, which also enabled growth on ethylene glycol (**Figure S5**). Notably, E6.2 was not able to utilize allantoin as sole nitrogen source, although it retained the ability to utilize it as a carbon source, either with or without the addition of ethylene glycol. This indicates that a regulatory cross-talk between allantoin and ethylene glycol metabolism exists. One could speculate that this cross-talk involves the PP_4296 gene (Göhler et al. 2011) whose expression might be affected by the mutation in E6.2.

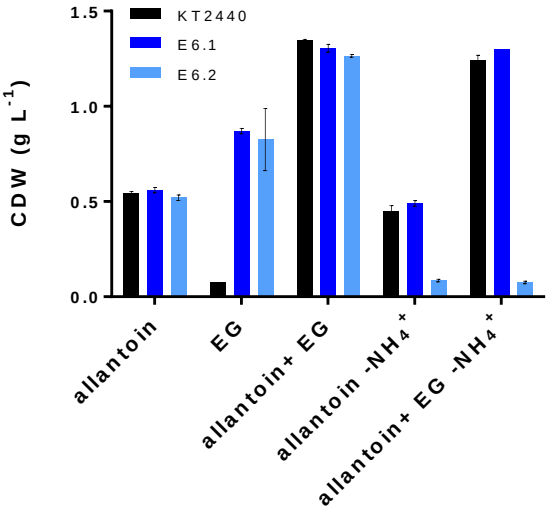


Figure 5 Biomass concentration of *P. putida* KT2440 and ALE strains E6.1 and E6.2 after 25 h in MSM containing 20 mM allantoin and/or 20 mM ethylene glycol (EG). The label '-NH_4^+' indicates that ammonium was omitted from the medium, leaving allantoin as a sole nitrogen source. Error bars indicate the deviation from the mean (n = 2).

In addition to the above mentioned mutations involving GclR, two additional mutations were found in the E6 strains (**Figure 4DE**). Both strains, E6.1 and E6.2, contain the same missense mutation (E34G) in the gene with locus tag PP_2046, encoding a LysR-type transcriptional regulator. This regulator likely controls the adjacent operon, which encodes enzymes involved in a beta-oxidation pathway including a 3-hydroxyacyl-CoA dehydrogenase (PP_2047), an acyl-CoA dehydrogenase (PP_2048), an alcohol dehydrogenase (PP_2049), a hypothetical protein (PP_2050) and a acetyl-CoA acetyltransferase (PP_2051- *fadA*) (Liu et al. 2011) (**Figure 4D**). Indeed, qRT-PCR shows a significantly higher expression of PP_2049 in E6.1 compared to E1.1 (**Figure 6**), indicating that the missense mutation in PP_2046 leads to an upregulation of this operon. Although the C2 compound ethylene glycol cannot be metabolized through beta-oxidation, the dehydrogenases may still affect its oxidation.

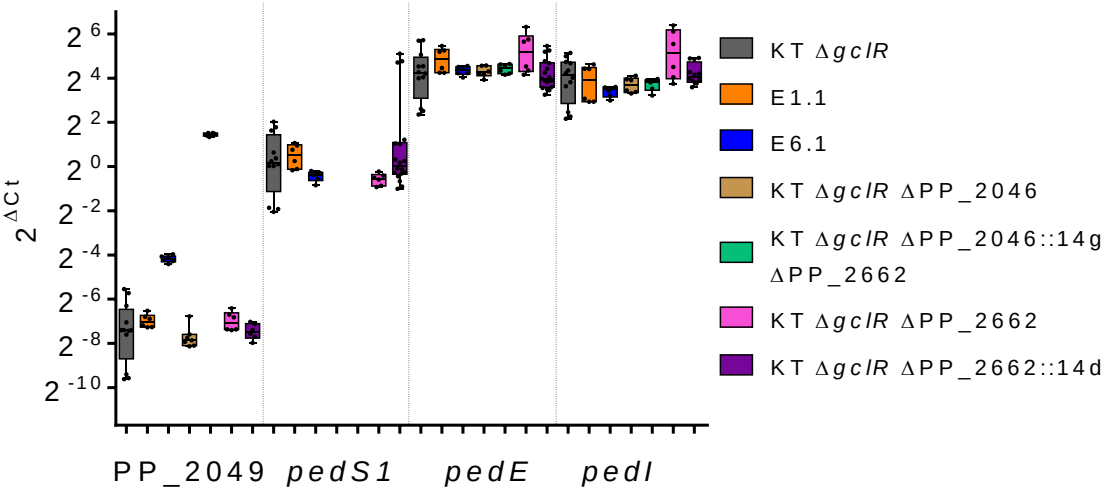


Figure 6 Box-and-whisker plot of relative expression levels of genes implicated in ethylene glycol metabolism in cells of *P. putida* KT2440 with several knockout background and cells of evolved mutants growing on 20 mM ethylene glycol and 40 mM acetate determined by qRT-PCR. The $2^{\Delta Ct}$ values were normalized to *rpoD*. Individual data points are plotted onto the graph, whiskers indicate minimum to maximum values.

Both E6 strains further contain an insertion in PP_2662, which was identified as a 17 kb Tn4652 transposon by read coverage analysis (**Supporting information 4**) and arbitrary-primed PCR (Martínez-García et al. 2014a). The PP_2662 gene encodes a putative porin and is located upstream of an operon encoding a putative two-component sensory system consisting of a hypothetical protein with a sensory domain (PP_2663) and a sensor histidine kinase PP_2664 (from now on referred to as PedS1 according to the gene nomenclature from Arias et al. (2008) (**Figure 4E**). These genes are co-regulated with a cluster of genes including the aldehyde and alcohol dehydrogenases *pedI*, *pedE,* and *pedH* in *P. putida* KT2440 in the presence of *n*-butanol (Vallon et al. 2015) and ethylene glycol (Mückschel et al. 2012). In *P. aeruginosa* the expression of a similar gene cluster is involved in the utilization of ethanol and is regulated by a complex hierarchy of sensory histidine kinases and transcriptional regulators, including *erbR* (previously *agmR*; PP_2665) and ErcS`, a homolog of PedS1 (Mern et al. 2010; Hempel et al. 2013; Promden et al. 2009). The Tn4652 transposon (PP_2964 – PP_5546) is known to be activated under stress conditions such as starvation (Ilves et al. 2001). It contains a predicted promoter at its 3'-end, facing the two-component regulator operon (**Supporting information 5-6**), and also is known for generating novel fusion promoters upon insertion into a new locus (Nurk et al. 1993). Thus, this transposon insertion into PP_2662 could affect ethylene glycol metabolism by disruption of the porin, and/or by upregulation of the *ped* cluster through overexpression of the sensor histidine kinase PedS1. The latter hypothesis is disproven by qRT-PCR analysis, which shows no significant difference between the expression of *pedS1*, *pedE*, and *pedI* in E1.1 and E6.1 (**Figure 6**). This indicates that it is mainly the disruption of the PP_2662 porin that led to the enhanced growth on ethylene glycol in E6.1. The exact role of this porin is unclear, but its disruption likely affects the exchange of ethylene glycol or its oxidation products across the outer membrane. The fact that E6.1 transiently accumulates much more glycolate than E1.1 (**Figure 3D**) suggests that the rate of ethylene glycol uptake or glycolate export from the periplasm is not affected, thus pointing to a possible transport effect on the aldehyde intermediates. That said, *pedE* and *pedI* are very highly expressed in both strains, and it is known that they are involved in the oxidation of a variety of alcohols and aldehydes including ethylene glycol (Mückschel et al. 2012; Arias et al. 2008). The essentiality of the Ped cluster was confirmed by the deletion of the *pedE-pedI* cluster (PP_2673-PP_2780) in the E6 strains, which eliminated their ability to grow on ethylene glycol (**Figure S6**).

Reverse engineering of ethylene glycol metabolism

Genomic analysis implicates mutations in the *gclR* gene as foundational to the growth of *P. putida* KT2440 on ethylene glycol as a sole carbon source, while the secondary mutations in PP_2046 and PP_2662 likely affect the rates of the initial oxidation reactions. To determine how these mutations assert their effect, we sought to replicate their phenotype in a reverse engineering approach. To this end, *gclR* was deleted in the wildtype of *P. putida* KT2440. The resulting Δ*gclR* strain grew readily on MSM with ethylene glycol as the sole carbon source (**Figure 7A**), confirming that the *gclR* mutations in the ALE strains were disruptive in nature, and supporting the hypothesis that GclR is a repressor of the genes encoding the Gcl pathway.

In order to assess the effect of the secondary mutations, a set of combinatorial strains was constructed, in which PP_2046 and PP_2662 were deleted in the Δ*gclR* background (**Table 1**). In addition, both genes were replaced by synthetic constitutive promoters. In the case of PP_2662, the 14d promoter of average strength (Zobel et al. 2015) was inserted facing PP_2663-*pedS1* to test the effect of overexpression of these downstream genes encoding a two-component regulator system. This was done to test the effect of the presence of an upstream promoter similar to the promoter-carrying transposon insertion in the E6 strains. In the case of PP_2046, the strong 14g promoter was inserted facing the downstream beta-oxidation operon which is upregulated in E6.1. Subsequently, the transcript levels of relevant genes were determined by qRT-PCR (**Figure 6**) and growth of all strains was compared in MSM with 120 mM EG (**Figure 7**).

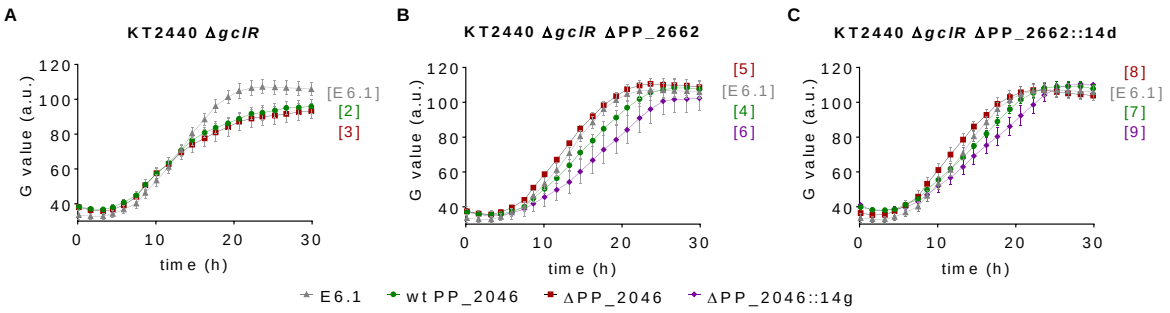


Figure 7 Growth comparison between E6.1 and reverse engineered *P. putida* KT2440 strains in MSM containing 120 mM ethylene glycol. Different colors indicate strains harboring the native PP_2046 (green, circle), the knockout of PP_2046 (red, squares), and the substitution of PP_2046 with the promotor 14g (purple, diamonds) in the background of KT2440 Δ*gclR* (**A**), KT2440 Δ*gclR* ΔPP_2662 (**B**) and KT2440 Δ*gclR* ΔPP_2662::14d (**C**). As a positive control, and for visual reference, growth of the evolved strain E6.1 is represented by grey triangles in each graph. Strain numbers next to the graphs refer to full strain names listed in **Table 1**. Growth was detected via a Growth Profiler in 24-square-well plates. Error bars depict the standard error of the mean (n ≥ 3).

Disruption of *gclR* indeed is essential for growth on EG, as without this mutation no growth was observed. Deletion of only PP_2046 in the Δ*gclR* background did not have any observable effect on the growth, nor did it change the transcript level of PP_2049. This indicates that PP_2046 is not a transcriptional repressor and suggests that the point mutations in the E6 strains were not disruptive in nature, rather causing constitutive activation. The disruption of PP_2662 in *P. putida* KT2440 Δ*gclR*, either with or without insertion of the 14d promoter, led to an improvement of the final biomass concentration compared to the Δ*gclR* progenitor, at the expense of the growth rate. Although the qRT-PCR analysis confirmed that the 14d promoter insertion increased the transcript level of the downstream *pedS1*, this overexpression didn't significantly affect transcript levels of *pedE* or *pedI*, nor did the promoter insertion affect the growth of the reverse engineered strain. This further confirms that the enhanced growth of the E6 strains was not caused by an altered expression of the *ped* genes. Upon combining PP_2046 and PP_2662 mutations in different genetic backgrounds, two trends became apparent. Strains with ΔPP_2046*::*14g grew worse than strains with the wildtype PP_2046 locus, which in turn were out-performed by strains harboring a ΔPP_2046 deletion without an additional promoter insertion. This is unexpected considering that in E6.1 the expression of PP_2049 was increased. The qRT-PCR analysis indicates that perhaps the heterologous overexpression with the 14g promoter was too strong, with transcript levels that were several orders of magnitude above that in strain E6.1 (**Figure 6**).

Of the reverse-engineered strains, *P. putida* KT2440 Δ*gclR* ΔPP_2046 ΔPP_2662 showed the fastest growth and highest final biomass (**Figure 7C**). Although the difference between reverse engineered strains with different modifications in the PP_2046 and PP_2662 loci is generally small, they mostly out-perform strain *P. putida* KT2440 Δ*gclR*, indicating the added benefit of these secondary mutations. Although the molecular basis underlying the genetic events that lead to the superior ethylene glycol metabolism of strain E6.1 and the engineered strains is still not completely understood, the results indicate a complex interplay between the two secondary mutations, likely affecting the balance between substrate transport across the outer membrane and alcohol and aldehyde oxidation. This hypothesis is in line with the results from a companion study, were we undertook an *ab initio* metabolic engineering approach to obtain a highly efficient ethylene glycol-utilizing strain of *P. putida* KT2440 (Franden et al. 2018). In that study, we could demonstrate that glycolaldehyde and glyoxal are indeed toxic to cells of *P. putida* KT2440 in concentrations above 4 mM, and that preventing the accumulation of these intermediates during ethylene glycol metabolism is crucial for efficient growth.

In all, we generated strains, using ALE and reverse engineering, that efficiently grow on ethylene glycol. These strains can be applied for the biotechnological conversion of pretreated waste streams that contain ethylene glycol or its derivatives. In this setting, strain stability is a key performance indicator. Indeed, the engineered strain KT2440 Δ*gclR* Δ PP_2046 Δ PP_2662 and the evolved strains E1.1 and E6.1 retained their ethylene glycol-metabolizing phenotype after more than 110 generations without selective pressure (**Figure S7**). This indicates that the implemented mutations pose no significant negative selection pressure under the conditions tested.

Conclusions

The metabolism of ethylene glycol and its derivatives plays a pivotal role in the biotechnological utilization of plastic waste and lignocellulose-derived streams, and its oxidation products glycolate and glyoxylate have a variety of value-added applications. The quantitative physiological characterization of ethylene glycol co-metabolism by *P. putida* KT2440 provides valuable insights for the production of value-added chemicals and identifies opportunities and bottlenecks for the use of ethylene glycol as a redox energy yielding co-substrate. ALE enabled *P. putida* KT2440 to utilize ethylene glycol as a sole carbon source. The characterization of evolved strains provided insights into the genetic and regulatory basis of their adapted metabolism. Based on those insights, the growth characteristics of evolved strains could be successfully reverse engineered in *P. putida* KT2440. Extensive transcriptional and biochemical analysis will be necessary to fully determine the combinatorial effect of the enzymes affected by these mutations. This manuscript along with the companion study (Franden et al. 2018), provides a foundation for such analyses, and enables the further development of *P. putida* as an applied synthetic biology workhorse in the aforementioned fields of application.

Experimental procedures

Strains and cultivation conditions

The chemicals used in this work were obtained from Carl Roth (Karlsruhe, Germany), Sigma-Aldrich (St. Louis, MO, USA), or Merck (Darmstadt, Germany) unless stated otherwise. Glycerol was kindly provided by Bioeton (Kyritz, Germany).

All bacterial strains used for this are listed in **Table 1**. For quantitative microbiology experiments, *P. putida* KT2440 strains were cultivated in mineral salt medium (MSM) (Hartmans S. et al. 1989) unless stated otherwise. Pre-cultures contained 20 mM glucose. For the cultivations with allantoin, 20 mM allantoin was dissolved in MSM and sterilized using filtration. *P. putida* transformants harboring pEMG derivatives were selected on LB agar plates with 50 mg L^{-1} kanamycin, for transformants harboring pSW2, 25 mg·L^{-1} gentamycin was added. Liquid cultivations were incubated at 30 °C, 200 rpm shaking speed with an amplitude of 50 mm in a Multitron shaker (INFORS, Bottmingen, Switzerland) using 500 mL non-baffled Erlenmeyer flasks with metal caps, containing 50 mL culture volume.

For adaptive laboratory evolution experiments in Aachen, a pre-culture of *P. putida* KT2440, cultivated in MSM with 20 mM glucose was used to inoculate 250 mL clear glass Boston bottles with Mininert valves (Thermo Fisher Scientific, Waltham, MA, USA) containing 10 mL MSM with 15 mM ethylene glycol (final OD_{600} of 0.01). Serial transfers were reinoculated six times after the cultures reached an OD_{600} of at least 0.5, with a starting OD_{600} of 0.1. After growth was detected (usually overnight), single colonies were isolated from ALE cultures by streaking samples on lysogeny broth containing 5 g L^{-1} NaCl with 1.5 % (w/v) agar-agar (LB agar plate). Two (E6.1 and E6.2) out of 36 colonies were chosen according to their growing behavior in MSM with 20 mM ethylene glycol determined using the Growth Profiler (described below).

For adaptive laboratory evolution experiments in Stuttgart, *P. putida* KT2440 was inoculated (final OD_{600} of 0.02) from an LB agar plate into three 250 mL Erlemeyer flasks without baffles containing 50 mL of a modified M9 salt medium (Wehrmann et al. 2017) with 20 mM ethylene glycol as carbon source. Two of the cultures were additionally supplemented with either 1 µg L^{-1} thiamin (E1.2) or 1 µg L^{-1} biotin (E1.3). Growth was monitored by OD_{600} measurements using a spectrophotometer. After

growth was detected, single colonies were isolated on LB agar plates. After three rounds of re-streaking of a single colony from each condition on fresh LB agar plates, a single colony was grown in 5 mL liquid LB medium (30°C, 180 rpm) overnight to prepare stock culture of strain E1.1, E1.2 and E1.3.

Chemostat experiments were carried out in DasBox mini reactors (Eppendorf, Hamburg, Germany) with a working volume of 100 mL and a dilution rate of 0.2 at 30 °C. The pH was kept at 7.0 by the automatic addition of NaOH and HCl, and partial oxygen pressure was kept at 30 % air saturation by automatic adjustment of the stirrer speed between 400 and 800 rpm. The three conditions (MSM with 30 mM sodium acetate, 30 mM sodium acetate and 30 mM ethylene glycol, and 30 mM sodium acetate and 30 mM glyoxylate) were tested subsequently in one experiment, waiting at least five volume changes to achieve a new steady state.

For online analysis of growth without offline sample analysis, a Growth Profiler 960 (Enzyscreen, Heemstede, The Netherlands) was used. This device analyzes cultures in microtiter plates with transparent bottoms by image analysis, and the resulting G-value correlates with cell density (REF Duetz). Pre cultures containing 4 mL MSM with 20 mM glucose in 24-well System duetz plates (Enzyscreen, Heemstede, The Netherlands) were cultivated in a Multitron shaker (INFORS, Bottmingen, Switzerland) with a 300 rpm shaking speed with an amplitude of 50 mm. Main cultures, either in a 96-well plates with 200 µL or in 24-well plates with 4 mL culture volume, using MSM with several concentrations of different carbon sources as indicated, were incubated at 30 °C, 225 rpm shaking speed with an amplitude of 50 mm in the Growth Profiler. In case of figure 7, 19-datasets may differ up to 20 min due to the fact that the data was generated in separate cultures on different days.

DNA methods

The construction of plasmids was performed either by standard restriction-ligation or Gibson assembly (Gibson et al. 2009) using the NEBuilder HiFi DNA Assembly (New England Biolabs, Ipswich, MA, USA). DNA modifying enzymes were purchased from New England Biolabs. Primers were purchased as unmodified DNA oligonucleotides from Eurofins Genomics (Ebersberg, Germany) and are listed in supporting information (**Table S3**). Clonal DNA sequences were amplified using the Q5 High-Fidelity Polymerase. Arbritrary-primed PCR was performed as describe by (Martínez-García et al. 2014a). For the transformation of DNA assemblies and purified plasmids into competent *E. coli* a heat shock protocol was performed (Hanahan 1983). For *P. putida* transformations either conjugational transfer or electroporation were performed as described by Wynands et al. (2018). Knockout strains were obtained using the pEMG system described by (Martínez-García and Lorenzo 2011) with a modified protocol described by Wynands et al. (2018). Plasmid inserts and gene deletions were confirmed by Sanger sequencing performed by Eurofins Genomics (Ebersberg, Germany).

To prepare *P. putida* cultures for RNA extraction, cells were grown overnight in M9 minimal medium containing 20 mM glucose in baffled shake flasks at 30°C, 225 rpm. Cells were then diluted and used to inoculate fresh cultures containing 20 mM ethylene glycol and 40 mM sodium acetate to an initial OD_{600} of 0.1. After incubation at 30°C with shaking at 225 rpm to mid-exponential growth phase (OD_{600} 0.8-1), 2x volume of Qiagen RNAprotect Bacteria Reagent was added to the cultures and allowed to mix for 5 minutes. Subsequently, cells were harvested by centrifugation at 5,000 x g for 15 min at 4°C. Supernatant was removed and cells were frozen and stored at -80°C until further analysis. Supernatants of cultures prior to addition of RNAprotect reagent was analyzed for acetate and ethylene glycol by HPLC that showed that substrate was still available. RNA was extracted from cells using Qiagen's RNeasy mini kit following manufacturer's instructions including a DNAse (Qiagen RNase-Free DNase) in column digestion for one hour at room temperature following the manufacturer's instructions. After one round of RNA isolation, a DNase digestion was performed (TURBO DNase; Ambion, Austin, TX, USA). After two hours incubation at 37°C, the DNase was removed from the RNA sample with an additional purification step using the Qiagen's RNeasy mini kit. cDNA was prepared from the purified RNA using an iScript Reverse Transcription supermix kit for RT-qPCR (Bio-Rad). The expression levels of seven genes were analyzed using primers designed by the RealTime qPCR tool for RT-qPCR (http://www.idtdna.com/scitools/Applications/RealTimePCR/) and listed in **Table S3**.

Quantitative RT-PCR was performed using iQ SYBR® Green Supermix (Bio-Rad) on a Bio-Rad CFX96 Touch™ Real-Time PCR Detection System (Bio-Rad Lab, Hercules, CA, USA). The reaction conditions were 10 min at 95°C, 39 × (15 s at 95°C, 45 s at 55°C, followed by a melting curve analysis: 1 min at 95°C, 81 × (30 s starting at 55°C, increasing 0.5°C per cycle, ending at 95°C). Experiments were performed in triplicate with biological duplicates. The gene expression levels were assessed by comparing the Ct value of the house keeping gene *rpo*D (Wang and Nomura, 2010) to the Ct value of the target gene using the following formula:

$$\text{Gene expression level} = 2^{\text{Ct}(rpoD)-\text{Ct}(\text{target})}$$

Ct values represent the first cycle at which the instrument can distinguish the fluorescence of nucleic acid amplification generated as being above the background signal. Final expression levels were averaged for each target gene and normalized to the expression level of the control (*P. putida* KT2440) strain.

Genomic DNA for resequencing was isolated through a High Pure PCR Template Preparation Kit (ROCHE life science, Basel, Switzerland). Sequencing and SNP/InDel calling was done by GATC (Konstanz, Germany) using Illumina technology as paired end reads of 125 base pairs. To map the reference sequence against the database, BWA with default parameters was used (Li and Durbin 2009). SNPs and InDels, analysed by GATK´s UnifiedGenotyper (DePristo et al. 2011; McKenna et al. 2010), were listed and visualized with the Integrative Genomics Viewer (IGV) (Thorvaldsdóttir et al. 2013). The sequences have been deposited in the Sequence Read Archive (SRA) with the accession number SRP148839**.**

Analytical methods

Bacterial growth was monitored as optical density at a wavelength of λ = 600 nm (OD_{600}) with an Ultrospec 10 Cell Density Meter (GE Healthcare, Little Chalfont, Buckinghamshire, UK).

For measuring extracellular metabolites, samples taken from liquid cultivation were centrifuged for 3 min at 17,000 ×g to obtain supernatant for High-Performance Liquid Chromatography (HLPC) analysis using a Beckman System Gold 126 Solvent Module equipped with a Smartline 2300 refractive index detector (Knauer, Berlin, Germany). Analytes were eluted using a 300 x 8 mm organic acid resin column together with a 40 x 8 mm organic acid resin precolumn (both from CS Chromatographie, Langerwehe, Germany) with 5 mM H_2SO_4 as mobile phase at a flow rate of 0.7 ml min^{-1} at 70 °C. (Mückschel et al. 2012) A list with retention times and detection limits are shown in **Table S4**.

Intracellular net fluxes were estimated by using the physiological data substrate uptake rates, metabolite secretion rates, and growth rates. Negative values represent substrate uptake rates. The *in vivo* redox cofactor regeneration rates in the ethylene glycol/ glyoxylate pathway were estimated by determining the sum of all *in vivo* reaction rates using the intracellular net fluxes and the redox cofactor stoichiometries of Figure 1, as previously described (Blank et al. 2008b).

The online analysis of growth using the Growth Profiler was analyzed using the Growth Profiler Control software V2_0_0. Cell densities are expressed as G-value, which is derived from imaging analysis of microtiter plates with transparent bottoms.

Statistical probability values were, if not stated otherwise, calculated using a paired Student's t-distribution test with homogeneity of variance (n = 2-3, significance level of 0.05).

Acknowledgments

The RWTH researchers acknowledge funding from the European Union´s Horizon 2020 research and innovation programme under grant agreement no. 633962 for the project P4SB. Nick Wierckx was supported by the German Research Foundation through the Emmy Noether project WI 4255/1-1. The laboratory of Lars M. Blank was partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s excellence strategy within the clusters of excellence

236 “TMFB – Tailor-Made Fuels from Biomass” and 2186 „FSC – The Fuel Science Center“. The work of Matthias Wehrmann and Janosch Klebensberger was supported by a research grant from the German Research Foundation (KL 2340/2-1). The NREL authors thank the U.S. Department of Energy (DOE), Energy Efficiency and Renewable Energy (EERE), Bioenergy Technologies Office (BETO) for funding this work via Contract No. DE-AC36-08GO28308 with NREL. We kindly acknowledge Dr. Juan Nogales for helpful discussions and thank Benedikt Wynands for advice in molecular biological work.

Conflict of Interest

Mary Ann Franden, Lahiru N. Jayakody, Gregg T. Beckham, Janosch Klebensberger, Nick Wierckx, Lars M. Blank and Wing-Jin Li have filed a patent application (US Patent App. 16/041,371) on the strains described in this manuscript.

References

Ahn S, Jung J, Jang I-A, Madsen EL, Park W (2016) Role of glyoxylate shunt in oxidative stress response. J Bio Chem 291:11928–11938

Arias S, Olivera ER, Arcos M, Naharro G, Luengo JM (2008) Genetic analyses and molecular characterization of the pathways involved in the conversion of 2-phenylethylamine and 2-phenylethanol into phenylacetic acid in *Pseudomonas putida* U. Environ Microbiol 10:413–432

Austin HP, Allen MD, Donohoe BS, Rorrer NA, Kearns FL, Silveira RL, Pollard BC, Dominick G, Duman R, El Omari K, Mykhaylyk V, Wagner A, Michener WE, Amore A, Skaf MS, Crowley MF, Thorne AW, Johnson CW, Woodcock HL, McGeehan JE, Beckham GT (2018) Characterization and engineering of a plastic-degrading aromatic polyesterase. Proceedings of the National Academy of Sciences of the United States of America 115:E4350-E4357

Belda E, van Heck RGA, José Lopez-Sanchez M, Cruveiller S, Barbe V, Fraser C, Klenk H-P, Petersen J, Morgat A, Nikel PI, Vallenet D, Rouy Z, Sekowska A, Martins Dos Santos VAP, Lorenzo V de, Danchin A, Médigue C (2016) The revisited genome of *Pseudomonas putida* KT2440 enlightens its value as a robust metabolic chassis. Environ Microbiol 18:3403–3424

Black BA, Michener WE, Ramirez KJ, Biddy MJ, Knott BC, Jarvis MW, Olstad J, Mante OD, Dayton DC, Beckham GT (2016) Aqueous stream characterization from biomass fast pyrolysis and catalytic fast pyrolysis. ACS Sustain Chem Eng 4:6815–6827

Blank LM, Ebert BE, Bühler B, Schmid A (2008a) Metabolic capacity estimation of *Escherichia coli* as a platform for redox biocatalysis. Constraint-based modeling and experimental verification. Biotechnol Bioeng 100:1050–1065

Blank LM, Ionidis G, Ebert BE, Bühler B, Schmid A (2008b) Metabolic response of *Pseudomonas putida* during redox biocatalysis in the presence of a second octanol phase. FEBS J 275:5173–5190

Bombelli P, Howe CJ, Bertocchini F (2017) Polyethylene bio-degradation by caterpillars of the wax moth *Galleria mellonella*. Curr Biol 27:R292-R293

Cartwright LN, Hullin RP (1966) Purification and properties of two glyoxylate reductases from a species of *Pseudomonas*. Biochem J 101:781–791

Cho C, Choi SY, Luo ZW, Lee SY (2015) Recent advances in microbial production of fuels and chemicals using tools and strategies of systems metabolic engineering. Biotechnol Adv 33:1455–1466

Cusa E, Obradors N, Baldomà L, Badia J, Aguilar J (1999) Genetic analysis of a chromosomal region containing genes required for assimilation of allantoin nitrgen and linked glyoxylate metabolism in *Escherichia coli*. J Bacteriol 181:7479-7484

Czernik S, Bridgwater AV (2004) Overview of applications of biomass fast pyrolysis oil. Energ Fuel 18:590–598

DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernytsky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, Daly MJ (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 43:491–498

Dobson S (2000) Ethylene glycol: environmental aspects. Concise international chemical assessment document 22. World Health Organization:1–24

Donald LJ, Hosefield D, Cuvelier SL, Ens W, Standing KG, Duckworth HW (2001) Mass spectrometric study of *Escherichia coli* repressor proteins, IclR and GclR, and their complexes with DNA. Protein Sci:1370–1380

Dragosits M, Mattanovich D (2013) Adaptive laboratory evolution - principles and applications for biotechnology. Microb Cell Fact 12:1–17

Franden MA, Jayakody LN, Li W-J, Wagner NJ, Cleveland NS, Michener WE, Hauer B, Blank LM, Wierckx N, Klebensberger J, Beckham GT (2018) Engineering *Pseudomonas putida* KT2440 for efficient ethylene glycol utilization. Metabolic engineering 48:197–207

Garcia JM, Robertson ML (2017) The furture of plastics recycling. Chemical advances are increasing the proportion of polymer waste that can be recycled. Sciences 358:870–872

Geyer R, Jambeck JR, Law KL (2017) Production, use, and fate of all plastics ever made. Science advances 3:e1700782 1-5

Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods 6:343–345

Göhler A-K, Kökpinar Ö, Schmidt-Heck W, Geffers R, Guthke R, Rinas U, Schuster S, Jahreis K, Kaleta C (2011) More than just a metabolic regulator--elucidation and validation of new targets of PdhR in Escherichia coli. BMC systems biology 5:197

Guzik MW, Kenny ST, Duane GF, Casey E, Woods T, Babu RP, Nikodinovic-Runic J, Murray M, O'Connor KE (2014) Conversion of post consumer polyethylene to the biodegradable polymer polyhydroxyalkanoate. Appl Microbiol Biot 98:4223–4232

Hanahan D (1983) Studies on transformation of *Escherichia coli* with plasmids. J Mol Biol 166:557–580

Hartmans S., Smiths JP, van der Werf MJ, Volkering F, Bont JAM de (1989) Metabolism of styrene oxide and 2-phenylethanol in the styrene-degrading *Xanthobacter s*train 124X. Appl Environ Microb 55:2850–2855

Hasegawa A, Ogasawara H, Kori A, Teramoto J, Ishihama A (2008) The transcription regulator AllR senses both allantoin and glyoxylate and controls a set of genes for degradation and reutilization of purines. Microbiology+ 154:3366–3378

Hazen TC (2018) Cometabolic Bioremediation. In: Steffan R (ed) Consequences of Microbial Interactions with Hydrocarbons, Oils, and Lipids: Biodegradation and Bioremediation. Springer International Publishing, Cham, pp 1–15

Ilves H, Horak R, Kivisaar M (2001) Involvement of sigmaS in starvation-induced transposition of *Pseudomonas putida* transposon Tn4652. J Bacteriol 183:5445–5448

Jambeck JR, Geyer R, Wilcox C, Siegler TR, Law KL (2015) Plastic waste inputs from land into the ocean. Science 347:764–768

Jayakody LN, Ferdouse J, Hayashi N, Kitagaki H (2017) Identification and detoxification of glycolaldehyde, an unattended bioethanol fermentation inhibitor. Crit Rev Biotechnol 37:177–189

Jayakody LN, Johnson CW, Whitham JM, Giannone RJ, Black BA, Cleveland NS, Klingeman DM, Michener WE, Olstad JL, Vardon DR, Brown RC, Brown SD, Hettich RL, Guss AM, Beckham GT (2018) Thermochemical wastewater valorization via enhanced microbial toxicity tolerance. Energy Environ. Sci. 311:484

Kenny ST, Runic JN, Kaminsky W, Woods T, Babu RP, Keely CM, Blau W, O'Connor KE (2008) Up-cycling of PET (polyethylene terephthalate) to the biodegradable plastic PHA (Polyhydroxyalkanoate). Environ Sci Technol 42:7696–7701

Kenny ST, Runic JN, Kaminsky W, Woods T, Babu RP, O'Connor KE (2012) Development of a bioprocess to convert PET derived terephthalic acid and biodiesel derived glycerol to medium chain length polyhydroxyalkanoate. Appl Microbiol Biot 95:623–633

Kumar S, Gupta RB (2008) Hydrolysis of microcrystalline cellulose in subcritical and supercritical water in a continuous flow reactor. Ind Eng Chem Res 47:9321–9329

Lau WWY, Armbrust EV (2006) Detection of glycolate oxidase gene glcD diversity among cultured and environmental marine bacteria. Environ Microbiol 8:1688–1702

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760

Liu Q, Luo G, Zhou XR, Chen G-Q (2011) Biosynthesis of poly(3-hydroxydecanoate) and 3-hydroxydodecanoate dominating polyhydroxyalkanoates by β-oxidation pathway inhibited *Pseudomonas putida*. Metab Eng 13:11–17

Loh K-C, Cao B (2008) Paradigm in biodegradation using *Pseudomonas putida*—A review of proteomics studies. Enzyme Microb Tech 43:1–12

Lu X, Yamauchi K, Phaiboonsilpa N, Saka S (2009) Two-step hydrolysis of Japanese beech as treated by semi-flow hot-compressed water. J Wood Sci 55:367–375

Martínez-García E, Aparicio T, Lorenzo V de, Nikel PI (2014a) New transposon tools tailored for metabolic engineering of gram-negative microbial cell factories. Front Bioeng Biotechnol 2:1–13

Martínez-García E, Lorenzo V de (2011) Engineering multiple genomic deletions in Gram-negative bacteria. Analysis of the multi-resistant antibiotic profile of *Pseudomonas putida* KT2440. Environ Microbiol 13:2702–2716

Martínez-García E, Nikel PI, Aparicio T, Lorenzo V de (2014b) Pseudomonas 2.0. Genetic upgrading of *P. putida* KT2440 as an enhanced host for heterologous gene expression. Microb Cell Fact 13:159

Mattioda G. CY (2000) Ullmann's Encyclopedia of Industrial Chemistry. Glyoxylic acid. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA (2010) The genome analysis toolkit. A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20:1297–1303

Mückschel B, Simon O, Klebensberger J, Graf N, Rosche B, Altenbuchner J, Pfannstiel J, Huber A, Hauer B (2012) Ethylene glycol metabolism by *Pseudomonas putida*. Appl Environ Microb 78:8531–8539

Narancic T, O'Connor KE (2017) Microbial biotechnology addressing the plastic waste disaster. Microb Biotechnol 10:1232–1235

Navone L, Macagno JP, Licona-Cassani C, Marcellin E, Nielsen LK, Gramajo H, Rodriguez E (2015) AllR controls the expression of S*treptomyces coelicolor* allantoin pathway genes. Appl Environ Microb 81:6649–6659

Nikel PI, Chavarría M, Fuhrer T, Sauer U, Lorenzo V de (2015) *Pseudomonas putida* KT2440 strain metabolizes glucose through a cycle formed by enzymes of the Entner-Doudoroff, Embden-Meyerhof-Parnas, and pentose phosphate pathways. J Biol Chem 290:25920–25932

Nikel PI, Martínez-García E, Lorenzo V de (2014) Biotechnological domestication of Pseudomonads using synthetic biology. Nat Rev Microbiol 12:368–379

Novichkov PS, Kazakov AE, Ravcheev DA, Leyn SA, Kovaleva GY, Sutormin RA, Kazanov MD, Riehl W, Arkin AP, Dubchak I, Rodionov DA (2013) RegPrecise 3.0--a resource for genome-scale exploration of transcriptional regulation in bacteria. BMC genomics 14:745

Nurk A, Tamm A, Hôrak R, Kivisaar M (1993) In-vivo-generated fusion promoters in *Pseudomonas putida*. Gene 127:23–29

Ogunola OS, Onada OA, Falaye AE (2018) Mitigation measures to avert the impacts of plastics and microplastics in the marine environment (a review). Environ Sci Pollut R:1–18

Park J-B, Bühler B, Panke S, Witholt B, Schmid A (2007) Carbon metabolism and product inhibition determine the epoxidation efficiency of solvent-tolerant *Pseudomonas* sp. strain VLB120ΔC. Biotechnol Bioeng 98:1219–1229

PlasticsEurope 2018 Plastics - The Facts 2017. An analysis of European plastics production, demand and waste data; https://www.plasticseurope.org/application/files/5715/1717/4180/Plastics_the_facts_2017_FINAL_for_website_one_page.pdf (from 31.08.2018)

Rintoul MR, Cusa E, Baldomà L, Badia J, Reitzer L, Aguilar J (2002) Regulation of the *Escherichia coli* allantoin regulon. Coordinated function of the repressor AllR and the activator AllS. J Mol Biol 324:599–610

Sajtos A. (1991) Process for the preparation of glyoxylic acid and glyoxylic acid derivates. US patent 5015,760; application year 1989

Schultz AC, Nygaard P, Saxild HH (2001) Functional analysis of 14 genes that constitute the purine catabolic pathway in *Bacillus subtilis* and evidence for a novel regulon controlled by the PucR transcription activator. J Bacteriol 183:3293–3302

Thorvaldsdóttir H, Robinson JT, Mesirov JP (2013) Integrative Genomics Viewer (IGV). High-performance genomics data visualization and exploration. Brief Bioinform 14:178–192

Vallon T, Simon O, Rendgen-Heugle B, Frana S, Mückschel B, Broicher A, Siemann-Herzberg M, Pfannstiel J, Hauer B, Huber A, Breuer M, Takors R (2015) Applying systems biology tools to study *n*-butanol degradation in *Pseudomonas putida* KT2440. Eng Life Sci 15:760–771

Vispute TP, Zhang H, Sanna A, Xiao R, Huber GW (2010) Renewable chemical commodity feedstocks from integrated catalytic processing of pyrolysis oils. Science 330:1222-1227

Ward PG, Goff M, Donner M, Kaminsky W, O'Connor KE (2006) A two step chemo-biotechnological conversion of polystyrene to a biodegradable thermoplastic. Environ Sci Technol 40:2433–2437

Wehrmann M, Billard P, Martin-Meriadec A, Zegeye A, Klebensberger J (2017) Functional role of lanthanides in enzymatic activity and transcriptional regulation of pyrroloquinoline quinone-dependent alcohol dehydrogenases in *Pseudomonas putida* KT2440. Mbio 8:1–14. e00570-17

Wei R, Zimmermann W (2017a) Biocatalysis as a green route for recycling the recalcitrant plastic polyethylene terephthalate. Microb Biotechnol 10:1302–1307

Wei R, Zimmermann W (2017b) Microbial enzymes for the recycling of recalcitrant petroleum-based plastics. How far are we? Microb Biotechnol 10:1308–1322

Wierckx N, Prieto MA, Pomposiello P, Lorenzo V de, O'Connor K, Blank LM (2015) Plastic waste as a novel substrate for industrial biotechnology. Microb Biotechnol 8:900–903

Wilkes RA, Aristilde L (2017) Degradation and metabolism of synthetic plastics and associated products by *Pseudomonas sp.* Capabilities and challenges. J Appl Microbl 123:582–593

Wynands B, Lenzen C, Otto M, Koch F, Blank LM, Wierckx N (2018) Metabolic engineering of *Pseudomonas taiwanensis* VLB120 with minimal genomic modifications for high-yield phenol production. Metab Eng 47:121–133

Yang J, Yang Y, Wu W-M, Zhao J, Jiang L (2014) Evidence of polyethylene biodegradation by bacterial strains from the guts of plastic-eating waxworms. Environ Sci Technol 48:13776–13784

Yoshida S, Hiraga, Kazimi, Takehara T, Oda K (2016) A bacterium that degrades and assimilates poly(ethylene terephthalate). Science 351:1196–1199

Yu Y, Lou X, Wu H (2008) Some recent advances in hydrolysis of biomass in hot-compressed water and its comparisons with other hydrolysis methods. Energ Fuel 22:46–60

Yue H, Zhao Y, Ma X, Gong J (2012) Ethylene glycol. Properties, synthesis, and applications. Chem Soc Rev 41:4218–4244

Zobel S, Benedetti I, Eisenbach L, Lorenzo V de, Wierckx N, Blank LM (2015) Tn7-based device for calibrated heterologous gene expression in *Pseudomonas putida*. ACS Synth Biol 4:1341–1351

Zobel S, Kuepper J, Ebert B, Wierckx N, Blank LM (2017) Metabolic response of *Pseudomonas putida* to increased NADH regeneration rates. Eng Life Sci 17:47–57

Table

Table 1. Bacterial strains used in this study

No.	strain	relevant characteristics	reference
	E. coli		
	DH5α	*supE44*, Δ*lacU169* (*φ80lacZ*ΔM15), *hsdR17* ($r_K^-m_K^+$), *recA1*, *endA1*, *gyrA96*, *thi-1*, *relA1*	Hanahan (1985)
	DH5αλpir	λ*pir* lysogen of DH5α; host for *oriV(*R6K) vectors	de Lorenzo and Timmis (1994)
	HB101 pRK2013	Sm^R, *hsdR-M*$^+$, *proA2*, *leuB6*, *thi-1*, *recA*; bearing plasmid pRK2013	Ditta et al. (1980)
	DH5α pSW-2	DH5a bearing plasmid pSW-2 encoding I-sceI nuclease, tool for genomic deletion	Martínez-García and de Lorenzo (2011)
	P. putida		
1	KT2440	cured, restriction-deficient derivative of *P. putida* mt-2	Bagdasarian (1981)
E6.1	KT2440 E6.1	Isolate after 6th transfer from laboratory evolution on ethylene glycol in Aachen, Germany	this study
E6.2	KT2440 E6.2	Isolate after 6th transfer from laboratory evolution on ethylene glycol in Aachen, Germany	this study
E1.1	KT2440 E1.1	Isolate after 1st transfer from laboratory evolution on ethylene glycol in Stuttgart, Germany	this study
E1.2	KT2440 E1.2	Isolate after 1st transfer from laboratory evolution on ethylene glycol in Stuttgart, Germany	this study
E1.3	KT2440 E1.3	Isolate after 1st transfer from laboratory evolution on ethylene glycol in Stuttgart, Germany	this study
2	KT2440 Δ*gclR*	Δ*gclR*	this study
3	KT2440 Δ*gclR* ΔPP_2046	Δ*gclR,* ΔPP_2046	this study
4	KT2440 Δ*gclR* ΔPP_2662	Δ*gclR,* ΔPP_2662	this study
5	KT2440 Δ*gclR* ΔPP_2046 ΔPP_2662	Δ*gclR,* ΔPP_2046, ΔPP_2662	this study
6	KT2440 Δ*gclR* ΔPP_2046*::*14g ΔPP_2662	Δ*gclR,* ΔPP_2046*::*14g; PP_2046 replaced by synthetic promotor 14g from Zobel et al. 2015, ΔPP_2662	this study
7	KT2440 Δ*gclR* ΔPP_2662*::*14d	Δ*gclR,* ΔPP_2662*::*14d; PP_2662 replaced by synthetic promotor 14d from Zobel et al. 2015)	this study
8	KT2440 Δ*gclR* ΔPP_2046 ΔPP_2662*::*14d	Δ*gclR,* ΔPP_2046, ΔPP_2662*::*14d	this study
9	KT2440 Δ*gclR* ΔPP_2046*::*14g ΔPP_2662*::*14d	Δ*gclR,* ΔPP_2046::14g, ΔPP_2662::14d	this study
10	KT2440 Δ*ped*	Δ*ped* (PP_2673-80)	this study
11	KT2440 E6.1 Δ*ped*	Δ*ped* (PP_2673-80)	this study
12	KT2440 E6.2 Δ*ped*	Δ*ped* (PP_2673-80)	this study

Supporting information for "Laboratory evolution reveals the metabolic and regulatory basis of ethylene glycol metabolism by *Pseudomonas putida* KT2440"

Wing-Jin Li[1], Lahiru N. Jayakody[2], Mary Ann Franden[2], Matthias Wehrmann[3], Tristan Daun[1], Bernhard Hauer[3], Lars M. Blank[1], Gregg T. Beckham[2*], Janosch Klebensberger[3*], Nick Wierckx[1,4*]

[1] Institute of Applied Microbiology-iAMB, Aachen Biology and Biotechnology-ABBt, RWTH Aachen University, Worringerweg 1, 52074 Aachen, Germany
[2] National Bioenergy Center, National Renewable Energy Laboratory, Golden, CO, 80401, USA
[3] University of Stuttgart, Institute of Biochemistry and Technical Biochemistry, Allmandring 31, 70569 Stuttgart, Germany
[4] Institute of Bio- and Geosciences IBG-1: Biotechnology, Forschungszentrum Jülich, 52425 Jülich, Germany

*Address correspondence to:
Nick Wierckx: n.wierckx@fz-juelich.de
Janosch Klebensberger: janosch.klebensberger@itb.uni-stuttgart.de
Gregg T. Beckham: gregg.beckham@nrel.gov

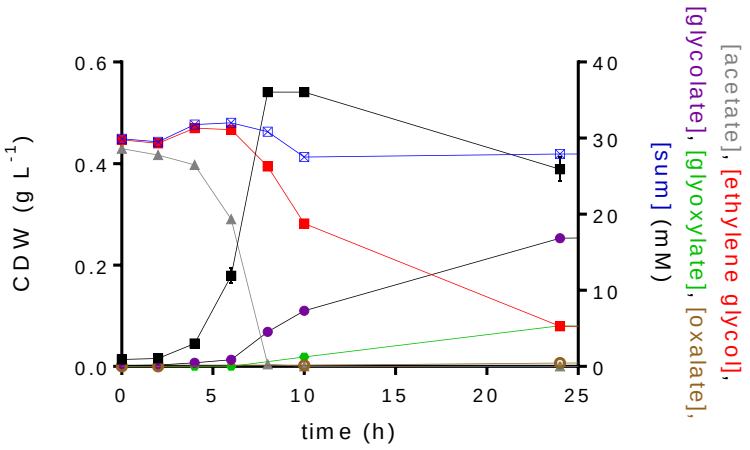


Figure S1 Growth of *P. putida* KT2440 in MSM medium containing 28.6 mM acetate and 29.8 mM ethylene glycol in duplicates. Compounds from ethylene glycol metabolism (ethylene glycol, glycolate, glyoxylate and oxalate) are shown individually and in sum. The culture was incubated at 30 °C and 200 rpm. Error bars indicate the deviation from the mean (n = 2) and according to error propagation.

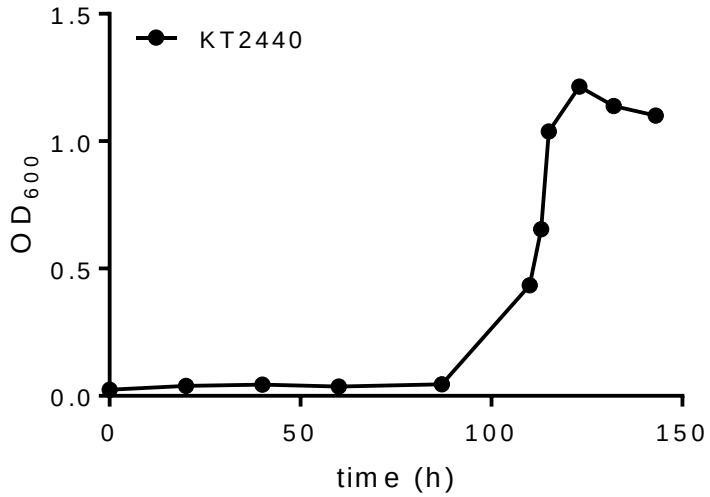


Figure S2 Growth of *P. putida* KT2440 in 50 mL (250 mL Erlenmeyer flask) of modified M9 medium (Wehrmann et al. 2017) containing 20 mM ethylene glycol. The culture was incubated at 30°C and 180 rpm and growth was determined by measuring OD_{600} in a photometer (Eppendorf Biophotometer). One exemplary culture is shown.

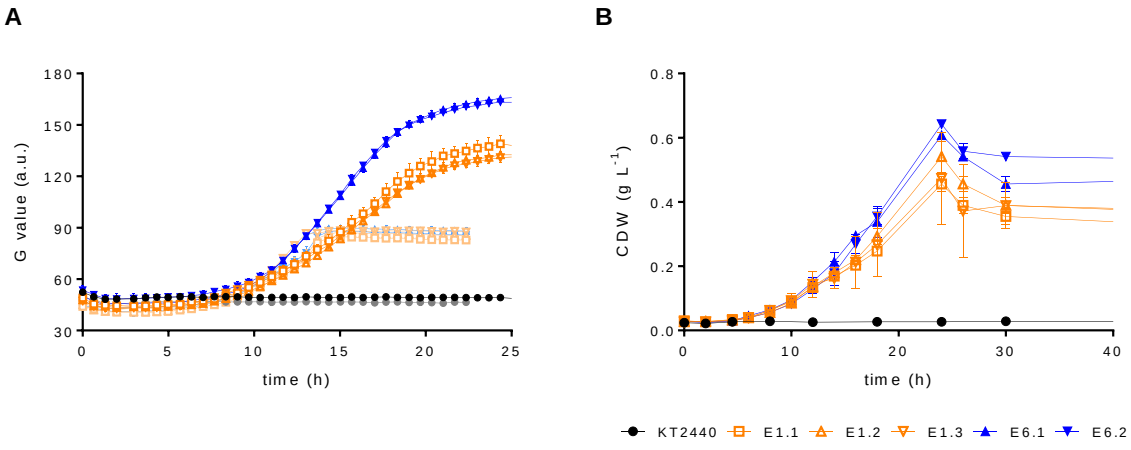


Figure S3 A: Growth comparison of *P. putida* KT2440 and all adapted strains in MSM containing 30 mM ethylene glycol (in light colors) and 120 mM ethylene glycol (in darker colors). Growth was detected via a Growth Profiler in 96-square-well plates. **B:** Biomass growth of the isolated ALE strains E1.1 and E6.1 growing on 30 mM ethylene glycol in a shake flask cultivation on MSM with 30 mM ethylene glycol. Error bars indicate the deviation from the mean (n = 2).

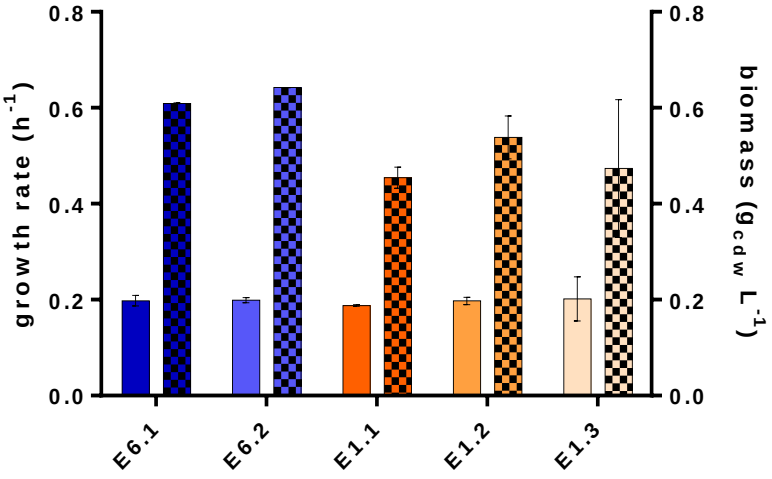


Figure S4 Comparison of growth rate and maximum biomass concentrations of *P. putida* KT2440 strains E6.1, E6.1, E1.1, E1.2 and E1.3 from the two evolution lines grown in a shake flask cultivation on MSM with 26.7 ± 0.4 mM ethylene glycol. Growth rate shown in filled pattern, maximum biomass concentrations are depicted in checked pattern. Error bars indicate the deviation from the mean (n = 2).

Table S1 List of all mutations found for all sequenced strains. Numbers of found Single Nucleotide Polymorphisms (SNP) and Insertion-Deletion polymorphisms (InDel) and their functional class, divided by the cause of mutation, are shown. KT[A] represents the *P. putida* KT2440 from Aachen. KT[S] describes *P. putida* KT2440 from Stuttgart.

		reference vs:		KT[A] vs:		KT[S] vs:		
		KT[A]	KT[S]	E6.1	E6.2	E1.1	E1.2	E1.3
sum of all mutations		83	80	13	6	4	4	4
type of mutation	SNP	55	54	13	5	4	3	3
	InDel	28	26	0	1	0	1	1
functional class (SNP/InDel)	intergenic	46/28	43/26	2/0	2/1	1/0	1/1	1/1
	silent	7/0	8/0	8/0	2/0	2/0	1/0	2/0
	missense	2/0	3/0	2/0	1/0	1/0	1/0	0/0
	nonsense	0/0	0/0	1/0	0/0	0/0	0/0	0/0

Supporting information 1: GclR binding sites (BS) mutation in *P. putida* E6.2

glcG←
GTTCATGGAAAGGTCCTCTTCTTGTTGTGAGAAGCCCTAAGGGCAGTTGAAAACCGATCGAAAAAATA

GAACACAACGAGCGATATTTT**TGTATACA**ATATTTTGAAACGATCGTATGCGATGGCGCAAACCTCCT
GclR BS1
TTTCACGGGCGCTCTGACGAAAGCCAGCTTAGCCGATGAAAACCCATTGACCTAAGCCGTCAGGCGTG
A
AATACACTCTGTCGCAAAGCAAGT**TGTATACA**ATTACAAAATCGATGAGGCACAAACCATGAGCAAAA
GclR BS2 →*gcl*
-35 Emergent promoter -10

Supporting information 2: Bprom analysis of mutated sequence in *P. putida* E6.2: (http://www.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfindb)

```
>test sequence:
TTGACCTAAGCCGTCAGGCGTGAATACACTCTGTCGCAAAGCAAGTTATATACAATTACAAAATCGATGAGGCAC
AAACC
 Length of sequence-        80
 Threshold for promoters -  0.20
 Number of predicted promoters -      1
 Promoter Pos:     61 LDF-  0.82
 -10 box at pos.     48 ATATACAAT Score    54
 -35 box at pos.     25 TACACT    Score     4

 Oligonucleotides from known TF binding sites:

 For promoter at     61:
       lexA:  TATATACA at position      47 Score -  13
```

Supporting information 3: Regprecise (http://regprecise.lbl.gov/RegPrecise) analysis of *gclR*:

Profile of regulator PA1520 in Pseudomonadaceae

Properties	
Regulator family:	GntR/Others
Regulation mode:	
Biological process:	Xanthine utilization; Allantoin utilization
Effector:	
Regulog:	PA1520 - Pseudomonadaceae

- By taxonomy - Pseudomonadaceae
- By TF family - GntR/Others
- By pathway - Xanthine utilization
- By pathway - Allantoin utilization

Table S2 List of GntR binding sites

PP_xxxx	Name	Position	Score	Binding site	Location
PP_4297	*gcl*	-35	4.7	TTGTATACAA	Upstream of *gcl* cluster
PP_4297	*gcl*	-175	4.7	TTGTATACAA	Upstream of *gcl* cluster
PP_4309	PA0476	-427	4.8	TTGTACACAA	Upstream of putative purine/allantoin permease and hydrantoin racemase
PP_4309	PA0476	-166	4.7	TTGTATACAA	Upstream of putative purine/allantoin permease and hydrantoin racemase
PP_4643	PA2938	-181	4.3	TTGTACACAC	Upstream of putative xanthine/uracil permease
PP_4643	PA2938	-78	4.7	TTGTATACAA	Upstream of putative xanthine/uracil permease
PP_4284	PA1519	-146	4.8	TTGTACACAA	Upstream of putative ureate permease
PP_4284	PA1519	-245	4.6	TTGTACACAG	Upstream of putative ureate permease
PP_4283	PA1520	-45	4.6	CTGTGTACAA	Upstream of *glcR* (Duplicate, reverse hit of PP_4284)
PP_4283	PA1520	-144	4.8	TTGTGTACAA	Upstream of *glcR* (Duplicate, reverse hit of PP_4284)
PP_4296	PA1503	-232	4.7	TTGTATACAA	Upstream of putative *glcG* (Duplicate, reverse hit of PP_4297)
PP_4296	PA1503	-92	4.7	TTGTATACAA	Upstream of putative *glcG* (Duplicate, reverse hit of PP_4297)
PP_4286	PA1517	-86	4.7	TTGTATACAA	Upstream of putative allantionase
PP_4286	PA1517	-291	4.8	TTGTACACAA	Upstream of putative allantionase
PP_4285	PA1518	-323	4.7	TTGTATACAA	Upstream of putative hydroxyisourate hydrolase (Duplicate, reverse hit of PP_4286)
PP_4285	PA1518	-118	4.8	TTGTGTACAA	Upstream of putative hydroxyisourate hydrolase (Duplicate, reverse hit of PP_4286)

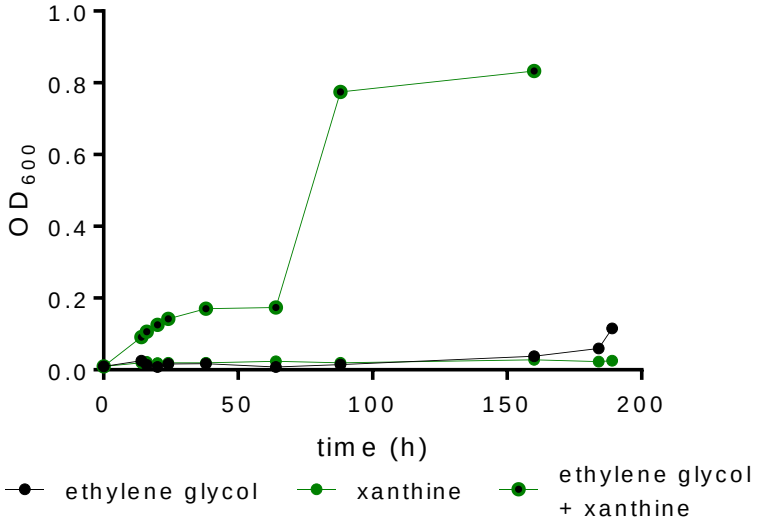
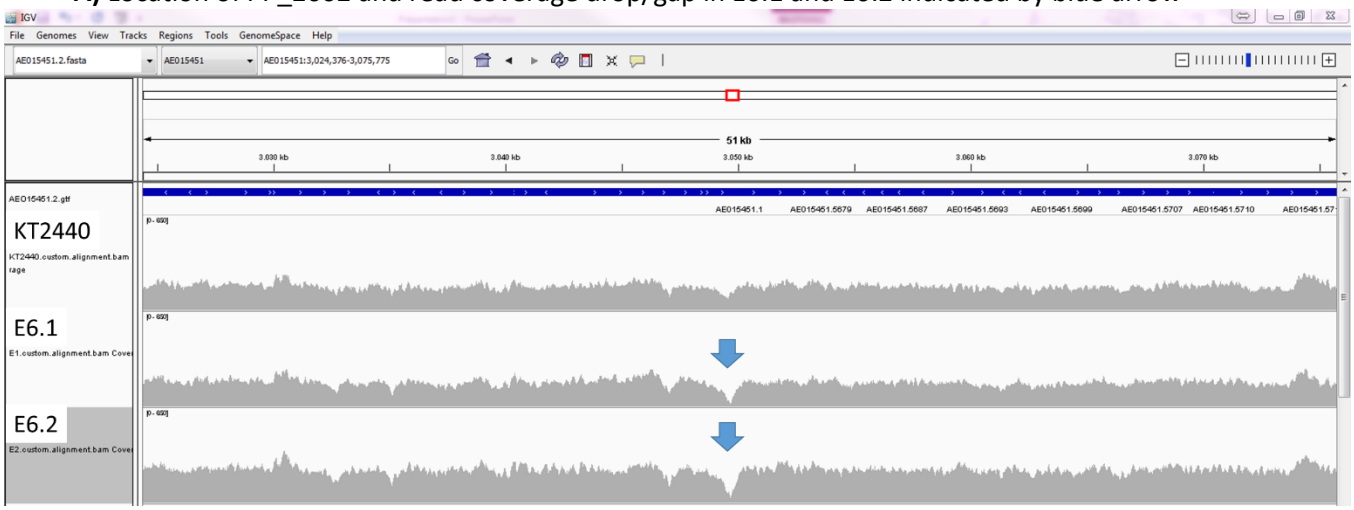


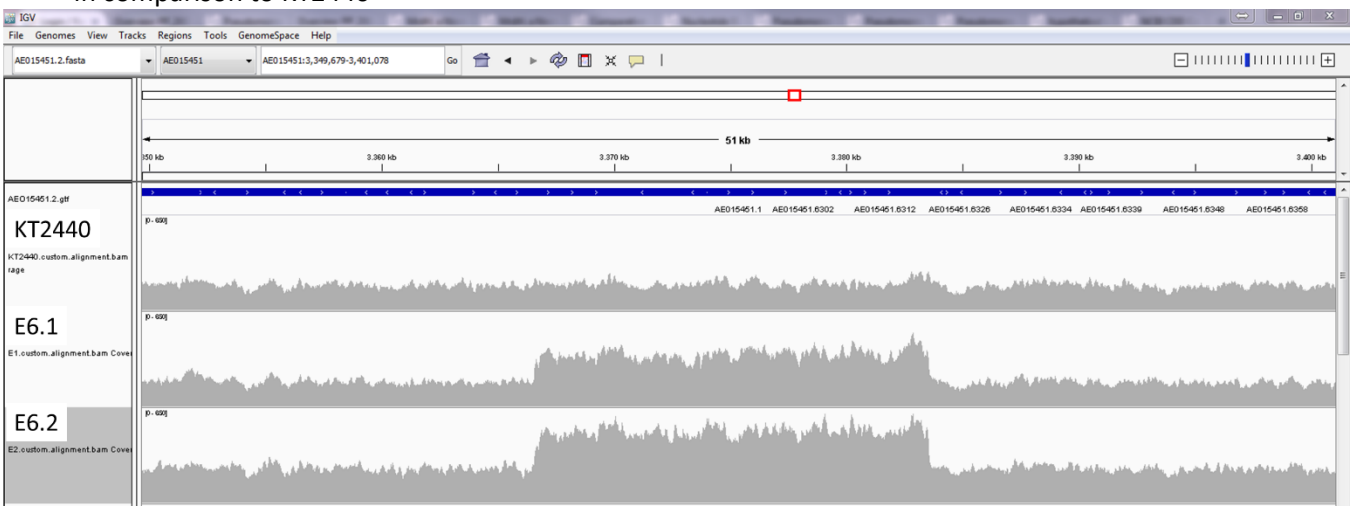
Figure S5 Growth of *P. putida* KT2440 in 20 mL (125 mL Erlenmeyer flask) of M9 medium containing 20 mM ethylene glycol (black circles), 60 mg L^{-1} xanthine (green circles) or a mixture of 20 mM ethylene glycol and 60 mg L^{-1} xanthine (green circles with black fill). The cultures were incubated at 30 °C and 180 rpm.

Supporting information 4: Locating Tn4652 using read coverage analysis

A) Location of PP_2662 and read coverage drop/gap in E6.1 and E6.2 indicated by blue arrow



B) Native location of Tn4652, showing twice approximately doubled read coverages in E6.1 and E6.2 in comparison to KT2440



Supporting information 5: Mapped transposon Tn4652 and the resulting putative promotor insertion

```
                                  PP_2662←
GGTGGGTTACGGGGTGCAGGCAAAGATGGGCGGCTGATGCCGAGATAAGGCAAAAATTAG

← 16.8kb Tn4652 →

CACCGATGCCATTGAACACATCCCCCAAGCGAGGCAAAAGCATCAGCATAGACGGCTAGC
                                        -35                -10
CAGACGGTGGATGACCAGCAAGCCACGGCCTCTCGAATAACCTTTTAAATCATATATTTA
                                          → PP_2662
CAGAACGAATGTCCTAATTTTTGCCTTATCTCGGCATGTCGCTGACCGGCTCGGGGTTCC
```

Supporting information 6: Bprom analysis of 3´ transposon site of Tn4652

```
>test sequence
CCCCCAAGCGAGGCAAAAGCATCAGCATAGACGGCTAGCCAGACGGTGGATGACCAGCAA
GCCACGGCCTCTCGAATAACCTTTTAAATCATATATTTACAGAACGAATGTCCTAATTTT
TGCCTTATCTCGGCAT
 Length of sequence-       136
 Threshold for promoters -  0.20
 Number of predicted promoters -      1
 Promoter Pos:    104 LDF-  1.39
 -10 box at pos.     89 TCATATATT Score    48
 -35 box at pos.     72 TCGAAT    Score    18

 Oligonucleotides from known TF binding sites:

 No such sites for promoter at    104
```

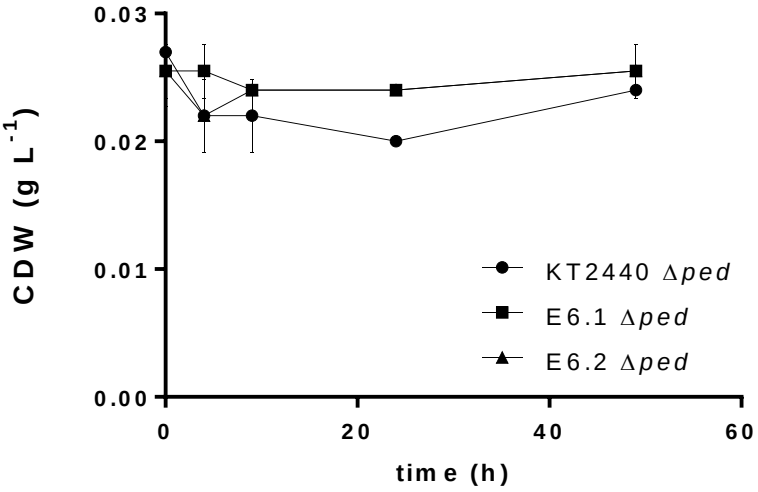


Figure S6 Growth of *P. putida* KT2440 Δ*ped* (PP_2673-80), E6.1 Δ*ped* (PP_2673-80) and E6.2 Δ*ped* (PP_2673-80) in 50 mL (500 mL Erlenmeyer flask) of MSM medium containing 30 mM ethylene glycol in duplicates. The cultures were incubated at 30 °C and 200 rpm.

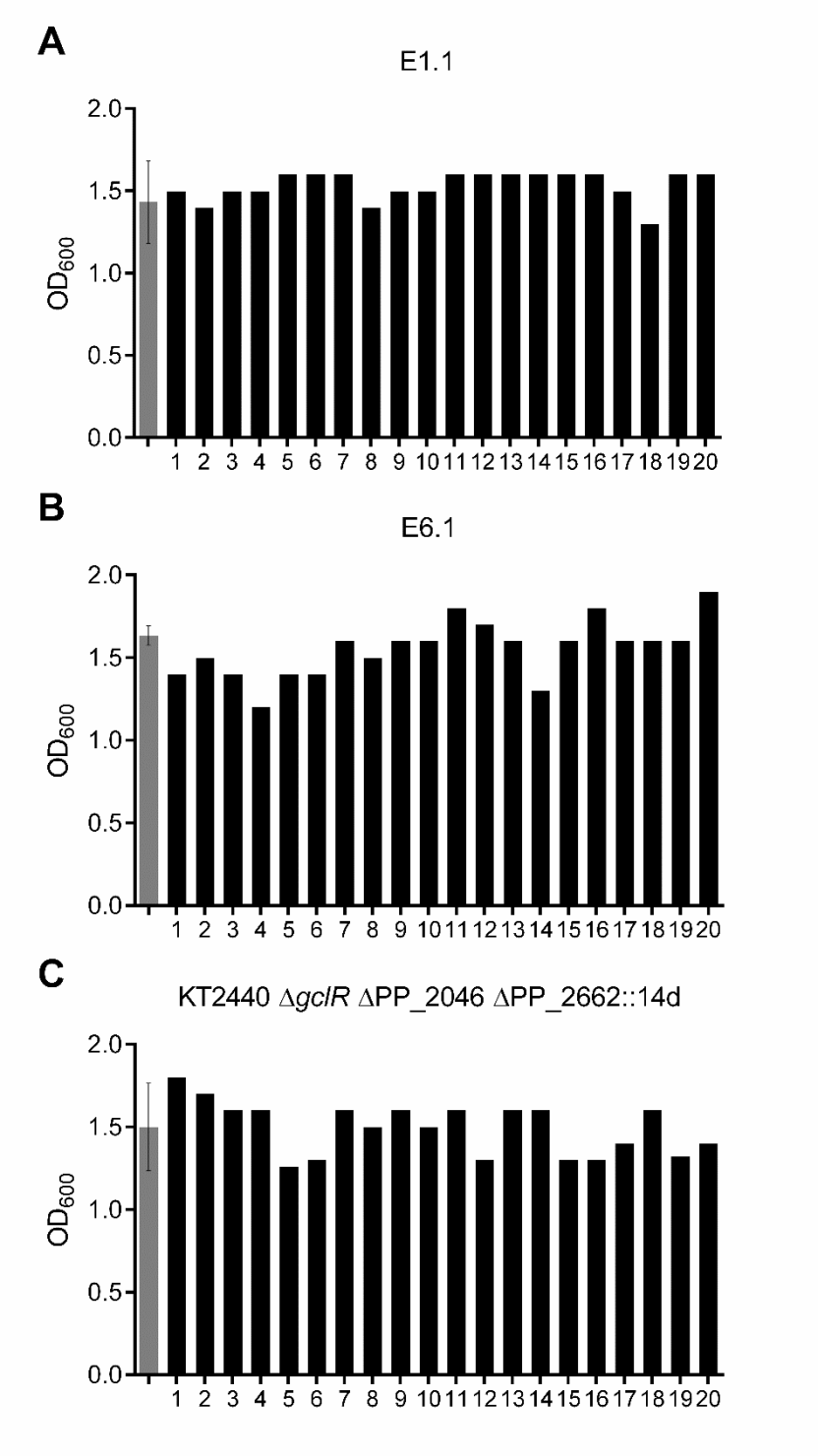


Figure S7 Long-term phenotypic robustness of E1.1, E6.1 and KT2440 Δ*gclR* ΔPP_2046 ΔPP_2662::14d. 20 single colonies, picked after 134 (E1.1), 116 (E6.1) and 134 (KT2440 Δ*gclR* ΔPP_2046 ΔPP_2662::14d) generations of cultivation in LB medium, were cultivated in MSM containing 30 mM ethylene glycol (black bars, n = 1) and compared to their corresponding parental strain (grey bars, error bars indicate the standard deviation; n = 3). Growth was detected via OD_{600} measurement after 30 h of cultivation in System Duetz 24-square-well plates. All picked clones grow on ethylene glycol, indicating that >95 % of the population from the cultures in unselective medium have retained their degradative phenotype.

Table S3 Primers used in this study

Name	Sequence (5´ → 3´)	Gene target	Description
WJ37.x	CTCTGAATTCGCCAGAACAGGCAATAGAG	PP_4297	TS1 Δ*gcl*
WJ38.2	TTCCTCTAGATGTGCCTCATCGATTTTGTAATTG	PP_4297	TS1 Δ*gcl*
WJ39	TTCCGGTACCGATCGCCTGACGCCCCCAGG	PP_4297	TS2 Δ*gcl*
WJ40	TTCCTCTAGACCCGCTTCGGTGGTGGTCAG	PP_4297	TS2 Δ*gcl*
WJ56	TCTCGGTACCCGCTGGCCCTTAACATTCCC	PP_2674	TS1 Δ*pedE*
WJ57	CCGCTCTAGAAATTTCCACCCGCTATTCAC	PP_2674	TS1 Δ*pedE*
WJ66	ATCGTCTAGAGCCCGCTCCCACAGGTTCAC	PP_2680	TS2 Δ*pedI*
WJ67	ATCGGTCGACGGCACCAAAGATGATTTCAG	PP_2680	TS2 Δ*pedI*
WJ70	TCTCGAATTCCCGATAGCAGCACCGATCAG	PP_4283	TS1 Δ*gclR*
WJ71	TCTCGGTACCATCGTTTGCCTGCGTGATCG	PP_4283	TS1 Δ*gclR*
WJ72	TCTCGGTACCTCTCGAGGCACGAAGAGAAAG	PP_4283	TS2 Δ*gclR*
WJ73	TCTCGTCGACCGATGATGCCTGCAGTCTTC	PP_4283	TS2 Δ*gclR*
WJ74	TCTCGAATTCTCCGGCATCCACCTGGCCTC	PP_2046	TS1 ΔPP_2046
WJ75	TCTCGGTACCAGCCATCAGGAAACGCGATAG	PP_2046	TS1 ΔPP_2046
WJ76	TCTCGGTACCTACCTTCGGCCTGCTTAGGG	PP_2046	TS2 ΔPP_2046
WJ77	TTCCTCTAGATCCAGGTCGATGCCCACCAC	PP_2046	TS2 ΔPP_2046
WJ78	CTGCGTGCACCAGCGTTAAG	PP_4281	Δ*gclR* Verification
WJ79	TGGCGATCAGCAGGAAGCAC	PP_4284	Δ*gclR* Verification
WJ80	TGAGGCTGACAGTGGCATTG	PP_2045	ΔPP_2046 Verification
WJ81	AGCGCATTATCGACCTGCAC	PP_2047	ΔPP_2046 Verification
WJ95	ACGCTCCTGCTTTCTTGTAG	Downstream PP_2046	ΔPP_2046 Verification
WJ97	GCGCAATGTGTCATACAACG	Downstream PP_4283	Δ*gclR* Verification
WJ110	GCAGCAACCGCACCGTTATG	PP_2661	PP_2662 Verification
WJ111	AGGCGTATGCCGCCAAAGTC	PP_2663	PP_2662 Verification
WJ112	TTCGCAGGCACTGGCTTC	PP_2661	PP_2662 Verification
WJ113	AGCTCGGTACCCGGGGATCCtCCCTGTTCGCACGAGGGCG	PP_2661	TS1 ΔPP_2662
WJ114	CCTAGCGGCCCCGGCCGCTTTGGCGTGGGGTGGCGA	Upstream PP_2662	TS1 ΔPP_2662
WJ115	TCGCCACCCCACGCCAAAGCGGCCGGGGCCGCTAGG	Downstream PP_2662	TS2 ΔPP_2662
WJ116	TGCATGCCTGCAGGTCGACTTCGGCACCGGTTACCACCAGCTTTTCC	PP_2663	TS2 ΔPP_2662
WJ117	CCTAGGTCAACTTATTATACAGTCGCGAATGTCGGATGTCAAGTAGATTAATTAATTTGGCGTGGGGTGGCGA	Upstream PP_2662	TS1 ΔPP_2662+ promotor insertion 14d

WJ118	TTAATTAATCTACTTGACATCCGACATTCGCGACTGTATAATAAGTTGACCTAGGGCGGCCGGGGCCGCTAGG	Downstream PP_2662	TS2 ΔPP_2662+ promotor insertion 14d
WJ119	AGCTCGGTACCCGGGGATCCtCCGGCATCCACCTGGCCTC	PP_2046	TS1 ΔPP_2046
WJ120	CCTAGGTCGTGCAATTATACCTGGCCGCGAGAGCCTTGTCAATGGGCTTAATTAAAGCCATCAGGAAACGCGATAG	Upstream PP_2046	TS1 ΔPP_2046+ promotor insertion 14g
WJ121	TTAATTAAGCCCATTGACAAGGCTCTCGCGGCCAGGTATAATTGCACGACCTAGGTACCTTCGGCCTGCTTAGGG	PP_2046	TS1 ΔPP_2046+ promotor insertion 14g
WJ122	TGCATGCCTGCAGGTCGACTTCCAGGTCGATGCCCACCAC	PP_2047	TS2 ΔPP_2046
WJ123	CGCCGGCATTGTTGGTGAAG	PP_2662	PP_2662 Verification
WJ124	GAACGGGTTGATGCCTTTGG	PP_2662	PP_2662 Verification
WJ125	TCCATTCGCGGTGTAGATTC	PP_2964	Tn4652 Verification
WJ126	CGCGCAGATCTACAGTTGGG	PP_2662	PP_2662 Verification
WJ127	ACTGACTCCACGGTAGAAAC	PP_2662	PP_2662 Verification
WJ128	CGTCGCTTTCCCTGTGTATC	PP_2984	Tn4652 Verification
oLP299	AGGCATTCGTGAAGTCATGG	PP_0387	*rpo*D qRT-PCR
oLP300	ATGTAACCGCTGAGAACGTC	PP_0387	*rpo*D qRT-PCR
oLP301	CTCGCCACTGGATCAACTG	PP_4297	*gcl* qRT-PCR
oLP302	GAACTGGAAGTCGTAGTCACC	PP_4297	*gcl* qRT-PCR
oLP303	TGCAGATCATGGAAGGTGAC	PP_4298	*hyi* qRT-PCR
oLP304	CAGGAAGCGGTAGTTGATCTC	PP_4298	*hyi* qRT-PCR
oLP305	AAAGAGGTTGCCCAGGAAG	PP_4299	*glxR* qRT-PCR
oLP306	CGAGCTCATGTCGATCACC	PP_4299	*glxR* qRT-PCR
oLP307	CCATCCTCAAACGCTACAAC	PP_4300	*ttuD* qRT-PCR
oLP308	TGGCGATCAACTGGAAGTG	PP_4300	*ttuD* qRT-PCR
oLP309	ACATCTTCCGCCTCAACTTC	PP_4301	*pykF* qRT-PCR
oLP310	TTGCAGGTCCATGAGGATG	PP_4301	*pykF*

			qRT-PCR
oLP325	AACTGAAGCTGATCCTGGTG	PP_0762	*hprA* qRT-PCR
oLP326	AGGGTATGCTGGGCTACA	PP_0762	*hprA* qRT-PCR
MAF_LP339	TGCCAAGCGTGTTTTGATTATC	PP_2049	PP_2049 qRT-PCR
MAF_LP340	CCGTAACCTCACTGAAGATCG	PP_2049	PP_2049 qRT-PCR
MAF_LP343	AAGAACTGGAACCCCATGG	PP_2674	*pedE* qRT-PCR
MAF_LP344	CTTGTAGTTCACTTCCTCGGTC	PP_2674	*pedE* qRT-PCR
MAF_LP345	ATCGAACAGAACCTGGAAGTG	PP_2680	*pedI* qRT-PCR
MAF_LP346	GCAAAATAGCGGAAGTGGTC	PP_2680	*pedI* qRT-PCR
MAF_LP351	GTTGTGATTGTTTCCTGCGG	PP_2663	PP_2663 qRT-PCR
MAF_LP352	AACTGTTCTCCGTAGGTGTTG	PP_2663	PP_2663 qRT-PCR
MAF_LP353	AGACCAATATCAACGGCCAG	PP_2664	PP_2664 qRT-PCR
MAF_LP354	GATCAGCACGAAGATACCCAC	PP_2664	PP_2664 qRT-PCR

Table S4 List of Retention time and detection limits for measuring extracellular metabolites using High-Performance Liquid Chromatography

Compound	**Retention time**	**Detection limit**
Oxalate	6.37 min	0.5 mM
Glucose	8.2 min	0.5 mM
Glyoxylate	8.7 min	0.5 mM
Glycolate	11.19 min	0.5 mM
Ethylene glycol	14.33	0.5 mM