Green Chemistry

Generation of Ionic Liquid Tolerant Pseudomonas putida KT2440 Strains via Adaptive Laboratory Evolution

Journal:	*Green Chemistry*
Manuscript ID	GC-ART-05-2020-001663.R1
Article Type:	Paper
Date Submitted by the Author:	23-Jul-2020

Complete List of Authors:	Lim, Hyun Gyu; University of California San Diego, Department of Bioengineering; Joint BioEnergy Institute Fong, Bonnie ; Joint BioEnergy Institute; Lawrence Berkeley National Laboratory, Biological Systems and Engineering Division Alarcon, Geovanni ; University of California San Diego, Department of Bioengineering Magurudeniya, Harsha; Joint BioEnergy Institute; Sandia National Laboratories California, Department of Biomass Science and Conversion Technology Eng, Thomas; Joint BioEnergy Institute, ; Lawrence Berkeley National Laboratory, Biological Systems and Engineering Szubin, Richard ; University of California San Diego, Department of Bioengineering Olson, Connor ; University of California San Diego, Department of Bioengineering Palsson, Bernhard; University of California San Diego, Department of Bioengineering; Joint BioEnergy Institute; Technical University of Denmark, Novo Nordisk Foundation Center for Biosustainability; University of California San Diego, Department of Pediatrics Gladden, John; Joint BioEnergy Institute, Deconstruction; Agile BioFoundry, Department of Energy; Sandia National Laboratories California, Department of Biomass Science and Conversion Technology Simmons, Blake; Joint BioEnergy Institute; Lawrence Berkeley National Laboratory, Biological Systems and Engineering Division Mukhopadhyay, Aindrila; Joint BioEnergy Institute; Lawrence Berkeley National Laboratory, Biological Systems and Engineering Division; Lawrence Berkeley National Laboratory, Environmental Genomics and Systems Biology Division Singer, Steven; Joint BioEnergy Institute; Lawrence Berkeley National Laboratory, Biological Systems and Engineering Division Feist, Adam; UCSD, Bioengineering; Joint BioEnergy Institute; Technical University of Denmark, Novo Nordisk Foundation Center for Biosustainability

Generation of Ionic Liquid Tolerant *Pseudomonas putida* KT2440 Strains via Adaptive Laboratory Evolution

Hyun Gyu Lim[a,b], Bonnie Fong[b,g], Geovanni Alarcon[a], Harsha D. Magurudeniya[b,f], Thomas Eng[b,g], Richard Szubin[a], Connor A. Olson[a], Bernhard O. Palsson[a,b,c,d], John M. Gladden[b,e,f], Blake A. Simmons[b,g], Aindrila Mukhopadhyay[b,g,h], Steven W. Singer[b,g], Adam M. Feist[a,b,c,*]

[a]Department of Bioengineering, University of California San Diego, 9500 Gilman Dr., La Jolla, CA 92093, USA

[b]Joint BioEnergy Institute, 5885 Hollis street, 4th floor, Emeryville, CA 94608, USA

[c]Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 2800 Kgs, Lyngby, Denmark

[d]Department of Pediatrics, University of California, San Diego, CA 92093, USA

[e]Department of Energy, Agile BioFoundry, Emeryville, CA 94608, USA.

[f]Department of Biomass Science and Conversion Technology, Sandia National Laboratories, 7011 East Ave, Livermore, CA 94550, USA

[g]Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94702, USA

[h]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94702, USA

*To whom correspondence should be addressed.

(Adam M. Feist)

afeist@ucsd.edu

Abstract

Although the use of ionic liquids (ILs) for the pretreatment of lignocellulosic biomass has been limited due to high costs, recent efforts to develop low-cost protic ILs show promise for achieving cost-effectiveness for biorefineries. However, an additional challenge remains in that ILs present in biomass hydrolysates are toxic to most microbial hosts, resulting in poor growth phenotypes. To address this issue, we applied an adaptive laboratory evolution (ALE) approach for tolerizing *Pseudomonas putida* KT2440, an industrially relevant bacterial host, to two low-cost ILs (triethanolammonium acetate [TEOH][OAc] and triethylammonium hydrogen sulfate [TEA][HS]). After continuous cultivations with gradually increased IL levels, we obtained evolved strains showing significant improvements in their growth performance under high concentrations of the ILs (maximum 4% [TEOH][OAc] and 8% [TEA][HS], in w/v) at which the wild-type strain cannot grow. Sequencing of evolved strains revealed multiple regions where mutations were associated with improved performance in minimal media conditions (*relA*, *gacS*, *oprB*/PP_1446, *fleQ*, *tktA*, and *uvrY*/PP_4100) and in IL-specific conditions (PP_5350, PP_4929/*emrE*, *oprD*, and PP_5324). We further validated the causality of the PP_5350 and *emrE* genes for improved IL tolerance via reverse engineering and transcriptomic analysis. A common mutation in the PP_5350 gene, encoding a RpiR family transcriptional regulator, was shown to significantly upregulate the glyoxylate cycle for efficient acetate catabolism. In addition, it was suggested that the *emrE* gene encodes an efflux pump which can export [TEA][HS]. Finally, the cultivation of two of the best performing evolved strains with IL-treated biomass hydrolysates demonstrated their considerable potential to be used as platform strains. Taken as a whole, this work provides strains for utilization of IL-treated biomass and a mechanistic understanding that could be further leveraged to develop efficient microbial bioprocesses.

Keywords

Pseudomonas putida, ionic liquid, tolerization, biorefinery, adaptive laboratory evolution, systems biology

Introduction

Lignocellulosic biomass, abundant in nature, is a promising feedstock for the microbial production of various biochemicals as it can be depolymerized into diverse carbon sources including hexoses (e.g., glucose, galactose), pentoses (e.g., xylose, arabinose), and aromatic compounds (e.g., coumaric acid, ferulic acid).[1] In this regard, during recent decades, microbial conversion of the lignocellulosic-sugars has been studied extensively using various types of microbial hosts for its efficient utilization. *Pseudomonas putida* KT2440 (KT2440) has garnered attention due to its high inherent tolerance and capability to utilize the aromatic compounds[2–6] which constitute 5-30% of total carbon in lignocellulosic biomass.[7]

One of the major difficulties in the bioprocessing of lignocellulosic biomass is due to its recalcitrance and irregular bond structure which limits the overall activity of depolymerizing enzymes. Previously, diverse pretreatment approaches (e.g., acid hydrolysis, steam explosion, ammonia fiber expansion, etc.) have been implemented;[8] however, disadvantages such as the formation of inhibitory sugar degradation products and a high-energy requirement limit their widespread adoption.[9] An alternative approach using an ionic liquid (IL)-based pretreatment has been reported.[10–13] ILs are mixtures of cation and anion salts that exist in liquid form at or near room temperature due to their low melting point. Because they can dissolve many organic chemicals, the robust structure of lignocellulosic biomass can be weakened, allowing efficient enzymatic hydrolysis even without acid pretreatment or a high-temperature (>180 ^{o}C) process.[14] Numerous studies, such as IL screening and condition optimization have been performed to realize the IL-based pretreatment of lignocellulosic biomass at an industrial scale.[8,15–19]

Even with great potential, the major obstacle in the use of ILs is their high cost.[8] While imidazolium-based ILs (e.g., 1-ethyl-3-methylimidazolium acetate, 1-butyl-3-

methylimidazolium chloride) have been extensively studied given their ability to obtain high sugar yields,[20] the cost was estimated at $20 - $100 per kg.[21,22] This high cost significantly hinders the commercial viability of IL-based processes. To address the issue, a few low-cost protic ILs have been reported. [21,23–26] These protic ILs can be prepared by simple mixing of low-cost bases (e.g., alkanolamines, alkylamines) and acids (e.g., acetic acid, sulfuric acid, phosphoric acid). It was demonstrated that a comparably high glucose yield and an even better thermal stability can be achieved.[8,21,23] Because their prices are estimated at $0.7 - $1.4 per kg,[23,25,27] significantly lower than imidazolium-based IL costs, the use of low-cost protic ILs suggests an economically feasible route towards the successful development of IL-based bioprocesses.

Having addressed economic viability, additional challenges remain because ILs in the hydrolysates can severely reduce the activity of host microorganisms.[16,28,29] This intolerance leads to costly purification processes (i.e., thorough washing of biomass with high water usage) which increase the overall cost.[19,30] While previous attempts have been made to overcome the IL-induced inhibition by isolating naturally IL-tolerant strains,[31,32] the use of a new species as production hosts require additional efforts to test their compatibilities (e.g., growth on sugars, genetic tractability, biosafety, etc.). Alternatively, the approach of Adaptive Laboratory Evolution (ALE), an enrichment process of strains with improved fitness under given selection pressures,[33,34] has been shown to have great potential in the generation of platform strains. Indeed, many important phenotypes for biomass processing such as improved substrate utilization or tolerance to toxic products have been obtained,[34–40] and improved biochemical production was achieved by using evolved strains.[39] Furthermore, related mechanisms have been suggested by subsequent sequencing analysis, enabling a better understanding of host

microorganisms[41,42]. Notably, there were successful ALE studies improving tolerance of *Escherichia coli* and *Saccharomyces cerevisiae* to imidazolium-based ILs.[28,43] Nevertheless, since the effect of ILs vary greatly depending on the type of IL and microorganisms,[44] a specific ALE study for tolerizing KT2440 to low-cost ILs was warranted.

In this study, an ALE approach was applied for tolerizing the wild-type KT2440 strain to low-cost ILs. We chose triethanolammonium acetate ([TEOH][OAc]) and triethylammonium hydrogen sulfate ([TEA)[HS]) as target ILs, given that their costs are expected to remain low and high glucose yields have been achieved[21,23,25]. For the tolerization, the cells were serially passed while maintaining the exponential phase with increased concentrations of each IL or their 1:1 mixture (by mass). After the tolerization ALE (i.e., TALE), we observed that isolates from evolved populations were able to grow at high concentrations of the ILs. Genomic sequencing revealed multiple key mutations were present in the adapted strains, and possible tolerance mechanisms were described based on information found in the literature and experimental validation. Finally, the cultivation of the evolved strains in IL-pretreated biomass hydrolysates confirmed that they would be potential platform hosts in an IL-based lignocellulosic biorefinery.

Experimental section

Bacterial strains, plasmids, oligonucleotides, biomass, and general reagents

Bacterial strains and plasmids used in this study are listed in Table S1. The wild-type KT2440 strain (ATCC 47054) and *E. coli* S-17 (ATCC 47055) stored at Joint BioEnergy Institute (Emeryville, CA, USA) were used. Oligonucleotides are listed in Table S2 and synthesized by Integrated DNA Technologies (Coralville, IA, USA). For routine PCR reactions, HotStarTaq DNA Polymerase (Qiagen, Hilden, Germany) was used. When fidelity is critical, Q5 polymerase (New

England Biolabs, Ipswich, MA, USA) was used. For plasmid cloning, NEBuilder® HiFi DNA Assembly Master Mix was used. All chemicals were purchased from Sigma Aldrich (St. Louis, MO, USA) unless mentioned otherwise. For the preparation of all reagents and media, purified water from a Milli-Q system (MilliporeSigma, Burlington, MA, USA) was utilized. The biomass utilized was Sorghum (*Sorghum bicolor*), which was a kind gift from Idaho National Labs (Idaho Falls, Idaho, USA). Commercial enzyme cocktails Cellic® CTec3 and HTec3 were generously provided by Novozymes (Davis, CA, USA).

Synthesis of the ionic liquids

Both [TEOH][OAc] and [TEA][HS] ILs were synthesized according to literature procedures.[26] In summary, alkyl amines and the corresponding acids were mixed in a stoichiometric molar ratio of 1:1. Before mixing, the amines were diluted with distilled water to approximately 75% v/v. The acids were then added slowly into the cognate solutions under continuous stirring at 0 ºC. Finally, distilled water was added to the mixture to adjust the concentration of the IL stock solutions to 50% w/v.

Routine cell culture methods

All microbial cultures were conducted at 30 ^{o}C using either the LB medium (10 g/L tryptone, 5 g/L yeast extract, and 10 g/L NaCl) or the modified minimal M9 medium (4 g/L glucose, 2 g/L $(NH_4)_2SO_4$, 6.8 g/L Na_2HPO_4, 3 g/L KH_2PO_4, 0.5 g/L NaCl, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 500 µL/L 2000X trace element solution). The composition of the trace element solution[28] is 4.5 g/L $ZnSO_4{\cdot}7H_2O$, 0.7 g/L $MnCl_2{\cdot}4H_2O$, 0.3 g/L $CoCl_2{\cdot}6H_2O$, 0.2 g/L $CuSO_4{\cdot}2H_2O$, 0.4 g/L $Na_2MoO_4{\cdot}2H_2O$, 4.5 g/L $CaCl_2{\cdot}2H_2O$, 3.0 g/L $FeSO_4{\cdot}7H_2O$, 1.0 g/L

H_3BO_3, 0.1 g/L KI, 15 g/L disodium ethylenediaminetetraacetate. The IL stock solutions were additionally added for the preparation of the IL-containing minimal media. When [TEA][HS] was added, the media were neutralized (pH 7) by the addition of 10 M NaOH solution.

For typical shake-flask cultures, a single colony was inoculated from an LB agar plate into 5 mL of the minimal medium contained in a 30-mL cylindrical flask. After overnight culture, the seed culture was refreshed by diluting into 15-mL of the fresh medium at 0.05-0.1 of optical density at 600 nm (OD_{600}). When OD_{600} reached 0.8-1, the culture was transferred to 15 mL of the same fresh medium at OD_{600} 0.05 to inoculate the main culture. For isolates from TALE5-8 and TALE13-16 showing a growth defect when they grew without acetate anions, 1 g/L acetate (pH 7) was supplemented to the medium. The medium was continuously stirred with a magnetic stir bar at 1,100 rpm and 30 ℃. OD_{600} was measured by using a Biomate 3S bench-top spectrophotometer from Thermo Fisher Scientific (Waltham, MA, USA). Cultures were conducted in triplicates otherwise noted.

When the cells were cultured in various conditions, BioscreenC from Growth Curves USA (Piscataway, NJ, USA) was used. The refreshed seed cultures were prepared identically with the flask-scale culture. However, main cultures were conducted in 200 μL of the minimal medium at the initial OD_{600} 0.1. The cultures were continuously shaken at a high level. The temperature of a plate lid was maintained at 31 ^{o}C to avoid evaporation of the medium. Cell growth was monitored by measuring OD_{600} with a 1-h interval.

Adaptive evolution of KT2440 using an automated ALE platform

The ALE experiments were carried out in the flask-scale cultures using the automated ALE platform.[28,33] Briefly, multiple colonies of KT2440 were independently inoculated into

initial ALE flasks with 15 mL of the minimal medium. After their overnight cultures, 750 μL of the grown cultures were transferred into the fresh medium containing initial concentrations of the ILs (Table S3). The OD_{600} values of each flask were periodically measured using a microtiter plate reader (Tecan, Männedorf, Switzerland) to monitor cell growth. The cultures were passed to next flasks when the OD_{600} value reached 0.3-0.4 (equivalent to OD_{600} 0.9-1.2 measured by the bench-top spectrophotometer using a 1-cm length cuvette) to maintain exponential growth phase. Typically, 150 μL was used as a passage volume (a 1/100 dilution). The concentration of each IL was increased stepwise by 20% of the initial concentration until the concentration reached 2.0% w/v. Then, their concentrations were increased by 0.4% w/v to reach a high concentration more rapidly. The frozen stocks of cultures were prepared regularly and kept at -80 ^{o}C until their analysis.

Genome and transcriptome analyses

The initial, intermediate, and end-point strains or populations were sequenced to investigate genomic variations. Specifically, end-point isolates of the TALE experiments were chosen by screening best colonies with the final concentrations of each IL. The intermediate isolates from the TALE experiments and all isolates from the control experiments were randomly chosen from streaked colonies on the M9 agar plates without any ILs.

For their resequencing, genomic DNA was extracted by using a Quick-DNA Fungal/Bacterial Miniprep Kit purchased from Zymo Research (Irvine, CA, USA). Then, pair-end resequencing libraries were prepared by using a Nextera XT kit from Illumina (San Diego, CA, USA). Raw sequencing data was processed by using a bioinformatics pipeline[45] which utilizes Breseq[46] (version 0.33.1) and Bowtie2[47] (version 2.3.4.1). A new reference genome was

created using gdtools of Breseq with applying the genome differences (Supplementary Data 1) between the newly sequenced genome and the reference genome of KT2440 (NCBI reference sequence: NC_002947.4). The mutations of evolved strains were investigated by uploading processed data on ALEdb[41] (version 1.0) based on the new reference genome. For populations samples, mutations with frequencies less than 0.5 were filtered out. Sequencing coverages of end-point isolates were more than 150X while coverages of other isolate and population samples were between 13-317X. The sequencing reads are available at the NCBI SRA database (BioProject number: PRJNA625743).

For transcriptomic analysis, cell cultures at the exponential phase (OD_{600} 0.4-0.6) were mixed with two volumes of RNA protect Bacteria Reagent (Qiagen, Hilden, Germany) for stabilization and their total RNA was extracted by using Quick-RNA Fungal/Bacterial Miniprep Kit (Zymo Research). Residual genomic DNA was removed by incubating the samples with DNase I (RNase-free, NEB). Ribosomal RNA was removed by hybridizing with 30-bp oligonucleotides designed to bind the ribosomal RNA of KT2440 specifically and treating with Thermostable RNase H purchased from Lucigen (Middleton, WI, USA). The remaining oligonucleotides were removed by incubating the resulting samples with DNase I (NEB). The samples were purified by using an RNA Clean & Concentrator Kit (Zymo Research). Paired-end RNA-Seq libraries were constructed using a KAPA RNA HyperPrep kit from Kapa Biosystems (Wilmington, MA, USA). Raw RNA-Seq reads were mapped to the original reference genome using Bowtie2[47] and transcript abundance was analyzed using summarizeOverlaps.[48] Calculation of transcripts per million (TPM) and prediction of differentially expressed genes were conducted using DESeq2.[49] RNA-Seq data files were deposited at Gene Expression Omnibus (accession number: GSE149827).

All libraries were sequenced on Illumina HiSeq or NextSeq platform at IGM Genomics Center (San Diego, CA, USA) and sample concentrations and qualities were checked by using Qubit (Thermo Fisher Scientific), 2100 Bioanalyzer (Agilent, Santa Clara, CA, USA), and 4200 Tapestation (Agilent).

Genome engineering of *Pseudomonas* strains

Pseudomonas strains were engineered for allelic exchange following the conjugation of plasmid DNA using selection with gentamicin followed by counterselection against the presence of heterologous expression of *sacB*.[50,51] For homologous recombination, 1-kb flanking sequences next to a target gene were chosen. Conjugation plasmids (pTE275_PP_5350' and pTE275_Δ*emrE*, Table S1) were prepared by using the Gibson assembly method. Briefly, to prepare pTE275_PP_5350' plasmid, the PP_5350 region of a mutated end-point isolate was amplified using the PP_5350_F and PP_5350_B primers. Then, the amplified DNA fragment was assembled with a vector fragment amplified with the PP_5350_V_F and PP_5350_V_B primers, and the pEX18GM plasmid (Table S1) which has essential elements for bacterial conjugation. To prepare the pTE275_Δ*emrE* plasmid, the upper and lower regions of *emrE* in the wild-type KT2440 strain were individually amplified using EmrE_F1, EmrE_B1, EmrE_F2, EmrE_B2 primer pairs. Then, the two fragments and a vector fragment amplified with the EmrE_V_F and EmrE_V_B primers and the pEX18GM plasmid as a template were assembled. All assembled mixtures were directly introduced into the *E. coli* S-17 strain using the conventional thermal transformation method and correct colonies were screened by colony PCR followed by sequencing of constructed plasmids.

For actual conjugation, *Pseudomonas* strains and *E. coli* S-17 strains (S17_PP_5350' and S17_Δ*emrE*, Table S1) with cognate plasmids were separately grown in 5-mL LB media. To maintain the plasmids in *E. coli*, gentamicin was included. When the cultures reached the exponential phase, they were mixed with a 1:5 ratio of *P. putida* to *E. coli*. A small volume of the culture equivalent to approximately 1.5 X 10^8 cells (OD_{600} ~0.1) was spotted on an LB agar plate. After 24 hours, the cell mass was resuspended in a 50 μL LB medium and streaked on a fresh LB agar plate containing gentamicin and chloramphenicol to isolate conjugated *Pseudomonas* strains only. The successful conjugation was confirmed by doing another colony PCR with the SacB_end_F primer that binds the *sacB* gene and either PP_5350_check_F or EmrE_check_F primer that bind the upper region of the target locus. The confirmed colony was inoculated in a 5-mL LB medium and grew until OD_{600} reached 0.6. Then, the culture was serially diluted and spread on LB agar plates with 250 g/L sucrose. A colony with a desired mutation was screened by doing another colony PCR with two outside primers (a pair of PP_5350_check_F and PP_5350_check_B or a pair of EmrE_check_F and EmrE_check_B). The concentrations of gentamicin and chloramphenicol were 30 μg/mL and 25 μg/mL, respectively.

Preparation of IL-treated biomass hydrolysates

The biomass sorghum was dried for 24 h in an oven at 40 ºC and knife-milled with a 2 mm screen (Thomas-Wiley Model 4, Swedesboro, NJ, USA). Subsequently, it was treated with each IL (20% (w/w) initial pretreatment slurry) in a 1 L Parr benchtop reactor from Parr Instrument Company (Moline, IL, USA). All the pretreatments were carried out at 140 °C for 3 h. After the pretreatment, the biomass-IL mixtures were diluted to achieve 4% and 5% (w/v) of the [TEOH][OAc] and [TEA][HS] by the addition of distilled water, and subsequently, the pH

was adjusted to ~ 5 by the addition of a small amount of either H_2SO_4 or NaOH. The pretreated biomass mixtures were hydrolyzed using a 9:1 (v/v) mixture of Cellic Ctec3 and Htec3 enzymes at an enzyme loading of 20 mg of enzyme per 1 g of starting biomass. Enzymatic hydrolysis experiments were conducted at 50 °C for 72 hours with constant agitation on a shaker. Next, the pH of the hydrolysates was adjusted to 7 by the addition of 10 M NaOH and concentrated medium salt solutions were added. The [TEA][HS]-treated hydrolysate was also further diluted 4 times and glucose was supplemented due to its unexpected toxicity. The hydrolysates were sterilized by filtration using 0.4-μm filters before cultivation and stored at 4 °C for further experiments.

Metabolite quantification

Concentrations of sugars, byproducts, and acetate were analyzed by using the 1260 Infinity II LC system (Agilent, Santa Clara, CA, USA) with the HPX-87H column (Biorad, Hercules, CA, USA) and 5 mM H_2SO_4 as a mobile phase. The flow rate of the mobile phase and the column temperature were maintained at 0.5 mL/min and 45 ℃, respectively. For quantifications, refractive index signals were monitored.

Results and discussion

Characterizing IL toxicity to wild-type *P. putida* KT2440 and the stress response

The toxicity of the chosen ILs was initially examined for KT2440 by monitoring cell growth (200 μL as a culture volume) with various IL concentrations (Fig. S1a and S1b). Similar to previous observations[28] with *E. coli* growing with imidazolium ILs, the growth rate of KT2440 was decreased as the concentration of the ILs were increased. At a concentration of

0.5% w/v of [TEOH][OAc], growth inhibition was observed, and growth was completely abolished at a concentration of 2% w/v for the wild-type strain. With [TEA][HS], although the extent of the inhibition was less than [TEOH][OAc], cellular growth was decreased at 2% w/v and no growth was observed with more than 5 w/v % [TEA][HS] during a 24 h observation time. Given that higher tolerance to IL is preferred to avoid the additional washing step and that other microorganisms have been evolved to grow above 5% w/v imidazolium IL,[28] the results indicated a necessity to improve the tolerance of KT2440.

To investigate IL-induced stress in wild-type KT2440, transcriptomic changes upon the addition of the growth-inhibiting concentration of each IL (0.5% [TEOH][OAc] or 2% [TEA][HS]) were analyzed. Differentially expressed genes (DEGs) from IL treatment were determined using the ratio of total normalized reads from the treated sample over the untreated control. This analysis revealed that the IL additions significantly affected the global gene expression profile; the expression profile of 632 genes (11.4% of the total genes, fold change > 2) were changed by either single IL out of which 59 genes were common to both (Fig. S2a and Supplementary Data 3). 19 and 14 genes showed large fold changes (> 10) by each IL respectively while there were no common genes (Fig. S2b and Table S4). These 632 genes were categorized into clusters of orthologous groups (COGs) based on their functional annotations (Fig. S2c and S2d). This analysis showed that genes involved in inorganic ion transport and metabolism [P], amino acid metabolism and transport [E], transcription [K], and general function prediction only [R] were mainly affected. Specifically, when [TEOH][OAc] was added, genes related to arginine/glutamate metabolism (*aruC*, PP_0269) and fusaric acid resistance (PP_1263, PP_1264, PP_1266) were significantly upregulated while genes related to acetate metabolism (*actP*-I) and benzoate metabolism (*benE*-I, *benD*, *benC,* PP_2036, PP_2037) were down-

regulated. On the other hand, when [TEA][HS] was added, large fold changes (>10) were only observed with upregulated genes: genes related to lipid metabolism (PP_3726, PP_3725, PP_3732), drug resistance (*mexC*, *emrE*, *mexD*), and stress response (*csbD*). However, a further detailed analysis was limited since ~40% of DEGs were for genes with limited conservation to known proteins and did not significantly cluster with characterized genes. Given the limitations of this RNA-Seq analysis of the ILs on gene expression, we turned to a TALE approach to efficiently generate sets of IL-tolerant strains based on the wild-type KT2440 starting strain. Furthermore, it was believed that this approach could provide an understanding of IL tolerance mechanisms and determine whether the identified DEGs were important.

ALE for tolerization of KT2440 to the ILs

To improve tolerance to the ILs, wild-type KT2440 clonal isolates were subjected to the TALE approach (Table S3) using an automated liquid handling platform. As a control experiment, KT2440 was cultivated in identical conditions without any ILs (lineages labeled as ALE1-4). For the TALE experiments, KT2440 was cultivated with continuous exposure to either [TEOH][OAc] (lineages labeled as TALE5-8) or [TEA][HS] (lineages labeled as TALE9-12). In addition, wild-type clonal isolates were also tolerized to a 1:1 mixture of the ILs (lineages labeled as TALE13-16). Thus, there were 12 independent TALE (Fig. 1) and 4 independent constant condition ALE experiments performed (Fig. S3). During the TALE experiments, the concentrations of ILs were gradually increased (20% of the starting concentrations for each step and a value of 0.4% w/v once the concentration reached 2% w/v, see Methods for details) from sub-lethal starting concentrations (0.5% [TEOH][OAc], 2% [TEA][HS], and 1% mixed ILs, in

w/v). Starting concentrations showed approximately 60% inhibition of growth rate for wild-type cells.

Notably, during the control experiments, the growth rates of evolved populations gradually increased from approximately 0.58 h^{-1} to 0.79 h^{-1} (on average, with a maximum of 0.87 h^{-1}) and were maintained for numerous generations at this apparent maximum growth rate, for the duration of the experiment (Fig. S3 and Table S3). In the cases of the TALE experiments (TALE5-16), the growth rates were generally decreased as the concentrations of ILs were increased. Specifically, the growth rates of the populations in TALE5-8 and TALE13-16 (tolerization to [TEOH][OAc] and the IL mixture) maintained their approximate initial growth rates (0.3 - 0.4 h^{-1}) until the concentrations reached 2% w/v. However, the population growth rates gradually began decreasing to below to 0.13 h^{-1} after the change in the step sizes to 0.4% w/v. In the case of the TALE experiments with an initial concentration of 2% (w/v) [TEA][HS] (TALE9-12), the growth rates consistently decreased from the start of the experiment, probably due to the larger step sizes from the beginning, and the final population growth rates were 0.17 h^{-1}, on average. Despite the reduced growth rates, the final concentrations of ILs reached 4.3% w/v for [TEOH][OAc], 7.3% w/v for [TEA][HS], and 4.9% w/v for the IL mixture – concentrations at which wild-type KT2440 cells cannot grow (Fig. 2a). The populations underwent from 335 to 609 generations and 1.64 - 3.10 x 10^{12} cumulative cell divisions (CCD)[52] depending on the evolution conditions (Table S3).

Physiological characterization of end-point strains from each condition

To confirm improved tolerance phenotypes, clonal isolates from the last flask of the independent evolution experiments were isolated (Table S1) and cultivated at a small-scale with

the absence and presence of the ILs (at 'low' and 'high' concentrations). The potential for cross-tolerance was also investigated by testing the isolates from all evolution experiments with both targeted ILs.

Evolved clones were screened under various concentrations of [TEOH][OAc] to understand the extent of their tolerance phenotypes. For [TEOH][OAc], 2% w/v and 4% w/v were chosen for screening at low and high concentration, respectively (Fig. 2a and Fig. S4a-e). Surprisingly, the cultivation revealed that all evolved strains, even from the control ALE experiments, showed improved tolerance to [TEOH][OAc]. Nevertheless, as expected, isolates from TALE5-8, which had been evolved with [TEOH][OAc] only, showed higher growth rates than strains from the other lineages (Fig. 2a). Furthermore, the isolates from TALE5-8 showed higher growth rates in the presence of [TEOH][OAc] than in an 'unstressed' condition (i.e., without the IL). Although it was less apparent, the final flask (i.e., end-point strain) TA_15_F84_I1 tolerized to the mixed IL also showed a similar improvement in growth rate (Fig. S4e), implying that the growth performance dependency is commonly present in the strains continuously exposed to [TEOH][OAc] either by itself or in a mixture. In case of the other strains, the addition of [TEOH][OAc] generally inhibited their growth rates.

A set of endpoint clones were also screened with two concentrations of [TEA][HS] to characterize their resistance phenotypes. In this case, 4% w/v and 8% w/v were chosen as the low and high concentrations, respectively (Fig. 2a and Fig. S5a-e). Similar to the results from cultivation with [TEOH][OAc], final flask isolates from TALE9-12 (exposed to the maximum 7.6% w/v [TEA][HS]) showed the highest growth rates among isolates from the other lineages in the presence of 8% w/v [TEA][HS]. In addition, most strains from the control ALE experiment and the experiment with the IL mixture also showed the capability to grow at this concentration.

However, in general, the strains evolved with [TEOH][OAc] showed reduced growth rates, even less than the wild-type strain (Fig. 2a), suggesting that the improved tolerance to [TEOH][OAc] of the TALE5-8 strains is specific to [TEOH][OAc] only. Additionally, all endpoint strains grew at the highest growth rates when ILs were absent in the medium, indicating the evolution approach did not induce any dependency on [TEA][HS].

The improved tolerance of the evolved clones was validated by the cultivation of the best performing end-point strains (i.e., strains with the highest growth rates with 4% w/v [TEOH][OAc] or 8% w/v [TEA][HS] IL at the small-scale cultivation) from each condition at a larger scale similar to the evolution conditions (Fig. 2b and 2c). Nearly all selected evolved strains consistently grew at the larger scale as well, whereas the wild-type KT2440 strain did not show any detectable growth (an exception was the TA16_F86_I1 strain which failed to grow with 8% [TEA][HS]). Specifically, the TA8_F83_I1 and the TA9_F60_I1 strains displayed the highest growth rates of 0.15 h^{-1} and 0.17 h^{-1} with [TEOH][OAc] and [TEA][HS], respectively. Additionally, the cultivation of the control ALE_1-4 strains at this larger format revealed an increase in growth rates to a maximum of 0.77 h^{-1} (Fig. S6), similar to those observed in their corresponding end-point populations (Table S3). This result contrasts with observed changes in growth rates for the same strains in the small-scale format. The lower growth rates during the small-scale cultivation might be due to decreased aeration; this observation suggests that benefits from acquired mutations are often affected by a given test condition.

Collectively, these results indicate that the ALE and TALE approaches were successful in adapting and tolerizing KT2440 to obtain improved fitness values within the given IL stress environments where the wild-type strain performs poorly. However, their different responses to the ILs suggested that the improved tolerance was achieved via multiple mechanisms depending

on the evolution conditions. Therefore, we proceeded to further investigate possible tolerance mutational mechanisms by sequencing obtained evolved strains.

Whole genome sequencing reveals 'key-3' genetic elements related to IL tolerance

To understand the mutational mechanisms enabling the acquired IL tolerance phenotypes, genome sequences of end-point strains and populations, as well as some intermediate strains, were analyzed (for TALE13-16, only end-point strains and populations were resequenced, Supplementary Data 2). Notably, sequence analysis revealed a number of mutations observed across multiple independent experiments (Fig. 3 and Table 1). When genomic loci were identified as mutated in three or more independent lineages and at least one end-point isolate, we defined these genetic features as 'key-3' regions. Furthermore, the 'key-3' regions were classified into two categories: key regions for culturing condition-specific adaptation, if observed in the control ALE experiments or other previous studies not related to ILs, and key regions for IL-specific adaptation if mutations were exclusively observed in the TALE experiments conducted here.

There were a number of culturing condition-specific 'key-3' regions identified in both the control and TALE experiments which were correlated with adaptation to the specific culturing conditions (e.g., medium). A total of six 'key-3' regions containing eight genes were identified (Table 1 and Fig. S7a-f). Of these genes, there were four genes encoding transcription regulators (*gacS*, *fleQ*, *uvrY*, and PP_4100), two genes encoding metabolic enzymes (*relA*, *tktA*), two genes encoding membrane or membrane-related proteins (*oprB*-II and PP_1446). Although mutations in this set of genes are not directly related to IL tolerance, they are related to glucose metabolism, biofilm formation, or general stress responses in which they can be speculated to

improve the general fitness of KT2440 (Fig. S6), thereby increasing tolerance to the IL or potentially any given stressor (See Supplementary Text 1 for detailed speculations about their roles).

A total of four IL-specific 'key-3' regions containing five genes were identified from evolutions performed under the TALE conditions (Table 1 and Fig. S7g-j). There were two genes encoding transcription regulators (PP_5350 and PP_4929), two genes encoding membrane proteins (*emrE* and *oprD*), and one gene encoding a sensor protein (PP_5324). Mutations in both PP_5350 and PP_4929/*emrE* regions were identified in all first isolates from the single-IL TALE experiments, indicating those regions were mutated early in the evolution. On the other hand, mutations in *oprD* and PP_5324 occurred in the middle or late stage of the TALE experiments. PP_5350 encodes a RpiR (ribose-phosphate-isomerase regulator)-family transcriptional factor. A recent study reported that its homolog in *P. aeruginosa* represses the expression of *aceA* (encoding isocitrate lyase) and *glcB* (encoding malate synthase),[53] key genes in the glyoxylate cycle.[54] In addition, from a database,[55] it was found that this PP_5350 potentially interacts with genes responsible for glucose metabolism (e.g., *glk* encoding glucokinase, *zwf* encoding glucose 6-phosphate 1-dehydrogenase, *eda* encoding KHG/KDPG aldolase, *edd* encoding phosphogluconate dehydratase). It was surprising to find that an identical 9-bp deletion was identified at the C-terminal of the PP_5350 coding region in all TALE experiments (8/8, 100%) from the [TEOH][OAc] set (Table 1). The deletion probably occurred in the identical region due to its repeated sequence (Fig. 4a), similar to the 82-bp deletion in *E. coli*.[56] As these mutations were highly specific to [TEOH][OAc], it strongly suggested linkage between the in-frame deletion in PP_5350 and the improved IL tolerance. Thus, the mutation was subjected to further characterization.

For TALE experiments that used [TEA][HS], the PP_4929-*emrE* region (encoding a LysR-family transcriptional regulator and a small multidrug resistance protein, respectively) was frequently mutated (8/8, 100%) (Fig. S7h and Table 1). Specifically, the upregulation of *emrE* was observed in the wild-type KT2440 strain upon the addition of [TEA][HS] (Table S4), suggesting that KT2440 has an inherent mechanism to reduce stress caused by this IL. In this manner, from the literature, it was found that *E. coli* K-12 MG1655 also has a similar EmrE and can export imidazolium ILs.[57] However, we found that the EmrE gene of KT2440 shares only 47% identity with its *E. coli* homolog (Fig. S8) and one isolate (TA13_F83_I1) from an IL mix TALE experiment acquired a frame-shift mutation in its *emrE*, which likely disrupts its activity. To confirm the role of the PP_4929-*emrE* region in IL tolerance, we performed reverse engineering and omics analysis on a strain containing a mutation uncovered in this study. Additionally, potential roles of the other two genes, *oprD* and PP_5324, are discussed in Supplementary Text 2.

Validation of causality for the PP_5350 RpiR transcription factor and its role in acetate utilization

The relationship between mutations in PP_5350 and tolerance of [TEOH][OAc] was characterized via reverse engineering, growth phenotyping, and transcriptomic analysis. Given that growth inhibition in KT2440 by [TEOH][OAc] was mainly caused by acetate ions (Fig. S9), an available carbon source, we hypothesized that mutations in PP_5350 enabled efficient utilization of acetate ions thereby reducing the concentration below toxic levels. Indeed, the TA8_F83_I1 strain grew better when neutralized acetate salts were supplemented in the medium while triethanolammonium did not affect the growth (Fig. S9), supporting the hypothesis. To

understand the effect of the 9-bp mutation in PP_5350, it was introduced into the KT2440 starting strain, resulting in the KT2440_PP_5350' strain (Table S1). Then, the KT2440, KT2440_PP_5350', and TA8_F83_I1 strains were cultivated in the M9 medium containing 2 g/L glucose and 2 g/L acetate as carbon sources (Fig. 4b-e). As expected, the KT2440_PP_5350' and TA8_F83_I1 strains, which have the PP_5350 mutation, showed higher growth rates and acetate consumption compared to the KT2440 strain.

Transcriptomic analysis of strains harboring the reproducibly occurring mutation in PP_5350 revealed that the mutation significantly activated genes responsible for the glyoxylate cycle, consistent with the deletion study[53] with *P. aeruginosa* (Fig. 4f). Specifically, the expression levels of *aceA* were the most differentially expressed in both KT2440_PP_5350' and the TA8_F83_I1 strains (196-fold and 122-fold, respectively) as compared to the KT2440 strain. The expression level of another essential gene in the glyoxylate cycle, *glcB*, was also up-regulated (15-fold and 11-fold, respectively), consistent with our hypothesis that the mutation in PP_5350 enabled rapid acetate assimilation, and that this was enabled by the activation of the glyoxylate cycle. Additionally, downregulated expression of genes related to glucose assimilation (e.g., *glk*, *edd*, Supplementary Data 4) was also observed; this finding can be associated with the observed low growth rates of strains with the mutation while grown without acetate ions. Further, genes related to the TCA cycle (e.g., *gltA* encoding citrate synthase, *sucAB* encoding 2-oxoglutarate decarboxylase and 2-oxoglutarate dehydrogenase, *sdhA* encoding succinate dehydrogenase) were mostly up-regulated. Although *zwf* (PP_5351, encoding one of three glucose 6-phosphate 1-dehydrogenases in KT2440) is divergently located next to PP_5350, its expression was not differentially changed. These regulated gene expressions likely resulted in the preferential utilization of acetate ions over glucose by the KT2440_PP_5350' or

TA8_F83_I1 strain (Fig. 4c). Further study is required to investigate whether PP_5350 directly or indirectly regulates all the aforementioned genes. Nevertheless, the results strongly indicated that the KT2440 strain was evolved for the rapid utilization of acetate and this was enabled by the novel 9-bp mutation in PP_5350 which activates the glyoxylate cycle.

Validation of causality for the EmrE as a potential exporter of IL and its regulation by PP_4929

The role of EmrE in IL tolerance was validated by the growth phenotyping of an EmrE-lacking strain. This validation was accomplished by deleting *emrE* in the TALE-derived TA9_F60_I1 strain. When the resulting strain (TA9_F60_I1_Δ*emrE*, Table S1) was cultivated with 8% [TEA][HS] (Fig. 5a), it showed a severe decrease in both growth and final OD. This finding confirmed that *emrE* was indeed required for the [TEA][HS] tolerance and EmrE likely confers as a functional efflux pump which exports the tested (and potentially additional) IL out of the cytoplasm.

We further investigated the relationship between PP_4929 and *emrE* by transcriptomic analysis. Specifically, transcript level changes of PP_4929 and *emrE* in the KT2440 strain and the TA9_F60_I1 strain were analyzed. The TA9_F60_I1 contains a premature stop codon mutation (E160*) in PP_4929 (Supplementary Data 2). It was observed that the expression levels of both PP_4929 and *emrE* genes were increased (a 11.7-fold increase and a 6.67-fold increase, respectively) (Fig. 5b and Supplementary Data 4). This observation was consistent with the annotation of PP_4929, a LysR-type transcriptional regulator. LysR-family transcription regulators have been characterized to regulate the expression of divergently located genes as well as their own expression.[58] These findings imply that PP_4929 is similarly a repressor for *emrE*

expression and represses its own expression. It is still not clear why the TA13_F83_I1 strain, evolved with the mixed IL, obtained a frame-shift mutation in *emrE* despite its positive role in IL tolerance. One possibility is that a potential burden exists if the *emrE* is highly expressed upon the addition of [TEA][HS], given that the expression of *emrE* was up-regulated in the wild-type KT2440 when 2% w/v [TEA][HS] was added (Table S4). The burden associated with membrane protein expression has been previously studied[59,60] and growth inhibitions were observed. The burden was possibly avoided from the inserted frameshift mutation, which resulted in a changed amino acid context. Nevertheless, the result obtained from the deletion of *emrE* in the TA9_F60_I1 strain suggests that EmrE is beneficial for improving tolerance to [TEA][HS] and its role is believed to facilitate the export of the IL from the cells.

Cultivation of evolved strains with IL-treated biomass hydrolysate

The applicability of the well-performing TALE-derived strains (A8_F83_I1 and A9_F60_I1 strains) was evaluated using biomass hydrolysates pretreated with either of the two targeted ILs (see Materials and Methods). When [TEOH][OAc] was utilized, sorghum was successfully pretreated and deconstructed, resulting in a hydrolysate containing 4% [TEOH][OAc] and a 72% and 71% yield of glucose and xylose, respectively (Table S5). In the case of [TEA][HS], a greater glucose yield was achieved (81%) whereas the xylose yield was low (16%). Two common sugar byproducts - furfural and 5-hydroxymethylfurfural (5-HMF) - were also quantified. Notably, furfural and 5-HMF were not detected (< 0.01 g/L) in the [TEOH][OAc]-treated hydrolysate. However, in the [TEA][HS]-treated hydrolysate, small amounts of furfural (1.75 g/L) and 5-HMF (0.35 g/L) were detected, likely due to a low pH condition caused by the formation of sulfuric acid in water. Unfortunately, the existence of the

sugar byproducts resulted in an additional toxicity,[61,62] requiring a further dilution of the [TEA][HS] hydrolysate to enable growth; the final concentration of [TEA][HS] was 1.5%. Nevertheless, after a 48-h cultivation in each hydrolysate, it was observed that both A8_F83_I1 and A9_F60_I1 displayed superior growth phenotypes (2.65-fold and 1.32-fold increases in the biomass formation) compared to the unevolved KT2440 strain (Fig. 6a and b). Although further optimization of the pretreatment process is required, these results indicate that the ALE technology successfully improved the fitness of KT2440 and generated platform strains relevant for lignocellulosic-biomass refining.

Conclusions

In this study, we demonstrated the tolerization of the industrially relevant microbe *P. putida* KT2440 to two toxic but low-cost ILs using an ALE approach to select for novel and beneficial mutations through natural selection. IL-tolerant strains were successfully obtained by systematically evolving independent lineages of the same starting strain under multiple conditions. Whole genome sequencing of evolved strains identified conserved genomic regions reproducibly mutated across multiple independent experiments. The recovered mutations from this process provided a basis for mechanistic insights on how IL tolerance can be generated in KT2440 under industrially relevant conditions. We provide rigorous experimental evidence to support genetic causality of two candidate regions with IL tolerance. Specifically, it was found that the 9-bp mutation in the PP_5350 gene activated the glyoxylate cycle for lowering acetate levels via assimilation and the *emrE* expression was upregulated for enhanced efflux of [TEA][HS]. The IL-tolerant strains generated from this study are thus beneficial towards

achieving cost-effectiveness of bioprocesses that convert IL-treated biomass by reducing costly purification steps and improving yields.

To realize full-scale lignocellulosic biomass processing with ILs, further development, as well as process evaluations (e.g., the mass balance of biomass), should be conducted. Because the composition and toxicity of biomass hydrolysates are greatly affected by factors such as biomass types, drying methods, pretreatment time, and temperature,[18,63–65] various parameter studies are needed to maximize the process efficiency. An important issue is to minimize the formation of the growth-inhibiting sugar byproducts when [TEA][HS] is used. This issue could be addressed by detailed optimization of pretreatment conditions including IL concentration, IL/biomass ratio, and temperature.[66] Additionally, while the burden of IL costs on overall process economics is lowered by using these low-cost ILs, its usage is still an expense that needs to be reduced. Therefore, IL-recycling before or after the microbial fermentation should be studied. In this regard, protic IL recycling was already successfully demonstrated in one recent study,[21] lending support to the industrial viability of these ILs.

This study is a clear example of how ALE can be efficiently utilized to generate strains with industrially relevant phenotypes by creating novel mutations in multiple genetic regions that are difficult to be rationally designed. Often it is believed that tolerance to toxic chemicals can be achieved by up-regulating the expression of genes encoding efflux pumps.[67,68] However, we lack a predictive understanding of which specific efflux pump will confer the strongest benefit for a specific stress or how to tune a given pump gene to enhance its activity. In this regard, the ALE approach, which does not rely on *a priori* assumptions as to what needs to mutate or change,[34] pinpoints the most relevant targets, not limited to efflux pumps, and suggests effective mutations for enhanced performance. Furthermore, mutations created by this ALE experimentation could

be useful for additional applications. For example, the strains with the identified 9-bp deletion in PP_5350 could show improved tolerance to other acetate-containing ILs, given that acetate is a common anion for ILs. In addition, the 9-bp deletion could be introduced into strains that are used in bioprocesses with feedstocks that contain acetate as one or a sole substrate.[54,69–71] Additionally, the *P. putida* strains obtained from the control experiment that possessed high growth rates in a minimal medium can be used in bioprocesses without ILs.

Finally, we believe that the ALE approach can be further applied to other less-characterized microorganisms with specialized metabolic lifestyles (e.g., syngas fermentation, plastic degradation) for the biorefinery.[72,73] ALE-based genetic perturbation followed by multi-omics analyses to understand beneficial mutations can create new knowledge and improve our understanding of novel host microorganisms. Such efforts will greatly strengthen our capability to develop microbial cell factories for accomplishing the ultimate goal of industrial biorefineries.

Conflicts of interests

The authors declare no competing interests.

Acknowledgments

This work conducted by the Joint BioEnergy Institute was supported by the Office of Science, Office of Biological and Environmental Research, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. We also appreciate Marc K. Abrams for editing the manuscript.

References

1. F. H. Isikgor and C. R. Becer, *Polym. Chem.*, 2015, **6**, 4497–4559.
2. J. I. Jiménez, B. Miñambres, J. L. García, and E. Díaz, *Environ. Microbiol.*, 2002, **4**, 824–841.
3. P. I. Nikel and V. de Lorenzo, *Metab. Eng.*, 2018, **50**, 142–155.
4. J. Nogales, J. Mueller, S. Gudmundsson, F. J. Canalejo, E. Duque, J. Monk, A. M. Feist, J. L. Ramos, W. Niu, and B. O. Palsson, *Environ. Microbiol.*, 2020, **22**, 255–269.
5. M. Park, Y. Chen, M. Thompson, V. T. Benites, B. Fong, C. J. Petzold, E. E. K. Baidoo, J. M. Gladden, P. D. Adams, J. D. Keasling, B. A. Simmons, and S. W. Singer, *ChemSusChem*, 2020, **13**, 1–14.
6. D. Salvachúa, A. Z. Werner, I. Pardo, M. Michalska, B. A. Black, B. S. Donohoe, S. J. Haugen, R. Katahira, S. Notonier, K. J. Ramirez, A. Amore, S. O. Purvine, E. M. Zink, P. E. Abraham, R. J. Giannone, S. Poudel, P. D. Laible, R. L. Hettich, and G. T. Beckham, *Proc. Natl. Acad. Sci. USA*, 2020, **117**, 9302–9310.
7. D. Kim, *Molecules*, 2018, **23**.
8. P. Halder, S. Kundu, S. Patel, A. Setiawan, R. Atkin, R. Parthasarthy, J. Paz-Ferreiro, A. Surapaneni, and K. Shah, *Renew. Sust. Energ. Rev.*, 2019, **105**, 268–292.
9. M. Mora-Pale, L. Meli, T. V. Doherty, R. J. Linhardt, and J. S. Dordick, *Biotechnol. Bioeng.*, 2011, **108**, 1229–1245.
10. G. Bokinsky, P. P. Peralta-Yahya, A. George, B. M. Holmes, E. J. Steen, J. Dietrich, T. S. Lee, D. Tullman-Ercek, C. A. Voigt, B. A. Simmons, and J. D. Keasling, *Proc. Natl. Acad. Sci. USA*, 2011, **108**, 19949–19954.

11. C. Li, B. Knierim, C. Manisseri, R. Arora, H. V. Scheller, M. Auer, K. P. Vogel, B. A. Simmons, and S. Singh, *Bioresour. Technol.*, 2010, **101**, 4900–4906.
12. A. M. Socha, R. Parthasarathi, J. Shi, S. Pattathil, D. Whyte, M. Bergeron, A. George, K. Tran, V. Stavila, S. Venkatachalam, M. G. Hahn, B. A. Simmons, and S. Singh, *Proc. Natl. Acad. Sci. USA*, 2014, **111**, E3587–95.
13. C. G. Yoo, Y. Pu, and A. J. Ragauskas, *Curr. Opin. Green Sustain. Chem.*, 2017, **5**, 5–11.
14. Z. Zhang, J. Song, and B. Han, *Chem. Rev.*, 2017, **117**, 6834–6880.
15. J. Amoah, N. Ishizue, M. Ishizaki, M. Yasuda, K. Takahashi, K. Ninomiya, R. Yamada, A. Kondo, and C. Ogino, *Bioresour. Technol.*, 2017, **245**, 1413–1420.
16. M. D. C. Portillo and A. Saadeddin, *Crit. Rev. Biotechnol.*, 2015, **35**, 294–301.
17. K. Ninomiya, A. R. I. Utami, Y. Tsuge, K. Kuroda, C. Ogino, T. Taima, J. Saito, M. Kimizu, and K. Takahashi, *Chem. Eng. J.*, 2018, **334**, 657–663.
18. K. M. Torr, K. T. Love, B. A. Simmons, and S. J. Hill, *Biotechnol. Bioeng.*, 2016, **113**, 540–549.
19. B. Neupane, N. V. S. N. M. Konda, S. Singh, B. A. Simmons, and C. D. Scown, *ACS Sustain. Chem. Eng.*, 2017, **5**, 10176–10185.
20. D. Fu, G. Mazza, and Y. Tamaki, *J. Agric. Food Chem.*, 2010, **58**, 2915–2922.
21. A. Brandt-Talbot, F. J. V. Gschwend, P. S. Fennell, T. M. Lammens, B. Tan, J. Weale, and J. P. Hallett, *Green Chem.*, 2017, **19**, 3078–3102.
22. F. Von Dissertation, Doctoral dissertation, RWTH Aachen University, 2014.
23. A. George, A. Brandt, K. Tran, S. M. S. N. S. Zahari, D. Klein-Marcuschamer, N. Sun, N. Sathitsuksanoh, J. Shi, V. Stavila, R. Parthasarathi, S. Singh, B. M. Holmes, T. Welton, B. A. Simmons, and J. P. Hallett, *Green Chem.*, 2015, **17**, 1728–1734.

24. F. J. V. Gschwend, A. Brandt-Talbot, C. L. Chambon, and J. P. Hallett, in *Ionic liquids: current state and future directions*, eds. M. B. Shiflett and A. M. Scurto, American Chemical Society, Washington, DC, 2017, vol. 1250, pp. 209–223.
25. J. Sun, N. V. S. N. M. Konda, R. Parthasarathi, T. Dutta, M. Valiev, F. Xu, B. A. Simmons, and S. Singh, *Green Chem.*, 2017, **19**, 3152–3163.
26. D. Simijonović, Z. D. Petrović, and V. P. Petrović, *J. Mol. Liq.*, 2013, **179**, 98–103.
27. L. Chen, M. Sharifzadeh, N. Mac Dowell, T. Welton, N. Shah, and J. P. Hallett, *Green Chem.*, 2014, **16**, 3098.
28. E. T. Mohamed, S. Wang, R. M. Lennen, M. J. Herrgård, B. A. Simmons, S. W. Singer, and A. M. Feist, *Microb. Cell Fact.*, 2017, **16**, 204.
29. J. I. Khudyakov, P. D'haeseleer, S. E. Borglin, K. M. Deangelis, H. Woo, E. A. Lindquist, T. C. Hazen, B. A. Simmons, and M. P. Thelen, *Proc. Natl. Acad. Sci. USA*, 2012, **109**, E2173–82.
30. N. M. Konda, J. Shi, S. Singh, H. W. Blanch, B. A. Simmons, and D. Klein-Marcuschamer, *Biotechnol Biofuels*, 2014, **7**, 86.
31. A. P. Reddy, C. W. Simmons, J. Claypool, L. Jabusch, H. Burd, M. Z. Hadi, B. A. Simmons, S. W. Singer, and J. S. VanderGheynst, *J. Appl. Microbiol.*, 2012, **113**, 1362–1370.
32. C. Yu, B. A. Simmons, S. W. Singer, M. P. Thelen, and J. S. VanderGheynst, *Appl. Microbiol. Biotechnol.*, 2016, **100**, 10237–10249.
33. R. A. LaCroix, B. O. Palsson, and A. M. Feist, *Appl. Environ. Microbiol.*, 2017, **83**.
34. T. E. Sandberg, M. J. Salazar, L. L. Weng, B. O. Palsson, and A. M. Feist, *Metab. Eng.*, 2019, **56**, 1–16.

35. M. Frederix, F. Mingardon, M. Hu, N. Sun, T. Pray, S. Singh, B. A. Simmons, J. D. Keasling, and A. Mukhopadhyay, *Green Chem.*, 2016, **18**, 4189–4197.
36. G. I. Guzmán, T. E. Sandberg, R. A. LaCroix, Á. Nyerges, H. Papp, M. de Raad, Z. A. King, Y. Hefner, T. R. Northen, R. A. Notebaart, C. Pál, B. O. Palsson, B. Papp, and A. M. Feist, *Mol. Syst. Biol.*, 2019, **15**, e8462.
37. R. M. Lennen, K. Jensen, E. T. Mohammed, S. Malla, R. A. Börner, K. Chekina, E. Özdemir, I. Bonde, A. Koza, J. Maury, L. E. Pedersen, L. Y. Schöning, N. Sonnenschein, B. O. Palsson, M. O. A. Sommer, A. M. Feist, A. T. Nielsen, and M. J. Herrgård, *BioRxiv*, 2019.
38. E. T. Mohamed, H. Mundhada, J. Landberg, I. Cann, R. I. Mackie, A. T. Nielsen, M. J. Herrgård, and A. M. Feist, *Microb. Cell Fact.*, 2019, **18**, 116.
39. T. P. Nguyen-Vo, Y. Liang, M. Sankaranarayanan, E. Seol, A. Y. Chun, S. Ashok, A. S. Chauhan, J. R. Kim, and S. Park, *Metab. Eng.*, 2019, **53**, 48–58.
40. T. Eng, P. Demling, R. A. Herbert, Y. Chen, V. Benites, J. Martin, A. Lipzen, E. E. K. Baidoo, L. M. Blank, C. J. Petzold, and A. Mukhopadhyay, *Microb. Cell Fact.*, 2018, **17**, 159.
41. P. V. Phaneuf, D. Gosting, B. O. Palsson, and A. M. Feist, *Nucleic Acids Res.*, 2019, **47**, D1164–D1171.
42. M. Dragosits and D. Mattanovich, *Microb. Cell Fact.*, 2013, **12**, 64.
43. K. B. Reed, J. M. Wagner, S. d' Oelsnitz, J. M. Wiggers, and H. S. Alper, *J. Ind. Microbiol. Biotechnol.*, 2019, **46**, 1715–1724.
44. I. R. Sitepu, L. L. Enriquez, V. Nguyen, C. Doyle, B. A. Simmons, S. W. Singer, R. Fry, C. W. Simmons, and K. Boundy-Mills, *Bioresource Technology Reports*, 2019, **7**, 100275.

45. S. Pabinger, A. Dander, M. Fischer, R. Snajder, M. Sperk, M. Efremova, B. Krabichler, M. R. Speicher, J. Zschocke, and Z. Trajanoski, *Brief. Bioinformatics*, 2014, **15**, 256–278.

46. D. E. Deatherage and J. E. Barrick, *Methods Mol. Biol.*, 2014, **1151**, 165–188.

47. B. Langmead and S. L. Salzberg, *Nat. Methods*, 2012, **9**, 357–359.

48. M. Lawrence, W. Huber, H. Pagès, P. Aboyoun, M. Carlson, R. Gentleman, M. T. Morgan, and V. J. Carey, *PLoS Comput. Biol.*, 2013, **9**, e1003118.

49. M. I. Love, W. Huber, and S. Anders, *Genome Biol.*, 2014, **15**, 550.

50. Y. Chen, J. Xia, Z. Su, G. Xu, M. Gomelsky, G. Qian, and F. Liu, *Appl. Environ. Microbiol.*, 2017, **83**.

51. T. T. Hoang, R. R. Karkhoff-Schweizer, A. J. Kutchma, and H. P. Schweizer, *Gene*, 1998, **212**, 77–86.

52. D.-H. Lee, A. M. Feist, C. L. Barrett, and B. Ø. Palsson, *PLoS One*, 2011, **6**, e26172.

53. S. K. Dolan, G. Pereira, R. Silva-Rocha, and M. Welch, *Microb. Biotechno.l*, 2020, **13**, 285–289.

54. H. G. Lim, J. H. Lee, M. H. Noh, and G. Y. Jung, *J. Agric. Food Chem.*, 2018, **66**, 3998–4006.

55. D. Szklarczyk, A. L. Gable, D. Lyon, A. Junge, S. Wyder, J. Huerta-Cepas, M. Simonovic, N. T. Doncheva, J. H. Morris, P. Bork, L. J. Jensen, and C. von Mering, *Nucleic Acids Res.*, 2019, **47**, D607–D613.

56. R. A. LaCroix, T. E. Sandberg, E. J. O'Brien, J. Utrilla, A. Ebrahim, G. I. Guzman, R. Szubin, B. O. Palsson, and A. M. Feist, *Appl. Environ. Microbiol.*, 2015, **81**, 17–30.

57. D. A. Higgins, J. M. Gladden, J. A. Kimbrel, B. A. Simmons, S. W. Singer, and M. P. Thelen, *J. Bacteriol.*, 2019, **201**.

58. S. E. Maddocks and P. C. F. Oyston, *Microbiology*, 2008, **154**, 3609–3623.

59. H. M. Jensen, T. Eng, V. Chubukov, R. A. Herbert, and A. Mukhopadhyay, *Sci. Rep.*, 2017, **7**, 13030.

60. S. Wagner, M. M. Klepsch, S. Schlegel, A. Appel, R. Draheim, M. Tarry, M. Högbom, K. J. van Wijk, D. J. Slotboom, J. O. Persson, and J.-W. de Gier, *Proc. Natl. Acad. Sci. USA*, 2008, **105**, 14371–14376.

61. J. R. M. Almeida, M. Bertilsson, M. F. Gorwa-Grauslund, S. Gorsich, and G. Lidén, *Appl. Microbiol. Biotechnol.*, 2009, **82**, 625–638.

62. E. Palmqvist and B. Hahn-Hägerdal, *Bioresour. Technol.*, 2000, **74**, 25–33.

63. A. Zoghlami and G. Paës, *Front. Chem.*, 2019, **7**, 874.

64. D. Essien and T. L. Richard, *Front. Bioeng. Biotechnol.*, 2018, **6**, 195.

65. M. Ambye-Jensen, K. S. Johansen, T. Didion, Z. Kádár, and A. S. Meyer, *Biotechnol Biofuels*, 2014, **7**, 95.

66. L. T. P. Trinh, Y.-J. Lee, J.-W. Lee, and W.-H. Lee, *Biotechnol. Bioprocess Eng.*, 2018, **23**, 1–10.

67. D. B. Kell, N. Swainston, P. Pir, and S. G. Oliver, *Trends Biotechnol.*, 2015, **33**, 237–246.

68. A. Mukhopadhyay, *Trends Microbiol.*, 2015, **23**, 498–508.

69. M. Jo, M. H. Noh, H. G. Lim, C. W. Kang, D.-K. Im, M.-K. Oh, and G. Y. Jung, *Microb. Cell Fact.*, 2019, **18**, 57.

70. M. H. Noh, H. G. Lim, S. H. Woo, J. Song, and G. Y. Jung, *Biotechnol. Bioeng.*, 2018, **115**, 729–738.

71. J. H. Lee, S. Cha, C. W. Kang, G. M. Lee, H. G. Lim, and G. Y. Jung, *Catalysts*, 2018, **8**, 525.

72. S. Yoshida, K. Hiraga, T. Takehana, I. Taniguchi, H. Yamaji, Y. Maeda, K. Toyohara, K. Miyamoto, Y. Kimura, and K. Oda, *Science*, 2016, **351**, 1196–1199.

73. I. S. Chang, B. H. Kim, R. W. Lovitt, and J. S. Bang, *Process Biochemistry*, 2001, **37**, 411–421.

74. D. Chatterji and A. K. Ojha, *Curr. Opin. Microbiol.*, 2001, **4**, 160–165.

75. J. M. Bergman, D. L. Hammarlöf, and D. Hughes, *PLoS One*, 2014, **9**, e90486.

76. O. N. Reva, C. Weinel, M. Weinel, K. Böhm, D. Stjepandic, J. D. Hoheisel, and B. Tümmler, *J. Bacteriol.*, 2006, **188**, 4079–4092.

77. G. J. Bentley, N. Narayanan, R. K. Jha, D. Salvachúa, J. R. Elmore, G. L. Peabody, B. A. Black, K. Ramirez, A. De Capite, W. E. Michener, A. Z. Werner, D. M. Klingeman, H. S. Schindel, R. Nelson, L. Foust, A. M. Guss, T. Dale, C. W. Johnson, and G. T. Beckham, *Metab. Eng.*, 2020, **59**, 64–75.

78. D. van den Broek, T. F. C. Chin-A-Woeng, G. V. Bloemberg, and B. J. J. Lugtenberg, *Microbiology*, 2005, **151**, 1403–1408.

79. E. Duque, J. de la Torre, P. Bernal, M. A. Molina-Henares, M. Alaminos, M. Espinosa-Urgel, A. Roca, M. Fernández, S. de Bentzmann, and J.-L. Ramos, *Environ. Microbiol.*, 2013, **15**, 36–48.

80. M. Martínez-Gil, M. I. Ramos-González, and M. Espinosa-Urgel, *J. Bacteriol.*, 2014, **196**, 1484–1495.

81. E. T. Mohamed, A. Z. Werner, D. Salvachúa, C. Singer, K. Szostkiewicz, M. Jiménez-Díaz, T. Eng, M. S. Radi, A. Mukhopadhyay, M. J. Herrgård, S. W. Singer, G. T. Beckham, and A. M. Feist, *Metab. Eng. Commun.*, submitted.

82. S. Chevalier, E. Bouffartigues, J. Bodilis, O. Maillot, O. Lesouhaitier, M. G. J. Feuilloley, N. Orange, A. Dufour, and P. Cornelis, *FEMS Microbiol. Rev.*, 2017, **41**, 698–722.

83. S. S. Fong, A. R. Joyce, and B. Ø. Palsson, *Genome Res.*, 2005, **15**, 1365–1372.

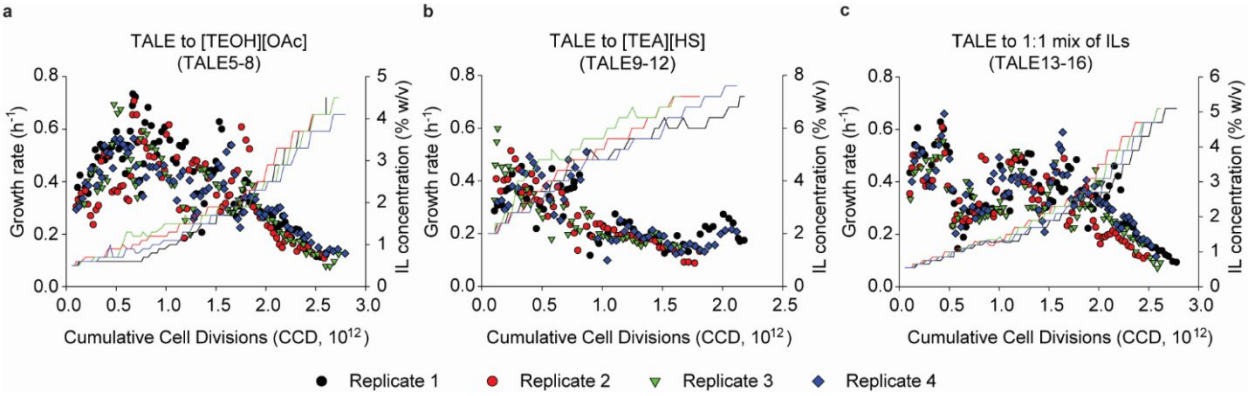
84. M. M. Ochs, M. P. McCusker, M. Bains, and R. E. Hancock, *Antimicrob. Agents Chemother.*, 1999, **43**, 1085–1090.

85. S. Follonier, I. F. Escapa, P. M. Fonseca, B. Henes, S. Panke, M. Zinn, and M. A. Prieto, *Microb. Cell Fact.*, 2013, **12**, 30.

86. J. J. Rodríguez-Herva, E. Duque, M. A. Molina-Henares, G. Navarro-Avilés, P. Van Dillewijn, J. De La Torre, A. J. Molina-Henares, A. S. La Campa, F. A. Ran, A. Segura, V. Shingler, and J.-L. Ramos, *Environ. Microbiol. Rep.*, 2010, **2**, 373–380.

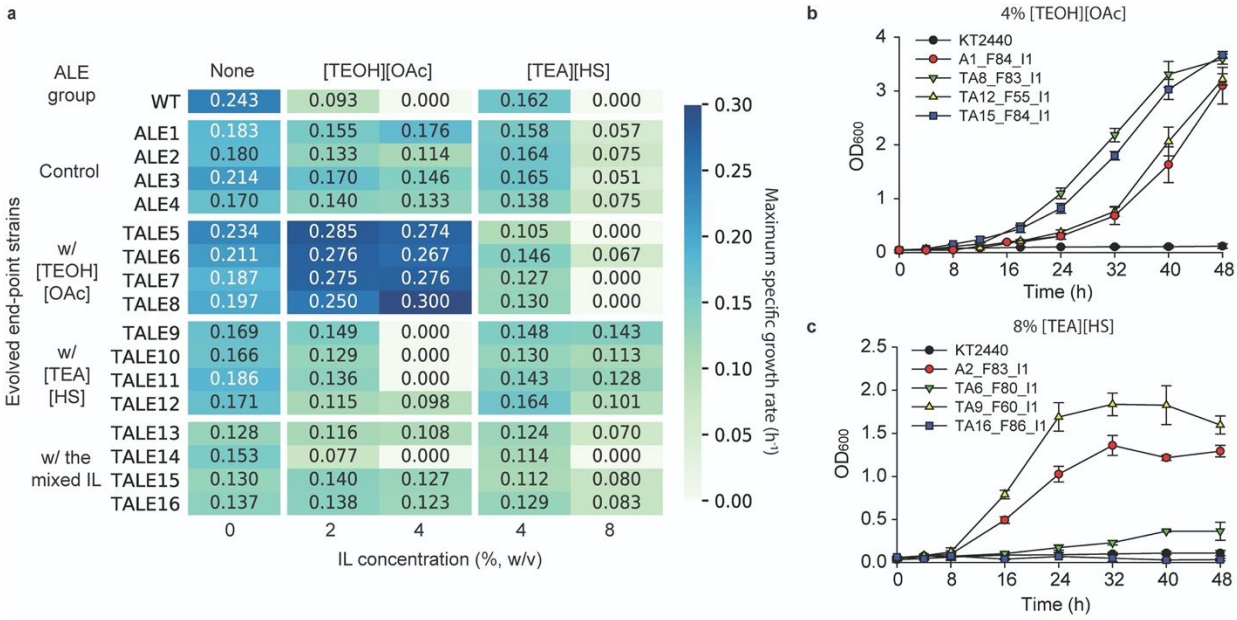
Figure captions

Fig. 1. Evolution trajectories of the TALE experiments.



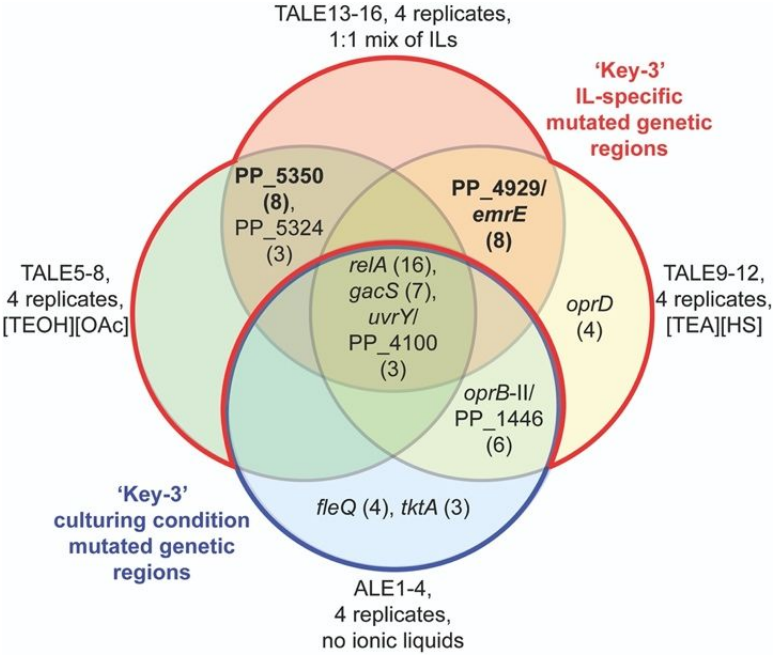
The tolerization ALE (TALE) experiments were conducted during a 3-month period with (a) triethanolammonium acetate ([TEOH][OAc]), (b) triethylammonium hydrogen sulfate ([TEA][HS]), and (c) their 1:1 mixture. *x*-axis indicates cumulative cell divisions (CCD).[52] Left and right *y*-axes indicate the approximate population growth rate and the IL concentration at each flask. Initial step sizes for decreasing the concentrations were 20% of the initial concentrations. For [TEOH][OAc], the step size was changed to 0.4% w/v when the concentration reached 2% w/v. It should be noted that the IL concentrations were occasionally lowered by the algorithm, designed to decrease the concentration when severe growth inhibition was observed.

Fig. 2. Growth characterizations of end-point strains with [TEOH][OAc] or [TEA][HS]



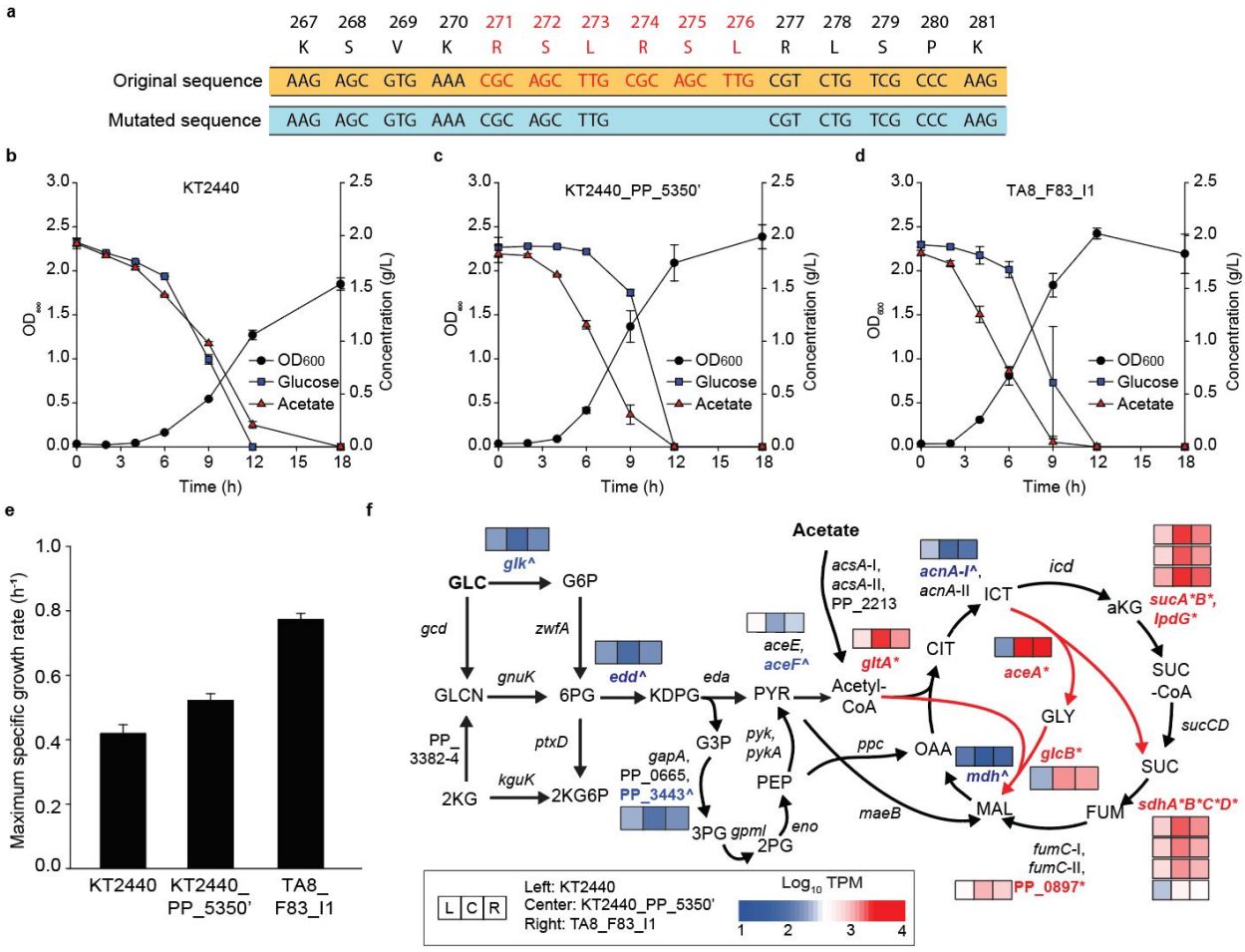
(a) Characterization of isolates from each end-point population from a small-scale cultivation (200 μL of the culture medium). Numbers indicate the maximum specific growth rates (h^{-1}) under the absence and presence of the IL. Right bar indicates the color-scale of the growth rate. (b and c) Flask-scale cultivation of the wild-type KT2440 strain and selected end-point strains representing each condition (15 mL of the culture medium). *x*-axis and *y*-axis indicate time (h) and optical density at 600 nm (OD_{600}), respectively. (b) A1_F84_I1, TA8_F83_I1, TA12_F55_I1, TA15_F84_I1 strains were cultivated with 4% [TEOH][OAc] and (c) A2_F83_I1, TA6_F80_I1, TA9_F60_I1, TA16_F86_I1 strains were cultivated with 8% [TEA][HS]. All cultures were conducted in biological triplicates (n=3).

Fig. 3. 'Key-3' regions related in the improved IL tolerance



A Venn diagram for 'key-3' regions that were mutated in three or more independent lineages and fixed in at least one endpoint isolate. Numbers in parentheses indicate the number of lineages where the gene or genetic region was mutated. Two types of mutations are identified, (1) 'Key-3' culturing condition mutated genetic regions (blue) which indicate mutations related to adaptation to the culturing conditions, and (2) 'Key-3' IL-specific mutated genetic regions (red) which indicate mutations related specifically to IL tolerance. Gene names: PP_5350, RpiR family transcriptional regulator; PP_5324, two-component system response regulator; PP_4929, LysR family transcriptional regulator; *emrE,* a small multidrug resistance protein; *oprD,* a basic amino acid specific porin; *uvrY*, BarA-UvrY two-component system response regulator; PP_4100, Cro-CI family transcriptional regulator; *relA*, ATP:GTP 3'-pyrophosphotransferase; *oprB*-II, carbohydrate-selective porin; PP_1446, TonB-dependent receptor; *fleQ*, a transcriptional regulator; *tktA*, transketolase A; *gacS*, sensor protein. '/' indicates an intergenic region.

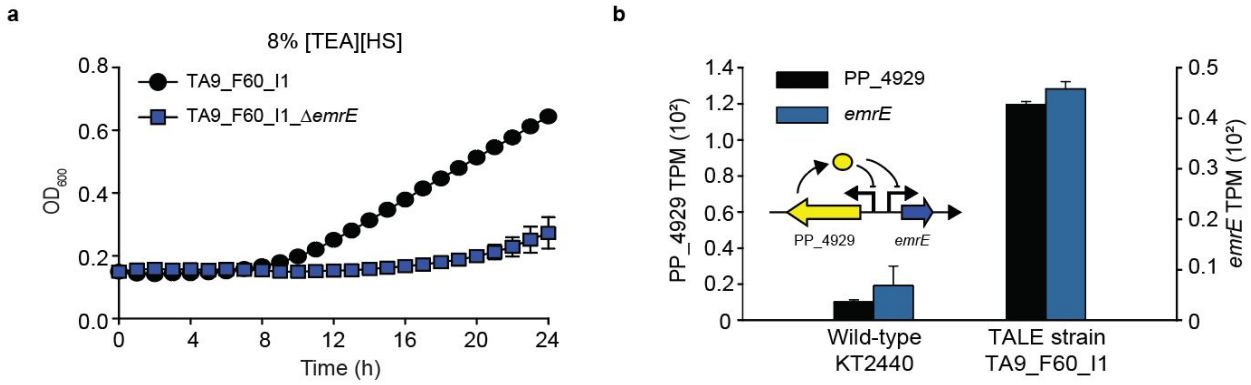
Fig. 4. Characteristics of the 9-bp mutation in PP_5350 and its effect on acetate utilization



(a) Comparison of the original and 9-bp deleted PP_5350 sequences (b-d) Growth and carbon consumption profiles of (b) the wild-type KT2440 strain, (c) the KT2440_PP_5350' strain, and (d) the TA8_F83_I1 strain in minimal medium with 2 g/L each of glucose and acetate and no ILs. *x*-axis indicate time (h). Left and right *y*-axes indicate OD_{600} and carbon concentration (g/L), respectively. Symbol: black circle, OD_{600}; blue square, glucose; red triangle, acetate. (e) Maximum specific growth rates on the glucose and acetate condition (h^{-1}) of the three strains. (b-e) Error bars indicate standard deviation of three independent cultures (*n*=3). (f) Changes in transcription levels of carbon metabolism related genes. Square boxes indicate averaged TPM values of the differentially expressed genes (marked with * for up and ^ for down) of wild-type

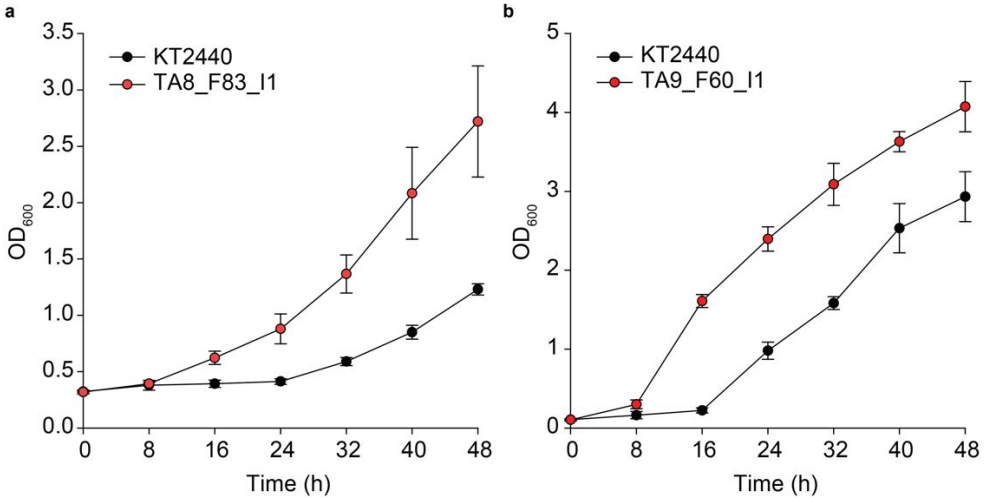
KT2440 (left), KT2440_PP_5350' (center), and TA8_F83_I1 (right) strains. Only genes differentially expressed in wild-type KT2440 and KT2440_PP_5350' strains were marked. The cultures were conducted in biological duplicates (n=2). Abbreviations: GLC, glucose; GLCN, gluconate; 2KG, 2-ketogluconate; G6P, glucose-6-phosphate(P); 6PG, 6-phosphogluconate; 2KG6P, 2-ketogluconate-6-P; KDPG, 2-keto-3-deoxy-6-phosphogluconate; G3P, glyceraldehyde-3-P; 3PG, glycerate-3-P; 2PG, 2-phosphoglycerate; PEP, phosphoenolpyruvate; PYR, pyruvate; Acetyl-CoA, acetyl coenzyme A (CoA), CIT, citrate; ICT, isocitrate; aKG, alpha-ketoglutarate; SUC-CoA, succinyl-CoA; SUC, succinate; FUM, fumarate; MAL, malate; OAA, oxaloacetate.

Fig. 5. Characterization of the role of EmrE and its relationship with PP_4929



(a) Growth profiles of the TA9_F60_I1 (PP_4929 has an early stop codon insertion mutation) and TA9_F60_I1_Δ*emrE* strains in the presence of 8% [TEA][HS]. *x*-axis and *y*-axis indicate time (h) and OD_{600}. Error bars indicate standard deviation of three independent cultures (*n*=3). (b) TPM of the PP_4929 and *emrE* genes of the TA9_F60_I1 and TA9_F60_I1_Δ*emrE* strains in the medium without the IL. Each *y*-axis indicates the TPM value of a gene. The TPM values were derived from four biological replicates of the KT2440 strain (*n*=4) and two biological replicates of the TA9_F60_I1 strain (*n*=2). The inset indicates a suggested regulation of the *emrE* gene expression by PP_4929.

Fig. 6. Cultivation of the wild-type KT2440 and TALE end-point strains in IL-treated biomass hydrolysate media



(a) The wild-type KT2440 and the TA8_F83_I1 strains were cultivated in [TEOH][OAc]-treated sorghum medium. (b) The wild-type KT2440 and the TA9_F60_I1 strains were cultivated in [TEA][HS]-treated sorghum medium. The concentrations in the media were 4% w/v and 1.5% w/v for [TEOH][OAc] and [TEA][HS], respectively. *x*-axis and *y*-axis indicate time (h) and OD_{600} during a 48-h cultivation. The cultures were conducted in biological triplicates (n=3).

Table 1. 'Key-3' regions in the evolved strains

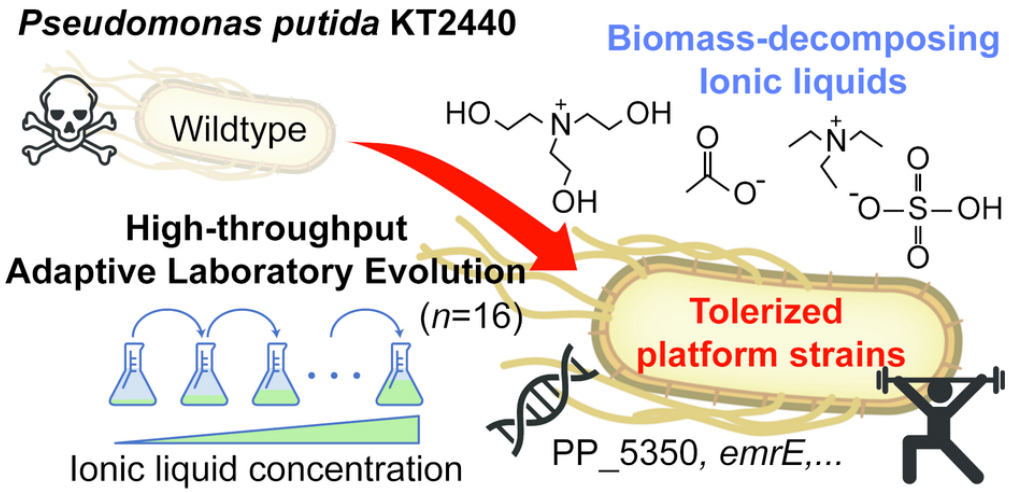
Region[a]	Description[b]	Number of lineages[c]	ALE condition	ALE number	Related study
Culturing condition mutated genetic regions					
relA	ATP:GTP 3'-pyrophosphotransferase	16	All	1-16	[74–76]
gacS	Sensor protein	7	[TEOH][OAc], [TEA][HS], the mixed IL	6, 7, 9, 10, 11, 15, 16	[77–81]
oprB-II / PP_1446	Carbohydrate-selective porin / TonB-dependent receptor	6	Control, [TEOH][OAc]	1, 2, 4, 6-8	[82]
fleQ	Transcriptional regulator	4	Control	1-4	[77,83]
tktA	Transketolase A	3	Control	1-3	
uvrY / PP_4100	BarA-UvrY two-component system response regulator / Cro-CI family transcriptional regulator	3	[TEOH][OAc], [TEA][HS], the mixed IL	5, 12, 13	[81]
IL-specific mutated genetic regions					
PP_5350	RpiR family transcriptional regulator	8	[TEOH][OAc], the mixed IL	5-8, 13-16	[53]
PP_4929 / *emrE*	LysR family transcriptional regulator / small multidrug resistance protein	8	[TEA][HS], the mixed IL	9-16	[57]

oprD	Basic amino acid specific porin	4	[TEA][HS]	9-12	[82,84]
PP_5324	Two-component system response regulator	3	[TEOH][OAc], the mixed IL	6, 8, 16	[85,86]

[a]*gacS* and *uvrY*/PP_4100 regions were classified as culturing condition mutated genetic regions because their mutations were observed in previous studies[77,81] that evolved KT2440 in a similar no-stress condition. All detailed mutations are listed in Supplementary Data 2. '/' indicates an intergenic region.

[b]Annotations were from Pseudomonas Genome DB (http://www.pseudomonas.com/).

[c]Number of lineages where mutations of the genomic region were observed.



78x38mm (300 x 300 DPI)