

Advancing the scale of synthetic biology via cross-species transfer of cellular functions enabled by iModulon engraftment

Received: 30 December 2023

Accepted: 29 February 2024

Published online: 15 March 2024

 Check for updates

Donghui Choe¹, Connor A. Olson¹, Richard Szubin¹, Hannah Yang¹,
Jaemin Sung¹, Adam M. Feist^{1,2} & Bernhard O. Palsson^{1,2} ✉

Machine learning applied to large compendia of transcriptomic data has enabled the decomposition of bacterial transcriptomes to identify independently modulated sets of genes, such as iModulons represent specific cellular functions. The identification of iModulons enables accurate identification of genes necessary and sufficient for cross-species transfer of cellular functions. We demonstrate cross-species transfer of: 1) the biotransformation of vanillate to protocatechuate, 2) a malonate catabolic pathway, 3) a catabolic pathway for 2,3-butanediol, and 4) an antimicrobial resistance to ampicillin found in multiple *Pseudomonas* species to *Escherichia coli*. iModulon-based engineering is a transformative strategy as it includes all genes comprising the transferred cellular function, including genes without functional annotation. Adaptive laboratory evolution was deployed to optimize the cellular function transferred, revealing mutations in the host. Combining big data analytics and laboratory evolution thus enhances the level of understanding of systems biology, and synthetic biology for strain design and development.

Synthetic biology aims to engineer new, or modify existing, cellular functions by designing, editing, and assembling the underlying genetic material. To achieve this goal, knowledge-based designs have been traditionally used that identify genes of known molecular functions to piece together the targeted cellular function. However, identifying all the necessary genetic components is not trivial and requires large-scale analysis such as genome-wide knockout studies and Tn-Seq^{1,2}. This approach also calls for serial steps of genetic transformation with separate testing of each individual gene of the targeted function³. For example, the development of the biosynthetic pathway of the antimalarial drug artemisinin demanded extensive research efforts and engineering^{4–6}. To circumvent the difficulties, computational tools have been developed to supplement experimental procedures and aid design^{7–9}. More importantly, recent advances in new data analytics applied to large transcriptomic datasets has enabled an alternative approach. Using Independent Component Analysis (ICA), we can now identify sets of independently

modulated genes (called iModulons) that constitute a particular cellular function¹⁰.

This advancement opens up the possibility to identify targeted functions in a particular strain and transfer them into alternate hosts. Successful cross-species transfer of desired functions would thus require: (i) the identification of the full genetic basis for the trait, (ii) the use of recombinant or DNA synthesis methods to capture these genes into a plasmid, (iii) the transfer of the plasmid into the target host, and (iv) making any needed changes to the new host that are critical to accommodate the transferred function. All these capabilities now exist, with (ii) and (iii) representing known approaches, while (i) and (iv) require novel transcriptomic analysis and the use of automated adaptive laboratory evolution.

ICA can be applied to find source regulatory signals in bacterial transcriptomes^{10–13}. iModulons are fundamental units of bacterial transcriptomes that have been found to represent the genetic basis for various cellular functions and are associated with particular

¹Department of Bioengineering, University of California San Diego, La Jolla, CA 92093, USA. ²Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Copenhagen, Denmark. ✉ e-mail: bpalsson@ucsd.edu

transcriptional regulator(s)^{14–16}. Many identified iModulons include genes that are distantly located on the genome, genes of unknown functions, or contain accessory genes that augment the targeted cellular function. iModulons thus represent an advanced scale of synthetic biology to transfer naturally evolved traits across species.

Here, we use iModulons to create cellular functions in a new host. We show that transferring iModulons is superior to using operons or single genes identified by genome annotation algorithms and rational design approaches^{9,17,18}. In some cases, the transferred function may not work optimally in the new host. In such cases, we can use adaptive laboratory evolution (ALE) to enable the host to optimally use the new function under selection pressure. We thus demonstrate that cross-species iModulon transfer is a versatile tool for synthetic biology.

Results

Cross-species transfer of *Pseudomonas* iModulons into *E. coli*

To initiate the project and prior to implementing cross-species iModulon transfer, we refactored a known cellular function within the original host as a proof of concept. Successful homologous refactoring and complementation of *E. coli*'s branched-chain amino acid (BCAA) metabolism was achieved (Supplementary Note, section 1 and Supplementary Fig. 1) to demonstrate identification, reconstruction, and transfer of genetic constituent of a biological function based on iModulon (i, ii, and iii). This motivated us to investigate the potential for transferring biological functions across species. Among the available species with iModulon structures in iModulonDB¹², *Pseudomonas* is well-known for its versatile metabolism to degrade and utilize diverse compounds, including aromatics^{19–21}. First, we chose to reconstruct and transfer a simple bioconversion process from *Pseudomonas putida*²⁵ to *E. coli* in order to examine iModulon's capability to rapidly identify genes associated with specific functions.

The VanR iModulon that is responsible for vanillate (VA) transport and conversion into protocatechuate (PCA) was chosen for our first cross-species iModulon transfer (Fig. 1A). It comprises three genes with annotated functions, *vanA*, *vanB*, *vanK*, and predicted porin-like *galP-IV* (Fig. 1B) in two converging operons (Fig. 1C). Notably, the iModulon exactly matches with the genes for the vanillate transport and metabolism^{22,23}. Four genes, *vanA*, *vanB*, *galP*, and *vanK* are functionally annotated to encode for vanillate O-demethylase oxidoreductase complex, outer-membrane porin, and a major facilitator superfamily transporter, respectively²². Although the function of the outer membrane OprD-domain containing *galP-IV* has never been addressed, it is hypothesized that it facilitates the diffusion of the ligand through the outer membrane^{23,24}. Since the mechanism of VanR regulation has not been established, the four genes constituting the VanR iModulon were cloned and heterologously expressed under the control of IPTG-inducible Trc promoter on a plasmid, pVanR_iM (Fig. 1D). When refactoring iModulons for heterologous expression, we tried to preserve native genetic arrangement, for VanR and following iModulons if possible, to ensure optimal expression levels of the gene members as demonstrated elsewhere^{25,26}.

E. coli carrying pVanR_iM converted VA into PCA up to 15.34 mg/l passively diffused to the supernatant^{27,28} during 48 h of fermentation in M9 glucose (4 g/l) medium supplemented with 100 mg/l VA, while the negative control carrying empty plasmid did not metabolize any VA (Fig. 1E). This first cross-species iModulon transplantation illustrates the rapid identification of enzymes required for biotransformation by ICA. Furthermore, iModulon engraftment provided a rapid way to biochemically verify a predicted pathway in a heterologous host.

Auxiliary genes may be needed for optimal function of cross-species transferred iModulons

Next, we chose to transfer an ampicillin resistance function of *Pseudomonas aeruginosa* to *E. coli*. *P. aeruginosa* displays beta-lactam

resistance with endogenous beta-lactamase, AmpC, and has an iModulon involved in the inducible ampicillin resistance¹⁶. Activity levels of the AmpC iModulon are highly induced against beta-lactam challenge, but not under other antibiotic treatments (Supplementary Fig. 2). In the previous iModulon engraftment examples, genes comprising an iModulon matched with the predicted genes necessary for building the desired function. However, identifying all the genes necessary to build a biological function may not be trivial, given previous characterization efforts. Many iModulons contain genes whose functions are unknown or are seemingly unrelated to the overall function being transferred.

The AmpC iModulon comprises class C beta-lactamase encoded by the *ampC* gene²⁹ that serves as a core for the functionality and six lesser characterized auxiliary genes, *carO* (PA0320), *creD* (PA0465), PA0466, PA0467, PA4111, and PA4112 (Fig. 2A). The seven iModulon genes are distributed across three genomic loci separated by over 4 Mb. *P. aeruginosa* readily becomes resistant to ampicillin by transcriptional activation of *ampC*³⁰. However, it is not known if the resistance trait is carried by this single gene. To examine if this resistance function is transferable across species, the constituent genes were refactored into a single operon (Fig. 2B). In addition, we constructed a plasmid that contained beta-lactamase alone to address any involvement of auxiliary factors in the function.

Ampicillin disc diffusion assay revealed that *E. coli* carrying the AmpC iModulon or *ampC* gene were resistant to ampicillin, while *E. coli* carrying empty plasmid were not (Supplementary Fig. 3). The source of AmpC iModulon, *P. aeruginosa*, showed ampicillin resistance with the minimum inhibitory concentration (MIC) of 2048 µg/ml (Fig. 2C). The MIC of ampicillin for laboratory *E. coli* strain MG1655 with empty plasmid was 16 µg/ml, which is comparable to previous reports^{31,32} (Fig. 2C, D). *E. coli* strain with the *P. aeruginosa* beta-lactamase showed a dramatic increase in ampicillin resistance with an MIC of 1024 µg/ml, while it was lower than that of the original host (Fig. 2D). Strikingly, *E. coli* harboring the entire AmpC iModulon, six auxiliary genes in addition to *ampC*, had an MIC of 4096 µg/ml, which was four times higher than that with *ampC* alone (Fig. 2D).

Although little is known about the molecular function of auxiliary genes, they were required to completely replicate the ampicillin resistance characteristics of *P. aeruginosa*. Previous reports have shown a decrease in beta-lactam resistance of the inner membrane protein *creD* knockout mutant of *P. aeruginosa*³³ and growth enhancement of *E. coli* by endogenous *creD* overproduction (shares 37.4% sequence identity; BLOSUM62)³⁴. Although the function of CreD is still elusive, reports indicate its relevance in biofilm development in *P. aeruginosa*³⁵ and envelope integrity in *Stenotrophomonas maltophilia*³⁶. Additionally, calcium-regulated oligonucleotide/oligosaccharide binding (OB)-fold protein CarO has been reported to be related to susceptibility to various stresses in bacteria³⁷. Also, it shares similarity with *Salmonella enterica* stress-related protein VisP (38% sequence identity), which binds to peptidoglycan and inhibits the lipid A modifying enzyme LpxO³⁸. Since lipid A is an anchor of lipopolysaccharide to the outer membrane and affects the properties of the outer membrane, expression of *carO* might be beneficial for cells to maintain structural integrity under cell wall deficient conditions induced by beta-lactam³⁹.

Engrafting *Pseudomonas* iModulons to *E. coli* highlighted critical properties of iModulon gene membership. Harnessing only core genes for transferred cellular function may not be sufficient, as auxiliary genes may be needed to reconstruct an optimal function. Full iModulon gene membership helps to recreate the targeted cellular function, even without a complete understanding of the molecular function of all the genes involved.

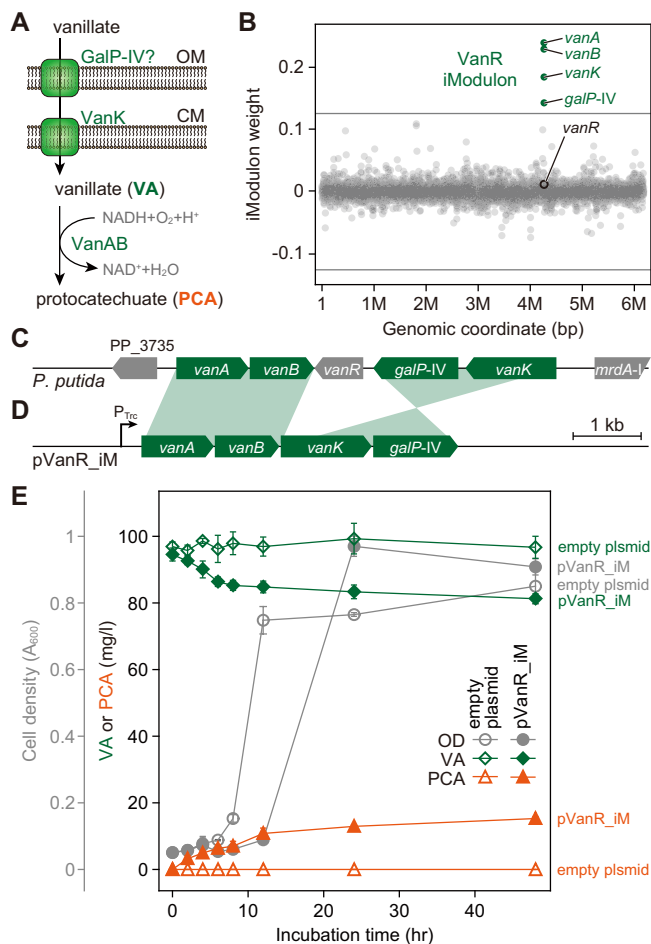


Fig. 1 | Cross-species transfer of the *Pseudomonas putida* VanR iModulon into *E. coli*. **A** Vanillate transport and conversion in *P. putida*. OM outer membrane. CM cytoplasmic membrane. **B** iModulon weights of genes in *P. putida*. Four genes (green circles) with high weighting constitute the VanR iModulon. Gray lines indicate thresholds for determining iModulon membership. Gray circles identify genes not in the iModulon. **C** Graphical representation of *vanR* locus on the *P. putida* chromosome. **D** The VanR iModulon was refactored in a single operon under the control of *trc* promoter (P_{Trc}), resulting in the pVanR_iM plasmid. Shades show genetic rearrangement for cloning purposes. **E** Vanillate (VA) conversion of *E. coli* carrying empty or pVanR_iM plasmid into protocatechuate (PCA). Gray circles, green diamonds, and orange triangles indicate cell density, VA, and PCA levels of the culture, respectively. Measurements from *E. coli* carrying empty or pVanR_iM plasmid are represented by hollow or filled symbols, respectively. Data were presented as mean values \pm SD. Error bars indicate the SD of three replicate cultures. Source data are provided as a Source Data file.

Complete iModulon gene membership is needed for successful cross-species transfer

As illustrated by the AmpC case, we further investigated the iModulon-based transfer of cellular traits and compared it to the alternative conventional methods. The 2,3-butanediol (2,3-BDO) iModulon was chosen to examine the role of iModulon genes of unknown functions. 2,3-BDO is a byproduct of bacterial fermentation processes that can be produced by a variety of microorganisms, including *Pseudomonas* species^{40–42}. In *Pseudomonas*, 2,3-BDO can serve as a carbon and energy source and is degraded by enzymes in the 2,3-BDO catabolic pathway⁴². This catabolic pathway involves the conversion of 2,3-BDO into acetoin, which is further converted into acetaldehyde and acetyl-CoA by butanediol dehydrogenase and acetoin dehydrogenase, respectively (Fig. 3A).

We transferred the 2,3-BDO iModulon of *P. putida* (called the AcoR iModulon¹⁵) to *E. coli*. The AcoR iModulon comprises *acoABC*

(encoding acetoin dehydrogenase complex), *bdhA* (encoding 2,3-BDO dehydrogenase), and a gene *acoX* (Fig. 3B). AcoX encodes for a protein of unknown function and co-exists with acetoin-utilizing genes in various bacteria^{41,43}. Operon prediction also suggests that the transcriptional unit contains *acoX* and two other hypothetical proteins (PP_0550 and PP_0551) in addition to characterized metabolic enzymes, *acoABC-bdhA* (Fig. 3C)^{18,44}.

To examine which genes are required for recreating the 2,3-BDO catabolic pathway, we built three different plasmid based on (1) operonic structure (Op353; *acoXABC-bdhA*-PP_0551-PP_0550), (2) iModulon structure (*acoXABC-bdhA*), and (3) four genes encoding enzymes predicted to be sufficient for converting 2,3-BDO into acetaldehyde and acetyl-CoA based on current gene annotations (pathway; *acoABC-bdhA*) (Fig. 3C). 2,3-BDO dehydrogenase activities of the source organism and *E. coli* strains carrying the three plasmids individually were examined during 96 h of batch cultivation in LB medium supplemented with 2 g/l of 2,3-BDO. The original strain, *P. putida* KT2440, showed 2,3-BDO utilization with a negligible level of acetoin (Fig. 3D). The negative control, *E. coli* MG1655 carrying an empty plasmid converted 0.77 g/l of 2,3-BDO into acetoin, possibly due to endogenous promiscuous alcohol dehydrogenase activity (Fig. 3E). On the other hand, the plasmids based on the pathway, operonic structure, and iModulon showed higher conversion of 2,3-BDO with amounts of 1.36, 1.75, and 1.96 g/l, respectively (Fig. 3E).

Interestingly, the strains showed varying levels of acetoin dehydrogenase activity. First, all the 2,3-BDO consumed by the negative control resulted in roughly the equimolar amount of acetoin; not surprising since there is no acetoin dehydrogenase introduced. The strain carrying the functional gene annotation-based pathway plasmid did not further convert acetoin into downstream products, even though it contained genes encoding for the acetoin dehydrogenase complex. Second, strains with the full operon or AcoR iModulon not only consumed more than 1.7 g/l of 2,3-BDO, but there was only a small amount of acetoin left in the medium, indicating conversion of acetoin by acetoin dehydrogenase. The difference between annotation-based and iModulon-based plasmid is the presence of *acoX* (Fig. 3C), a gene encoding a predicted small molecule kinase that has been reported to have no acetoin, NAD, or pyruvate kinase activity⁴⁵. However, *acoX* was critical for acetoin dehydrogenase activity.

Although the *acoX* product has no known function in acetoin metabolism, it is conserved and colocalizes on the genome with the acetoin dehydrogenase in several acetoin-utilizing bacteria from multiple phyla, such as *P. aeruginosa* (76% sequence identity) and *Clostridium magnum* (32% sequence identity)⁴². However, there is no significant match of AcoX from the BLASTP search on other acetoin-utilizing bacteria such as *Bacillus subtilis*, *Klebsiella pneumoniae*, and *Pelobacter carbinolicus*. Therefore, the requirement of AcoX in acetoin metabolism is species-specific and could not be determined by analyzing the genome sequence context.

When the iModulon and operonic constructs were compared, the iModulon construct performed better than the operonic construct for 2,3-BDO degradation (Fig. 3E). Two additional genes in the operonic construct encode the predicted membrane occupation and recognition nexus (MORN) domain-containing peptidase and a NAD(P)-binding oxidoreductase, whose relation with 2,3-BDO metabolism is unknown. These two genes were irrelevant for function. Instead, expression of the hypothetical proteins reduced 2,3-BDO degradation, possibly by imposing an unnecessary transcriptional burden on the cell. The iModulon gene membership provided information on the necessary genes to support a 2,3-BDO catabolic process that would not have been found using only functional gene annotation. This example illustrates the unique advantages of using the iModulon structure for cross-species transfer of the full genetic basis for a desired integrated function.

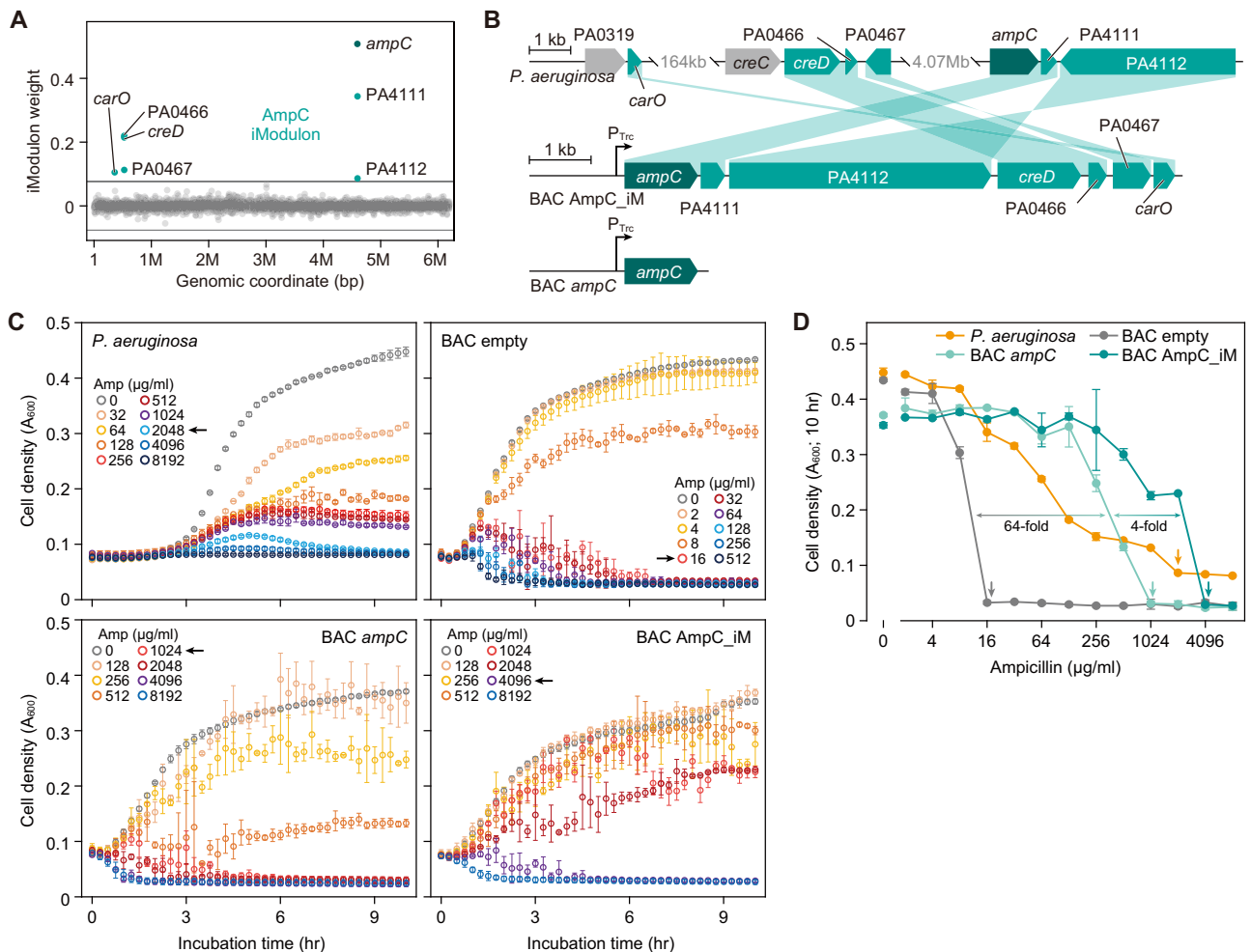


Fig. 2 | *E. coli* carrying the *Pseudomonas aeruginosa* AmpC iModulon confers better ampicillin resistance than cells expressing beta-lactamase alone.

A iModulon weights of genes in *P. aeruginosa*. Seven genes constitute the AmpC iModulon (blue circles). Gray lines indicate thresholds for determining iModulon membership. Gray circles identify genes not in the iModulon. **B** Refactoring the *P. aeruginosa* AmpC iModulon on bacterial artificial chromosome (BAC). Genes are expressed with the *trc* promoter (P_{Trc}). Shades show genetic rearrangement for cloning purposes. **C** Dose-kill curves of *P. aeruginosa* and *E. coli* carrying empty BAC, *BAC_ampC*, or *BAC_AmpC_iM*. Data were presented as mean values \pm SD. Error

bars indicate the SD of biological replicates ($n = 3$). Note that the range of ampicillin concentration (Amp) is different, due to the huge difference in ampicillin tolerance. **D** Cell density of cultures treated with different ampicillin concentrations after 10 h of incubation. Data for *P. aeruginosa* and *E. coli* carrying empty BAC, *BAC_ampC*, or *BAC_AmpC_iM* are in orange, gray, light blue, and blue, respectively. Arrows indicate the minimum inhibitory concentration (MIC). Data were presented as mean values \pm SD. Error bars indicate the SD of biological replicates ($n = 3$). Source data are provided as a Source Data file.

ALE optimizes the functionality of catabolic iModulons

Lastly, we chose the MdcR iModulon from *P. aeruginosa*¹⁶ to transfer into *E. coli* that, again, comprises genes identical to a reported set for malonate transport and utilization. The MdcR iModulon comprises seven subunits of malonate decarboxylase complex²¹ and two putative membrane proteins, MadL-MadM (Fig. 4A and Supplementary Fig. 4). Although the function of these membrane proteins have not been elucidated in *P. aeruginosa*, MadL and MadM have 71 and 81% of sequence homology to malonate transporters in *Malonomonas rubra*⁴⁶, respectively, suggesting a potential malonate uptake function. These genes are encoded in a single operon on the *P. aeruginosa* genome, and thus the entire operon was subjected to cross-species transfer.

The operon was cloned and heterologously expressed under the control of a *Trc* promoter on a plasmid, named pMdcR_iM (Fig. 4B). Malonate is a non-native nutrient for *E. coli*, thus it is expected that a strain with the pMdcR_iM alone would then enable growth in M9 malonate medium as the breakdown product of the pathway, acetate, can support growth⁴⁷. We experimented with varying levels of

expression using different concentrations of the inducer (IPTG) to activate the MdcR iModulon. *E. coli* could slowly utilize (doubling time of 11.2 ± 0.6 h; over the course of 72 h of fermentation in M9 malonate medium) malonate as a carbon source only at weak expression level (Fig. 4C). In contrast to complete utilization of malonate by *P. aeruginosa* within 12 h of fermentation (Fig. 4C), the observed slow utilization by *E. coli* suggests a potential metabolic imbalance in *E. coli*, perturbed by and unable to accommodate the malonate pathway.

Therefore, we implemented adaptive laboratory evolution to allow *E. coli* to rebalance and optimize its metabolism with malonate as a substrate. The *E. coli* strain carrying the pMdcR_iM was grown in an M9 malonate medium and evolved using serial passaging that imposes growth rate selection pressure (Fig. 4D) on an automated ALEbot⁴⁸. After 21 passages, populations showed faster growth with a short lag phase compared to their ancestor (Fig. 4E). The evolved populations fully consumed malonate within 40 hrs. Subsequently, three clones were isolated from each replicate evolved population, and they all displayed a faster growth rate than the ancestor (Fig. 4F).

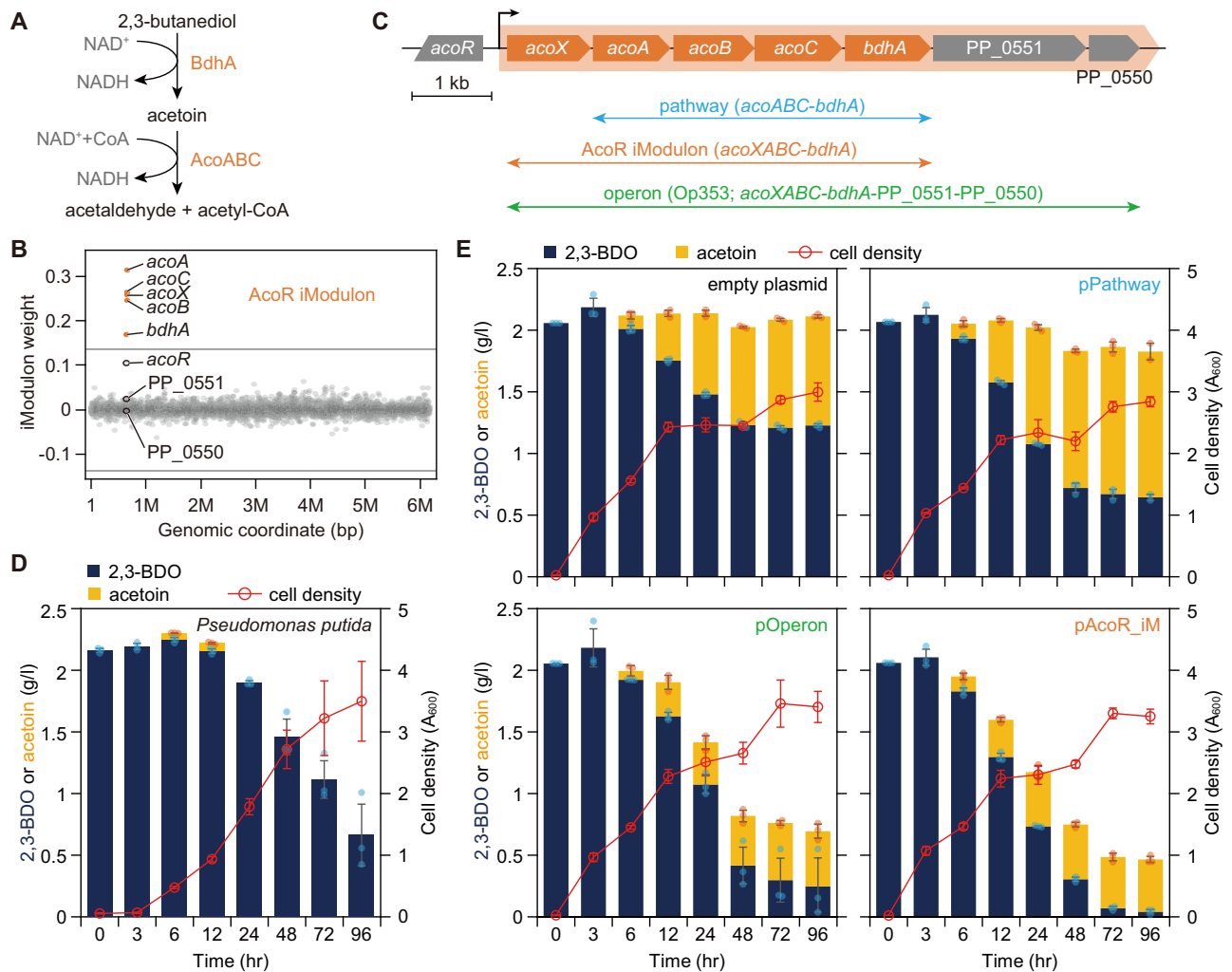


Fig. 3 | Cross-species transfer of 2,3-butanediol (2,3-BDO) utilization iModulon of *Pseudomonas putida* in *E. coli*. **A** A pathway responsible for 2,3-BDO utilization. **B** Scatter plot shows weights of genes in *P. putida* to AcoR iModulon. Gray lines indicate thresholds for determining iModulon membership. Five genes constitute the AcoR iModulon (orange circles). Gray circles identify genes not in the iModulon. Black circles are three neighboring genes. **C** Genomic structure of the AcoR iModulon. Orange shade shows predicted operonic structure. Genes in the iModulon are in orange. Arrows indicate three different plasmid constructs for cross-species transfer. **D** 2,3-BDO degradation by *P. putida*. The formation of acetoin was negligible. Blue and yellow boxes represent 2,3-BDO and acetoin in the culture

medium. Red circles show cell density. Dots indicate individual data points. Data were presented as mean values \pm SD. Error bars indicate the SD of the three biological replicates. **E** 2,3-BDO and acetoin degradation by *E. coli* carrying empty plasmid or one of the three constructs. 2,3-BDO was added at the start of the culture and the remaining amount and acetoin formation was measured. Blue and yellow boxes represent 2,3-BDO and acetoin in the culture medium. Red circles show cell density. Dots indicate individual data points. Data were presented as mean values \pm SD. Error bars indicate SD of the three biological replicates. Source data are provided as a Source Data file.

To understand the genetic bases of improved growth, we resequenced the genome of the evolved clones (Supplementary Table 1). All the evolved clones carried mutations on DNA polymerase I, encoded by *polA*, which is required for plasmid maintenance (Fig. 4G)⁴⁹. Previous studies reported a change of plasmid copy number induced by *polA* mutation⁵⁰. Quantitative measurement of plasmid copy number indicated a reduction of plasmid copy number, which led to a reduction of MdcR iModulon expression (Fig. 4H). Thus, the initial metabolic failure was likely due to the sub-optimal expression of the MdcR iModulon (Supplementary Note, section 2), which could be optimized by ALE.

Engraftment of the MdcR iModulon, in addition to three other iModulons, demonstrated cross-species iModulon transfer as a rapid way of creating new functionality in bacteria with minimal engineering. We found that the overall behavior of the iModulon interferes with the host factors that require modifications to optimally support the system. This optimization could be rapidly achieved by ALE that identified

few genetic changes in the host, while the transferred genes acquired no adaptive mutations.

Discussion

Historically synthetic biology has built multigenic functions in a serial manner. Individual genes are introduced to the host, and their function is assessed. iModulons, obtained through big data analysis of transcriptome compendia, describe sets of co-expressed genes that constitute independent cellular functions, suggesting that multigenic traits can be captured and transferred. Here we demonstrate that this is possible through cross-species transfer of cellular functions from *Pseudomonas* species into *E. coli*. We successfully transfer three metabolic traits and an antimicrobial resistance trait between these Gram-negative species.

iModulons lead to the identification of genes that are not functionally annotated but essential for recreating a cellular function in a new host. Thus, these genes gain a network-level functional

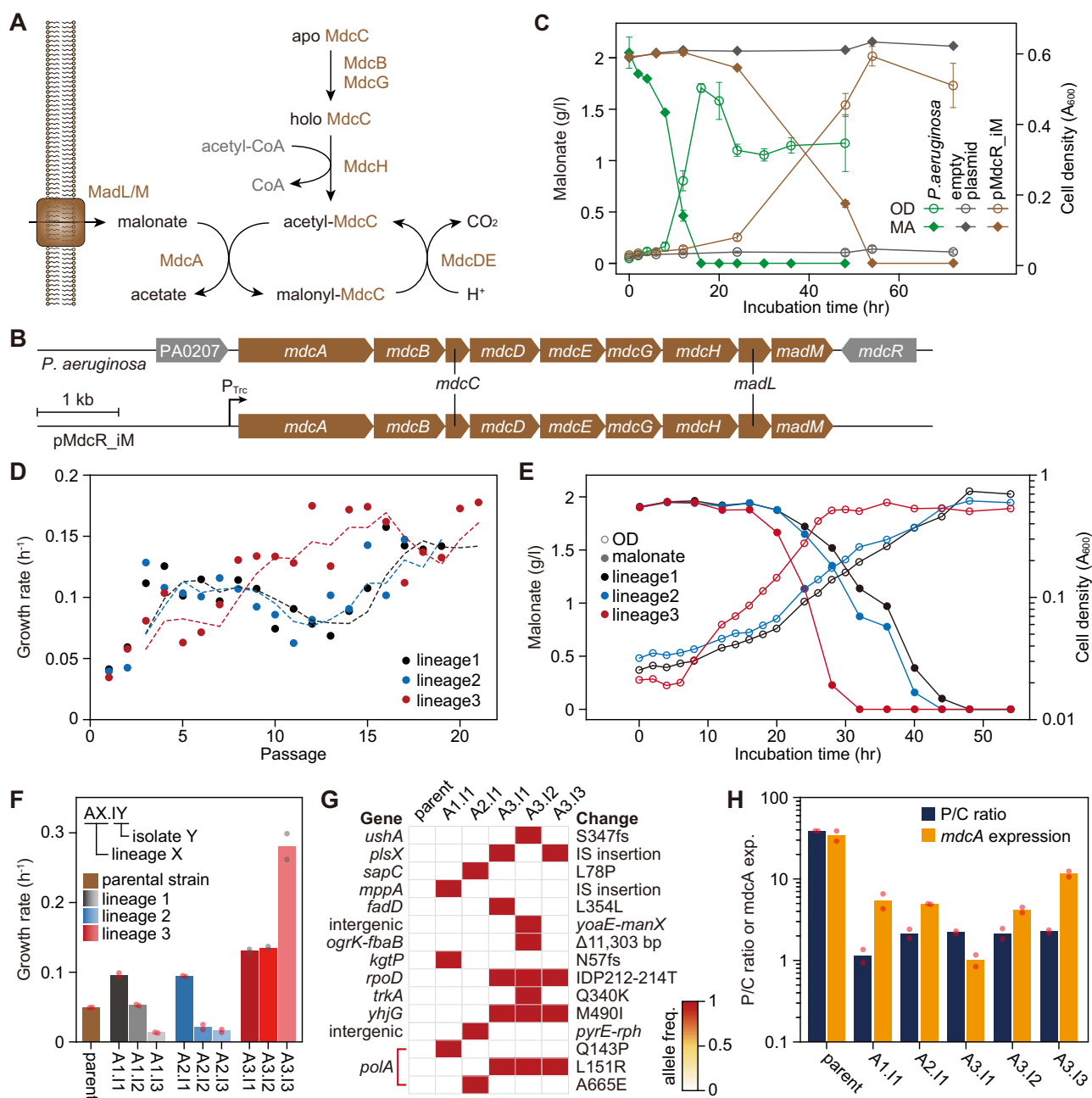


Fig. 4 | Adaptive laboratory evolution improves functionality of *Pseudomonas aeruginosa* MdcR iModulon in *E. coli*. **A** Malonate catabolic pathway in *P. aeruginosa*. **B** Genetic structure of the malonate catabolic operon of *P. aeruginosa* cloned in a heterologous expression plasmid, pMdcR_iM. Brown genes constitute MdcR iModulon. **C** Malonate utilization of *P. aeruginosa*, *E. coli* carrying empty plasmid, and pMdcR_iM. Cells were incubated for up to 72 h in M9 malonate (2 g/l) media. Circles and diamonds show cell density and malonate concentration in culture, respectively. Green, gray, and brown lines represent *P. aeruginosa*, *E. coli* carrying empty plasmid, and pMdcR_iM plasmid, respectively. Data were presented as mean values \pm SD. Error bars indicate the SD of three replicated cultures. **D** Growth rates of *E. coli* carrying the MdcR iModulon over the course of evolution. Dashed lines are moving averages of three individual ALE lineages. Growth rates for each ALE lineage are colored differently. **E** Malonate utilization and growth of three evolved populations. Circles represent cell density, with the solid circles being

extracellular malonate concentrations. Measurements for each ALE lineage are colored differently. Data were presented as mean values of two replicated cultures. **F** Growth rates of clones isolated from malonate-evolved populations in M9 malonate medium. Data were presented as mean values of two replicated cultures. Dots show individual data points. Strain names are given as AX.IY. X is the ALE lineage number and Y is an arbitrary identifying number for the clonal isolate from the same ALE lineage. **G** Adaptive mutations in the ALE endpoint clones that are not present in the parental strain. fs, frameshift mutation. The heatmap shows allele frequencies colored as in the provided color key. **H** Plasmid-to-chromosome copy number ratio (P/C ratio) and expression level of *mdcA* of unevolved parent strain and evolved clones. Blue and orange boxes represent the P/C ratio and *mdcA* expression level, respectively. Data were presented as mean values in two biologically replicated cultures. Dots are individual data points, each of which is composed of two technical duplicates. Source data are provided as a Source Data file.

annotation, as opposed to classical molecular-level functional annotation. We also find that iModulon genes may have an auxiliary function, i.e., they enhance the cellular functions being transferred. Finally, we find that laboratory evolution can improve the transferred cellular function in the new host. Sequencing laboratory-evolved strains can identify host factors that enable or enhance the cellular function encoded on the transferred construct.

Taken together, these factors lead to an advancement of synthetic biology, including the identification of all genes needed to constitute a cellular function and revealing host factors that need modification to optimize the engineered strain. This acceleration and higher predictability in strain design and construction should accelerate the development of synthetic biology and its deployment for practical purposes such as biomanufacturing.

Methods

Bacterial strains and culture conditions

E. coli strain K-12 substrain MG1655 is used as a recipient of iModulon. Cells were grown in LB medium (Novagen, 71753) or M9-defined medium (47.75 mM Na₂HPO₄, 22.04 mM KH₂PO₄, 8.56 mM NaCl, 18.70 mM NH₄Cl, 2 mM MgSO₄, 0.1 mM CaCl₂, and trace elements). Trace elements were prepared in 2000× concentrated solution (100 mM FeCl₃, 9.54 mM ZnCl₂, 8.41 mM CoCl₂, 8.27 mM Na₂MoO₄, 0.75 mM CaCl₂, 0.91 mM CuCl₂, and 0.5 mM H₃BO₃ in 3.7% (w/w) hydrochloric acid solution). Branched-chain amino acids were supplemented with a final combined concentration of 0.15 mM. To construct the ΔBCAA strain (MG1655 Δ*ilvLGMEDAYC::ilvY*, Δ*leuABCD*, *thrL-119::sacB-cat*), sequential lambda-recombinations were performed⁵¹. First, *ilvLGMEDAYC* was replaced with a kanamycin resistance cassette flanked by two flippase recognition sequences. The cassette was removed by flippase-mediated excision, and a DNA cassette containing *ilvY-phe^S*-nptII* was introduced to replace the flippase recombination scar. Then *phe^S*-nptII* dual-selectable marker was again removed by the lambda recombination. A successful recombinant was selected by *p*-chloro-phenylalanine counter selection⁵² on an LB-agar plate containing 2 mM *p*-chloro-phenylalanine. The genomic locus containing *leuABCD* was removed by introducing the kanamycin cassette, followed by flippase recombination. Lastly, a *sacB-cat* dual selection marker was inserted upstream of *thrL* serving as a landing pad for REXER. Yeast was grown in YPDA (10 g/l yeast extract, 20 g/l peptone, 30 mg/l adenine sulfate, and 20 g/l dextrose) or synthetic His dropout medium (1.7 g/l yeast nitrogen base without amino acid and ammonium sulfate, 1.9 g/l yeast synthetic drop-out medium supplements without histidine, 5 g/l ammonium sulfate, 20 g/l agar, and 20 g/l dextrose) at 30 °C. MdcR and AmpC iModulons were sourced from *Pseudomonas aeruginosa* strain K2733¹⁶. AcoR and VanR iModulons were amplified from *Pseudomonas putida* strain KT2440. The fermentation experiment was done in a 15 ml liquid culture contained in a 30-ml test tube. Cells were incubated at 37 °C on a heat-block and stirred at 600 rpm with a cross-shaped magnetic stir bar. To induce heterologous iModulons, (IPTG) was added to cultures to 0.1 or 1 mM of final concentration. All reagents were purchased from Sigma-Aldrich unless stated otherwise.

Investigation of target iModulons and transfer design

Target iModulons for transfer were selected from a catalog of iModulons available through iModulonDB (<https://imodulondb.org/>)¹². Initially, iModulons of *P. aeruginosa* and *P. putida* were inspected by their predicted functional annotations. iModulons associated with functions that are uncharacterized, already present in *E. coli* (e.g., translation), or challenging to experimentally screen (e.g., quorum sensing) were excluded. Subsequently, the functions of iModulons were further investigated by a comprehensive examination of individual gene functions using gene annotation databases (Biocyc¹⁸ and Pseudomonas Genome Database⁵³) and a literature search.

The functional annotation of an iModulon was revised in this step if needed. For instance, the *P. aeruginosa* MdcR iModulon was initially annotated as a malonate biosynthetic process, but we corrected its function to malonate catabolism as it encodes the malonate catabolic enzyme complex²¹. Finally, iModulon activities across diverse experimental conditions were examined using the built-in graph function of iModulonDB in the individual iModulon page, which can also be obtained by downloading two tables, namely Experimental Conditions and iModulon Activity. In this step, we took a rational approach to test if changes in iModulon activity under specific conditions were consistent with established biological knowledge. For example, activities of BCAA biosynthetic iModulons and AmpC iModulon are induced in the absence of BCAA and the presence of beta-lactam in the media (Supplementary Figs. 1C, 2), affirming the functional annotations.

The strategy for iModulon transfer was established by addressing the genetic context of iModulons. Firstly, gene members of iModulons were extracted by selecting genes with their absolute iModulon weight exceeding the threshold (Figs. 1B, 2A, 3B and Supplementary Figs. 1A, 4) (Gene Table in an iModulon page). Next, the genetic organization (location, orientation, and operonic structure) was assessed through Biocyc¹⁸ and Pseudomonas Genome Database⁵³. If genes were found in a single operon, the entire operon was directly amplified and cloned. If genes were located in multiple genomic loci, they were refactored into an operon, while preserving the genetic order to ensure an optimal expression level, as demonstrated previously^{25,26}.

DNA assembly and cloning for capturing iModulon

The bacterial artificial chromosome (BAC) encoding BCAA biosynthetic iModulon was constructed by assembling PCR amplified pCAP_BAC (Addgene plasmid #120229)⁵⁴ fragments, *nptII* promoter, *pheS-nptII*, *thr*, *ilvG*, *ilvMEDA*, *ilvC*, and *leu* loci using transformation-associated recombination (TAR) cloning⁵⁵ in *Saccharomyces cerevisiae* strain VL6-48 (ATCC MYA-3666; MATα his3-Δ200 trp1-Δ1 ura3 – 52 lys2 ade2 – 1 met14 cir0). To perform TAR cloning, yeast VL6-48 was grown in 50 ml YPDA broth at 30 °C until A₆₀₀ of 0.4. Cells were collected by centrifuging at 3000 × *g* for 5 min. Then the cell pellet was washed with 10 ml of buffer 1 (100 mM LiAc, 10 mM Tris-HCl (pH8.0), and 1 mM EDTA (pH8.0)) and resuspended with 0.5 ml of buffer 1. Cell suspension (100 μl) was combined with 100 ng each of the DNA fragments, 100 μg of denatured salmon sperm DNA (Sigma-Aldrich, D9156), and 600 μl of buffer 2 (100 mM LiAc, 10 mM Tris-HCl (pH8.0), 1 mM EDTA (pH8.0), 40% (w/v) PEG-3350 (Sigma-Aldrich, P4338)). The mixture was mixed thoroughly and incubated at 30 °C for 30 min. Then, the mixture was incubated at 42 °C for 15 min after the addition of 70 μl of DMSO. The cell was harvested by centrifugation at 16,000 × *g* for 1 min after incubation on ice for 1 min. Liquid supernatant was discarded, and the cell pellet was resuspended with 1 ml of YPDA broth. After incubation at 30 °C for 2 h, the cell was harvested again with centrifugation at 16000 × *g* for 1 min and resuspended with 100 μl of sterile water. Successful clones were screened on a solid synthetic His dropout medium. A spacer array targeting both the landing pad and the BAC was constructed by annealing of two primers, RX4_1_F and RX4_1_R followed by primer extension using RX4_2_F and RX4_2_R. The spacer array was cloned in a pUC plasmid containing the crRNA leader sequence. MdcR and AcoR iModulons were PCR amplified from *P. aeruginosa* and *P. putida* genomic DNA. DNA fragments had 15–20 nt homology to pTrcHis2A (Invitrogen, V36520) plasmid and cloned into the linearized plasmid backbone using In-Fusion Cloning Kit (Takara Bio, 638948). Two converging operons encoding VanR iModulons were PCR amplified separately from *P. putida* genomic DNA and cloned into pTrcHis2A plasmid backbone in three-fragments ligation reaction using In-Fusion Cloning Kit resulting in pVanR_iM. AmpC iModulon was PCR amplified as six different fragments from *P. aeruginosa* genomic DNA. Together with the Trc promoter fragment amplified from pTrcHis2A plasmid and pCAP_BAC, entire DNA fragments were

assembled using TAR cloning. Assembled BAC was extracted using Gentra Puregene Yeast/Bact. Kit (Qiagen, 158567). Plasmids and BACs were electro-transformed to *E. coli* NEB10 β (New England Biolabs, C3020K) for propagation and MG1655 for downstream experiments. Primer sequences for PCR amplification were summarized in Supplementary Table 2. All the plasmids were sequence confirmed by whole-plasmid sequencing (Primordium Labs).

REXER

The refactored BCAA biosynthetic iModulon was introduced into the genome using the replicon excision enhanced recombination (REXER) method^{56,57} with minor modification. First, BAC_BCAA_Biosynthetic_iM plasmid containing BCAA biosynthetic iModulons was introduced using electroporation to *E. coli* Δ BCAA strain carrying pREDCas9 plasmid and *sacB-cat* dual-selection landing pad upstream of *thrl* gene. The recipient cell was grown in terrific broth (TB) containing 25 μ g/ml chloramphenicol, 50 μ g/ml kanamycin, 100 μ g/ml carbenicillin, and 30 mM L-arabinose at 37 °C to cell density (A_{600}) of 0.8. Two micrograms of PCR-amplified linear guide-RNA construct, targeting both the landing pad and the BAC, were introduced using electroporation, and cells were screened on NaCl-free LB-agar plate (10 g/l tryptone, 5 g/l yeast extract, and 1.5% agar) containing 100 μ g/ml carbenicillin, 50 μ g/ml kanamycin, and 10% sucrose. Successful recombinants were further screened by chloramphenicol sensitivity and PCR genotyping.

HPLC

Culture supernatant was collected by filtering 200 μ l of the crude culture through a 96-Well PVDF Filtration Plate (0.2 μ m; Agilent, 203980-100). Exo-metabolites were analyzed from 10 μ l of the sample using 1260 Infinity II HPLC System (Agilent) equipped with Multi-sampler (G7167A), HIP Degasser (G4225A), Binary Pump (G1312C), and Refractive Index Detector (G1362A), Thermostatted Column Compartment (G1316A), and Aminex HPX-87H HPLC Column (300 \times 7.8 mm; BioRad, 1250140). Five millimolar sulfuric acid was used as a mobile phase, and the detector temperature was maintained at 30 °C. Vanillate, protocatechuate, acetoin, and 2,3-butanediol were detected at a column temperature of 65 °C with a mobile phase flow rate of 0.6 ml/min. For detection of malonate column temperature was maintained at 45 °C.

Quantitative PCR

To measure the plasmid-to-chromosome ratio and the expression level of the *mdcR* iModulon construct, cells were grown in 15 ml M9 malonate medium and sampled at the mid-log phase. Total DNA was extracted from 1 ml of the culture using a Quick-DNA Miniprep Kit (Zymo Research, D3024) as instructed by the manufacturer. 100 ng of DNA extract was subject to quantitative PCR in a 20 μ l reaction containing AccuPower PCR 2 \times Master Mix (Bioneer, K-2018), 10 μ M each of primers, SYBR Green I Nucleic Acid Gel Stain (Invitrogen, S7563). Fluorescence signals were monitored by the CFX Duet Real-Time PCR System (BioRad, 12016265). The amount of beta-lactamase (*bla*) gene on the plasmid was compared to the *alaA* gene on the chromosome to estimate the relative number of the plasmid-to-chromosome. PCR efficiencies were measured from three twofold dilution series of each gene. Plasmid-to-chromosome ratio (P/C ratio) is calculated by

$$P/C \text{ ratio} = \frac{E_{alaA}^{Cq-alaA}}{E_{bla}^{Cq-bla}} \quad (1)$$

where E_{alaA} and E_{bla} are the PCR efficiencies of *alaA* gene and *bla* gene, respectively. $Cq-alaA$ and $Cq-bla$ are the quantification of cycles of respective genes.

To estimate the expression level of MdcR iModulon, the relative expression level of *mdcA* to 16S rRNA was measured. RNA was extracted from 14 ml of the culture using Quick-RNA Fungal/Bacterial

Miniprep Kit (Zymo Research, R2014) as instructed by the manufacturer. Residual DNA was removed by incubating 2 μ g of RNA extract at 37 °C for 30 min in a 50 μ l reaction containing 2 U of RNase-free DNase I (New England Biolabs, M0303) followed by a purification using RNA Clean & Concentrator Kit (Zymo Research, R1013). cDNA was synthesized from 300 ng of the DNA-depleted RNA sample using the SuperScript II First-Strand Synthesis System (Invitrogen, 11904018) as instructed by the manufacturer. About 1 μ l of cDNA synthesis reaction was subject to quantitative PCR in a 20 μ l reaction containing AccuPower PCR 2 \times Master Mix, 10 μ M each of primers, SYBR Green I Nucleic Acid Gel Stain. The relative amount of *mdcA* transcript was quantified by the $\Delta\Delta Cq$ method using *rrsA* transcript as a reference, after adjusting PCR efficiencies that were measured from three twofold dilution series of each gene. All the primers were designed by Primer-BLAST⁵⁸ to have no predicted cross-reactivity. Primer sequences are summarized in Supplementary Table 2.

Antibiotics sensitivity assay

For disc diffusion assay, overnight grown *E. coli* culture was diluted to 5×10^8 cells/ml and 500 μ l of the diluted culture was spread on an LB-agar plate. Thirty microliters of ampicillin solutions with different concentrations were dropped on sterilized filter paper disks (9 mm diameter; Sigma-Aldrich, 1703932). After drying for 15 min, the disks were placed on the LB-agar plate and the plate was incubated at 37 °C overnight. For the dose-killing assay, overnight grown *E. coli* culture was diluted to 5×10^8 cells/ml in LB medium containing an appropriate concentration of ampicillin, and 100 μ l of the diluted culture was transferred to a 96-well microplate. Cells were incubated at 37 °C with agitation in Infinite 200 Pro microplate reader (Tecan) and A_{600} was monitored every 15 min up to 10 h.

Adaptive laboratory evolution

E. coli K-12 MG1655 carrying pMdcR_iM plasmid was evolved via serial propagation of 150 μ l into 15 ml M9 malonate (2 g/l) minimal medium containing 100 μ g/ml carbenicillin in three biologically replicated cultures. Cultures were incubated at 37 °C, aerated by magnetic stirring, and reinoculated at A_{600} of 0.3 (Tecan Sunrise microplate reader; equivalent to an A_{600} of 0.750 on a conventional spectrophotometer with a path length of 10 mm) using an automated system. After 40 days of propagation, evolved cultures were subjected to a malonate utilization assay, and three clones were isolated from each culture.

DNA sequencing and analysis

Genomic DNA was isolated using Quick-DNA Miniprep Kit as instructed by the manufacturer. Whole-genome DNA-seq libraries were generated with a NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs, E7645) and run on an Illumina NovaSeq X Plus with 100 cycles pair-ended recipe. The sequencing results were processed with the in-house pipeline⁵⁹ that incorporates the BreSeq pipeline to identify mutations⁶⁰. *E. coli* K-12 MG1655 genome sequence (National Center for Biotechnology Information accession no. NC_000913.3) was used as a reference sequence.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The genome resequencing data generated in this study have been deposited in the EMBL Nucleotide Sequence Database (ENA) under accession code [PRJEB69310](https://doi.org/10.1038/s41467-024-46486-3). The iModulon activity and gene membership data used in this study are available in the iModulonDB [<https://imodulondb.org/>]. Predicted operonic structures were from the Biocyc database [<https://www.biocyc.org/>]. Genome sequence and genetic organization information were assessed through the

Pseudomonas Genome Database [<https://www.pseudomonas.com/>]. All other data are available in the article, Supplementary items, or Source Data file. Source data are provided with this paper.

Code availability

The mutation detection from the DNA sequencing results was performed with the ALEdb (v1.1.0) pipeline [<https://www.aledb.org/>] that incorporates the BreSeq pipeline^{59,60}.

References

- Campos, M. et al. Genomewide phenotypic analysis of growth, cell morphogenesis, and cell cycle events in *Escherichia coli*. *Mol. Syst. Biol.* **14**, e7573 (2018).
- Griffin, J. E. et al. High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. *PLoS Pathog.* **7**, e1002251 (2011).
- Tang, Y. & Gomer, R. H. An improved shotgun antisense method for mutagenesis and gene identification. *Biotechniques* **68**, 163–165 (2020).
- Bertea, C. M. et al. Identification of intermediates and enzymes involved in the early steps of artemisinin biosynthesis in *Artemisia annua*. *Planta Med.* **71**, 40–47 (2005).
- Dietrich, J. A. et al. A novel semi-biosynthetic route for artemisinin production using engineered substrate-promiscuous P450(BM3). *ACS Chem. Biol.* **4**, 261–267 (2009).
- Paddon, C. J. & Keasling, J. D. Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. *Nat. Rev. Microbiol.* **12**, 355–367 (2014).
- Blin, K. et al. antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic Acids Res.* **51**, W46–W50 (2023).
- Paley, S. M. & Karp, P. D. Evaluation of computational metabolic-pathway predictions for *Helicobacter pylori*. *Bioinformatics* **18**, 715–724 (2002).
- Nielsen, A. A. K. et al. Genetic circuit design automation. *Science* **352**, aac7341 (2016).
- Sastry, A. V. et al. The *Escherichia coli* transcriptome mostly consists of independently regulated modules. *Nat. Commun.* **10**, 5536 (2019).
- Lamoureux, C. R. et al. A multi-scale expression and regulation knowledge base for *Escherichia coli*. *Nucleic Acids Res.* **51**, 10176–10193 (2023).
- Rychel, K. et al. iModulonDB: a knowledgebase of microbial transcriptional regulation derived from machine learning. *Nucleic Acids Res.* **49**, D112–D120 (2021).
- Saelens, W., Cannoodt, R. & Saeys, Y. A comprehensive evaluation of module detection methods for gene expression data. *Nat. Commun.* **9**, 1090 (2018).
- Shin, J., Rychel, K. & Palsson, B. O. Systems biology of competency in *Vibrio natriegens* is revealed by applying novel data analytics to the transcriptome. *Cell Rep.* **42**, 112619 (2023).
- Lim, H. G. et al. Machine-learning from *Pseudomonas putida* KT2440 transcriptomes reveals its transcriptional regulatory network. *Metab. Eng.* **72**, 297–310 (2022).
- Rajput, A. et al. Advanced transcriptomic analysis reveals the role of efflux pumps and media composition in antibiotic responses of *Pseudomonas aeruginosa*. *Nucleic Acids Res.* **50**, 9675–9688 (2022).
- Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).
- Karp, P. D. et al. The BioCyc collection of microbial genomes and metabolic pathways. *Brief. Bioinform.* **20**, 1085–1093 (2019).
- Molina-Henares, M. A. et al. Functional analysis of aromatic biosynthetic pathways in *Pseudomonas putida* KT2440. *Microb. Biotechnol.* **2**, 91–100 (2009).
- Assinder, S. J. & Williams, P. A. The TOL plasmids: determinants of the catabolism of toluene and the xylenes. *Adv. Microb. Physiol.* **31**, 1–69 (1990).
- Maderbocus, R. et al. Crystal structure of a *Pseudomonas* malonate decarboxylase holoenzyme hetero-tetramer. *Nat. Commun.* **8**, 160 (2017).
- Wada, A. et al. Characterization of aromatic acid/proton symporters in *Pseudomonas putida* KT2440 toward efficient microbial conversion of lignin-related aromatics. *Metab. Eng.* **64**, 167–179 (2021).
- Brunel, F. & Davison, J. Cloning and sequencing of *Pseudomonas* genes encoding vanillate demethylase. *J. Bacteriol.* **170**, 4924–4930 (1988).
- Yoshihara, E., Yoneyama, H., Ono, T. & Nakae, T. Identification of the catalytic triad of the protein D2 protease in *Pseudomonas aeruginosa*. *Biochem. Biophys. Res. Commun.* **247**, 142–145 (1998).
- Nagy-Staron, A. et al. Local genetic context shapes the function of a gene regulatory network. *Elife* **10**, e65993 (2021).
- Boulas, I. et al. Assessing in vivo the impact of gene context on transcription through DNA supercoiling. *Nucleic Acids Res.* **51**, 9509–9521 (2023).
- Vermaas, J. V. et al. Passive membrane transport of lignin-related compounds. *Proc. Natl Acad. Sci. USA* **116**, 23117–23123 (2019).
- Pugh, S., McKenna, R., Osman, M., Thompson, B. & Nielsen, D. R. Rational engineering of a novel pathway for producing the aromatic compounds p-hydroxybenzoate, protocatechuate, and catechol in *Escherichia coli*. *Process Biochem.* **49**, 1843–1850 (2014).
- Naas, T. et al. Beta-lactamase database (BLDB) - structure and function. *J. Enzym. Inhib. Med. Chem.* **32**, 917–919 (2017).
- Bagge, N. et al. Constitutive high expression of chromosomal beta-lactamase in *Pseudomonas aeruginosa* caused by a new insertion sequence (IS1669) located in *ampD*. *Antimicrob. Agents Chemother.* **46**, 3406–3411 (2002).
- Ibache-Quiroga, C., Oliveros, J. C., Couce, A. & Blázquez, J. Parallel evolution of high-level aminoglycoside resistance in under low and high mutation supply rates. *Front. Microbiol.* **9**, 427 (2018).
- Tashiro, Y., Eida, H., Ishii, S., Futamata, H. & Okabe, S. Generation of small colony variants in biofilms by *Escherichia coli* harboring a conjugative F plasmid. *Microbes Environ.* **32**, 40–46 (2017).
- Moya, B. et al. β -Lactam resistance response triggered by inactivation of a nonessential penicillin-binding protein. *PLoS Pathog.* **5**, e1000353 (2009).
- Takebayashi, Y., Taylor, E. S., Whitwam, S. & Avison, M. B. CreC sensor kinase activation enhances growth of *Escherichia coli* in the presence of cephalosporins and carbapenems. *Antimicrob. Agents Chemother.* **63**, e00846–19 (2019).
- Zamorano, L. et al. The *Pseudomonas aeruginosa* CreBC two-component system plays a major role in the response to β -lactams, fitness, biofilm growth, and global regulation. *Antimicrob. Agents Chemother.* **58**, 5084–5095 (2014).
- Huang, H.-H. et al. Expression and functions of CreD, an inner membrane protein in *Stenotrophomonas maltophilia*. *PLoS ONE* **10**, e0145009 (2015).
- Fukushima, K., Kumar, S. D. & Suzuki, S. YgiW homologous gene from *Pseudomonas aeruginosa* 25W is responsible for tributyltin resistance. *J. Gen. Appl. Microbiol.* **58**, 283–289 (2012).
- Moreira, C. G. et al. Virulence and stress-related periplasmic protein (VisP) in bacterial/host associations. *Proc. Natl Acad. Sci. USA* **110**, 1470–1475 (2013).
- Murtha, A. N. et al. High-level carbapenem tolerance requires antibiotic-induced outer membrane modifications. *PLoS Pathog.* **18**, e1010307 (2022).
- Ali, N. O., Bignon, J., Rapoport, G. & Debarbouille, M. Regulation of the acetoin catabolic pathway is controlled by sigma L in *Bacillus subtilis*. *J. Bacteriol.* **183**, 2497–2504 (2001).

41. Liu, Q. et al. 2,3-Butanediol catabolism in *Pseudomonas aeruginosa* PAO1. *Environ. Microbiol.* **20**, 3927–3940 (2018).
42. Liu, Y. et al. Dehydrogenation mechanism of three stereoisomers of butane-2,3-diol in *Pseudomonas putida* KT2440. *Front. Bioeng. Biotechnol.* **9**, 728767 (2021).
43. Priefert, H. et al. Identification and molecular characterization of the *Alcaligenes eutrophus* H16 *aco* operon genes involved in acetoin catabolism. *J. Bacteriol.* **173**, 4056–4071 (1991).
44. Romero, P. R. & Karp, P. D. Using functional and organizational information to improve genome-wide computational prediction of transcription units on pathway-genome databases. *Bioinformatics* **20**, 709–717 (2004).
45. Xiao, Z. J., Huang, Y. L., Zhu, X. K. & Qiao, S. L. Functional study of AcoX, an unknown protein involved in acetoin catabolism. *Adv. Mat. Res.* **393–395**, 776–779 (2011).
46. Schaffitzel, C., Berg, M., Dimroth, P. & Pos, K. M. Identification of an Na⁺-dependent malonate transporter of *Malonomonas rubra* and its dependence on two separate genes. *J. Bacteriol.* **180**, 2689–2693 (1998).
47. Kumari, S., Tishel, R., Eisenbach, M. & Wolfe, A. J. Cloning, characterization, and functional expression of *acs*, the gene which encodes acetyl coenzyme A synthetase in *Escherichia coli*. *J. Bacteriol.* **177**, 2878–2886 (1995).
48. LaCroix, R. A. et al. Use of adaptive laboratory evolution to discover key mutations enabling rapid growth of *Escherichia coli* K-12 MG1655 on glucose minimal medium. *Appl. Environ. Microbiol.* **81**, 17–30 (2015).
49. Kogoma, T. & Maldonado, R. R. DNA polymerase I in constitutive stable DNA replication in *Escherichia coli*. *J. Bacteriol.* **179**, 2109–2115 (1997).
50. Yang, Y. L. & Polisky, B. Suppression of ColE1 high-copy-number mutants by mutations in the *polA* gene of *Escherichia coli*. *J. Bacteriol.* **175**, 428–437 (1993).
51. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl Acad. Sci. USA* **97**, 6640–6645 (2000).
52. Fredens, J. et al. Total synthesis of *Escherichia coli* with a recoded genome. *Nature* **569**, 514–518 (2019).
53. Winsor, G. L. et al. Enhanced annotations and features for comparing thousands of *Pseudomonas* genomes in the *Pseudomonas* genome database. *Nucleic Acids Res.* **44**, D646–D653 (2016).
54. Bauman, K. D. et al. Refactoring the cryptic streptophenazine biosynthetic gene cluster unites phenazine, polyketide, and non-ribosomal peptide biochemistry. *Cell Chem. Biol.* **26**, 724–736.e7 (2019).
55. Gietz, R. D. Yeast transformation by the LiAc/SS carrier DNA/PEG method. *Methods Mol. Biol.* **1205**, 1–12 (2014).
56. Wang, K. et al. Defining synonymous codon compression schemes by genome recoding. *Nature* **539**, 59–64 (2016).
57. Robertson, W. E. et al. Creating custom synthetic genomes in *Escherichia coli* with REXER and GENESIS. *Nat. Protoc.* **16**, 2345–2380 (2021).
58. Ye, J. et al. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinforma.* **13**, 134 (2012).
59. Phaneuf, P. V., Gosting, D., Palsson, B. O. & Feist, A. M. ALEdb 1.0: a database of mutations from adaptive laboratory evolution experimentation. *Nucleic Acids Res.* **47**, D1164–D1171 (2018).
60. Deatherage, D. E. & Barrick, J. E. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods Mol. Biol.* **1151**, 165–188 (2014).

Acknowledgements

This work was funded by the Novo Nordisk Foundation grant NNF20CC0035580 and the Y.C. Fung Endowed Chair in Bioengineering at UC San Diego to B.O.P. We thank Ms. Emily Armbruster and Dr. Joe Pogliano for providing *Pseudomonas aeruginosa* strain K2733. We thank Marc Abrams for editing the manuscript.

Author contributions

B.O.P. conceived and supervised the study. D.C., A.M.F., and B.O.P. designed the experiments. D.C., C.A.O., R.S., H.Y. and J.S. performed the experiments. D.C., A.M.F. and B.O.P. analyzed the data. D.C., A.M.F. and B.O.P. wrote the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-46486-3>.

Correspondence and requests for materials should be addressed to Bernhard O. Palsson.

Peer review information *Nature Communications* thanks Kechun Zhang, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024