Title

Systems biology of competency in Vibrio natriegens is revealed by applying novel data analytics to the transcriptome

Journal

Cell Reports, 42(6)

ISSN

2639-1856

Authors

Shin, Jongoh

Rychel, Kevin

Palsson, Bernhard O

Publication Date

2023-06-01

DOI

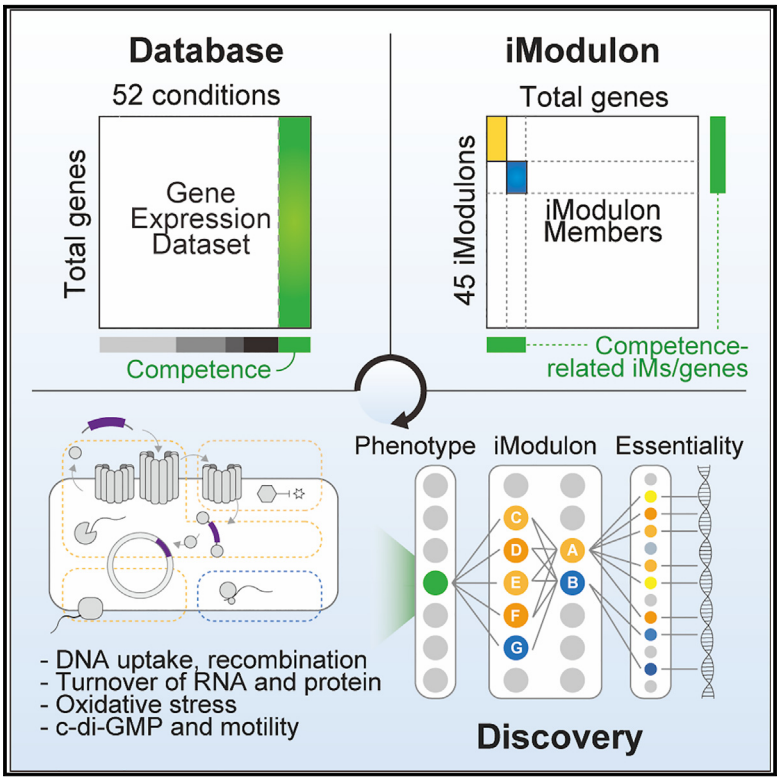
10.1016/j.celrep.2023.112619

Cell Reports

Resource

Systems biology of competency in *Vibrio natriegens* is revealed by applying novel data analytics to the transcriptome

Graphical abstract



Authors

Jongoh Shin, Kevin Rychel, Bernhard O. Palsson

Correspondence

palsson@ucsd.edu

In brief

Shin et al. generate a transcriptome database, which includes various conditions including competence, and identify the significance of diverse regulons in competence by utilizing machine learning and functional screening. This study can serve as a valuable resource for further studies on natural competence.

Highlights

- Transcriptomic and iModulon database with various conditions including competence
- iModulon-discovery cycle highlights importance of various regulons for competence
- ROS formation and inhibition affect natural competence

Shin et al., 2023, Cell Reports *42*, 112619
June 27, 2023
https://doi.org/10.1016/j.celrep.2023.112619

CellPress

Cell Reports

CellPress
OPEN ACCESS

Resource

Systems biology of competency in *Vibrio natriegens* is revealed by applying novel data analytics to the transcriptome

Jongoh Shin,[1] Kevin Rychel,[1] and Bernhard O. Palsson[1,2,3,4,*]
[1]Department of Bioengineering, University of California San Diego, La Jolla, CA 92093, USA
[2]Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 2800 Lyngby, Denmark
[3]Department of Pediatrics, University of California, San Diego, La Jolla, CA, USA
[4]Lead contact
*Correspondence: palsson@ucsd.edu
https://doi.org/10.1016/j.celrep.2023.112619

SUMMARY

Vibrio natriegens regulates natural competence through the TfoX and QstR transcription factors, which are involved in external DNA capture and transport. However, the extensive genetic and transcriptional regulatory basis for competency remains unknown. We used a machine-learning approach to decompose *Vibrio natriegens*'s transcriptome into 45 groups of independently modulated sets of genes (iModulons). Our findings show that competency is associated with the repression of two housekeeping iModulons (iron metabolism and translation) and the activation of six iModulons; including TfoX and QstR, a novel iModulon of unknown function, and three housekeeping iModulons (representing motility, polycations, and reactive oxygen species [ROS] responses). Phenotypic screening of 83 gene deletion strains demonstrates that loss of iModulon function reduces or eliminates competency. This database-iModulon-discovery cycle unveils the transcriptomic basis for competency and its relationship to housekeeping functions. These results provide the genetic basis for systems biology of competency in this organism.

INTRODUCTION

Natural competence for transformation is a process by which prokaryotes import extracellular DNA from their environment and recombine it into their genome.[1,2] *Vibrio cholerae* (*Vc*) and other *Vibrio* species can enter a state of competence under specific conditions.[3,4] Recent studies have shown that a complex transcriptional regulatory network (TRN) of competence in *Vc* strains is activated by environmental signals[4–9]: starvation (detected by the cyclic adenosine monophosphate [cAMP]-cAMP receptor protein complex [CRP]), presence of chitin (detected by ChiS and TfoS), high cell density (detected by the quorum sensing system), and high concentration of cytidine (detected by CytR). The detection of the presence of chitin induces the production of TfoX,[5] the main regulator of competence. TfoX overexpression was found to eliminate the necessity for chitin induction.[5,10,11] TfoX activates chitin catabolism, a set of competence genes including DNA-uptake pilus complex (type IV pilus),[5,6] and a second major competence regulator, QstR.[12] Under high cell density conditions, HapR, together with TfoX, induces the expression of QstR,[12,13] which directly induces the expression of a group of known competence-related genes, such as the inner membrane translocator ComEC, the cytoplasm DNA translocator ComF, and the hexameric helicase ComM.[12,14] Under the absence of cytidine conditions, CytR can promote the expression of DNA-uptake pilus genes.[15] Additional studies further revealed that the antibacterial type VI secretion system (T6SS) is part of TfoX regulon in *Vc*.[14]

Recent studies have demonstrated that TfoX-based competency can be activated in seven other *Vibrio* species besides *Vc* (Table S1).[6,11] Among these, *V. natriegens* (*Vn*) has emerged as a novel chassis organism for molecular cloning and biotechnological applications.[16] Its non-pathogenic nature, rapid growth, versatile metabolism, and genetic tools render *Vn* a promising model organism suitable for various applications, such as producing diverse chemicals, proteins, biopolymers, and nanoparticles.[17,18] Furthermore, its natural competence offers an alternative approach for genome engineering[10,19] and strain enhancement through adaptation and evolution facilitated by horizontal gene transfer. However, a comprehensive understanding of *Vn*'s metabolism, cell functions (including competency), and regulatory mechanisms remains limited, with current TRN knowledge derived from well-studied relatives like *Vc* or *Vibrio fischeri*.

Genome-scale TRN reconstruction requires extensive experiments for probing binding sites and measuring transcriptional regulator activities.[20] Independent component analysis (ICA) has addressed this challenge by efficiently identifying quantitative strain-specific TRNs from transcriptomic datasets alone. ICA decomposes the mixed signals that comprise the transcriptomic dataset into independently modulated groups of genes (iModulons) and their condition-dependent activity levels.

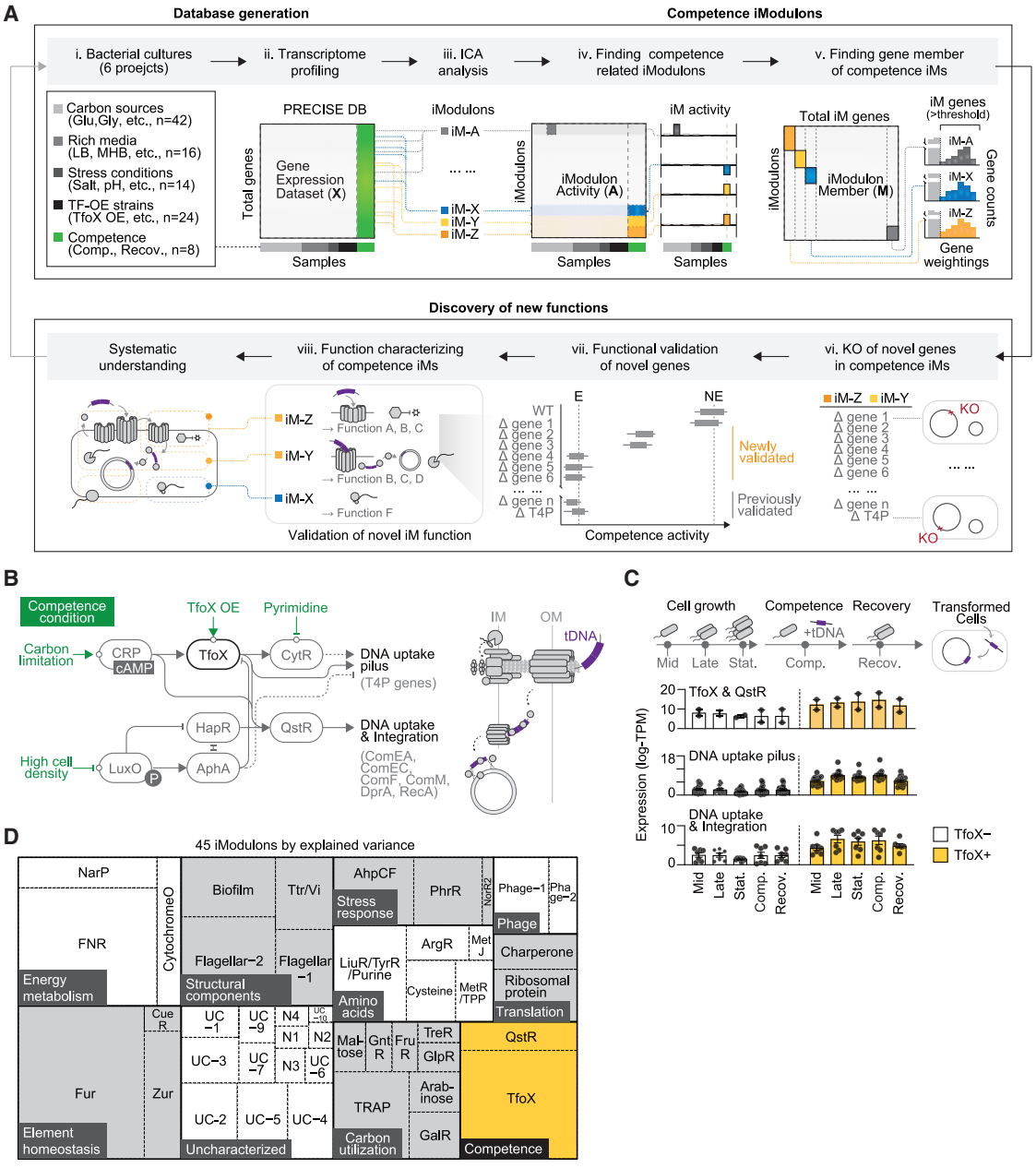


Figure 1. Database-iModulon-discovery cycle in *Vn*

(A) Overall approach of this study. (i and ii) To determine iModulons in *Vn*, we generated 104 RNA-seq datasets, including the competence condition. Each column in the X matrix is an expression profile from a specific sample. (iii) We performed ICA to decompose the X matrix into two matrices, M and A, in each sample. The relationship between the three matrices is X = M × A. (iv) The A matrix contains the condition-dependent activities of the iModulons. Using the A matrix, we found competence iModulons that showed significant differences in activity in the samples obtained under competency conditions compared with the control samples. (v) Each iModulon in the M matrix contains a gene weighting for total genes. Using the M matrix, we could find competence genes, which make up the competence iModulons. (vi and vii) Based on the genes constituting competency iModulons, new genes or operons are discovered that have not been previously reported to contribute to the natural competency process. E, essential for natural transformation; NE, non-essential for natural transformation. (viii) The function of competency iModulons for natural transformation is further evaluated.

(B) Overall transcriptional regulation of *Vn*'s natural competence was estimated by comparative genomic analysis with previously reported competency regulons. tDNA, transforming DNA.

(C) The change of transcript level of competency genes during cell growth and competency assay. Overview of samples during cell growth, competence, and recovery steps (top panel). The expression level of TFs (TfoX, QstR), DNA uptake (type IV DNA pilus and minor pilins), and integration-related genes (*comEC*,

(legend continued on next page)

As the most effective signal deconvolution algorithm[21] for TRN inference[22] among 42 examined, ICA offers advantages like unsupervised learning, noise robustness, quantitative analysis, novel regulon discovery, and applicability to various organisms. It has been used to establish strain-specific TRNs for model[23,24] and understudied bacteria,[25–29] enabling deep TRN studies, including iModulon-regulon relationships and the discovery of novel regulons and stimulons.[23,24,29] However, previous ICA studies primarily focus on species with extensive transcriptomic data and well-characterized regulons.

Here, we conducted the entire ICA-based database-iModulon-discovery cycle to elucidate *Vn*'s global TRN and genetic basis, involving transcriptomic database generation, ICA decomposition, and knockout screening. With a new database of 104 RNA sequencing (RNA-seq) samples, ICA identified 45 iModulons and their activity levels across various conditions. Focusing on competence iModulons, we discovered iModulons for TfoX and QstR that were specifically activated under competence conditions. We identified six additional competence-related iModulons and the list of genes that they represent. Finally, we show that competent cells globally orchestrate iModulons to enhance antioxidant defense mechanisms, highlighting the crucial role of regulating oxidative stress levels during the competence process.

RESULTS

Transcriptomic data generation for iModulon identification

Our objective was to identify iModulons involved in competency using the database-iModulon-discovery cycle (Figure 1A). We first generated competence-associated RNA-seq samples of *Vn*, inducing competence in the TfoX+ strain (Vn-PTrc-TfoX strain showing overexpression of TfoX; Figures S1A–S1D) at a stationary phase (i.e., high cell density condition) with transforming DNA (tDNA) and IPTG (Figure 1B). Transformation efficiency (TE) values were determined by calculating the ratio of colony-forming units (CFUs) observed on selective media to those observed on non-selective media (see STAR Methods). The competence samples showed 2.69 × 10^{-3} average TE in *Vn* (Figure S1E), a similar level to that of other *Vibrio* strains.[10,11] The transcript levels of the DNA-uptake pilus and DNA-uptake integration genes were highly increased (average $\log_2$ 8.7-fold and $\log_2$ 4.4-fold increases, respectively) in the TfoX+ strain under the competence condition compared with the TfoX– strain (*Vn*-pTrc strain containing backbone vector) and then subsequently decreased (0.05- and 0.21-fold, respectively) under the recovery condition (rich media, low cell density, and no IPTG condition) (Figure 1C).

iModulon analysis performs best when it can learn transcriptional modules from a diverse condition space. Therefore, to collect a wide range of transcriptomes, we generated new RNA-seq datasets from *Vn* exposed to various conditions (Table S1). We also collected 42 *Vn* RNA-seq profiles available from public sources. Using an ICA pipeline,[30] we compiled a high-quality transcriptome compendium (Pearson's r = 0.98 between replicates) containing 104 RNA-seq samples (newly generated 96 samples + 8 public data) from 52 unique conditions across 6 projects (Figures S2A–S2C). These samples included the use of 16 carbon sources (carbon project), cell growth under 4 rich media and the expression of 4 competence-related transcription factors (TFs; rich/TF-overexpression project), 5 stress conditions (stress I, II, and III projects), and 2 competence-related conditions (competence project) (Table S1).

We named this compendium natPRECISE104 (natriegens precision RNA-seq expression compendium for independent signal exploration containing 104 profiles), following previous transcriptomic compendia naming conventions (see https://imodulondb.org/). ICA of this transcriptome compendium (X matrix) yielded 45 iModulons (M matrix), revealing their activities under each condition (A matrix)[23] (Figures 1A and 1D; Table S2). The iModulons, representing statistically independent gene expression signals, included 851 genes (19% total coding sequences) and explained 68.6% of the variation in natPRECISE104's gene expression (Figure S2D).

While principle-component analysis (PCA) captures variance in a dataset, it lacks direct biological interpretation. In contrast, ICA extracts conserved, reproducible regulatory signals annotated with genetic, biochemical, and biological information.[31] Thus, the 68.6% variation explained with iModulons corresponds to biological interpretation.

iModulon classification

The 45 iModulons were grouped into three categories (regulatory, functional, and uncharacterized) as previously described.[23,24,30] First, "regulatory" iModulons that statistically overlapped with previously known regulons derived from traditional bottom-up experiments were identified (see STAR Methods and Table S2). The 17 regulatory iModulons identified (37% of total) were consistent with known regulons in *Vc* (e.g., ferric uptake regulator R [FurR] and TreR) (Figure S2E; Table S2). Following an established procedure,[30] the relationship between the regulatory iModulons and linked regulons was evaluated by quantifying the fraction of known regulon genes in the iModulon (iModulon recall [MR]) and the fraction of the regulon covered by the iModulon (regulon recall [RR]). Based on MR and RR values, the regulatory iModulons were classified as well matched, regulon subset, regulon discovery, or poorly matched[23,24] (Figure S2E).

Among the remaining 28 iModulons, 15 were "functional" (33% of total) based on Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment or functional annotation (e.g., ribosomal proteins and flagellar) (Table S2). The other 13 iModulons were classified as "uncharacterized," indicating unknown gene function or regulation. The iModulons were further categorized into ten types based on shared functions, providing

comEA, *comF*, *comM*, *priA*, *dprA*, and *recA*) at each sampling point (bottom panel). TfoX– and TfoX+ indicate Vn-PTrc and Vn-PTrc-TfoX strains, respectively. The bar plot and error bar indicate the mean and standard error of the mean values.
(D) Treemap of the 45 iModulons. The size of each rectangle indicates the proportion of explained variances of each iModulon.
See also Figures S1 and S2 and Table S1 and S2.

a systems-level overview (Figure 1D). Detailed information on the 45 iModulons can be found in the iModulon database (https://imodulondb.org/).[32] These iModulons represent the first quantitative reconstruction of the TRN in *Vn*.

Competency iModulons identified in *Vn*

After establishing the iModulon structure for the *Vn* transcriptome, we sought to identify the iModulons involved in competence by analyzing condition-dependent activities. In the TfoX+ strain, various iModulons were activated or repressed throughout the natural transformation process (Figure S3A). In the competence condition, traditional differentially expressed genes analysis identified 1,475 significantly regulated genes ($|\log_2$ fold change$| > 1$, adjusted p value [p_{adj}] < 0.05) in the TfoX+ strain compared with TfoX–, making the meaningful interpretation of transcriptomic changes difficult without ICA (Figure S3B).

Conversely, ICA revealed eight iModulons (explaining 23% RNA-seq data variance) showing significant activity differences in the competent state (false discovery rate [FDR] < 0.05, an activity threshold of 10) (Figure 2A). The activity comparison of all iModulons on all samples revealed that TfoX and QstR are "competence-specific" iModulons (Figure 2B), expressed during competence and recovery, and in corresponding TF-overexpression strains. The remaining six differentially activated iModulons, including UC-5 (uncharacterized-5), ArgR, Flagellar-1, AhpCF, Fur, and ribosomal protein iModulons, are categorized as "competence-related" iModulons (Figure 2B); the first four are activated, while the latter 2 are deactivated. The competence-related iModulons exhibited significant activity changes in non-competence conditions, such as in the UC-5 iModulon during various growth stages with different carbon sources (glucose, ethanol, glucosamine, and N-acetylglucosamine).

We then compared *Vn*'s competence-specific iModulons (TfoX and QstR) to *Vc*'s regulons. Unlike other regulons, the TfoX regulon of *Vc* was only reconstituted from RNA-seq data[7] and not from TF-gene interaction information (Table S2). The TfoX iModulon and its regulon shared 17 overlapping genes, including *tfoX*, DNA-uptake pilus, and *dprA* (Figures 2C and S3C). The QstR iModulon contained ten genes, including the *qstR*, *comEA*, *comF*, and *comM* genes[33] but not *comEC*. Although *Vn*'s QstR and TfoX iModulons do not perfectly match *Vc*'s regulons, the overlaps are significant (Table S2; FDRs $< 10^{-8}$), covering all known genes considered important for competence except for *comEC*.[31] These differences present opportunities for discovery, as iModulons detect expression-based signals, while regulons are defined by TF-binding sites, and this difference is observed for many iModulons (https://imodulondb.org/).[32]

Next, we evaluated the activity relationships between TfoX and QstR iModulons. As the QstR regulon is a sub-regulon of TfoX,[12,13] the QstR iModulon activity positively correlated and depended on TfoX iModulon activity (Pearson's R = 0.5, $p = 0.6 \times 10^{-9}$) (Figures 3D and S3D). We also identified a negative correlation (Pearson's R = −0.3 to 0.4) between ribosomal protein iModulons and TfoX/ArgR/AhpCF/UC-5 iModulons (Figure S3D). This correlation suggests resource allocation between natural competence and optimal cell growth, similar to the conserved fear (stress hedging) versus greed (growth) trade-off observed in many bacteria.[25,28,29]

Furthermore, we investigated the relationship between the competence iModulon's activity and the associated TF's expression. Consistent with the previously confirmed regulon in *Vc* (Figure S1C), the expression of positive regulators TfoX, HapR, QstR, and CytR was significantly increased (436.0-, 9.4-, 198-, and 4.1-fold increased, respectively) in competence cells (TfoX+) compared with the control (TfoX−) (Figure S3E), while the negative regulator AhpA was significantly decreased (0.3-fold decreased). The neutral TFs, which activate competence through conformational change by substrate binding (e.g., CRP) or phosphorylation (e.g., LuxO), maintained consistent expression (1.1- and 0.9-fold, respectively). Under the recovery condition, the regulation of positive and negative regulators disappeared (Figure S3E), indicating that QstR iModulon is coregulated by CRP/cAMP and HapR.[13,35,36] Interestingly, TfoX overexpression suppressed the *ahpA* expression for all samples (0.3- to 0.6-fold decreased), suggesting that TfoX iModulon genes negatively regulate *ahpA* expression.

Overall, while *Vn* iModulon and *Vc* regulon compositions differ, key competence genes and related TFs' expression in both systems are consistent.

Construction of gene deletion mutants in competency iModulons and evaluation of their phenotype

The presence of uncharacterized competency iModulons and uncharacterized genes in six activated competency iModulons highlights the potential for discovering unknown gene functions related to competence (Figures 2B–2C). Thus, we generated 83 deletion mutants targeting uncharacterized genes to investigate their role in natural transformation (Figures S4A–S4D). The mutants covered all genes in the TfoX, QstR, UC-5, and AhpCF iModulons except for 18 previously validated genes, such as DNA-uptake pilus and other known genes. Notably, the UC5-Δ1 mutant is included in both the TfoX and UC-5 iModulons since both contain the RS02830-RS02835 operon. As Flagellar-1 and ArgR iModulon functions are well known, we only generated one mutant for each (ArgR-Δ1 and Fla1-Δ1) by deleting the *argBCGH* or *fliEFGHIJ* operon, respectively. The ArgR-Δ1 and Fla1-Δ1 mutants exhibited arginine auxotroph in minimal medium and significantly reduced mobility, respectively (Figures S4E–4F).

Next, we measured the TE of all mutants (Figure 3A) with TfoX+ and TfoX− controls (Figure S4B) using tDNA-Kan^R, which contains ~2 kb homology of the *endA* gene and kanamycin resistance cassette. For intuitive explanation, the obtained TE was further standardized as relative TE (rTE; %) to represent the increase or decrease of TE compared with the wild-type TfoX+. The mutants were classified into three gene groups based on rTE: essential (<1% rTE; 6 mutant strains covering 11 genes), quasi-essential (≥1% and <50% rTE; 41 mutant strains covering 70 genes), and non-essential (≥50% rTE; 37 mutant strains covering 62 genes) (Figures 3A–3C). As expected, *pilABC* genes (TfoX-Δ1) were classified as essential and showed rTE values similar to the negative control, consistent with previous results in *Vc*.[37] *Vn*'s slight kanamycin resistance[34] may result in a few colonies with tDNA-Kan^R in negative controls (TfoX− and

CellPress
OPEN ACCESS

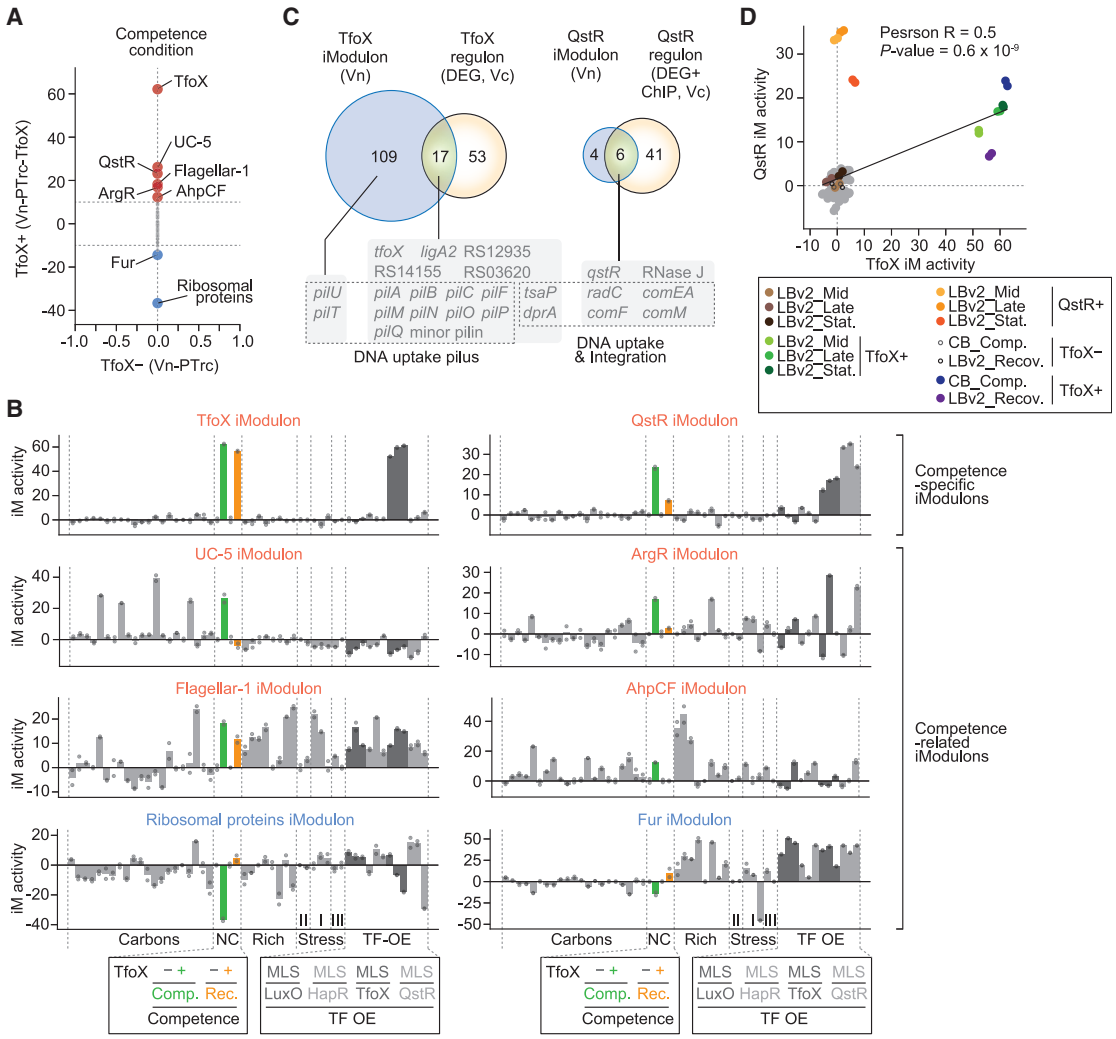


Figure 2. Discovery of competency iModulons

(A) Differential iModulon activity graph for the competence (TfoX+) and control samples (TfoX−) under the competence condition. The iModulons with a change in activity level >10 and FDR <0.05 for a given pair of conditions were considered significant.

(B) The activity of competency iModulons. iModulon activities in total samples from all projects. Gray dots represent individual samples (n = 2). Vertical gray lines separate different projects (carbons, NC [competence], rich, TF overexpression [OE], stress I [salinity/temperature], stress II [high salt, antibiotic], and stress III [microgravity]) in the natPRECISE104. Red and blue colors indicate iModulon significantly increasing or decreasing with iModulon activity under competence, respectively. iM, iModulon.

(C) Comparison of TfoX and QstR iModulons with the previously known regulons from the literature (RNA-seq and chromatin immunoprecipitation [ChIP]-seq). The numbers denote the number of genes contained in each subset.

(D) Activity correlation between the TfoX iModulon and QstR iModulon across all conditions. These two iModulons exhibited a positive correlation between their activity (Pearson's $R = 0.5$, $p < 0.6 \times 10^{-9}$). The plot indicated that the QstR iModulon was only activated under the QstR-OE (QstR+) and TfoX-OE strains (TfoX+). See also Figure S3 and Table S2.

TfoX-Δ1), resulting in a TE of $\sim 10^{-6}$[10]; however, these are not actual transformants (Figure S5A). rTE measurements using tDNA-Spec^R and spectinomycin, a substance that does not produce spontaneous resistant mutants, showed significant similarity (Pearson's $r = 0.89$, $p < 0.0001$), indicating high data reproducibility and reliability (Figures S5B and S5C). The lack of a correlation between specific growth rate and TE (Pearson's $r = 0.07$, $p < 0.73$) suggests that poor growth is not the cause of the low TE (Figures S5D).

Overall, our data indicated that at least 52% of the 83 deleted genes in competency iModulons significantly contribute to the competence process (Figure 3C). Essential genes for competence tend to be more significantly conserved in the other 17 *Vibrio* species than non-essential genes (Figure 3D). Additionally, the distribution of quasi-essential genes reveals the high frequency of core (10/41) and unique (8/41) genes, suggesting that the beneficial proteins for competence vary substantially at the strain level.

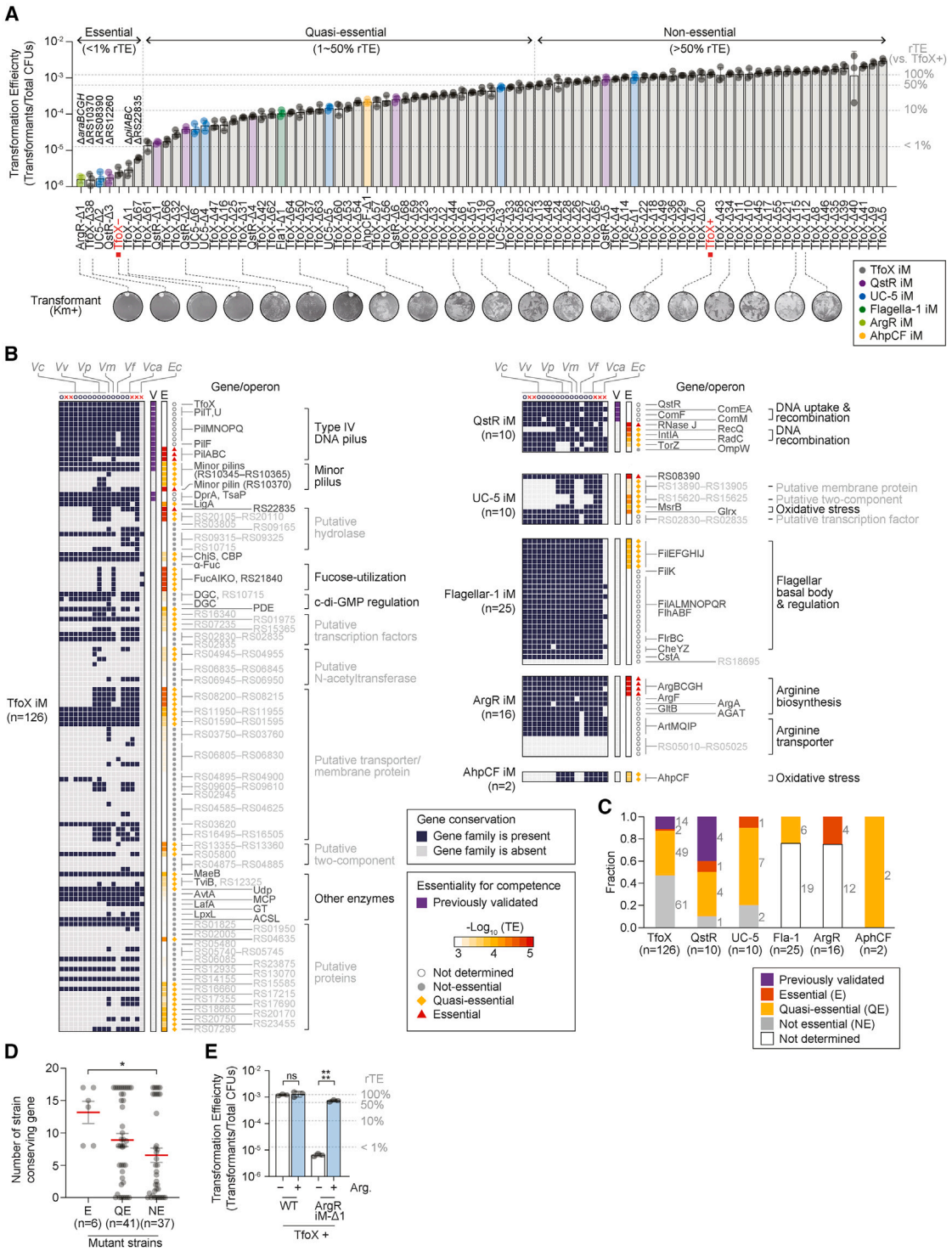


Figure 3. Determination of the essentiality of genes for natural transformation

(A) Transformation efficiency (TE) in 2 control (in red) and 83 mutant strains lacking a gene/operon in competency iModulons. The rTE values were calculated by normalizing TE with the positive control (TfoX+). Each dot indicates an individual value of TE of three independent biological replicates. The bar plot and error bar indicate mean and standard deviation values, respectively. Tested mutant strains were classified as essential (<1% rTE [relative TE]), quasi-essential (≥1% and <50% rTE), and non-essential (≥50% rTE) based on the rTE values. Please note that *Vn* has a low-level natural resistance to kanamycin.[34] When using kanamycin as a selective agent, only a few colonies may form in negative controls such as TfoX− and TfoX-Δ1, but they are not actual transformants (Figure S5A). Thus, the data suggested that there are actual transformants when the TE value is >2.6 × 10^{-6} (>0.2% rTE) while using kanamycin.

(legend continued on next page)

Essential genes for natural competence in *Vn*

In the TfoX, QstR, UC-5, and ArgR iModulons, we found eight essential genes in addition to known genes (Figure 3C). First, the one minor pilin gene of *Vn* (RS10370, TfoX-Δ38) is essential for competence (Figures 3A–3B). The minor pilins can promote pilus assembly initiation and non-specific DNA binding.[38] *Vn*'s minor pilin genes consist of one operon (for *Vc*, VC0858–VC0861; for *Vn*, RS10350–RS10365) and one gene (for *Vc*, VC0857; for Vn, RS10370) identical to *Vc*, but protein identities were relatively low (average 35.5%), except for one gene (VC0861 for *Vc*, 62.7%) (Figures S1B and S6A). Previous studies in *Vc* reported that all minor pilin mutants, including an operon[37,38] and another gene,[37] showed significantly reduced TE, indicating their critical role in DNA uptake. Our data classified the minor pilin operon (TfoX-Δ37, 9.55% rTE) as quasi-essential and another minor pilin gene (TfoX-Δ38, 0.13% rTE) as essential (Figures 3A–3B, S5E, and S5F). Although the rTE of TfoX-Δ37 measured with tDNA-Spec^R was much lower at 0.1% compared with the rTE measured with tDNA-Kan^R, it was still possible to confirm the formation of transformants in TfoX-Δ37, unlike in TfoX-Δ38 (Figures S5B, S5E, and S5F). This suggests that a single minor pilin protein may play a relatively more important role in linear tDNA binding in *Vn* compared with the proteins present in the minor pilin operon.

Additionally, we found that hypothetical haloacid dehydrogenase (HAD) family hydrolase (TfoX-Δ67, 0.46% rTE) was also essential for competence (Figures 3A–3B). The HAD family exhibits a wide range of biological functions, from amino acid biosynthesis to detoxification[39] (Figure S6B). However, the exact role of the HAD family during competence in *Vibrio* has not been elucidated.

In the QstR iModulon, RNase J (QstR-Δ3, 0.14% rTE), regulated by QstR,[7] was classified as essential (Figures 3A–3B). Since RNase J possesses both 5′-3′ exo- and endo-nuclease functions,[40] it can degrade structured RNA fragments more effectively than typical bacterial mRNA degradation by skipping the polyadenylation-assisted decay pathway, a mechanism similar to mRNA decay of *Bacillus subtilis*[41] (Figure S6C). Although RNase J contains β-lactamase and β-CASP domains without a C-terminal domain, previous results showed that the RNase J Δc-terminal domain still maintains endo- and exo-RNase activity.[40] Considering that the M48 family metallopeptidase (TfoX-Δ61, 1.11% rTE) was identified as quasi-essential, these findings suggest that facilitating mRNA and protein decay is essential for fast resource allocation during the competence process.

The ArgR-Δ1 strain (Δ*argBCGH*) showed the lowest TE (0.13% rTE) (Figures 3A–3B), suggesting that arginine biosynthesis is vital for competence. When arginine (1 mM) was supplemented during competence, the TE was restored to 60% rTE (Figure 3E). Arginine may be essential for the protein synthesis process during competence.

The UC-5 iModulon has one essential gene, RS08390, encoding a TVP38/TMEM64 family membrane protein (Figure S6D). The UC5-Δ2 strain lacking the RS08390 showed 0.14% rTE (Figures 3A–3B). While the protein's function is unclear, it is well conserved in the 17 *Vibrio* genomes, suggesting conserved roles for competence.

The competency-specific and -related iModulons provide a productive approach to reveal the genetic basis for competence. We can identify the essential genes and explore their individual functions. Some are known, and some need functional characterization. However, as far as the systems-level mechanisms are concerned, we can be understood by showing that iModulon inactivation leads to competence inactivation.

Quasi-essential genes for natural competence in *Vn*

The large number of quasi-essential genes in the competency iModulons is notable. Although 48% of proteins' functions (91/189) are unknown (66% of TfoX iModulon, 80% of UC-5 iModulon), knockout screening showed that various putative proteins are significantly involved in competence (Figure 3B). Based on known-function genes, we identified 68 quasi-essential genes involved in DNA processing, recombination, 3′,5′-cyclic diguanylate monophosphate (c-di-GMP) cycle, and multiple defense mechanisms against oxidative damage.

The TfoX and QstR iModulons contain one (*ligA*) and three (*radC*, *recQ*, and *intI*) quasi-essential genes for DNA processing and recombination, respectively. The *ligA* and *radC* genes were regulated by QstR[7] (Figure 2C). The *ligA* (RS06105; TfoX-Δ25, 5.07% rTE) encoding ATP-dependent DNA ligase was quasi-essential (Figures 3A–3B). Since bacteria typically use NAD-dependent DNA ligase (RS09615, *ligN*) for ligation (Figure S6E), *ligA*'s quasi-necessity in competence and the high conservation

(B) Gene essentiality for natural competence. The dark blue circle indicates *Vibrio* strains whose natural competence activity was previously validated by chitin- or TfoX-dependent activation. Gene presence or absence is presented in the left heatmap in dark blue and gray, respectively. The purple color shown in the "V" column indicates the previously validated competence gene in *Vibrio* species. Each TE value for a gene is shown as a heatmap in the "E" column. If genes are deleted as an operon, the TE values and essentiality of a mutant strain were equally shown in each gene. Putative and unknown function genes were grayed out with gene name or locus tag. The used *Vibrio* strains for genome comparison are shown in order from left to right as follows: 7 *Vibrio chorale* (*Vc*) strains (A1552, N16961, MO10, c6706, c6709, P27459, E7946), 3 *Vibrio vulnificus* (*Vv*) strains (CMCP6, MO6-24/O, ATCC 27562), *Vibrio parahaemolyticus* RIMD 2210633 (*Vp*), *Vibrio mimicus* SCCF01 (*Vm*), *Vibrio fischeri* ES114 (*Vf*), 4 *Vibrio campbellii* (*Vca*) strains (DS40M4, NBRC 15631, BB120, HY-01), and *Escherichia coli* MG1655 (*Ec*).

(C) Gene fraction at the levels of essentiality classification from previously validated genes to non-essential genes.

(D) Conservation of essential, quasi-essential, and non-essential genes for competence among 17 *Vibrio* species. If genes were deleted with an operon, the average number of shared strains of each gene in the operon was shown. Red and black lines indicate the mean and standard error of the mean values. Significance was assessed by the Wilcoxon-Mann-Whitney test (*$p < 0.05$).

(E) TE change of ArgRΔ-1 strain with and without arginine. Each dot indicates an individual value of TE of three biological replicates. The bar plot and error bar indicate mean and standard deviation values, respectively. In the row named Arg., − and + symbols indicate that no arginine and 1 mM arginine was added into the competence buffer, respectively. Significance was assessed by a two-tailed Student's t test (****$p < 0.0001$).

See also Figure S4–S6 and Table S1 and S2.

in competence *Vibrio* (13/13 strains) is noteworthy. LigA's signal peptide domain may facilitate ligation outside the cytosol (Figure S6E), as previously suggested.[42]

The QstR-Δ1 (Δ*recQ*), QstR-Δ2 (Δ*intI*), and QstR-Δ4 (Δ*radC*) strains showed rTEs of 1.27%, 2.98%, and 6.79%, respectively (Figures 3A–3B). The RecQ family ATP-dependent DNA helicase likely unwinds DNA during competence with ATP and single-stranded DNA (ssDNA)-binding protein.[43] The *intI* gene encodes superintegron integrase, preferentially binding ssDNA.[44] *radC* encodes the RecG-like DNA recombination/repair function. IntI[45,46] and RadC[7,47,48] are induced in naturally transformable bacteria, but their exact roles remain unknown.

The TfoX iModulon contained two diguanylate cyclases (DGCs) and one phosphodiesterase (PDE) involved in c-di-GMP generation and degradation[49] (Figures 3A–3B). c-di-GMP regulates processes, including motility (low c-di-GMP),[50] T6SS-mediated killing (low c-di-GMP),[51] and biofilm formation (high c-di-GMP)[50] in *Vc* (Figure S6F). Interestingly, two DGC mutants (TfoX-Δ8 and TfoX-Δ52) were non-essential (50.68%–124.30% rTE), while the PDE mutant (TfoX-Δ64) showed a significant reduction in TE (8.60% rTE). In line with the importance of reduction of c-di-GMP concentration for TfoX-dependent T6SS activation through the expression of PDE,[51] the quasi-essentiality of PDE suggests that lowering intracellular c-di-GMP is crucial for competence in *Vn*. Moreover, the activation of HapR, which is the major inhibitor of biofilm formation,[52] and activation of the Flagellar-1 iModulon (Fla1-Δ1, 8.26% rTE) indicate that inhibiting biofilm formation and increasing motility are important for natural transformation.

Overall, the 15 newly identified essential and quasi-essential genes, deleted in nine mutant strains, show increased expression during competence (Figure S5G) and significantly contribute to natural transformation activity. We further validated that the TE deficiencies were significantly restored by complementing the expression of the identified genes in each deletion mutant (Figure S5H). Thus, elucidating these essential and quasi-essential genes provides a roadmap to the genetic basis of competence and underlying biochemical and regulatory functions. While our study reveals causality at the systems level, further detailed studies are needed to elucidate the molecular mechanisms.

Competency iModulons contain multiple defense mechanisms that mitigate oxidative damage

We found that the TfoX, ArgR, UC-5, AhpCF, and Fur iModulons contained several genes involved in reducing oxidative stress, classified as quasi-essential genes.

First, an efficient metabolic strategy to avoid increasing oxidative stress may involve reducing electron flow through the electron transport chain by decreasing NADH and $FADH_2$ levels.[53] Oxidative stress triggers reactions that raise NADPH pools, providing reducing power for the antioxidant system, such as thioredoxins, glutaredoxins, and alkyl hydroperoxide reductase.[54] We observed activation of malate dehydrogenase MaeB (TfoX-Δ50, 9.10% rTE) in the TfoX iModulon and arginine biosynthesis (ArgR-Δ1, 0.13% rTE) in the ArgR iModulon (Figure 4A). MaeB provides an alternative route for pyruvate synthesis with NADPH generation without NADH generation.[55] Activation of arginine biosynthesis can promote the conversion of α-ketoglutarate to glutamate, absorbing one molecule of NADH and reducing oxidative TCA cycle flux. Indeed, metabolomic studies of *Mycobacterium tuberculosis* lacking arginine biosynthetic metabolism demonstrated an enhanced flux to the TCA cycle and oxidative phosphorylation followed by reactive oxygen species (ROS) generation.[56] These processes may allow for NADPH production (Figure 4A) and reduce metabolic flux of NADH and $FADH_2$ generation. Arginine generated via the ArgR iModulon can further convert to polyamine (such as putrescine), protecting bacterial cells from oxidative stress.[57]

Second, the UC5 and AhpCF iModulons contain four quasi-essential genes involved in ROS detoxifying and regeneration of scavenging molecules (Figure 4B). The UC-5 iModulon contains methionine sulfoxide reductase B (*msrB*; UC5-Δ5, 11.44% rTE) and glutaredoxin (*glrX*; UC5-Δ6, 3.09% rTE) genes, which remove ROS by reducing cysteine or methionine residues oxidized by ROS with NADPH.[58] The AhpCF iModulon contains alkyl hydroperoxidase (*ahpC* and *ahpF*, AhpCF-Δ1, 17.03% rTE), the primary scavenger of hydrogen peroxide in bacteria.[59] These results align with previous oxidative stress responses in *Escherichia coli*.[60] Furthermore, the overexpression of genes involved in UC-5 and AhpCF iModulons during competence increased viability (i.e., total CFUs) by an average of over 15.7-fold while reducing TE by an average of 0.11-fold at a 0.001% arabinose concentration compared with TfoX+ (Figure S6G). This suggests that maintaining an appropriate oxidative stress level is necessary, with the ArgR and AhpCF iModulons playing a crucial role in inhibiting ROS overgeneration during competence.

Third, iron excess can cause oxidative stress in bacteria through the Fenton reaction (Figure 4C).[61] To reduce oxidative stress, bacteria use a Fur to repress the expression of iron uptake proteins and activate the storage proteins for Fe^{3+}.[62,63] Consistent with these previous studies, the Fur iModulon showed significantly decreased expression of iron-complex transport systems (*fecABCDE*, *feoABC*, *fhuABCDE*, and *vctABCD*; $p < 0.0001$) in the TfoX+ strain, while expression of ferric iron-storage enzymes (Bfr and Bfd) significantly increased ($p = 0.028$) in the TfoX+ strain compared with the TfoX− strain during competence (Figure 4C). These results indicate additional oxidative stress reduction mechanisms through iron depletion during competence.

Taken together, these results reveal that broad ROS damage mitigation is a key characteristic of the competence process. Further analysis of housekeeping competency-related ArgR and AphCF iModulons reveals an even broader nature of the ROS response.

The ArgR and AhpCF iModulons control oxidative stress during the competence process

To further examine ROS damage responses, we used six strains, including wild-type (WT)-TfoX−, WT-TfoX+, ArgRΔ1-TfoX–, ArgRΔ1-TfoX+, AhpCFΔ1-TfoX–, and AhpCFΔ1-TfoX+ to investigate intracellular ROS levels and susceptibility of oxidative damage during the competence.

First, we measured the internal ROS levels in these strains under the competence condition using the ROS fluorescence sensor CM-H_2DCFDA[64] to evaluate arginine metabolism's contribution to ROS formation inhibition. We discovered that

CellPress
OPEN ACCESS

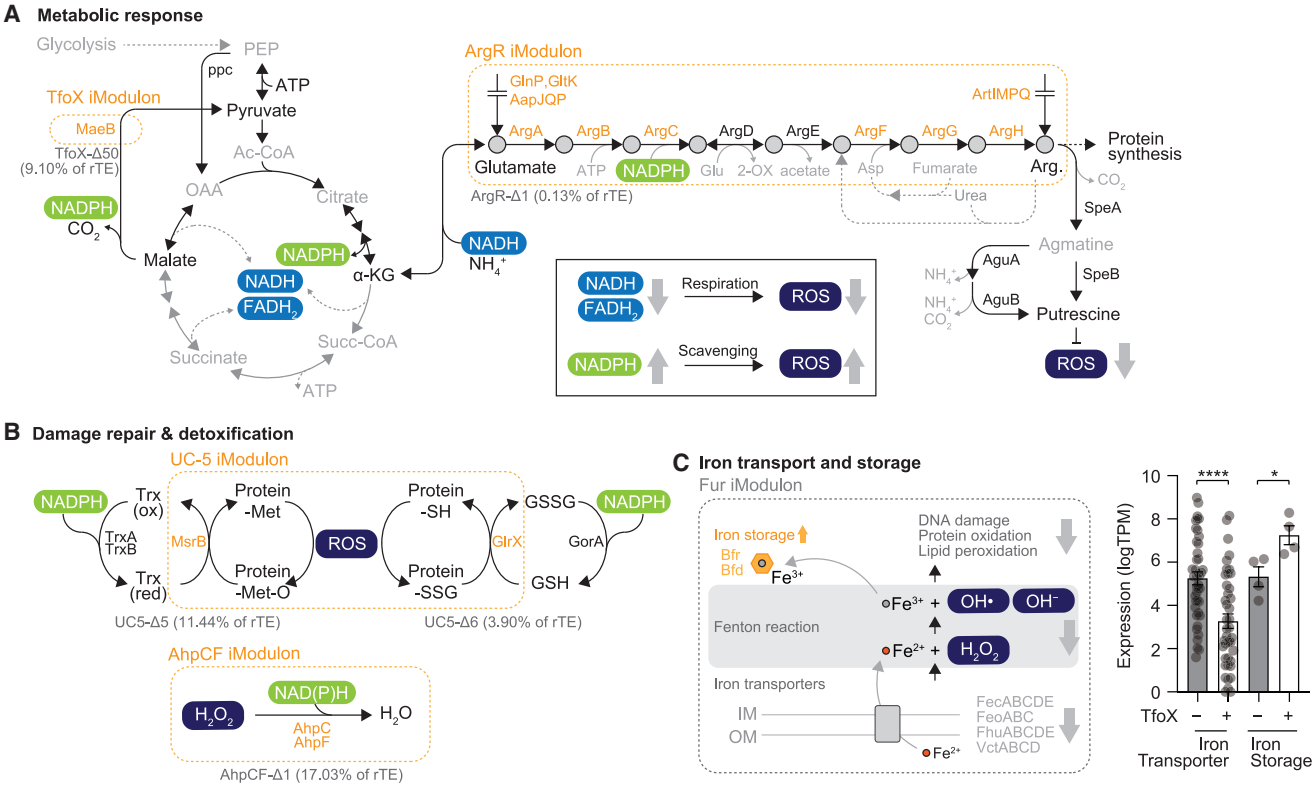


Figure 4. Functional delineation of competency iModulons to mitigate ROS damage

(A–C) Schematic representation of multiple defense mechanisms of competence iModulons (TfoX, ArgR, UC-5, AhpCF, and Fur iModulons) against oxidative damage, such as (A) metabolic conversion of NADH into NADPH, production of polyamide from arginine, (B) ROS-detoxifying enzymes, redox proteins for oxidative damage repair, and (C) the repression of iron-complex transport systems. RNA levels are shown as $\log_2$ transcripts per million (log-TPM). The bar plot and error bar indicate mean and standard error of the mean values (iron transporter, n = 42; iron storage, n = 4), respectively. For intuitive explanation and understanding, the obtained TE value of the mutant was further standardized as rTE values (%), which represent the increase or decrease of TE with respect to TfoX+. TfoX– and TfoX+ indicate the Vn-PTrc and Vn-PTrc-TfoX strains, respectively. Significance was assessed by Wilcoxon-Mann-Whitney test (*$p < 0.05$; ****$p < 0.0001$).

See also Figure S4 and S5 and Table S2.

ROS levels in WT-TfoX+ (3.5 relative fluorescence intensity [RFI]) were significantly higher ($p = 0.005$) than in WT-TfoX– (2 RFI) during competence progression (Figure 5A). We observed increased ROS generation during competence in ArgRΔ1-TfoX+ (average 1.6-fold increased, $p < 0.002$, at all time points) and AhpCFΔ1-TfoX+ (average 2.2-fold increased, $p < 0.026$, at all time points), indicating TfoX-dependent ROS generation. Additionally, ROS levels in AhpCFΔ1-TfoX+ (5.5 RFI) and ArgR-Δ1-TfoX+ (4.7 RFI) were significantly higher (>1.4-fold increased, $p < 0.013$) than that of WT-TfoX+ at all time points (Figure 5A).

Next, we measured total antioxidant capacity by measuring the viability of various concentrations of external hydrogen peroxide (H_2O_2) under the competence condition. WT-TfoX+ showed higher sensitivity and susceptibility to oxidative stress ($p < 0.04$) under 0.1–0.5 mM H_2O_2 compared with WT-TfoX– (Figure 5B). Consistently, ArgRΔ1-TfoX+ ($p < 0.041$) and AhpCFΔ1-TfoX+ ($p < 0.016$) strains showed increased susceptibility to oxidative stress under 0.1–0.5 mM H_2O_2 when expressing *tfoX*. Moreover, the ArgRΔ1-TfoX+ ($p < 0.038$) and the AhpCFΔ1-TfoX+ ($p < 0.014$) strains exhibited greater H_2O_2 vulnerability than WT-TfoX+ under 0.1 and 0.2 mM H_2O_2 (Figure 5B), suggesting that ArgR and AhpCF iModulons contribute to the inhibition of intracellular ROS generation during competence.

Collectively, our data showed that natural transformation leads to oxidative stress accumulation and that five iModulons (TfoX, ArgR, UC-5, AhpCF, and Fur) are involved in multiple defense mechanisms against oxidative damage. These mechanisms include NADH to NADPH metabolic conversion, ROS-detoxifying enzymes, redox proteins for oxidative damage repair, and repression of iron-complex transport systems. Overall, our findings show that broad and coordinated ROS damage mitigation is a key feature of the competence process, requiring a high level of coordination within the complexities of systems biology.

DISCUSSION

Vn is a promising model organism for molecular cloning and biotechnological applications due to its non-pathogenic nature, rapid growth, and versatile metabolism. Natural competence,

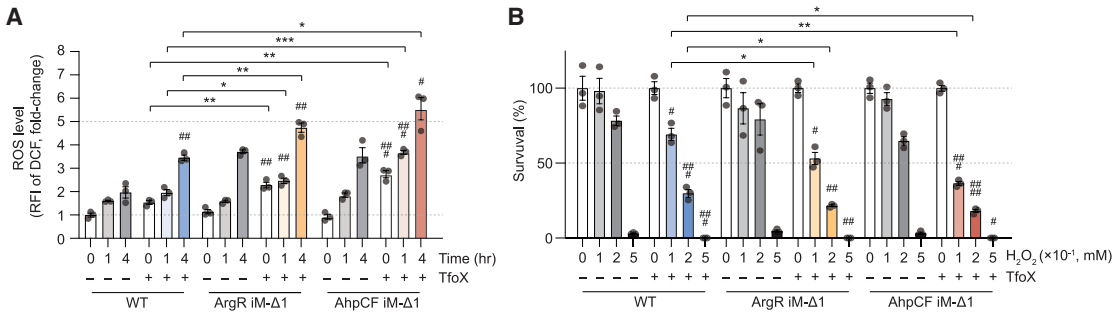


Figure 5. ROS levels and H_2O_2 susceptibility in *Vn* during natural competence

(A) Measurement of ROS levels in WT-TfoX−, WT-TfoX+, ArgRΔ1-TfoX−, ArgRΔ1-TfoX+, AhpCFΔ1-TfoX−, and AhpCFΔ1-TfoX+ strains. Relative fluorescence intensity (RFI) is presented as a fold change relative to the WT-tfoX− group (n = 3). The bar plot and error bar indicate the mean and standard error of the mean values, respectively. Significance was assessed by two-tailed Student's t test (*p < 0.05; **p < 0.01; ***p < 0.001); "strain name"-TfoX– versus "strain name"-TfoX+ groups at same time points (#p < 0.05; ##p < 0.01; ###p < 0.001).

(B) H_2O_2 susceptibility assays for WT-TfoX−, WT-TfoX+, ArgRΔ1-TfoX−, ArgRΔ1-TfoX+, AhpCFΔ1-TfoX−, and AhpCFΔ1-TfoX+ strains during natural competence. Various H_2O_2 concentrations (0, 0.1, 0.2, and 0.5 mM) were treated in each transformation reaction. The bar plot and error bar indicate the mean and standard error of the mean values (n = 3), respectively. Significance was assessed by two-tailed Student's t test (*p < 0.05; **p < 0.01); "strain name"-TfoX– versus "strain name"-TfoX+ groups at same time points (#p < 0.05; ##p < 0.01; ###p < 0.001; ####p < 0.0001).

See also Figure S4–S6 and Table S2.

one of the pathways for introducing external genes, can be harnessed for genome engineering. Using a strain as a foundation allows for the efficient production of desired biomolecules by removing or modifying genes deemed unnecessary within the genome.[65] Gaining a deeper understanding of the genes involved in essential cellular functions and metabolism, including natural competence, can greatly benefit these processes.

However, a comprehensive understanding of its metabolism, cell functions including competency, and regulatory mechanisms remains limited compared with well-studied relatives like *Vc* or *V. fischeri*. ICA-derived iModulons provide a biologically meaningful and unsupervised method for studying TRNs, facilitating the identification of new regulons and offering insights into condition-specific gene expression regulation, in contrast to GO or KEGG enrichment analyses. By employing the database-iModulon-discovery cycle, we generated transcriptomic databases, identified 45 iModulons, and uncovered competence-related findings: (1) repression of housekeeping iModulons (Fur and ribosomal proteins), (2) activation of six iModulons (containing 189 genes) with TfoX, QstR, and novel iModulon (UC-5), and (3) activation of three housekeeping iModulons (Flagella-1, ArgR, and AhpCF). This analysis provides a comprehensive genetic and transcriptomic basis for competency in *Vn*.

The competency process interacts with five housekeeping functions: iron depletion and ribosomal protein deactivation reduce oxidative stress and prevent growth during competence, and upregulated functions include motility, arginine metabolism, and the *ahpCF* operon, increasing chances of encountering nutrients and DNA, reducing ROS generation, and presumably minimizing damage to ssDNA during integration.

Our approach also reveals that there are three cellular responses related specifically to the competency process itself. Two of these are related to the known TFs TfoX and QstR. The TfoX and QstR iModulons include known competency-related genes, except *comEC*. A third response was identified that was unknown and represented by the UC-5 iModulon. Our results thus uncover new features in these competency iModulons that play important roles, falling into five functional categories (Figure 6).

First, our data identified 16 unknown membrane proteins and two component genes as essential or quasi-essential within the TfoX and UC-5 iModulons, with the TMEM64 family protein in the UC-5 iModulon being essential for competence. The mechanisms underlying their involvement in DNA uptake warrant further investigation.

Second, we discovered that the three DNA processing and recombination enzymes RecQ, RadC, and IntIA, in the QstR iModulon, are quasi-essential for *Vn* competency, consistent with the activation in other bacteria during competence.[46,66] Surprisingly, we observed a significant decrease in rTE with both linear tDNAs (1.3%–6.8% rTE, tDNA-Kan^R; 3.6%–42.5% rTE, tDNA-Spec^R) and the pUC-Spec^R plasmid (0.03%–16.64% rTE) in these three mutant strains (Figure S5B), indicating that these genes may have roles beyond genome integration during *Vn* competence.

Third, RNA and protein turnover appears to be crucial for resource reallocation during the reversion of competence phenotype. Reactivation of the ribosomal protein iModulon during recovery suggests resource suppression for cell growth in favor of transient competency (Figure S3A). Our data also highlighted the necessity of RNase J1 and M48 family metallopeptidase, supporting the idea that competence systems impose a significant energetic burden on the cell.[67,68]

Fourth, we found a PDE involved in c-di-GMP degradation to be quasi-essential for competency in the TfoX iModulon, indicating that suppressing biofilm formation and promoting motility via low c-di-GMP is critical, consistent with Flagellar-1 iModulon activation.

Fifth, *Vn* is highly susceptible to oxidative stress during competency, with five iModulons emphasizing the importance of inhibiting ROS formation and maintaining oxidative scavenging capacity. Although a recent study in *Acinetobacter baylyi*

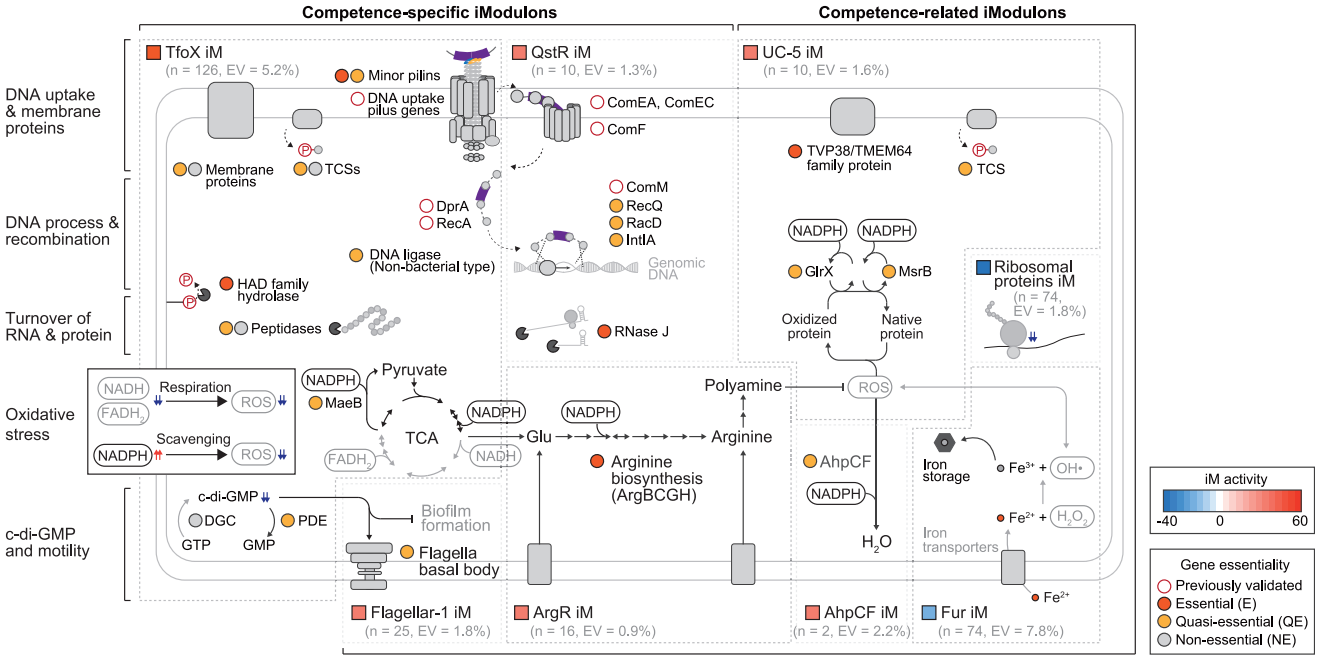


Figure 6. Overview of the systems biology of natural competence in *Vn* summarized by eight competence iModulons

The summary is represented by only essential and quasi-essential genes with known functions in eight competence iModulons. iM, iModulon; TCSs, two-component systems; EV, explained variance.

See also Table S2.

demonstrated that the accumulation of ROS increases competency by increasing membrane permeability,[69] further studies are needed to elucidate the multi-dimensional ROS response and its functions in the *Vibrio* species competency process.

The exact role of ROS in bacterial competence is unclear, but previous studies suggest possible involvement. First, ROS may act as signaling molecules during competence, as seen in *Streptococcus pneumoniae*.[70] Second, appropriate ROS levels might trigger DNA repair mechanisms,[71] allowing bacteria to uptake external DNA to fix damaged genomic regions. However, the ROS's role in competence may differ among species, and further research is needed to clarify ROS generation mechanisms and their functions during competence.

This study has limitations, including the possibility that identified competence iModulons and gene essentiality may not be representative of all *Vibrio* species or their competence regulation. As noted by a previous study,[11] natural transformation frequency and mechanisms can vary significantly among species, and gene composition for competence may depend on each strain's genomic context. Contrary to our data in *Vn*, deletion of *ligA* did not significantly reduce TE in *Vc* and *Haemophilus influenzae*.[7,42] Inactivation of RadC did not significantly affect competence in *Vc*, *Streptococcus pneumoniae*, and *H. influenzae*.[7,42,72] Also, the distribution of the conservation of quasi-essential genes shows that 34 genes are conserved only in *Vibrio* strains other than *Vc*, and 9 genes are *Vn*-specific genes. As a representative example, when the fucose degradation pathway was removed from *Vn*, a significant decrease in TE occurred (TfoX-Δ66, 1.38% rTE), but the pathway was conserved only in *V. vulnificus* MO6-24/O and *V. mimicus* SCCF01 (Figures 3B and S6H). Thus, some competence proteins' functions in transformable bacteria may differ, and their exact functions remain to be determined.

Furthermore, the essentiality for competence can be specific to a given condition, as TE values can be influenced by various factors like the form of tDNA. Some deletion mutation strains showed significant TE changes depending on tDNA form, such as linear tDNA (tDNA-Spec^R, for genome integration) and a plasmid (pUC-Spec^R, for replicating plasmid), (Figures S5B and S5C). For instance, the UC5-Δ2 strain, lacking the TVP38/TMEM64 family membrane protein, showed a significant increase in TE only when using pUC-Spec^R as tDNA (0.14% rTE, tDNA-Kan^R; 10.8% rTE, tDNA-Spec^R; 77.6% rTE, pUC-Spec^R; $p < 0.013$).

Additionally, four mutant strains (TfoX-Δ10 [lacking hypothetical TF], TfoXΔ15 [lacking membrane protein], TfoX-Δ17 [lacking hypothetical two-component system], and TfoXΔ36 [lacking hypothetical transporter]) did not show a drastic decrease in TE (89.7%–119.7% rTE, tDNA-Kan^R; 61.1%–133.6% rTE, tDNA-Spec^R) when using linear tDNA but showed a significant decrease (5%–45.2% rTE, $p < 0.011$) when using pUC-Spec^R. Given that the majority of these genes are hypothetical membrane proteins, it is possible that some factors are necessary depending on the form of tDNA during DNA uptake. Therefore, genes considered non-essential within competence iModulons may still be required under specific circumstances. Further research is needed to fully understand the natural competence system and evaluate gene function within competence iModulons under diverse conditions.

We present the first comprehensive systems biology description of competence, using the database-iModulon-discovery cycle. This approach elucidates competence in *Vn* at genetic, systems, and phenotypic levels. We note that even if the specific molecular function of a gene product is not known, we do know that the gene is a part of a cellular process represented by an iModulon and that the deactivation of the iModulon reduces competency. Our findings pave the way for future studies that should culminate in a full understanding of the detailed roles that the constituent genes play in competence. This system's approach can be applied to other bacteria, providing insights into the structure and function of competence across the phylogenetic tree.

Limitation of the study

While our study offers valuable insights into the systems biology of competence in *Vn*, some limitations must be acknowledged. The identified competence iModulons and gene essentiality may not be universally representative for all *Vibrio* species due to potential variability in natural transformation mechanisms and genomic contexts among different strains. The essentiality of certain genes for competence might be condition specific, and the influence of various factors such as the form of the transforming DNA warrants further investigation. The exact functions and mechanisms of certain unknown membrane proteins and two-component genes, as well as the role of ROS in bacterial competence, remain largely unexplored. As our study primarily focuses on *Vn*, extrapolating these findings to other species necessitates careful consideration and further research.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2023.112619.

ACKNOWLEDGMENTS

We would like to thank Marc Abrams (Systems Biology Research Group, University of California San Diego) for assistance with paper editing. This work was supported by The Novo Nordisk Foundation (NNF) Center for Biosustainability (CfB) at the Technical University of Denmark (NNF20CC0035580) and the Y.C. Fung Endowed Chair in Bioengineering at the University of California San Diego.

AUTHOR CONTRIBUTIONS

B.O.P. conceived and supervised the study. J.S. and B.O.P. designed the experiments. J.S. performed the experiments. J.S., K.R., and B.O.P. analyzed the data. J.S., K.R., and B.O.P. wrote the manuscript. All authors read and approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: December 7, 2022
Revised: April 27, 2023
Accepted: May 22, 2023
Published: June 6, 2023

REFERENCES

1. Johnston, C., Martin, B., Fichant, G., Polard, P., and Claverys, J.-P. (2014). Bacterial transformation: distribution, shared mechanisms and divergent control. Nat. Rev. Microbiol. *12*, 181–196. https://doi.org/10.1038/nrmicro3199.
2. Lorenz, M.G., and Wackernagel, W. (1994). Bacterial gene transfer by natural genetic transformation in the environment. Microbiol. Rev. *58*, 563–602. https://doi.org/10.1128/mr.58.3.563-602.1994.
3. Blokesch, M. (2016). Natural competence for transformation. Curr. Biol. *26*, R1126–R1130. https://doi.org/10.1016/j.cub.2016.08.058.
4. Seitz, P., and Blokesch, M. (2013). Cues and regulatory pathways involved in natural competence and transformation in pathogenic and environmental Gram-negative bacteria. FEMS Microbiol. Rev. *37*, 336–363. https://doi.org/10.1111/j.1574-6976.2012.00353.x.
5. Meibom, K.L., Blokesch, M., Dolganov, N.A., Wu, C.-Y., and Schoolnik, G.K. (2005). Chitin induces natural competence in Vibrio cholerae. Science *310*, 1824–1827. https://doi.org/10.1126/science.1120096.
6. Matthey, N., and Blokesch, M. (2016). The DNA-uptake process of naturally competent Vibrio cholerae. Trends Microbiol. *24*, 98–110. https://doi.org/10.1016/j.tim.2015.10.008.
7. Jaskólska, M., Stutzmann, S., Stoudmann, C., and Blokesch, M. (2018). QstR-dependent regulation of natural competence and type VI secretion in Vibrio cholerae. Nucleic Acids Res. *46*, 10619–10634. https://doi.org/10.1093/nar/gky717.
8. Haycocks, J.R.J., Warren, G.Z.L., Walker, L.M., Chlebek, J.L., Dalia, T.N., Dalia, A.B., and Grainger, D.C. (2019). The quorum sensing transcription factor AphA directly regulates natural competence in Vibrio cholerae. PLoS Genet. *15*, e1008362. https://doi.org/10.1371/journal.pgen.1008362.

9. Metzger, L.C., and Blokesch, M. (2016). Regulation of competence-mediated horizontal gene transfer in the natural habitat of Vibrio cholerae. Curr. Opin. Microbiol. *30*, 1–7. https://doi.org/10.1016/j.mib.2015.10.007.

10. Dalia, T.N., Hayes, C.A., Stolyar, S., Marx, C.J., McKinlay, J.B., and Dalia, A.B. (2017). Multiplex genome editing by natural transformation (MuGENT) for synthetic biology in Vibrio natriegens. ACS Synth. Biol. *6*, 1650–1655. https://doi.org/10.1021/acssynbio.7b00116.

11. Simpson, C.A., Podicheti, R., Rusch, D.B., Dalia, A.B., and van Kessel, J.C. (2019). Diversity in natural transformation frequencies and regulation across Vibrio species. mBio *10*, 027888-19–e2819. https://doi.org/10.1128/mBio.02788-19.

12. Lo Scrudato, M., and Blokesch, M. (2013). A transcriptional regulator linking quorum sensing and chitin induction to render Vibrio cholerae naturally transformable. Nucleic Acids Res. *41*, 3644–3658. https://doi.org/10.1093/nar/gkt041.

13. Lo Scrudato, M., and Blokesch, M. (2012). The regulatory network of natural competence and transformation of Vibrio cholerae. PLoS Genet. *8*, e1002778. https://doi.org/10.1371/journal.pgen.1002778.

14. Borgeaud, S., Metzger, L.C., Scrignari, T., and Blokesch, M. (2015). The type VI secretion system of *Vibrio cholerae* fosters horizontal gene transfer. Science *347*, 63–67. https://doi.org/10.1126/science.1260064.

15. Watve, S.S., Thomas, J., and Hammer, B.K. (2015). CytR is a global positive regulator of competence, type VI secretion, and chitinases in Vibrio cholerae. PLoS One *10*, e0138834. https://doi.org/10.1371/journal.pone.0138834.

16. Thoma, F., and Blombach, B. (2021). Metabolic engineering of Vibrio natriegens. Essays Biochem. *65*, 381–392. https://doi.org/10.1042/EBC20200135.

17. Xu, J., Yang, S., and Yang, L. (2022). Vibrio natriegens as a host for rapid biotechnology. Trends Biotechnol. *40*, 381–384. https://doi.org/10.1016/j.tibtech.2021.10.007.

18. Hoff, J., Daniel, B., Stukenberg, D., Thuronyi, B.W., Waldminghaus, T., and Fritz, G. (2020). Vibrio natriegens: an ultrafast-growing marine bacterium as emerging synthetic biology chassis. Environ. Microbiol. *22*, 4394–4408. https://doi.org/10.1111/1462-2920.15128.

19. Stukenberg, D., Hoff, J., Faber, A., and Becker, A. (2022). NT-CRISPR, combining natural transformation and CRISPR-Cas9 counterselection for markerless and scarless genome editing in Vibrio natriegens. Commun. Biol. *5*, 265. https://doi.org/10.1038/s42003-022-03150-0.

20. Santos-Zavaleta, A., Sánchez-Pérez, M., Salgado, H., Velázquez-Ramírez, D.A., Gama-Castro, S., Tierrafría, V.H., Busby, S.J.W., Aquino, P., Fang, X., Palsson, B.O., et al. (2018). A unified resource for transcriptional regulation in Escherichia coli K-12 incorporating high-throughput-generated binding data into RegulonDB version 10.0. BMC Biol. *16*, 91. https://doi.org/10.1186/s12915-018-0555-y.

21. Liebermeister, W. (2002). Linear modes of gene expression determined by independent component analysis. Bioinformatics *18*, 51–60. https://doi.org/10.1093/bioinformatics/18.1.51.

22. Saelens, W., Cannoodt, R., and Saeys, Y. (2018). A comprehensive evaluation of module detection methods for gene expression data. Nat. Commun. *9*, 1090–1112. https://doi.org/10.1038/s41467-018-03424-4.

23. Sastry, A.V., Gao, Y., Szubin, R., Hefner, Y., Xu, S., Kim, D., Choudhary, K.S., Yang, L., King, Z.A., and Palsson, B.O. (2019). The Escherichia coli transcriptome mostly consists of independently regulated modules. Nat. Commun. *10*, 5536–5614. https://doi.org/10.1038/s41467-019-13483-w.

24. Rychel, K., Sastry, A.V., and Palsson, B.O. (2020). Machine learning uncovers independently regulated modules in the Bacillus subtilis transcriptome. Nat. Commun. *11*, 6338–6410. https://doi.org/10.1038/s41467-020-20153-9.

25. Poudel, S., Tsunemoto, H., Seif, Y., Sastry, A.V., Szubin, R., Xu, S., Machado, H., Olson, C.A., Anand, A., Pogliano, J., et al. (2020). Revealing 29 sets of independently modulated genes in Staphylococcus aureus, their regulators, and role in key physiological response. Proc. Natl. Acad. Sci. USA *117*, 17228–17239. https://doi.org/10.1073/pnas.2008413117.

26. Chauhan, S.M., Poudel, S., Rychel, K., Lamoureux, C., Yoo, R., Al Bulushi, T., Yuan, Y., Palsson, B.O., and Sastry, A.V. (2021). Machine learning uncovers a data-driven transcriptional regulatory network for the crenarchaeal thermoacidophile sulfolobus acidocaldarius. Front. Microbiol. *12*, 753521. https://doi.org/10.3389/fmicb.2021.753521.

27. Yoo, R., Rychel, K., Poudel, S., Al-Bulushi, T., Yuan, Y., Chauhan, S., Lamoureux, C., Palsson, B.O., and Sastry, A. (2022). Machine learning of all Mycobacterium tuberculosis H37Rv RNA-seq data reveals a structured interplay between metabolism, stress response, and infection. mSphere *7*, e0003322. https://doi.org/10.1128/msphere.00033-22.

28. Rajput, A., Tsunemoto, H., Sastry, A.V., Szubin, R., Rychel, K., Sugie, J., Pogliano, J., and Palsson, B.O. (2022). Machine learning from Pseudomonas aeruginosa transcriptomes identifies independently modulated sets of genes associated with known transcriptional regulators. Nucleic Acids Res. *50*, 3658–3672. https://doi.org/10.1093/nar/gkac187.

29. Lim, H.G., Rychel, K., Sastry, A.V., Bentley, G.J., Mueller, J., Schindel, H.S., Larsen, P.E., Laible, P.D., Guss, A.M., Niu, W., et al. (2022). Machine-learning from Pseudomonas putida KT2440 transcriptomes reveals its transcriptional regulatory network. Metab. Eng. *72*, 297–310. https://doi.org/10.1016/j.ymben.2022.04.004.

30. Sastry, A.V., Poudel, S., Rychel, K., Yoo, R., Lamoureux, C.R., Chauhan, S., Haiman, Z.B., Al Bulushi, T., Seif, Y., and Palsson, B.O. (2021). Mining all publicly available expression data to compute dynamic microbial transcriptional regulatory networks. Preprint at bioRxiv. https://doi.org/10.1101/2021.07.01.450581.

31. Sastry, A.V., Hu, A., Heckmann, D., Poudel, S., Kavvas, E., and Palsson, B.O. (2021). Independent component analysis recovers consistent regulatory signals from disparate datasets. PLoS Comput. Biol. *17*, e1008647. https://doi.org/10.1371/journal.pcbi.1008647.

32. Rychel, K., Decker, K., Sastry, A.V., Phaneuf, P.V., Poudel, S., and Palsson, B.O. (2021). iModulonDB: a knowledgebase of microbial transcriptional regulation derived from machine learning. Nucleic Acids Res. *49*, D112–D120. https://doi.org/10.1093/nar/gkaa810.

33. Seitz, P., Pezeshgi Modarres, H., Borgeaud, S., Bulushev, R.D., Steinbock, L.J., Radenovic, A., Dal Peraro, M., and Blokesch, M. (2014). ComEA is essential for the transfer of external DNA into the periplasm in naturally transformable Vibrio cholerae cells. PLoS Genet. *10*, e1004066. https://doi.org/10.1371/journal.pgen.1004066.

34. Yin, M., Ye, B., Jin, Y., Liu, L., Zhang, Y., Li, P., Wang, Y., Li, Y., Han, Y., Shen, W., and Zhao, Z. (2020). Changes in Vibrio natriegens growth under simulated microgravity. Front. Microbiol. *11*, 2040. https://doi.org/10.3389/fmicb.2020.02040.

35. Wu, R., Zhao, M., Li, J., Gao, H., Kan, B., and Liang, W. (2015). Direct regulation of the natural competence regulator gene tfoX by cyclic AMP (cAMP) and cAMP receptor protein (CRP) in Vibrios. Sci. Rep. *5*, 14921–15015. https://doi.org/10.1038/srep14921.

36. Blokesch, M. (2012). Chitin colonization, chitin degradation and chitin-induced natural competence of Vibrio cholerae are subject to catabolite repression. Environ. Microbiol. *14*, 1898–1912. https://doi.org/10.1111/j.1462-2920.2011.02689.x.

37. Seitz, P., and Blokesch, M. (2013). DNA-uptake machinery of naturally competent Vibrio cholerae. Proc. Natl. Acad. Sci. USA *110*, 17987–17992. https://doi.org/10.1073/pnas.1315647110.

38. Ellison, C.K., Dalia, T.N., Vidal Ceballos, A., Wang, J.C.-Y., Biais, N., Brun, Y.V., and Dalia, A.B. (2018). Retraction of DNA-bound type IV competence pili initiates DNA uptake during natural transformation in Vibrio cholerae. Nat. Microbiol. *3*, 773–780. https://doi.org/10.1038/s41564-018-0174-y.

39. Burroughs, A.M., Allen, K.N., Dunaway-Mariano, D., and Aravind, L. (2006). Evolutionary genomics of the HAD superfamily: understanding the structural adaptations and catalytic diversity in a superfamily of phosphoesterases and allied enzymes. J. Mol. Biol. *361*, 1003–1034. https://doi.org/10.1016/j.jmb.2006.06.049.

40. Raj, R., Nadig, S., Patel, T., and Gopal, B. (2020). Structural and biochemical characteristics of two Staphylococcus epidermidis RNase J paralogs RNase J1 and RNase J2. J. Biol. Chem. *295*, 16863–16876. https://doi.org/10.1074/jbc.ra120.014876.

41. Bechhofer, D.H. (2009). Chapter 6 messenger RNA decay and maturation in Bacillus subtilis. In Progress in Molecular Biology and Translational Science (Academic Press), pp. 231–273. https://doi.org/10.1016/S0079-6603(08)00806-4.

42. Sinha, S., Mell, J.C., and Redfield, R.J. (2012). Seventeen Sxy-dependent cyclic AMP receptor protein site-regulated genes are needed for natural transformation in Haemophilus influenzae. J. Bacteriol. *194*, 5245–5254. https://doi.org/10.1128/JB.00671-12.

43. Killoran, M.P., and Keck, J.L. (2006). Sit down, relax and unwind: structural insights into RecQ helicase mechanisms. Nucleic Acids Res. *34*, 4098–4105. https://doi.org/10.1093/nar/gkl538.

44. Francia, M.V., Zabala, J.C., de la Cruz, F., and García Lobo, J.M. (1999). The IntI1 integron integrase preferentially binds single-stranded DNA of the attC site. J. Bacteriol. *181*, 6844–6849. https://doi.org/10.1128/JB.181.21.6844-6849.1999.

45. Guerin, E., Cambray, G., Sanchez-Alberola, N., Campoy, S., Erill, I., Da Re, S., Gonzalez-Zorn, B., Barbé, J., Ploy, M.-C., and Mazel, D. (2009). The SOS response controls integron recombination. Science *324*, 1034. https://doi.org/10.1126/science.1172914.

46. Baharoglu, Z., Krin, E., and Mazel, D. (2012). Connecting environment and genome plasticity in the characterization of transformation-induced SOS regulation and carbon catabolite control of the Vibrio cholerae integron integrase. J. Bacteriol. *194*, 1659–1667. https://doi.org/10.1128/JB.05982-11.

47. Redfield, R.J., Cameron, A.D.S., Qian, Q., Hinds, J., Ali, T.R., Kroll, J.S., and Langford, P.R. (2005). A novel CRP-dependent regulon controls expression of competence genes in Haemophilus influenzae. J. Mol. Biol. *347*, 735–747. https://doi.org/10.1016/j.jmb.2005.01.012.

48. Peterson, S.N., Sung, C.K., Cline, R., Desai, B.V., Snesrud, E.C., Luo, P., Walling, J., Li, H., Mintz, M., Tsegaye, G., et al. (2004). Identification of competence pheromone responsive genes in Streptococcus pneumoniae by use of DNA microarrays. Mol. Microbiol. *51*, 1051–1070. https://doi.org/10.1046/j.1365-2958.2003.03907.x.

49. Jenal, U., Reinders, A., and Lori, C. (2017). Cyclic di-GMP: second messenger extraordinaire. Nat. Rev. Microbiol. *15*, 271–284. https://doi.org/10.1038/nrmicro.2016.190.

50. Liu, X., Beyhan, S., Lim, B., Linington, R.G., and Yildiz, F.H. (2010). Identification and characterization of a phosphodiesterase that inversely regulates motility and biofilm formation in Vibrio cholerae. J. Bacteriol. *192*, 4541–4552. https://doi.org/10.1128/JB.00209-10.

51. Metzger, L.C., Stutzmann, S., Scrignari, T., Van der Henst, C., Matthey, N., and Blokesch, M. (2016). Independent regulation of type VI secretion in Vibrio cholerae by TfoX and TfoY. Cell Rep. *15*, 951–958. https://doi.org/10.1016/j.celrep.2016.03.092.

52. Hammer, B.K., and Bassler, B.L. (2003). Quorum sensing controls biofilm formation in Vibrio cholerae. Mol. Microbiol. *50*, 101–104. https://doi.org/10.1046/j.1365-2958.2003.03688.x.

53. Van Acker, H., and Coenye, T. (2017). The role of reactive oxygen species in antibiotic-mediated killing of bacteria. Trends Microbiol. *25*, 456–466. https://doi.org/10.1016/j.tim.2016.12.008.

54. Prinz, W.A., Åslund, F., Holmgren, A., and Beckwith, J. (1997). The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the Escherichia coliCytoplasm. J. Biol. Chem. *272*, 15661–15667. https://doi.org/10.1074/jbc.272.25.15661.

55. Wang, B., Wang, P., Zheng, E., Chen, X., Zhao, H., Song, P., Su, R., Li, X., and Zhu, G. (2011). Biochemical properties and physiological roles of NADP-dependent malic enzyme in Escherichia coli. J. Microbiol. *49*, 797–802. https://doi.org/10.1007/s12275-011-0487-5.

56. Tiwari, S., van Tonder, A.J., Vilchèze, C., Mendes, V., Thomas, S.E., Malek, A., Chen, B., Chen, M., Kim, J., Blundell, T.L., et al. (2018). Arginine-deprivation-induced oxidative damage sterilizes Mycobacterium tuberculosis. Proc. Natl. Acad. Sci. USA *115*, 9779–9784. https://doi.org/10.1073/pnas.1808874115.

57. Chattopadhyay, M.K., Tabor, C.W., and Tabor, H. (2003). Polyamines protect Escherichia coli cells from the toxic effect of oxygen. Proc. Natl. Acad. Sci. USA *100*, 2261–2265. https://doi.org/10.1073/pnas.2627990100.

58. Ezraty, B., Gennaris, A., Barras, F., and Collet, J.-F. (2017). Oxidative stress, protein damage and repair in bacteria. Nat. Rev. Microbiol. *15*, 385–396. https://doi.org/10.1038/nrmicro.2017.26.

59. Seaver, L.C., and Imlay, J.A. (2001). Alkyl hydroperoxide reductase is the primary scavenger of endogenous hydrogen peroxide in Escherichia coli. J. Bacteriol. *183*, 7173–7181. https://doi.org/10.1128/JB.183.24.7173-7181.2001.

60. Seo, S.W., Kim, D., Szubin, R., and Palsson, B.O. (2015). Genome-wide reconstruction of OxyR and SoxRS transcriptional regulatory networks under oxidative stress in Escherichia coli K-12 MG1655. Cell Rep. *12*, 1289–1299. https://doi.org/10.1016/j.celrep.2015.07.043.

61. Bradley, J.M., Svistunenko, D.A., Wilson, M.T., Hemmings, A.M., Moore, G.R., and Le Brun, N.E. (2020). Bacterial iron detoxification at the molecular level. J. Biol. Chem. *295*, 17602–17623. https://doi.org/10.1074/jbc.REV120.007746.

62. Seo, S.W., Kim, D., Latif, H., O'Brien, E.J., Szubin, R., and Palsson, B.O. (2014). Deciphering Fur transcriptional regulatory network highlights its complex role beyond iron metabolism in Escherichia coli. Nat. Commun. *5*, 4910. https://doi.org/10.1038/ncomms5910.

63. Mey, A.R., Wyckoff, E.E., Kanukurthy, V., Fisher, C.R., and Payne, S.M. (2005). Iron and Fur regulation in Vibrio cholerae and the role of Fur in virulence. Infect. Immun. *73*, 8167–8178. https://doi.org/10.1128/IAI.73.12.8167-8178.2005.

64. Dong, T.G., Dong, S., Catalano, C., Moore, R., Liang, X., and Mekalanos, J.J. (2015). Generation of reactive oxygen species by lethal attacks from competing microbes. Proc. Natl. Acad. Sci. USA *112*, 2181–2186. https://doi.org/10.1073/pnas.1425007112.

65. LeBlanc, N., and Charles, T.C. (2022). Bacterial genome reductions: tools, applications, and challenges. Front. Genome Ed. *4*, 957289. https://doi.org/10.3389/fgeed.2022.957289.

66. Krin, E., Pierlé, S.A., Sismeiro, O., Jagla, B., Dillies, M.-A., Varet, H., Irazoki, O., Campoy, S., Rouy, Z., Cruveiller, S., et al. (2018). Expansion of the SOS regulon of Vibrio cholerae through extensive transcriptome analysis and experimental validation. BMC Genom. *19*, 373. https://doi.org/10.1186/s12864-018-4716-8.

67. Bacher, J.M., Metzgar, D., and de Crécy-Lagard, V. (2006). Rapid evolution of diminished transformability in Acinetobacter baylyi. J. Bacteriol. *188*, 8534–8542. https://doi.org/10.1128/JB.00846-06.

68. Utnes, A.L.G., Sørum, V., Hülter, N., Primicerio, R., Hegstad, J., Kloos, J., Nielsen, K.M., and Johnsen, P.J. (2015). Growth phase-specific evolutionary benefits of natural transformation in Acinetobacter baylyi. ISME J. *9*, 2221–2231. https://doi.org/10.1038/ismej.2015.35.

69. Wang, Y., Lu, J., Engelstädter, J., Zhang, S., Ding, P., Mao, L., Yuan, Z., Bond, P.L., and Guo, J. (2020). Non-antibiotic pharmaceuticals enhance the transmission of exogenous antibiotic resistance genes through bacterial transformation. ISME J. *14*, 2179–2196. https://doi.org/10.1038/s41396-020-0679-2.

70. Bättig, P., and Mühlemann, K. (2008). Influence of the spxB gene on competence in Streptococcus pneumoniae. J. Bacteriol. *190*, 1184–1189. https://doi.org/10.1128/JB.01517-07.

71. Gupta, A., and Imlay, J.A. (2022). Escherichia coli induces DNA repair enzymes to protect itself from low-grade hydrogen peroxide stress. Mol. Microbiol. *117*, 754–769. https://doi.org/10.1111/mmi.14870.

72. Attaiech, L., Granadel, C., Claverys, J.-P., and Martin, B. (2008). RadC, a misleading name? J. Bacteriol. *190*, 5729–5732. https://doi.org/10.1128/JB.00425-08.

73. Lee, H.H., Ostrov, N., Wong, B.G., Gold, M.A., Khalil, A.S., and Church, G.M. (2019). Functional genomics of the rapidly replicating bacterium Vibrio natriegens by CRISPRi. Nat. Microbiol. *4*, 1105–1113. https://doi.org/10.1038/s41564-019-0423-8.

74. Choe, D., Szubin, R., Poudel, S., Sastry, A., Song, Y., Lee, Y., Cho, S., Palsson, B., and Cho, B.-K. (2021). RiboRid: a low cost, advanced, and ultra-efficient method to remove ribosomal RNA for bacterial transcriptomics. PLoS Genet. *17*, e1009821. https://doi.org/10.1371/journal.pgen.1009821.

75. Page, A.J., Cummins, C.A., Hunt, M., Wong, V.K., Reuter, S., Holden, M.T.G., Fookes, M., Falush, D., Keane, J.A., and Parkhill, J. (2015). Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics *31*, 3691–3693. https://doi.org/10.1093/bioinformatics/btv421.

76. Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25. https://doi.org/10.1186/gb-2009-10-3-r25.

77. Wang, L., Wang, S., and Li, W. (2012). RSeQC: quality control of RNA-seq experiments. Bioinformatics *28*, 2184–2185. https://doi.org/10.1093/bioinformatics/bts356.

78. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930. https://doi.org/10.1093/bioinformatics/btt656.

79. Ewels, P., Magnusson, M., Lundin, S., and Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics *32*, 3047–3048. https://doi.org/10.1093/bioinformatics/btw354.

80. Hyvärinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. IEEE Trans. Neural Network. *10*, 626–634. https://doi.org/10.1109/72.761722.

81. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: machine learning in Python. J. Mach. Learn. Res. *12*, 2825–2830.

82. EsterKriegel, Sander, and Xu. (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise (KDD).

83. Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S.K., Cook, H., Mende, D.R., Letunic, I., Rattei, T., Jensen, L.J., et al. (2019). eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res. *47*, D309–D314. https://doi.org/10.1093/nar/gky1085.

84. Erian, A.M., Freitag, P., Gibisch, M., and Pflügl, S. (2020). High rate 2,3-butanediol production with Vibrio natriegens. Bioresource Technology Reports *10*, 100408. https://doi.org/10.1016/j.biteb.2020.100408.

85. Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., 3rd, and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods *6*, 343–345. https://doi.org/10.1038/nmeth.1318.

86. Hulsen, T., Huynen, M.A., de Vlieg, J., and Groenen, P.M.A. (2006). Benchmarking ortholog identification methods using functional genomics data. Genome Biol. *7*, R31. https://doi.org/10.1186/gb-2006-7-4-r31.

87. McConn, J.L., Lamoureux, C.R., Poudel, S., Palsson, B.O., and Sastry, A.V. (2021). Optimal dimensionality selection for independent component analysis of transcriptomic data. BMC Bioinf. *22*, 584. https://doi.org/10.1186/s12859-021-04497-7.

88. Novichkov, P.S., Kazakov, A.E., Ravcheev, D.A., Leyn, S.A., Kovaleva, G.Y., Sutormin, R.A., Kazanov, M.D., Riehl, W., Arkin, A.P., Dubchak, I., and Rodionov, D.A. (2013). RegPrecise 3.0–a resource for genome-scale exploration of transcriptional regulation in bacteria. BMC Genom. *14*, 745. https://doi.org/10.1186/1471-2164-14-745.

89. Karp, P.D., Billington, R., Caspi, R., Fulcher, C.A., Latendresse, M., Kothari, A., Keseler, I.M., Krummenacker, M., Midford, P.E., Ong, Q., et al. (2019). The BioCyc collection of microbial genomes and metabolic pathways. Brief. Bioinform. *20*, 1085–1093. https://doi.org/10.1093/bib/bbx085.

90. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. *44*, D457–D462. https://doi.org/10.1093/nar/gkv1070.

91. UniProt Consortium (2021). UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Res. *49*, D480–D489. https://doi.org/10.1093/nar/gkaa1100.

92. Gene Ontology Consortium (2021). The Gene Ontology resource: enriching a GOld mine. Nucleic Acids Res. *49*, D325–D334. https://doi.org/10.1093/nar/gkaa1113.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
Escherichia coli NEB 10-beta	New England Biolabs	Cat#C3020K
Vn (Wild-type *Vibrio natrigens* ATCC 14048)	ATCC	ATCC 14048
Vn-P_{Trc} (Vn derivative, pTrc; CmR)	This study	N/A
Vn-P_{Trc}-TfoX (Vn derivative, pTrc-tfoX; Cm^R)	This study	N/A
Vn-P_{Trc}-QstR (Vn derivative, pTrc-qstR; Cm^R)	This study	N/A
Vn-P_{Trc}-HapR (Vn derivative, pTrc-hapR; Cm^R)	This study	N/A
Vn-P_{Trc}-LuxO (Vn derivative, pTrc-luxO; Cm^R)	This study	N/A
Vn-P_{Trc}-TfoY (Vn derivative, pTrc-tfoY; Cm^R)	This study	N/A
Vn-P_{BAD}-TfoX (Vn derivative, pBAC-tfoX; Cm^R)	This study	N/A
Vn-P_{BAD}-QstR (Vn derivative, pBAC-qstR; Cm^R)	This study	N/A
Vn-P_{BAD}-HapR (Vn derivative, pBAC-hapR; Cm^R)	This study	N/A
Vn-P_{BAD}-LuxO (Vn derivative, pBAC-luxO; Cm^R)	This study	N/A
Vn-P_{BAD}-TfoY (Vn derivative, pBAC-tfoY; Cm^R)	This study	N/A
TfoX-Δ1; Vn-Δ(RS01335, RS01340, RS01345):$Carb^R$, carrying pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ2; Vn-Δ(RS01560,RS01555):$Carb^R$, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ3; Vn-Δ(RS01590,RS01595):$Carb^R$, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ4; Vn-ΔRS01825:$Carb^R$, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ5; Vn-ΔRS01950:$Carb^R$, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ6; Vn-ΔRS01975:$Carb^R$, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ7; Vn-ΔRS02005:$Carb^R$, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ8; Vn-ΔRS02075:$Carb^R$, pTrc-*tfoX* vector; Cm^R	This study	N/A
UC5-Δ1; Vn-Δ(RS02830,RS02835):$Carb^R$, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ9; Vn-ΔRS02925:$Carb^R$, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ10; Vn-ΔRS02935:$Carb^R$, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ11; Vn-ΔRS02945:$Carb^R$, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ12; Vn-ΔRS03620:$Carb^R$, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ13; Vn-Δ(RS03750,RS03755,RS03760):$Carb^R$, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ14; Vn-ΔRS03805:$Carb^R$, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ15; Vn-Δ(RS04585,RS04605,RS04610,RS04615, RS04625):$Carb^R$, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ16; Vn-ΔRS04635:$Carb^R$, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ17; Vn-Δ(RS04875,RS04880,RS04885):$Carb^R$, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ18; Vn-Δ(RS04895,RS04900):$Carb^R$, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ19; Vn-Δ(RS04945,RS04950,RS04955):$Carb^R$, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ20; Vn-ΔRS05160:$Carb^R$, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ21; Vn-ΔRS05480:$Carb^R$, pTrc-*tfoX* vector; Cm^R	This study	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
TfoX-Δ22; Vn-Δ(RS05740,RS05745):Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ23; Vn-ΔRS05800:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ24; Vn-ΔRS06085:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ25; Vn-ΔRS06105:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ26; Vn-Δ(RS06805,RS06810,RS06815, RS06820,RS06825,RS06830):Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ27; Vn-ΔRS23875:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ28; Vn-Δ(RS06835,RS06840,RS06845):Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ29; Vn-Δ(RS06945,RS06950):Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ30; Vn-ΔRS07235:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ31; Vn-ΔRS07295:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ32; Vn-Δ(RS08200,RS08205,RS08210,RS08215): Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ33; Vn-ΔRS08710:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ34; Vn-ΔRS09165:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ35; Vn-Δ(RS09315,RS09320,RS09325):Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ36; Vn-Δ(RS09605,RS09610):Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ37; Vn-Δ(RS10345,RS10350,RS10355,RS10360, RS10365):Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ38; Vn-ΔRS10370:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ39; Vn-ΔRS10715:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ40; Vn-ΔRS11630:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ41; Vn-ΔRS11725:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ42; Vn-Δ(RS11950,RS11955):Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ43; Vn-ΔRS12295:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ44; Vn-Δ(RS12325,RS12330):Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ45; Vn-ΔRS12935:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ46; Vn-ΔRS13070:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ47; Vn-Δ(RS13355,RS13360):Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ48; Vn-ΔRS13470:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ49; Vn-ΔRS14155:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ50; Vn-ΔRS14960:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ51; Vn-ΔRS15365:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ52; Vn-Δ(RS15470,RS15475):Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ53; Vn-ΔRS15585:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ54; Vn-ΔRS16340:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ55; Vn-Δ(RS16495,RS16500,RS16505):Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ56; Vn-ΔRS16660:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ57; Vn-ΔRS17215:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
TfoX-Δ58; Vn-ΔRS17355:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ59; Vn-ΔRS17690:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ60; Vn-ΔRS18665:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ61; Vn-Δ(RS20105,RS20110):Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ62; Vn-ΔRS20170:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ63; Vn-ΔRS20750:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ64; Vn-ΔRS21085:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ65; Vn-ΔRS21100:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ66; Vn-Δ(RS21820,RS21825,RS21830,RS21835, RS21840):Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ67; Vn-ΔRS22835:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ68; Vn-ΔRS23455:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
TfoX-Δ25+ CP; Vn-ΔRS06105:Carb^R, pTrc-tfoX vector, Cm^R; pUC-SpecR-TfoX-Δ25, Spec^R	This study	N/A
TfoX-Δ32+ CP; Vn-Δ(RS08200,RS08205,RS08210, RS08215):Carb^R, pTrc-tfoX vector, Cm^R; pUC-SpecR-TfoX-Δ32, Spec^R	This study	N/A
TfoX-Δ42+ CP; Vn-Δ(RS11950,RS11955):CarbR, pTrc-tfoX vector, Cm^R; pUC-SpecR-TfoX-Δ42, Spec^R	This study	N/A
QstR-Δ1; Vn-ΔRS02195:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
QstR-Δ2; Vn-ΔRS04580:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
QstR-Δ3; Vn-ΔRS12260:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
QstR-Δ4; Vn-ΔRS12610:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
QstR-Δ5; Vn-ΔRS16805:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
QstR-Δ6; Vn-ΔRS16880:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
QstR-Δ1+CP; Vn-ΔRS02195:Carb^R, pTrc-tfoX vector; Cm^R, pUC-SpecR-QstR-Δ1, Spec^R	This study	N/A
QstR-Δ2+CP; Vn-ΔRS04580:Carb^R, pTrc-tfoX vector; Cm^R, pUC-SpecR-QstR-Δ2, Spec^R	This study	N/A
QstR-Δ3+CP; Vn-ΔRS12260:Carb^R, pTrc-tfoX vector; Cm^R, pUC-SpecR-QstR-Δ3, Spec^R	This study	N/A
QstR-Δ4+CP; Vn-ΔRS12610:Carb^R, pTrc-tfoX vector; Cm^R, pUC-SpecR-QstR-Δ4, Spec^R	This study	N/A
Fla1-Δ1; Vn-Δ(RS02635,RS02640,RS02645,RS02650, RS02655, RS02660):Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
ArgR-Δ1; Vn-Δ(RS15005,RS15010,RS15015,RS15020): Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
ArgR-Δ1-TfoX–; Vn-Δ(RS15005,RS15010,RS15015, RS15020):Carb^R	This study	N/A
UC5-Δ2; Vn-ΔRS08390:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
UC5-Δ3; Vn-Δ(RS13890,RS13900,RS13905):Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
UC5-Δ4; Vn-Δ(RS15620,RS15625):Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
UC5-Δ5; Vn-ΔRS17200:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
UC5-Δ6; Vn-ΔRS22995:Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
UC5-Δ2+CP; Vn-ΔRS08390:Carb^R, pTrc-tfoX vector; Cm^R, pUC-SpecR-UC5-Δ2, Spec^R	This study	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
UC5-Δ3+CP; Vn-Δ(RS13890,RS13900,RS13905):Carb^R, pTrc-tfoX vector; Cm^R, pUC-SpecR-UC5-Δ3, Spec^R	This study	N/A
AhpCF-Δ1; Vn-Δ(RS16355,RS16360):Carb^R, pTrc-*tfoX* vector; Cm^R	This study	N/A
AhpCF-Δ1-TfoX–; Vn-Δ(RS16355,RS16360):Carb^R	This study	N/A
Chemicals, peptides, and recombinant proteins		
LB Miller broth	Sigma-Aldrich	Cat#71753-6
Brain Heart Infusion broth	Sigma-Aldrich	Cat#53286
Mueller Hinton Broth	Sigma-Aldrich	Cat#70192
Nutrient Broth	BD Difco™	Cat#234000
Critical commercial assays		
DCFDA/H2DCFDA - Cellular ROS Assay Kit	Abcam	Cat#ab113851
Deposited data		
Genome sequences of *V. natriegens*	Lee et al.[73]	GeneBank: GCA_001456255.1
All RNA-Seq raw data generated in this study	This study	NCBI SRA: PRJNA880800
RNA-Seq data for Salinity and Temperature Variations	CBMSE, Naval Research Laboratory	NCBI SRA: PRJNA763369
RNA-Seq data for Microgravity	Yin et al.[34]	NCBI SRA: PRJNA624924
Oligonucleotides		
See Table S3 for oligonucleotides used in this study	This study	N/A
Recombinant DNA		
pACYC184; for Cloning and protein expression (p15A origin; Cm^R, Tc^R)	New England Biolabs	Cat#C3020K
pTrcHis2A; for Cloning and protein expression (pBR322 origin; Trc and lacO promoter; Amp^R)	ThermoFisher Scientific	Cat#V36520
pBAD/Myc-HisC; for Cloning and protein expression (pBR322 origin; araBAD promoter; Amp^R)	ThermoFisher Scientific	Cat#V44001
pBAD; for Arabinose-inducible protein expression (pACYC derivative containing araC-araBAD promoter fragment; Cm^R)	This study	N/A
pBAC-tfoX; for Arabinose-inducible protein expression (pBAD derivative encoding endogenous tfoX gene under control of araBAD promoter)	This study	N/A
pBAC-tfoY; for Arabinose-inducible protein expression (pBAD derivative encoding endogenous tfoY gene under control of araBAD promoter)	This study	N/A
pBAC-qstR; for Arabinose-inducible protein expression (pBAD derivative encoding endogenous qstR gene under control of araBAD promoter)	This study	N/A
pBAC-hapR; for Arabinose-inducible protein expression (pBAD derivative encoding endogenous hapR gene under control of araBAD promoter)	This study	N/A
pBAC-luxo; for Arabinose-inducible protein expression (pBAD derivative encoding endogenous luxO gene under control of araBAD promoter)	This study	N/A
pTrc; for IPTG-inducible protein expression (pACYC derivative containing lacI-Trc promoter fragment; CmR)	This study	N/A
pTrc-tfoX; for IPTG-inducible protein expression (pBAD derivative encoding endogenous tfoX gene under control of Trc promoter)	This study	N/A
pTrc-tfoY; for IPTG-inducible protein expression (pBAD derivative encoding endogenous tfoY gene under control of Trc promoter)	This study	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
pTrc-qstR; for IPTG-inducible protein expression (pBAD derivative encoding endogenous qstR gene under control of Trc promoter)	This study	N/A
pTrc-hapR; for IPTG-inducible protein expression (pBAD derivative encoding endogenous hapR gene under control of Trc promoter)	This study	N/A
pTrc-luxO; for IPTG-inducible protein expression (pBAD derivative encoding endogenous luxO gene under control of Trc promoter)	This study	N/A
ptDNA-TfoX-Δ1; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ1 strain)	This study	N/A
ptDNA-TfoX-Δ2; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ2 strain)	This study	N/A
ptDNA-TfoX-Δ3; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ3 strain)	This study	N/A
ptDNA-TfoX-Δ4; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ4 strain)	This study	N/A
ptDNA-TfoX-Δ5; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ5 strain)	This study	N/A
ptDNA-TfoX-Δ6; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ6 strain)	This study	N/A
ptDNA-TfoX-Δ7; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ7 strain)	This study	N/A
ptDNA-TfoX-Δ8; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ8 strain)	This study	N/A
ptDNA-UC5-Δ1; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating UC5-Δ1 strain)	This study	N/A
ptDNA-TfoX-Δ9; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ9 strain)	This study	N/A
ptDNA-TfoX-Δ10; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ10 strain)	This study	N/A
ptDNA-TfoX-Δ11; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ11 strain)	This study	N/A
ptDNA-TfoX-Δ12; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ12 strain)	This study	N/A
ptDNA-TfoX-Δ13; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ13 strain)	This study	N/A
ptDNA-TfoX-Δ14; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ14 strain)	This study	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
ptDNA-TfoX-Δ15; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ15 strain)	This study	N/A
ptDNA-TfoX-Δ16; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ16 strain)	This study	N/A
ptDNA-TfoX-Δ17; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ17 strain)	This study	N/A
ptDNA-TfoX-Δ18; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ18 strain)	This study	N/A
ptDNA-TfoX-Δ19; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ19 strain)	This study	N/A
ptDNA-TfoX-Δ20; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ20 strain)	This study	N/A
ptDNA-TfoX-Δ21; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ21 strain)	This study	N/A
ptDNA-TfoX-Δ22; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ22 strain)	This study	N/A
ptDNA-TfoX-Δ23; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ23 strain)	This study	N/A
ptDNA-TfoX-Δ24; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ24 strain)	This study	N/A
ptDNA-TfoX-Δ25; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ25 strain)	This study	N/A
ptDNA-TfoX-Δ26; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ26 strain)	This study	N/A
ptDNA-TfoX-Δ27; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ27 strain)	This study	N/A
ptDNA-TfoX-Δ28; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ28 strain)	This study	N/A
ptDNA-TfoX-Δ29; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ29 strain)	This study	N/A
ptDNA-TfoX-Δ30; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ30 strain)	This study	N/A
ptDNA-TfoX-Δ31; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ31 strain)	This study	N/A
ptDNA-TfoX-Δ32; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ32 strain)	This study	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
ptDNA-TfoX-Δ33; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ33 strain)	This study	N/A
ptDNA-TfoX-Δ34; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ34 strain)	This study	N/A
ptDNA-TfoX-Δ35; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ35 strain)	This study	N/A
ptDNA-TfoX-Δ36; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ36 strain)	This study	N/A
ptDNA-TfoX-Δ37; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ37 strain)	This study	N/A
ptDNA-TfoX-Δ38; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ38 strain)	This study	N/A
ptDNA-TfoX-Δ39; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ39 strain)	This study	N/A
ptDNA-TfoX-Δ40; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ40 strain)	This study	N/A
ptDNA-TfoX-Δ41; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ41 strain)	This study	N/A
ptDNA-TfoX-Δ42; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ42 strain)	This study	N/A
ptDNA-TfoX-Δ43; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ43 strain)	This study	N/A
ptDNA-TfoX-Δ44; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ44 strain)	This study	N/A
ptDNA-TfoX-Δ45; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ45 strain)	This study	N/A
ptDNA-TfoX-Δ46; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ46 strain)	This study	N/A
ptDNA-TfoX-Δ47; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ47 strain)	This study	N/A
ptDNA-TfoX-Δ48; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ48 strain)	This study	N/A
ptDNA-TfoX-Δ49; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ49 strain)	This study	N/A
ptDNA-TfoX-Δ50; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ50 strain)	This study	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
ptDNA-TfoX-Δ51; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ51 strain)	This study	N/A
ptDNA-TfoX-Δ52; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ52 strain)	This study	N/A
ptDNA-TfoX-Δ53; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ53 strain)	This study	N/A
ptDNA-TfoX-Δ54; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ54 strain)	This study	N/A
ptDNA-TfoX-Δ55; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ55 strain)	This study	N/A
ptDNA-TfoX-Δ56; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ56 strain)	This study	N/A
ptDNA-TfoX-Δ57; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ57 strain)	This study	N/A
ptDNA-TfoX-Δ58; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ58 strain)	This study	N/A
ptDNA-TfoX-Δ59; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ59 strain)	This study	N/A
ptDNA-TfoX-Δ60; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ60 strain)	This study	N/A
ptDNA-TfoX-Δ61; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ61 strain)	This study	N/A
ptDNA-TfoX-Δ62; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ62 strain)	This study	N/A
ptDNA-TfoX-Δ63; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ63 strain)	This study	N/A
ptDNA-TfoX-Δ64; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ64 strain)	This study	N/A
ptDNA-TfoX-Δ65; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ65 strain)	This study	N/A
ptDNA-TfoX-Δ66; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ66 strain)	This study	N/A
ptDNA-TfoX-Δ67; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ67 strain)	This study	N/A
ptDNA-TfoX-Δ68; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating TfoX-Δ68 strain)	This study	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
ptDNA-QstR-Δ1; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating QstR-Δ1 strain)	This study	N/A
ptDNA-QstR-Δ2; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating QstR-Δ2 strain)	This study	N/A
ptDNA-QstR-Δ3; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating QstR-Δ3 strain)	This study	N/A
ptDNA-QstR-Δ4; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating QstR-Δ4 strain)	This study	N/A
ptDNA-QstR-Δ5; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating QstR-Δ5 strain)	This study	N/A
ptDNA-QstR-Δ6; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating QstR-Δ6 strain)	This study	N/A
ptDNA-Fla1-Δ1; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating Fla1-Δ1 strain)	This study	N/A
ptDNA-ArgR-Δ1; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating ArgR-Δ1 strain)	This study	N/A
ptDNA-UC5-Δ2; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating UC5-Δ2 strain)	This study	N/A
ptDNA-UC5-Δ3; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating UC5-Δ3 strain)	This study	N/A
ptDNA-UC5-Δ4; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating UC5-Δ4 strain)	This study	N/A
ptDNA-UC5-Δ5; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating UC5-Δ5 strain)	This study	N/A
ptDNA-UC5-Δ6; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating UC5-Δ6 strain)	This study	N/A
ptDNA-AhpCF-Δ1; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating AhpCF-Δ1 strain)	This study	N/A
ptDNA-EndA-KanR; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating ΔendA strain with kanamycin resistance gene)	This study	N/A
ptDNA-EndA-SpecR; for transfer DNA generation (pACYC derivative containing transfer DNA fragment for generating ΔendA strain with spectinomycin resistance gene)	This study	N/A
pUC-SpecR; pUC origin; SpecR	This study	N/A
pUC-SpecR-PRUT; pUC-specR derivative containing promoter (J23106), RBS (B0033m), and terminator (Bba_0012)	This study	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
pUC-SpecR-TfoX-Δ25; complementing the expression of the deleted gene(s) in TfoX-Δ25 mutant	This study	N/A
pUC-SpecR-TfoX-Δ42; complementing the expression of the deleted gene(s) in TfoX-Δ42 mutant	This study	N/A
pUC-SpecR-TfoX-Δ32; complementing the expression of the deleted gene(s) in TfoX-Δ32 mutant	This study	N/A
pUC-SpecR-UC5-Δ2; complementing the expression of the deleted gene(s) in UC5-Δ2 mutant	This study	N/A
pUC-SpecR-UC5-Δ3; complementing the expression of the deleted gene(s) in UC5-Δ3 mutant	This study	N/A
pUC-SpecR-QstR-Δ3; complementing the expression of the deleted gene(s) in QstR-Δ3 mutant	This study	N/A
pUC-SpecR-QstR-Δ1; complementing the expression of the deleted gene(s) in QstR-Δ1 mutant	This study	N/A
pUC-SpecR-QstR-Δ2; complementing the expression of the deleted gene(s) in QstR-Δ2 mutant	This study	N/A
pUC-SpecR-QstR-Δ4; complementing the expression of the deleted gene(s) in QstR-Δ4 mutant	This study	N/A
Software and algorithms		
RiboRid	Choe et al.[74]	https://github.com/SBRG/RiboRid_Design
R	The R Project for Statistical Computing	https://cran.r-project.org/
Python	Python Software Foundation	https://www.python.org/
GenBank2gff3	BioPerl	https://github.com/bioperl/bioperl-live
Roary	Page et al.[75]	https://sanger-pathogens.github.io/Roary/
fasterq-dump	NCBI	https://github.com/ncbi/sra-tools/
Modulome workflow	Sastry et al.[30]	https://github.com/avsastry/modulome-workflow
Trim Galore	Babraham Bioinformatics	https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/
FastQC	Babraham Bioinformatics	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
Bowtie	Langmead et al.[76]	https://bowtie-bio.sourceforge.net/index.shtml
sam2bam	Samtools	http://www.htslib.org/
RSEQC	Wang et al.[77]	https://rseqc.sourceforge.net/
FeatureCounts	Liao et al.[78]	https://subread.sourceforge.net/
MultiQC	Ewels et al.[79]	https://multiqc.info/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
FastICA	Hyvärinen[80]	https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.FastICA.html
Scikit-Learn	Pedregosa et al.[81]	https://scikit-learn.org/stable/index.html
DBSCAN	Ester et al.[82]	https://scikit-learn.org/stable/modules/generated/sklearn.cluster.DBSCAN.html
Pymodulon	Sastry et al.[30]	https://github.com/SBRG/pymodulon
EggNOG mapper	Huerta-Cepas et al.[83]	https://github.com/eggnogdb/eggnog-mapper
GraphPad Prism	GraphPad	https://www.graphpad.com/scientific-software/prism/
Other		
iModulon data of *V. natrigens*	This study	https://imodulondb.org/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by Dr. Bernhard O. Palsson (palsson@ucsd.edu).

Materials availability

This study did not generate new unique reagents and physical samples. Strains used in this study are available upon request from Dr. Bernhard O. Palsson. All computational results of this study are described below in the "Data and code availability" section.

Data and code availability

- All RNA-seq data generated in this study have been deposited at NCBI SRA and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. All iModulon data produced in this study is publicly available at iModulon DB web portal (https://imodulondb.org/)
- All original code used in this study is freely and fully available through several venues (DOIs also listed in the key resources table).
- Any additional data reported in this study will be available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Bacterial strains and growth conditions

Unless stated otherwise, chemical reagents for cell culture were purchased from Sigma-Aldrich (Burlington, MA). Bacterial strains used/generated in this study are described in Table S3. *Vibrio natriegens* ATCC 14048 (*Vn*) were used as the wild type strains. *Vn* was routinely cultured aerobically at 30°C in LBv2 (25 g/L LB Miller broth, 200 mM NaCl, 4.2 mM KCl, and 23.14 mM $MgCl_2$) with agitation at 180 rpm or LBv2 agar plates (LBv2 supplemented with 1.5% wt/vol agar). For the construction of 102 RNA-Seq data, *Vn* strains were cultivated at 30°C under the different media types as follows: LBv2 (25 g/L LB, 200 mM NaCl, 4.2 mM KCl, and 23.14 mM $MgCl_2$); BHIN (37 g/L Brain Heart Infusion broth [#53286], 1.5% wt/vol NaCl); MHBN (21 g/L Mueller Hinton Broth [#70192], 1.5% wt/vol NaCl); NBON (8 g/L Nutrient Broth [BD Difco, Franklin Lakes, NJ, #234000], 1.5% wt/vol NaCl); E2xYT (32 g/L tryptone, 20 g/L yeast extract, 17 g/L NaCl, 0.4% wt/vol glucose, 17.6 mM Na_2HPO4, pH 7.4); M9Na (M9 minimal medium with additional NaCl [42 mM Na_2HPO_4, 22 mM KH_2PO_4, 258.5 mM NaCl, 18.6 mM NH_4Cl, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$]) supplemented with 0.5–1% carbon source. Based on the previous reports,[16,84] we selected different carbon sources for the growth of *Vn* in the M9Na minimal medium. For M9Na minimal medium, the carbon sources were added at the following concentrations: 0.5% wt/vol glucose, 1.0% wt/vol arabinose, 1.0% wt/vol galactose, 1.0% wt/vol N-acetylglucosamine (GlcNAc), 1.0% wt/vol gluconate, 1.0% wt/vol glucosamine, 1.0% wt/vol glycerol, 1.0% wt/vol maltose, 1.0% wt/vol mannitol, 1.0% wt/vol trehalose, 1.0% wt/vol ethanol, 1.0% wt/vol fructose, 1.0% wt/vol fumarate, 1.0% wt/vol ribose, 1.0% wt/vol succinate, and 1.0% wt/vol sucrose. General LB media was used for low-salt conditions. High-salt stress and low pH conditions were prepared by adding V3 salts (475 mM NaCl, 9.7 mM KCl, and 54 mM $MgCl_2$) or HCl (up to pH 5.5). 25 μg/mL kanamycin[34] was added into the M9Na medium supplemented with 0.5% glucose for generating antibiotic stress conditions. Since *Vn* has a natural resistance to kanamycin slightly, they can grow slowly at this concentration. When the antibiotic selection of *Vn* was required, the antibiotics were used at the following

concentrations: 50 μg/mL carbenicillin (Carb), 10 μg/mL chloramphenicol (Cm), 100 μg/mL kanamycin (Kan), or 360 μg/mL spectinomycin (Spec). For plasmid cloning, NEB 10-beta Competent *Escherichia coli* (New England BioLabs, Ipswich, MA) was used and cultivated aerobically at 37°C in LB Miller broth (LB, #71753-6) with continuous shaking at 180 rpm or LB agar (LB with 1.5% wt/vol agar). When *E. coli* harbored a plasmid, appropriate antibiotics were used as follows: 100 μg/mL ampicillin/carbenicillin, 25 μg/mL chloramphenicol, or 50 μg/mL kanamycin.

METHOD DETAILS

Plasmid construction

Oligonucleotides used in this study are listed in Table S3. PrimeSTAR GXL DNA Polymerase (Takara Bio, Shiga, Japan) or Q5 High-Fidelity DNA Polymerase (New England Biolabs) were used for all high-fidelity PCR amplifications and genotype analysis. All oligonucleotides were synthesized by Integrated DNA Technologies (Coralville, IA). Restriction enzymes were purchased from New England Biolabs unless stated otherwise. PCR products and plasmids from *E. coli* were purified using the DNA Clean & Concentrator (Zymo Research, Irvine, CA) and Monarch Plasmid DNA Miniprep Kit (New England Biolabs), respectively. Plasmids were constructed by Gibson assembly[85] using NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs). We engineered the pACYC184 plasmid by deleting the tetracycline resistance gene using primer pair pACYC-cat-ori_fwd/rev. For arabinose- and isopropyl β-D-1-thiogalactopyranoside (IPTG)-inducible expression in *Vn*, the lacI-Trc promoter (induced by IPTG) and araC-araBAD promoter (induced by arabinose) were amplified from pBAD/Myc-HisC and pTrcHis2A using primer pairs ptrc_fwd–ptrc_rev, respectively. Then, the two types of promoters were cloned into the linearized pACYC plasmid using the NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs), resulting in pBAC and pTRC. These plasmids allow inducible expression from the araBAD or Trc promoter upon arabinose or IPTG provision. For the construction of pTrc (or pBAD)-tfoX, -tfoY, -qstR, -hapR, and -luxO plasmids, the transcription factors were amplified from *Vn* genomic DNA using primer pairs "*target*"_Trc_(or BAD)_fwd–"*target*"_Trc_(or BAD)_rev, and cloned into the linearized pTrc (or pBAD) plasmid (amplified by pTrc [or BAD]_Cir_F–pTRC [or BAD]_Cir_R primers) using the NEBuilder HiFi DNA Assembly Master Mix. For complementation experiments, linearized pUC (pUC-ori_fwd/rev) and spectinomycin resistance cassette (Sm_first_For/Rev, Sm_second_For/Rev, Sm_third_For/Rev serially) were amplified by PCR, and the two DNA fragments were assembled using NEBuilder HiFi DNA Assembly Master Mix, resulting in pUC-Spec^R plasmid. A synthetic PRUT cassette was then assembled, consisting of a promoter (J23106), an RBS (B0033m), and a terminator (Bba_0012), using the primer-pair (J23106_F/Bba_B0012_R). This cassette was then cloned into the linearized pUC-Sm plasmid (pUC-Sm_cassette_fwd/_rev), resulting in the generation of pUC-Spec^R-PRUT. To achieve ectopic expression of newly identified competence genes, each target gene was amplified using primer pairs "KO-strain-name"_CP_fwd and "KO-strain-name"_CP_rev from genomic DNA of *Vn*, and then cloned into linearized pUC-SpecdR-PRUT (pUC-Sm_PRUT_fwd/_rev), resulting in 9 pUC-SpecR-"KO-strain-name" plasmids. All plasmids were validated by Sanger sequencing (Eton Bioscience, San Diego, CA).

Electroporation

Electrocompetent *Vn* was prepared similarly to a previously described.[73] Briefly, *Vn* was grown LBv2 overnight at 30°C with continuous shaking at 180 rpm. On the following day, the *Vn* cells were pelleted at 5,000 × g for 5 min at 4°C, washed in LBv2, and used to inoculate a new culture at 1:100 dilution in 50 mL LBv2. The cell culture was grown at 30°C until an OD_{600} of 0.4 at 180 rpm. The cells were centrifuged at 5,000 × g for 5 min at 4°C; then, cells were washed by resuspension in cold 1M sorbitol. The washing step was repeated two times. After the wash, the cells were resuspended in 250 μL cold 1M sorbitol, and 50 μL of cells were aliquoted into 1.5 mL microcentrifuge tubes. For electrotransformation, the amount of DNA (total 50–200 ng in 2–3 μL) was added to the aliquot of electrocompetent cells and then transferred to a 0.2 mm gap size Gene Pulser cuvette (Bio-Rad Laboratories). Using Bio-Rad Gene Pulser, the cells were paused at 800 V, 25 μF, and 1000 Ω. The cells were immediately recovered in 1mL LBv2 media for 1 h at 30 °C at 180 rpm. The recovered cells were plated on an LBv2 agar plate (1.5% agar) containing an appropriate antibiotic as needed. Plates were incubated for at least 12 h at room temperature or at least 8 h at 30°C.

Transforming DNA (transfer DNA) preparation

For generation of two transforming DNA (tDNA-Kan^R and tDNA-Carb^R), we linearized pACYC184 by PCR amplification using primer pair pACYC_Vec_For/Rev. For the construction of Carb^R and Kan^R antibiotic resistance cassettes, beta-lactamase gene (carbenicillin resistance) and kanamycin resistance genes were amplified using primer pair (Amp_For/Rev or Km_For/Rev primer pairs) with the strong synthetic constitutive promoter (BBa_J23111), RBS (BBa_B0034), and terminator (BBa_B1010) sequence. Homology of about 2 kb on either side of the mutation results in the highest transformation efficiencies.[10] Homology arm (HA) of about 2 kb on both sides of the target operon was amplified from *Vn* genomic DNA using the "KO-strain-name"_Left/Right_For/Rev primer set. For subcloning of transfer DNA, each left HA, right HA, and antibiotic cassette (Kan^R for deletion of *dns* (*endA*) gene, Carb^R for other genes) were cloned into the linearized pACYC184 DNA fragment using the NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs), resulting ptDNA-endA-Kan^R or ptDNA-"KO-strain-name" plasmids. For generation of tDNA-Spec^R, ptDNA-endA was linearized by PCR using primer pair ptDNA-endA_For/Rev. For the construction of Spec^R antibiotic resistance cassettes, spectinomycin resistance gene was amplified using three primer pair serially (Sm_first_For/Rev, Sm_second_For/Rev, Sm_third_For/Rev) with the strong synthetic constitutive promoter (BBa_J23111), RBS (BBa_B0034), and terminator (BBa_B1010) sequence. For subcloning of

tDNA-Spec^R, the antibiotic cassette (Spec^R for deletion of *dns* (*endA*) gene) were cloned into the linearized ptDNA-endA fragment using the NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs), resulting in ptDNA-endA-Spec^R. All subcloned transfer DNA plasmids were validated by Sanger sequencing (Eton Bioscience). Then, the tDNAs were PCR-amplified with primer pair "KO-strain-name"_Left_For/Right_Rev using each KO plasmid as a template. The PCR products were digested by DpnI (New England Biolabs) and purified using the DNA Clean & Concentrator (Zymo Research).

Natural transformation

Natural transformation assays were carried out in a similar manner to a previously reported method.[10] For the activation of the natural competence system, Vn-PTrc-TfoX strain carrying pTrc-tfoX plasmid (inducible expression of tfoX from Trc promoter upon provision of IPTG) was grown overnight (12−18 h) in LBv2 media supplemented with 10 μg/mL chloramphenicol and 100 μM IPTG in a roller drum at 30°C. When the expression of TfoXs from an araBAD promoter was required, Vn-PBAD-TfoX strain carrying plasmid pBAD-tfoX (induced by 0.2% wt/vol arabinose) was used for natural competence. For each transformation reaction, the optical density (OD_{600}) of the cell culture was adjusted to 4.0, and this 5 μL of culture was transferred to 350 μL of competence buffer (referred to as CB, 28 g/L of Instant Ocean Sea Salt, www.instantocean.com) supplemented with 0.2% arabinose or 500 μM IPTG. The amount of transforming DNA (transfer DNA, total 50 ng) was added to the cells and gently inverted to mix. The competence reactions were incubated at 30°C for 4 h without agitation. The reactions were recovered by transferring them to 1 mL of fresh LBv2 and incubating them for 2 h at 30°C with agitation at 180 rpm. Before cell plating, samples were centrifuged for 5 min at 5,000 × g at 4°C, the supernatant was removed, and cells were then resuspended in 200 μL of LBv2. To determine the transformation efficiency, 1:2 and 1:10 dilution factors are typically used to select transformants that select for integration of selected product LBv2 supplemented with 50 μg/mL Carb, 100 μg/mL Kan, or 360 μg/mL Spec) and 1:10000 and 1:100000 dilution are used to determine the total colony-forming units (CFUs) on non-selective LBv2 agar plates. Transformation efficiency was calculated as the number of transformants divided by the total CFUs. Deletion of the target gene was verified by PCR with purified genomic DNA using "KO-strain-name"_Val_For/Rev in Table S3.

To compare transformation efficiency among the 83 deletion strains (showing Carb^R phenotype), we carried out 83 unique transformation assays with each deletion strain as described above. For the activation of the natural competence system, each deletion strain containing pTrc-tfoX plasmid was grown overnight (12−18 h) in LBv2 media supplemented with 10 μg/mL Cm, 50 μg/mL Carb, and 100 μM IPTG in a roller drum at 30°C. For the competence assay, 50 ng of linear *dns* (*endA*) Kan^R transfer DNA (tDNA-Kan^R), linear *dns* (*endA*) Spec^R transfer DNA (tDNA-Spec^R), or circular pUC-Spec^R was consistently used for all strains. The transformed cells were plated onto a selective agar plate (LBv2 supplemented with 50 μg/mL Carb and 100 μg/mL Kan or 360 μg/mL Spec) and non-selective agar plate (LBv2 supplemented with 50 μg/mL Carb), resulting in the number of transformants and the total number of CFUs, respectively.

Comparative genomics

For comparative genome analysis, a total of 18 *Vibrio* strains whose natural competence activity was confirmed by chitin (or -dependent or chitin-independent [i.e., TfoX-dependent]) were selected (Table S1) as follows: seven *V. cholerae* (*Vc*) strains (A1552, c6706, c6709, E7946, P27459, N16961, MO10), three *V. vulnificus* (*Vv*) strains (CMCP6, MO6-24/O, ATCC 27562), *V. parahaemolyticus* (*Vp*) RIMD 2210633, *V. mimicus* (*Vm*) SCCF01, *Vn* ATCC 14048, *V. fischeri* (*Vf*)ES114, and four *V. campbellii* (*Vca*) strains (NBRC 15631, DS40M4, BB120, HY-01). Each genome was acquired in GenBank format from the public database NCBI (https://www.ncbi.nlm.nih.gov/guide/howto/dwn-genome/). The downloaded Genebank files were converted into GFF files with embedded nucleotide and amino acid sequences using BioPerl script (bp_GenBank2gff3.pl, https://github.com/bioperl/bioperl-live). We then generated a gene presence/absence matrix by Markov clustering based on an all-against-all bidirectional BLAST with a 50% minimum identity[86] through Roary v3.11.2.[75]

Generation of RNA-Seq compendium

A total of 102 RNA-Seq data accounting for "Carbons", "Rich/TF-OE/Stress-II", and "Competence" projects was newly generated in this study. For generating transcriptomic data, we extracted total RNAs for 48 unique conditions, including various media types (LBv2, BHIN, MHBN, NBON, and M9Na), various growth phases (mid-exponential, late-exponential, stationary phage), stress conditions (low-salt, high-salt, low pH, and antibiotic), various carbon sources (glucose, arabinose, galactose, GlcNAc, gluconate, glucosamine, glycerol, maltose, mannitol, trehalose, ethanol, fructose, fumarate, ribose, succinate, and sucrose), overexpression of competence-related TFs (TfoX, TfoY, QstR, HapR, and LuxO), and natural competence (TfoX+ and TfoX− for competent and incompetence control, respectively). The biological duplicate cells were collected under all conditions. When OD_{600} reached the desired value (Table S1), samples were then centrifuged for 10 min at 5,000 × g at 4°C, and the supernatant was removed. RNA-Seq libraries were constructed similarly to the previously reported.[28] Briefly, total RNA was extracted from the collected cells using a Quick-RNA Fungal/Bacterial Microprep Kit (Zymo Research) according to the manufacturer's protocols. For sequence-specific digestion of ribosomal RNAs in total RNAs, we designed anti-rRNA oligonucleotide probes for *Vn* using an in-house Python script (https://github.com/SBRG/RiboRid_Design). All sequences for the 65 anti-rRNA oligonucleotide probes are available in Table S3. As previously described, all ribosomal RNAs and genomic DNA contaminants were removed from 1 μg total RNA using the RiboRid method.[74] Depletion of rRNAs was confirmed using 4150 TapeStation System (Agilent, Santa Clara, CA) with High Sensitivity

RNA ScreenTape. The RNA-Seq libraries were constructed with the resulting rRNA-depleted RNA using a KAPA RNA HyperPrep kit (Roche, Basel, Switzerland) according to the manufacturer's instructions. The RNA-Seq libraries were quality-checked using 4150 TapeStation System (Agilent) with D1000 ScreenTape, and the libraries were quantified using the Qubit 2.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA) with Qubit dsDNA HS Assay Kit. The libraries were pooled and sequenced by the 100 bp single-end recipe on an Illumina NovaSeq 6000 platform at the UC San Diego IGM Genomics Center.

Compilation of natPRECISE104 dataset

In addition to 102 RNA-Seq data generated in this study, 42 RNA-Seq samples were also obtained from the NCBI Sequence Read Archive (released before Jan 01, 2022) with fasterq-dump software (https://github.com/ncbi/sra-tools/), resulting a total of 144 *Vn* RNA-Seq samples. Before ICA, data processing and quality control were carried out using previously reported Modulome workflow (https://github.com/avsastry/modulome-workflow).[30] Briefly, Trim Galore (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) and FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) were used for raw read trimming. High-quality reads were then mapped to the *Vn* reference genome (GenBank: GCA_001456255.1)[73] using Bowtie.[76] After the read mapping, generated SAM files were converted to BAM files using the sam2bam function of Samtools (http://www.htslib.org/). The read count of each gene in each library was calculated using RSEQC[77] and FeatureCounts.[78] For a high-quality expression dataset, all data quality metrics (Trim Galore, FastQC, Bowtie, RSEQC, and FeatureCounts) were compiled using MultiQC,[79] and all samples that did not fulfill these four criteria were rejected: per_base_sequence_quality, per_sequence_quality_scores, per_base_n_content, adapter_content. Samples were also discarded if sequencing depth was insufficient ($<4 \times 10^5$ reads mapped to coding sequence region). To reduce technical noise, we discarded samples if (a) a sample did not conform to a general expression pattern (determined by hierarchical clustering), (b) a sample showed poor correlation within biological replicates ($R^2 < 0.90$), and (c) samples has no biological replicate. Read counts were converted into $\log_2$ transcripts per million (log-TPM) (Table S2). We named this compendium natPRECISE104 (natriegens Precision RNA-seq Expression Compendium for Independent Signal Exploration). The natPRECISE104 includes 104 samples generated under 52 specific cell growth conditions within 6 projects. Among the 6 projects, four projects (n = 96) containing Carbon, Rich/TF-OE/, Stress-I, and Competence projects were created in this study, and other two projects (n = 8), including Stress-II and Stress-III, were obtained from NCBI SRA. First, in the case of the Carbon project, the transcriptome change according to the carbon sources was measured in a minimal medium. The Rich/TF-OE project measured transcriptome changes according to the rich media conditions and cell growth stage of overexpression strains of various competence-related TFs under rich media. In addition, there are three additional stress-related projects. The Stress-I contains transcriptome changes under high-salt, antibiotic (kanamycin), and low pH (pH 5.5) conditions. In addition, the Stress-II and Stress-III projects are derived from NCBI SRA and show changes in the transcriptome according to the environment of salinity/temperature and microgravity,[34] respectively. To eliminate potential batch effects, including technical and non-biological factors, we normalized each transcriptomic profile based on reference condition values within each project (Table S1) in the same manner as previously described.[23,28,29] This step assured that most independent components resulted from biological variation instead of technical variance. This normalization enables the comparison of gene expression and iModulon activity within a project to a reference condition but not across projects. After quality control, the final RNA-Seq data used for ICA is a total of 104 samples, which consists of 5 projects: Carbons (n = 44), Rich/TF-OE/Stress-II (n = 44), Competence (n = 8), Salinity/Temperature (Stress-I) (n = 4), and Microgravity (Stress-III) (n = 4). Three projects (n = 96) containing Carbon, Rich/TF-OE/Stress-II, and Competence projects were created in this study, and the other two projects (n = 8) were obtained from NCBI SRA (see Table S1 for accession numbers).

Independent component analysis (ICA)

ICA was used for decomposing transcriptomic data matrix (X, 4515 genes by 104 conditions) into two subcomponents (M and A, independent components [iModulons] and their activity, respectively) in each sample, where $X = M \times A$. To define the independent components, the ICA was conducted on the transcriptomic data according to a previously reported workflow (see OptICA step in modulone workflow, https://github.com/avsastry/modulome-workflow) with default parameters.[23,30] For ICA, independent components (ICs) were calculated through 100 iterations using the FastICA algorithm[80] and Scikit-Learn,[81] and final robust ICs were identified by clustering of resulting ICs using the Scikit-Learn implementation of the DBSCAN.[82] To find the optimal dimension for ICs,[87] we performed the clustering iteratively on the transcriptomic data with a step size of 10 for dimensions between 10 and 100 in the same manner as previously described.[23] Among these, we selected the optimal dimension of 90, where the number of non-single gene components and the number of final components (the number of non-single gene ICs found in the largest possible dimension) were the same in the corresponding dimension. Finally, we obtained the M matrix containing 45 robust ICs (i.e., iModulons) and the A matrix containing their activity for each condition in each sample by decomposing the transcriptomic data matrix (X). Each iModulon in the M matrix contains a weighting for each gene, and most gene weights in an iModulon are insignificant. To identify iModulon genes with significant weightings, we determined the optimal threshold for each iModulon. The genes with absolute weightings larger than the threshold are the sets of genes that make up iModulons. For identification iModulon genes with significant weighting, the optimal threshold for each iModulon is determined by computing the D'Agostino K^2 test of Scikit-Learn[81] as described in previous reports (https://github.com/avsastry/modulome-workflow).[29,30] For each iModulon from the M matrix, the optimal threshold value was determined by iteratively eliminating the gene with the highest absolute gene weight and calculating the K^2 value. When the remaining genes showed a normal distribution close enough to zero (D'agostino K^2 test statistic <500), the removed genes and gene

weight values were selected as the significant genes and the iModulon thresholds, respectively. For the Flagellar-1, TfoX, Ribosomal proteins, UC-7 iModulons, their thresholds were manually modified to the threshold values to include genes in the same operon or transcription unit as follows: Flagellar-1 (original value: 0.097, adjusted value: 0.069), TfoX (original value: 0.040, adjusted value: 0.038), Ribosomal proteins (original value: 0.066, adjusted value: 0.043), and UC-7 iModulons (original value: 0.061, adjusted value: 0.056). Total iModuon threshold values are listed in the M matrix (Table S2). As the M matrix (iModulon genes) and A matrix (iModulon activity) are independent components, manually changing the thresholds does not influence analyses related to iModulon activity.

Characterization of iModulon function

The Pymodulon package (https://github.com/SBRG/pymodulon) was used to characterize the 45 iModulons.[30] The function of 45 iModulons was characterized as described in previous reports.[29,30] First, the function of iModulons was characterized by enriching the transcriptional regulatory network (TRN). Since the transcriptional regulation of *Vn* has not been elucidated so far, we first collected primary transcription factor (TF)-gene interactions from the closely related *Vibrio* species *V. parahaemolyticus* RIMD 2210633 and *Vc* strains (E7946, N16961, C6706, and A1552) based on RegPrecise v3.2,[88] BioCyc Database Collection,[89] and previously published literature. Then, 931 TF-gene interactions were reconstructed by the comparative genomic approach with the corresponding strains. All defined TRN and their sources are listed in Table S2. Transcription regulators of each iModulon were inferred by performing Fisher's Exact Test with Benjamini-Hochberg correction (false discovery rate [FDR] $< 10^{-5}$). The iModulons that significantly overlapped with the TRN were named with TFs related to each TRN. Second, the iModulon functions were inferred by annotating the genes against the Kyoto Encyclopedia of Genes and Genomes (KEGG)[90] and Cluster of Orthologous Groups (COG). The information on KEGG and COG was obtained by using the EggNOG mapper.[83] The KEGG modules with Benjamini-Hochberg corrected FDR $<10^{-2}$ (Fisher's Exact test) were considered statistically significant. Uniprot IDs were obtained using the Uniprot ID mapper,[91] and operon information was obtained from Biocyc.[89] Gene ontology (GO) annotations were obtained from AmiGO2.[92] iModulons were further named with significantly enriched functional characteristics.

Differential activation analysis

Differences in iModulon activities were calculated as previously described.[24,29,30] Briefly, the differences in iModulon activity between biological replicates for each iModulon were fit to a log-norm distribution. For statistical analysis, the differences in the mean activity were calculated and compared against all iModulons to calculate a p-value. For the multiple hypothesis testing, the above p-values were further adjusted with Benjamini-Hochberg correction. The iModulons with change in activity level >10 and FDR <0.05 for a given pair of conditions were considered significant. Activity clustering of iModulons was conducted as previously described.[30]

Construction of deletion mutant strains

To efficiently construct 83 mutants from parental *Vn* WT strain, a gene-deletion method based on natural transformation was used.[10] For these, a total of 83 tDNAs were individually generated by subcloning and PCR amplification, as mentioned above. These tDNAs were delivered into the WT Vn strain via natural transformation, as described above. The transformed cells were selected on LBv2 agar plates supplemented with both Cm and Carb. Following natural transformation, colonies were identified for each gene deletion using genomic DNA extraction and deletion-confirming PCR. Genomic DNA was purified from the collected cells using Gentra Puregene Yeast/Bact. Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocols. Deletion of the target gene was confirmed by PCR using ''KO-strain-name''_Val_For/Rev. If the DNA size of PCR product between WT and deletion strain is similar, BbsI restriction enzyme was used to PCR product for determination of genotype of the colonies. Genotype and plasmid DNA were further validated by Sanger sequencing (Eton Bioscience). All subcloning plasmids and primers used to generate mutant constructs are listed in Key resources table and Table S3.

Motility assay

Briefly, 1 mL of cultures grown overnight for each strain were centrifuged and re-suspended in their supernatants, and their final OD_{600} was set to 4. The motility phenotypes were quantified by spotting 2 μL of these OD-adjusted cultures on soft LB agar (0.3%) plates. Before measurement, plates were incubated for 24 h at room temperature. The swarming diameter (in millimeters) is the average of three independent biological repeats.

H_2O_2 resistance assays

To test the effect of the deletion of arginine metabolism and AhpCF on hydrogen peroxide (H_2O_2) resistance, we carried out each natural transformation reaction with the treatment of various H_2O_2 concentrations (0, 0.1, 0.2, and 0.5 mM). After the transformation reaction in the competence buffer, the H_2O_2 was eliminated by centrifuging at 4,000 × g for 10 min. The reactions were recovered by transferring them to 1 mL of fresh LBv2 and incubating them for 2 h at 30°C with agitation at 180 rpm. The samples were then diluted, plated on LBv2 agar with chloramphenicol, and incubated for 10–12 h at 30°C. Cell survival was determined by counting CFUs and was normalized based on not treated samples (0 mM). Survival (%) is shown as the average of three independent biological repeats.

Fluorescence dye-based reactive oxygen species (ROS) detection

For fluorescence-based ROS detection, H_2DCFDA was used with some modifications.[64] CM-H_2DCFDA (20 mM) was purchased from Abcam (ab113851) and diluted in dimethyl sulfoxide (DMSO) to make a 1 mM stock solution (1000 ×). To obtain sufficient cell density for determining ROS level under the competence condition, we carried out each transformation reaction by scaling up 30 times (10.5 mL/each reaction) more than the general method (0.35 mL) prepared above. At each time point, cells were washed with CB and resuspended in 1 mL of competence buffer. The cell suspension samples were then mixed with 1 μL of CM-H_2DCFDA stock solution (1 mM), resulting in a final concentration of 1 μM H_2DCFDA, and incubated for 30 min at 30°C. Subsequently, 100 μL of the resulting samples were transferred to a black 96-well microplate. The fluorescence was measured using the Infinite 200 Pro Plate Reader (Tecan, Männedorf, Switzerland) with 485/535 nm excitation/emission wavelengths, and the fluorescence signals were normalized with OD_{600} value.

QUANTIFICATION AND STATISTICAL ANALYSIS

In addition to the transcriptome and ICA analysis, additional data statistical testing (Pearson's correlation coefficient, two-tailed Student's t-test, and two-tailed Wilcoxon-Mann-Whitney test) was performed in GraphPad Prism v8 software (GraphPad, San Diego, CA, USA). p-values <0.05 were considered statistically significant.