Experimental evolution reveals nitrate tolerance mechanisms in *Desulfovibrio vulgaris*

Bo Wu[1, 2], Feifei Liu[2, 3], Aifen Zhou[2], Juan Li[4], Longfei Shu[1], Megan L. Kempher[2], Xueqin Yang[1], Daliang Ning[2], Feiyan Pan[2], Grant M. Zane[5], Judy D. Wall[5], Joy D. Van Nostrand[2], Philippe Juneau[6], Shouwen Chen[7, 8*], Qingyun Yan[1, 2*], Jizhong Zhou[2, 9, 10*], and Zhili He[1, 2, 4*]

[1]Environmental Microbiomics Research Center, School of Environmental Science and Engineering, Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Sun Yat-sen University, Guangzhou 510006, China; [2]Institute for Environmental Genomics and Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK 73019, USA; [3]Guangdong Provincial Key Lab of Microbial Culture Collection and Application, Guangdong Institute of Microbiology and State Key Laboratory of Applied Microbiology Southern China, Guangzhou 510070, China; [4]College of Agronomy, Hunan Agricultural University, Changsha 410128, China; [5]Departments of Biochemistry and Molecular Microbiology & Immunology, University of Missouri-Columbia, Columbia, MO 65211, USA; [6]Département des Sciences Biologiques, TOXEN, Ecotoxicology of Aquatic Microorganisms Laboratory, Université du Québec à Montréal, Montréal, Canada;[7]State Key Laboratory of Agricultural Microbiology, Huazhong Agricultural University, Wuhan 430070, China; [8]Hubei Collaborative Innovation Center for Green Transformation of Bio-Resources, College of Life Sciences, Hubei University, Wuhan 430062, China; [9]Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA; [10]School of Environment, Tsinghua University, Beijing 100084, China

[*]**Corresponding author**: Zhili He (132 East Circle, University Town, Guangzhou 510006, China; email: hezhili@mail.sysu.edu.cn; phone: +86-020-3106-5837); Qingyun Yan (132 East Circle, University Town, Guangzhou 510006, China; email: yanqy6@mail.sysu.edu.cn; Shouwen Chen (No. 368, Youyi Road, Wuhan 430062, China; email: chenshouwen@mail.hzau.edu.cn); or Jizhong Zhou (101 David L. Boren Blvd., Norman, OK 73019; email: jzhou@ou.edu; phone: 405-325-6073).

Running title: Nitrate tolerance mechanisms in *D. vulgaris*

Conflict of interest: The authors declare no conflict of interest

Abstract

Elevated nitrate in the environment inhibits sulfate reduction by important microorganisms of sulfate-reducing bacteria (SRB). Several SRB may respire nitrate to survive under elevated nitrate, but how SRB that lack nitrate reductase survive to elevated nitrate remains elusive. To understand nitrate adaptation mechanisms, we evolved 12 populations of a model SRB (i.e., *Desulfovibrio vulgaris* Hildenborough, DvH) under elevated $NaNO_3$ for 1000 generations, analyzed growth and acquired mutations, and linked their genotypes with phenotypes. Nitrate-evolved (EN) populations significantly ($p < 0.05$) increased nitrate tolerance, and whole genome resequencing identified 119 new mutations in 44 genes of 12 EN populations, among which six functional gene groups were discovered with high mutation frequencies at the population level. We observed a high frequency of nonsense or frameshift mutations in nitrosative stress response genes (NSR: DVU2543, DVU2547 and DVU2548), nitrogen regulatory protein C family genes (NRC: DVU2394-2396, DVU2402 and DVU2405), and nitrate cluster (DVU0246-0249 and DVU0251). Mutagenesis analysis confirmed that loss-of-functions of NRC and NSR increased nitrate tolerance. Also, functional gene groups involved in fatty acid synthesis, iron regulation, and two-component system (*LytR*/*LytS*) known to be responsive to multiple stresses, had a high frequency of missense mutations. Mutations in those gene groups could increase nitrate tolerance through regulating energy metabolism, barring entry of nitrate into cells, altering cell membrane characteristics or conferring growth advantages at the stationary phase. This study advances our understanding of nitrate tolerance mechanisms and has important implications for linking genotypes with phenotypes in DvH.

Introduction

Sulfate-reducing bacteria (SRB) carry out dissimilatory sulfate reduction, and play an important role in the biogeochemical cycling of sulfur, carbon, nitrogen, and metals [1, 2]. They are ubiquitous in anaerobic environments such as marine and lacustrine sediments [3], where nitrate and sulfate are commonly found [4-6]. Nitrate concentrations in natural environments are generally less than 2 mM [7, 8], while in contaminated sites, nitrate concentrations can be greater than 100 mM [9]. Elevated nitrate could inhibit sulfate reduction performed by SRB due to its competition for organic electron donors [6], which has a special importance in contaminated or engineered systems since it mitigates sulfide production in the environment [10-12]. To confer an advantage under elevated nitrate, several SRB can use nitrate as an alternative electron acceptor [13]. However, the model sulfate-reducing bacterium, *Desulfovibrio vulgaris* Hildenborough (DvH) lacks a functional nitrate reductase [14, 15], and how it responds to elevated nitrate and thrives in such environments for a long time exposure remains elusive [16]. Therefore, it is important to understand nitrate tolerance mechanisms, and we used DvH to advance our knowledge about SRB responses and functions in the environment in this study.

Previous studies showed complex responses of DvH to elevated nitrate by different approaches, such as iTRAQ peptide tags [17], microarrays [4], transposon liquid enrichment sequencing (TnLE-seq) [18], and mutagenesis [19]. For example, transcriptomic analysis of DvH showed that elevated nitrate up-regulated key functional genes involved in osmotic, ionic, and nitrite stress responses, and few common patterns of gene expression were observed in response to elevated nitrate, nitrite, or salt [4]. Further analyses indicated that a specific nitrate cluster (DVU0245-DVU0251) was related to increased nitrate tolerance, but not for nitrite tolerance in DvH [18], and elevated nitrate and nitrite had distinct effects on inhibition of DvH by mutagenesis

[19]. Although these studies have provided new insights into our understanding of the response of DvH to short-term exposure to elevated nitrate at the protein, transcriptional and gene levels, more studies are needed to reveal the genetic basis of nitrate tolerance when DvH is exposed to elevated nitrate that mimics an environmental stress, especially for a long time.

Experimental evolution of asexual microbial populations coupled with whole-genome resequencing provides a powerful approach to identify adaptive mutations and elucidate genotype-phenotype relationships [20, 21]. Beneficial mutations typically emerge after an adaptive evolution of microorganisms [20, 22], which has been used to discover adaptive changes under specific conditions for prokaryotes such as *Escherichia coli*, *Pseudomonas aeruginosa*, and *Desulfovibrio vulgaris* [22-26], phages ϕ6 and T4 [27, 28], and eukaryotes such as *Saccharomyces cerevisiae* and *Chlamydomonas reinhardtii* [29, 30]. For example, a previous study indicated a rapid selective sweep of mutations occurred in *D. vulgaris* populations under NaCl stress within the first 100 generations, and four mutations contributed to shortened lag phases, increased growth rates and biomass yield [23]. Also, identical or similar genetic changes (e.g., mutations occurred in same/similar genes with similar phenotypic changes) in replicate populations are considered as a strong indicator of beneficial mutations resulting from stress adaptation in the experimental evolution [31-34]. Parallel evolutionary losses of ribose catabolic function occurred in 12 *E. coli* populations during 2000 generations in a glucose minimal medium to gain beneficial fitness in glucose-limited environments [35], and such genetic parallelism at the population level was also observed in the gene Dde_2265 encoding sulfate adenylythransferase (*sat*) in *Desulfovibrio alaskensis* evolved under perchlorate stress for about 117 generations, leading to increased perchlorate tolerance [36]. Therefore, it is expected that experimental evolution and resequencing

of whole genomes provide powerful approaches to link genotypes to phenotypes, and explore nitrate tolerance mechanisms in DvH.

In this study, we aimed to understand nitrate tolerance mechanisms in DvH through experimental evolution, which could provide a comprehensive understanding of genetic basis of the adaptation evolution in DvH. We hypothesized that a set of beneficial genetic changes would emerge in DvH populations after exposure to elevated nitrate ($NaNO_3$) for 1000 generations [22, 23], and that some of these mutations would be parallel among evolved populations [20, 32, 34, 36]. We evolved 12 populations of DvH under elevated nitrate for 1000 generations, compared genotypic and phenotypic characteristics of ancestral (AN) populations and EN populations, linked their genotypes with phenotypes by mutagenesis, and explored its nitrate tolerance mechanisms. We found that six functional groups with a high frequency of mutations, which were directly or indirectly involved in nitrate tolerance. Most importantly, we showed new evidence that nonsense and frameshift mutations in key nitrate associated genes conferred nitrate tolerance and genetic parallelism in nitrate-evolved populations. This study advances our understanding of nitrate tolerance mechanisms in DvH.

Materials and methods

Strains, cultivation conditions and propagation

We used 12 AN populations (AN1-12) of *D. vulgaris* Hildenborough (ATCC 29579) from a previous study [37], and a clone of DvH was the founder of 12 populations. The first subculture was non-stressed for all 12 populations, and they were frozen to serve as 12 AN populations, from which we randomly chose three AN populations (AN2, AN8 and AN11) to determine a suitable $NaNO_3$ concentration for experimental evolution, which allowed AN populations to reach a

stationary phase within 48 h (Fig. S1). AN populations were grown at 37°C and propagated in the defined LS4D medium [38] amended with 10 mM $NaNO_3$ for 500 generations, then with 30 mM $NaNO_3$ for another 500 generations to keep ongoing nitrate selective pressure for DvH populations. Cultures were propagated by transferring 1% (100 μL) of the final volume (10 mL) into a fresh medium every 48 h, and populations were archived every 100 generations at -80ºC for later studies. Every 48 hours, the OD_{600} was recorded for each population before transfer (Fig. S2), and the statistical significance of biomass differences (OD_{600}) of 6.7 generations (after the first transfer) vs. 1000 generations (after the last transfer) was based on ANOVA tests.

Growth phenotypes of ancestral and evolved populations

The growth of both AN and nitrate-evolved (EN) populations in the LS4D medium with 30 mM or 100 mM $NaNO_3$ or without additional $NaNO_3$ was measured by a spectrophotometer (Thermo Spectronic 20D+, Waltham, Massachusetts, USA) with three replicates for each population, and OD_{600} values of all populations were recorded every 2-3 h over a period of 70 h. Lag phase was defined as the time after-inoculation to $OD_{600} = 0.2$; growth rate was 2.303 × the slope of the linear portion of growth curve by plotting Ln (OD_{600}) as the y-axis and time as the x-axis; biomass yield was the maximum OD_{600} as previously described [37]. Statistical significance of multiple pairwise-comparison was based on the ANOVA test. In addition, extracellular nitrate concentration of AN and EN populations was measured by ion chromatography (ICS-90A, Dionex, USA).

Genomic DNA extraction and whole genome resequencing of AN and EN populations

Genomic DNA (gDNA) of all AN and EN populations was extracted with a GenElute Bacterial Genomic DNA kit (Sigma) following the manufacturer's instructions. The quality of gDNA was assessed with a Nanodrop ND-1000 (Thermo Fisher Scientific, Inc., Wilmington, DE, USA).

Illumina HiSeq sequencing of gDNA samples was conducted at the Oklahoma Medical Research Foundation (Oklahoma City, OK, USA) following standard library preparation protocols and Illumine sequencing protocols. The sequencing data have been deposited into NCBI Sequence Read Archives and are accessible through Bioproject PRJNA630554.

Mutation calling

Mutations in all AN and EN populations were called by our in-house pipeline (http:// http://zhoulab5.rccc.ou.edu:8080/) with the DvH genome sequence (NC_002937.3) as the reference. Btrim was used for quality control with filter parameter settings, including windows size as 5, average quality score as 30 and minimal insert size as 50 [39]. The alignment was generated by Bowtie2, and the default settings of bowtiw2-build and bowtie2-x were used for database building and alignments [40]. Samtools was used to call mutations including single nucleotide polymorphisms (SNPs) and insertions/deletions (INDELs); Samtools-mplieup was used to filter variants, probabilistic realignment for computation of base alignment quality was disabled to reduce false SNPs caused by misalignments, minimum mapping quality for alignment was 0, and minimum base quality for a base was 13 [41]. The output variant call format (VCF) file was generated by Bcftools [42]. Phred scale quality score was used as the filter score. Breseq was also used to call mutations with the default setting [43], and new junctions/insertions in output/index.html and output/marginal.html (allele frequency > 10%) were added into the mutation data analysis. Compared to the original NCBI reference DvH strain, the ancestral DvH strain used in this study already has 22 genomic differences [23], which were removed from our mutation analysis. The allele frequency of mutations was calculated based on DP4 values: forward reference alleles, reverse reference alleles, forward alter alleles, and reverse alter alleles by Samtools [41].

Mutation validation

To validate mutations called in this study, a subset of regions bearing mutations with different filter values were chosen (Table S1). Genomic DNA from AN or EN populations was extracted and amplified by PCR, and PCR amplicons were sequenced by a Sanger platform at the Oklahoma Medical Research Foundation to verify those mutations identified by the Illumina sequencing platform. All the primers used in this study are listed in Table S1.

Mutation data analysis and prediction of mutated protein structure

Mutations were assigned to a gene if they occurred within a coding region or within a 100-bp upstream of open reading frames [44]. We combined output VCF files of whole genome sequence data from Bcftools [42] to generate SNPs/INDELs tables for AN or EN populations (Table S2 and Table S3). Briefly, mutation data analysis included five major steps: (i) all mutation information was extracted from output VCF files of AN and EN populations, including gene name, position for the mutation, mutated bases and reference bases; (ii) proteins encoded by these genes and Clusters of Orthologous Groups (COGs) of the proteins [45] were added into the table manually; (iii) as mutations or polymorphisms were detected in the ancestor [23], we only analyzed newly acquired mutations in the EN population (defined as newly acquired mutations), including new junctions occurred in coding genes that predicted by breseq (Table S4); (iv) to focus on mutated functional genes with newly acquired mutations in EN populations, a gene identified with newly acquired mutations was labeled with the highest allele frequency of the mutation occurring in the gene for each EN population (Table S5); and (v) gene co-expression prediction was based on information in operon and regulon prediction in Microbesonline (http://www.microbesonline.org/), and experimental evidence was based on publications about mutated genes, especially genes that were co-mentioned in same publication identified by String [46].

185 Types of detected mutations in this study include: (i) nonsense mutation: one-base change in
186 a DNA sequence that results in a premature stop codon; (ii) frameshift mutation: an insertion or
187 deletion in a DNA sequence that often results in a non-functional protein (e.g., truncated proteins)
188 [47]; (iii) missense mutation: one-base change in a DNA sequence that results in a different amino
189 acid in the mature protein; and (iv) silent mutation: one-base substitution in a DNA sequence that
190 does not involve amino acid replacement in the coded protein.

191 We also used Phyre2 [48] to predict changes in the tertiary structure of HcpR encoded by
192 DVU2547, one of the most frequently mutated genes in EN populations with the highest hits (>99%
193 confidence) for predicted models (Fig. S3).

194 **Mutagenesis and growth phenotypes of deletion mutants**

195 DvH parental strain JZ001 (Δ*upp*) is the founder strain of AN populations, and was used as the
196 parental strain to generate marker exchange (ME) mutants of genes by a marker replacement
197 approach [49]. The growth of parental strain and ME mutants of NSR was tested in LS4D, LS4D
198 + 30 mM $NaNO_3$ and LS4D + 0.15 mM $NaNO_2$. The growth phenotypes of the parental strain and
199 ME mutants of NRC genes were tested in LS4D, LS4D + 30 mM $NaNO_3$, and LS4D + 100 mM
200 NaCl. Genetic complementation was done for all the mutants (ME2394, ME2395, ME2396,
201 ME2405, ME2543, ME2547 and ME2548) and EN7 variant complemented with the plasmid
202 containing a native promoter, ribosome binding site and wild type gene (Table 1). Further reverse-
203 transcription PCR was done for the mutants and their complemented strains at the exponential
204 phase to confirm the expression of complemented genes. The growth of complemented strains was
205 tested in LS4D and LS4D + 30 mM $NaNO_3$. Statistical significance of growth parameters was
206 based on the T-test. All DvH strains used in this study are listed in Table 1.

Results

Nitrate tolerance increased in nitrate-evolved populations

To choose a suitable nitrate concentration for our experimental evolution, three ancestral (AN) populations (AN2, AN8 and AN11) were randomly selected and grown under different nitrate concentrations. Their lag phase increased while growth rates and biomass gradually decreased as nitrate concentrations increased from 0 mM to 30 mM (Fig. S1). We evolved 12 populations under elevated nitrate (10 or 30 mM $NaNO_3$), which is higher than nitrate concentrations in natural environments for 1000 generations, and all 12 nitrate-evolved (EN) populations reached a higher biomass compared to AN populations (Fig. S2).

EN populations had a shorter lag phase (Fig.1a) and a faster growth rate (Fig.1b) compared to AN populations under normal growth conditions and showed significantly ($p < 0.05$) higher nitrate tolerance compared to AN populations (with 30 mM $NaNO_3$) (Fig. 1). With 100 mM additional $NaNO_3$, the growth of EN populations was repressed but they could still grow relatively well while AN populations barely grew within 70 h (Fig. 1). The nitrate concentration (30 mM) did not change at the end of the cultivation period for AN or EN populations, indicating that a change in nitrate concentration was not detected after DvH was evolved for 1000 generations under elevated nitrate.

Genes and functional gene groups with new mutations in EN populations

Sequencing reads of AN and EN populations covered about 99% of the DvH reference genome sequence (NC_002937.3) with an average sequencing depth of 88 × for AN populations, and 92 × for EN populations. A total of 58 and 167 mutations were respectively detected in AN and EN populations, and the number of mutations ranged from 8 to 26 for each AN population (Table S2),

while the number of mutations ranged from 13 to 33 for each EN population (Table S3). We focused on newly acquired mutations and associated genes in the EN populations, including 119 newly acquired mutations (containing nine large deletions and one insertion predicted using breseq) in 44 genes associated with energy, amino acid, carbohydrate, nucleotide, lipid and inorganic ion metabolism, cellular process and signal, and information storage and processing (Table S4).

Among those 44 genes with newly acquired mutations, we focused on genes with high mutation frequency at the population level (>33%, refs to genes bearing mutations in more than four of 12 EN populations), and six functional gene groups of 19 genes were found based on gene co-expression prediction or experimental evidence from previous studies (Table S4 and Table S5): (a) nitrosative stress response genes (NSR: DVU2543, DVU2547 and DVU2548): DVU2547 (hybrid cluster protein regulator, *hcpR*) harboring nine mutations in seven of 12 EN populations, regulates the expression of DVU2543 (*hcp*, identical mutation in two EN populations) [14, 50] and exists in the same operon with DVU2548 (*acpD*, five mutations in three EN populations) [51]; (b) nitrogen regulatory protein C family genes (NRC: DVU2394-2396, DVU2402 and DVU2405): a two component system DVU2394 (*ntrC*)/DVU2395 (response regulator/histidine kinase) in the operon DVU2394-DVU2396 (with 18 mutations in 11 EN populations) regulates the operon DVU2397-2405 [52], in which DVU2402 mutated in one EN population and DVU2405 harboring three mutations in two EN populations; (c) nitrate cluster (DVU0246-0249 and DVU0251) [18] with 17 mutations in all 12 EN populations; (d) fatty acid synthesis genes (DVU1204 and DVU1208): DVU1204 (3-oxoacyl-(acyl-carrier-protein (ACP) synthase II, *fabF*) harboring six mutations in seven EN populations and DVU1208 (phosphate acyltransferase, *plsX*) harboring an identical mutation in two EN populations are predicted in one operon involved in fatty acid synthesis [23, 51]; (e) iron regulatory genes (DVU0942 and DVU2571): DVU0942 (*fur*) harboring

four mutations in six EN populations is a regulator of ferrous iron transport protein B gene DVU2571 (*feoB*) [53], which mutated in the other six EN populations; and (f) a two component system DVU0596/DVU0597 (*lytR*/*lytS*) harboring five mutations in six EN populations regulates the expression of carbon starvation protein genes DVU0598 (*cstA*) and DVU0599 (*cstB*) [52] (Table 2).

A large proportion (84/119) of mutations occurred in those 19 genes of six functional groups in EN populations (Table 2 and Table S4). For example, seven EN populations had mutations in DVU2547, including an identical mutation occurred in EN7 and EN8; nine EN populations had mutations in the two component system DVU2394/DVU2395; all 12 EN populations harbored mutations in nitrate cluster. For the other 27 mutated genes with 35 newly acquired mutations, we found seven genes with 17 new mutations acquired in at least two EN populations (Table S6), and 18 genes acquired new mutations in only one EN population (Table S5). In the following analysis, we focused on those highly mutated six functional groups (with mutation frequency $\geqslant$ 50% at the populations level) to link their genotypes with phenotypes towards understanding nitrate tolerance mechanisms in DvH.

Six functional gene groups were highly mutated in EN populations

We further analyzed potential effects of mutations on functions of their associated genes/proteins in the six functional groups. First, nonsense and frameshift mutations were identified in the first three functional gene groups, presumably resulting in premature or truncated proteins. (i) Nonsense and frameshift mutations were identified in the NSR group of three genes with allele frequencies ranged from 17 – 100% (an average of 72%). Seven of 17 mutations in this group were fixed or nearly fixed (more than 90%), and interestingly, five of them were nonsense and frameshift mutations. Also, the same mutation was identified in this group, such as a nonsense mutation in

coordinate 2654808 in DVU2543 (Hybrid cluster protein, *hcp*) in EN4 and EN5, and in position 2660403 of DVU2547 (Hcp regulator, *hcpR*) in EN7 and EN8 (Table 3). (ii) Among 22 mutations that occurred in the NRC group, 11 nonsense and frameshift mutations were identified in this group with four of them fixed in all EN populations, and the allele frequency of mutations ranged from 11–100% with an average of 58% (Table 4). (iii) Nonsense and frameshift mutations were also observed in three genes (DVU0246, DVU0247 and DVU0248) of nitrate cluster with a mutation allele frequency ranged from 13–77% (an average of 36%). All the mutations occurred in DVU0246 (Pyruvate phosphate dikinase, *ppdk*) and DVU0247 (Response regulator, *cheY* like/RR) could result in truncated proteins. Identical frameshift mutation was identified at the coordinate 279997 of DVU0246 in EN1 and EN8, and the same missense mutation occurred at the coordinate 285948 of DVU0251 (Transmembrane protein, *tauE* like) in EN 7 and EN 12 (Table 5). Interestingly, eight of 10 mutations detected using breseq were located in these three functional groups (NSR, NRC and nitrate cluster), which were nonsense and frameshift mutations (Table S4).

Second, missense mutations occurred in the other three gene groups, potentially resulting in functional changes in their encoding proteins. Mutations in fatty acid synthesis and iron regulatory genes had higher allele frequencies (43-100% and 88-100% with an average of 92% and 99%, respectively) compared to other gene groups (Table S4). For fatty acid synthesis genes, mutations in DVU1204 (*fabF*) were observed in seven out of 12 EN populations, which were fixed or nearly fixed, and the same SNP was detected in DVU1204 in EN4 and EN5, while a same silent mutation was found in DVU1208 (*plsX*) in EN8 and EN10 (Table S7). Also, all EN populations harbored mutations in iron regulatory genes, and the same mutation occurred in DVU0942 (*fur*) in EN4, EN5 and EN9 (Table S8). Additionally, six mutations were detected in *lytR/lytS* with the lowest allele frequency (11 – 40% with an average of 19%) among those six gene groups, but all of the

298 mutations were missense except an intergenic mutation at the upstream of DVU0597 start codon
299 in EN3 (Table S9).

300 The results showed that nonsense, frameshift and missense mutations occurred in those six
301 gene groups, and they may result in prematurely terminated proteins, truncated proteins, or
302 function-changed proteins, thus affecting functions of DvH in response to elevated nitrate.
303 Furthermore, as those mutations appeared to parallelly occur in EN populations, it is expected that
304 such observed mutations may be related to increased nitrate tolerance in DvH.

305 **Deletion mutants of NSR and NRC genes resulted in improved nitrate tolerance**

306 To determine if nonsense and frameshift mutations are involved in nitrate tolerance in EN
307 populations and to understand possible nitrate tolerance mechanisms in DvH, we generated
308 representative deletion mutants of NRC and NSR genes, and analyzed their growth under elevated
309 nitrate, NaCl, or nitrite. The NSR group includes three genes: *hcp* (hybrid cluster protein,
310 DVU2543), *hcpR* (hcp regulator, DVU2547) and *acpD* (DVU2548) (Table 2). DVU2547 encoding
311 HcpR was one of the most frequently mutated genes in seven out of 12 EN populations (Table 3).
312 Also, these mutations were identified in key domains of *hcpR* (DVU2547), including DNA-
313 binding domain (EN1-1, EN1-2), cAMP-binding domain (EN1-2), and dimer interface (EN3-1,
314 EN3-2, EN6, EN9) (Fig. S3), and the same nonsense SNP was fixed in EN7 and nearly fixed in
315 EN8. Analysis of those mutants indicated that the growth of parental strain JZ001 and NSR gene
316 deletion mutants was quite similar without nitrate or nitrite addition (Fig. 2a), but these mutants
317 grew faster than JZ001 with 0.15 mM nitrite added although a lower final biomass was obtained
318 for the mutants (Fig. 2b). Especially, when 30 mM was added, all three mutants had shorter lag
319 phases, higher growth rates, and higher biomass yield than JZ001 (p <0.05), indicating increased
320 nitrate tolerance in those mutants (Fig. 2c). In addition, a similar growth was observed for JZ001P

(JZ001 with empty vector pMO719) and the complemented strains (ME2543 and ME2547) without nitrate (data not shown), while decreased nitrate tolerance was observed in the complemented strains (Fig. 2d) and EN7 variant (Fig. S4) under elevated nitrate, possibly due to the overexpression of complemented genes. It is noted that we could not obtain a complemented strain for ME2548 after several attempts. The results demonstrated that nonsense and frameshift mutations in the NSR group (Table 3) could improve nitrate tolerance in EN populations.

NRC group contains a two component system DVU2394 (Response Regulator, RR)/DVU2395 (Histidine Kinase, HK), which regulates the expression of operon DVU2397-2405 encoding alcohol dehydrogenase and heterodisulfide reductase (Table 2). Marker exchange (ME) mutants of DVU2394, DVU2395, DVU2396 and DVU2405 in the NRC group were generated. While no significant growth differences were observed between ME mutants and JZ001 in LS4D (Fig. 3a) or LS4D with 100 mM NaCl (Fig. 3b), all NRC mutants had higher growth rates (p <0.05) in LS4D amended with 30 mM nitrate except ME2405, which had the shortest lag phase compared to both JZ001 and other ME mutants (Fig. 3c). As expected, the parental strain (with empty plasmid pMO719)-like phenotypes were observed for all complemented strains of NRC mutants (Fig. 3d). The results confirmed that nonsense and frameshift mutations in the NRC gene group (Table 4) likely led to increased nitrate tolerance in EN populations.

Discussion

Experimental evolution coupled with whole-genome resequencing enable us to link genotype with phenotype, understand nitrate tolerance mechanisms, and facilitate the advancement of our knowledge about sulfate reducing bacteria (SRB) responses to elevated nitrate in the environment. In this study, we found that most newly acquired mutations occurred in six gene groups of DvH

through experimental evolution under elevated nitrate, and deletion mutants of NSR and NRC genes confirmed their roles in nitrate tolerance, indicating that those nonsense and frameshift mutations were beneficial. Also, identical or similar mutations were observed in EN populations evolved independently, suggesting genetic parallelism of adaptive selection at the population level.

Beneficial mutations in stress-evolved populations are considered as drivers of adaptation [21, 32]. Previous studies of adaptive evolution in *Desulfovibrio* species showed beneficial mutations emerged along the experimental evolution and conferred a fitness advantage under stress conditions [23, 36]. In this study, we found a set of high-frequency mutations in six gene groups of EN populations, and they conferred nitrate tolerance in DvH, appearing to be beneficial. However, native functions of those gene groups are not likely involved with nitrate metabolism but largely by gene regulations, such as three two-component systems (DVU2394/DVU2395, DVU0596/DVU0597 and DVU0246/DVU0247) and two global transcriptional regulators/regulons (DVU2547/DVU2543 and DVU0942/DVU2571) [14, 18, 52-54]. Those regulatory systems may perceive nitrate stress and control downstream responses, and based on functions of mutated genes and the fact that no nitrate was consumed, our results suggest that adaptive mechanisms of DvH under elevated nitrate may include (i) shifting of energy metabolism, (ii) barring entry of nitrate into the cell, (iii) altering cell membrane characteristics, and (iv) conferring growth advantages at the stationary phase.

First, a shift of energy metabolism could provide physiological advantages for DvH to survive nitrate stress [4], thus increasing nitrate tolerance after experimental evolution. A previous study indicated that a deletion mutant of DVU0916 (*rex*), a repressor of sulfate adenylyl transferase increased nitrate tolerance [18]. We found that a large number of mutations were identified in NSR and NRC genes involved in energy metabolism, and functional loss of those genes conferred nitrate

tolerance. DVU2547 (*hcpR*) of NSR was proposed to be involved in energy metabolism through repressing sulfate reduction pathways and activating expression of DVU2543 (*hcp*), which responds to reactive nitrogen species and was up-regulated under elevated nitrite or nitrate [4, 55-57]. A possible explanation for growth advantages of an *hcpR* mutation was suggested to be derepression of sulfate-reducing genes resulting from the loss of the repressor HcpR (Fig. 4a). Also, we found NRC genes conferred nitrate tolerance, which has not been reported before. In this gene group, DVU2405 (alcohol dehydrogenase, *adh*) regulated by a two-component system (DVU2394/DVU2395) was shown to be one of the most highly expressed genes contributing to energy metabolism with different electron donors in DvH [58]. Thus functional loss of genes in the NRC group may lead to a shift to methyl/S-adenosyl-methionine metabolism due to a lack of alcohol oxidation pathways catalyzed by alcohol dehydrogenase encoded by DVU2405 [4, 58] (Fig. 4b). Our results suggested that those frameshift and nonsense mutations in the NSR and NRC gene groups could relieve nitrate stress through shifting energy metabolisms, and those new findings advance our understanding the role of NSR and NRC groups in nitrate tolerance in DvH.

Second, blocking nitrate entry into the cell could confer nitrate tolerance in DvH. The nitrate cluster genes were proposed to allow non-specific nitrate transport, while transposon mutants of nitrate cluster genes were found to have a growth advantage under nitrate stress [18]. Also, DVU0249 encodes an outer membrane-associated homodimer [59], and DVU0251 is a TauE like transmembrane protein [18]. Therefore, we speculate that those nitrate cluster genes may be involved in nitrate uptake or transport, and the functional loss of those genes may prevent nitrate entry into the DvH cell [18], which is consistent with our study, showing that nonsense and frameshift mutations occurred in nitrate cluster (Fig. 4c).

Third, altering membrane lipid composition is crucial for bacterial survival and adaptation under environmental stresses [60]. Membrane fatty acid synthesis was shown to be involved in heat stress, salt stress, and oxidative stress adaptation in *E.coli* [61]. Previous studies also indicated that the percentage of unsaturated fatty acid in cell membrane was affected under salt stress, while the formation of acyl-accepter was catalyzed by DVU1204 (*fabF*), a possible essential gene in DvH [23, 62], and it was up-regulated under heat stress [63]. We found that the gene *fabF* was high-frequently mutated in EN populations, which could increase unsaturated fatty acid percentages in cell membrane, thus increasing nitrate tolerance possibly by relieving osmotic stress (Fig. 4d).

Fourth, growth advantages at the stationary phase (GASP) could be beneficial for bacteria to adapt to different stresses due to competitive ability conferred by GASP mutations (e.g., in the gene *rpoS*) emerged during a long-term stationary phase [64]. A previous study with *E. coli* indicated the transcriptional regulator RpoS exhibited GASP mutations and responded to multiple stresses [65]. However, DvH does not have an annotated *rpoS* gene. Alternatively, iron homeostasis maintained by DVU2571 (*feoB*) and up-regulation of DVU0598 and DVU0599 encoding carbon starvation proteins regulated by *lytR*/*lytS*, which suggested being involved in stresses response as DvH cells transitioned into the stationary phase [66]. Furthermore, DVU2571 and *lytR*/*lytS* were also mutated in serial transfers with or without salt stress [23]. Such an experimental evolution may emerge mutations in those genes up-regulated at the stationary phase, as evolved DvH populations were kept at the stationary phase for ~ 24 h in every transfer (48 h). In this case, it is reasonable to assume that newly acquired mutations in those GASP-related genes may increase nitrate tolerance through relieving possible environmental stresses like nutrient depletion (Fig. 4e and 4f).

411 Notably, some beneficial mutations occurred in evolved populations may just increase the
412 overall growth (e.g., shortened lag phase, increased growth rate, increased biomass) without direct
413 effects for a specific stress during an experimental evolution [67]. For example, our previous study
414 showed that a mutation in DVU0597 was related to increased growth and biomass as well as
415 shortened lag phase while another mutation in DVU2472 was related to increased biomass and
416 shortened lag phase in salt-evolved populations (ES) of DvH [23]. In this study, we found
417 shortened lag phase and increased growth rate for EN populations with or without nitrate stress,
418 which is similar to a previous study, showing that both control-evolved populations (EC) and ES
419 had a shorter lag phase with or without salt stress compared to their ancestral populations (without
420 evolution) in DvH [37]. Also, a few shared mutated genes (DVU1204, DVU2571, and
421 DVU0596/DVU0597) were found in those populations (EN, EC and ES), which may confer an
422 increase of general stress tolerance [23, 60, 62-64, 66]. Furthermore, the fact that high
423 concentrations of iron could be regulated by DVU2571 during the lag phase emphasized the role
424 of iron accumulation in the early stage of bacterial growth [68]. Therefore, our results highlighted
425 potential roles of general stress response genes, especially GASP associated genes (DVU0597 and
426 DVU2571) in promoting the overall growth in EN populations.
427 In addition, parallel genetic changes were observed at the mutation, gene, and gene group
428 levels in independent populations. At the mutation level (identical mutations among EN
429 populations), identical mutations were observed in six gene groups and other less defined genes,
430 and they were nearly fixed or fixed. For example, identical nonsense and frameshift mutations
431 occurred in the NSR gene or the nitrate cluster of independent EN populations, resulting in
432 truncated proteins and increased nitrate resistance. Identical fixed or nearly fixed missense
433 mutations were also observed at the coordinate 1034711 in DVU0924 in EN4, EN5 and EN9

(Tables S3 and S8). At the gene level (different mutations in the same gene across EN populations), nonsense and frameshift mutations were observed for genes in the NSR, NRC, and nitrate cluster, which may result in nonfunctional proteins and similar phenotypic changes that conferred nitrate tolerance. Moreover, parallel genetic changes were also reflected at the gene group level (genetic changes occurred in functional gene groups of independent populations), leading to similar phenotypic shifts, especially in NSR, NRC and nitrate cluster. For example in the NRC group, loss-of-function mutations were identified in DVU2394 (RR) in EN5 and EN9, in DVU2395 (HK) in EN5, EN9, EN10 and EN12, in DVU2396 in EN1, EN2 and EN6, indicating that these genetic changes were similar and contributed to nitrate tolerance in independent populations. Notably, in this study, parallel evolution was more likely to happen via loss-of-function mutations. Therefore, such genetic parallelism involved in SNP/INDELs coupled with their high allele frequencies under elevated nitrate may be largely beneficial mutations in EN populations.

In summary, this study reveals nitrate tolerance mechanisms largely associated with six gene groups by experimental evolution in DvH. When DvH was exposed to elevated nitrate for 1000 generations, a number of beneficial mutations occurred. Nonsense and frameshift mutations were identified in NSR, NRC and nitrate clusters, resulting in increased nitrate tolerance through regulating energy metabolism and barring entry of nitrate into DvH cells, while missense mutations in fatty acid synthesis genes, iron regulatory genes and *lytR*/*lytS* may confer general stress tolerance by experimental evolution under elevated nitrate through altering cell membrane characteristics and conferring growth advantages at the stationary phase. Also, genetic parallelism is reflected at the mutation, gene and gene group levels, allowing us to further understand DvH responses to elevated nitrate. This study unravels genetic basis of evolutionary changes in nitrate-evolved populations and provides a comprehensive understanding of nitrate tolerance mechanisms,

457 which has important implications for linking genotypes with phenotypes in DvH. Future studies
458 may further understand possible indirect nitrate tolerance mechanisms using a coinciding evolved
459 control without additional nitrate and explore the contribution of important individual mutations
460 and other highly mutated genes to nitrate tolerance in EN populations.

461

462 **ACKNOWLEDGEMENTS**

463 The experimental work of this study (October 2013 to June 2017) was largely supported by
464 ENIGMA-Ecosystems and Networks Integrated with Genes and Molecular Assemblies
465 (http://enigma.lbl.gov), a Scientific Focus Area Program at Lawrence Berkeley National
466 Laboratory is based upon work supported by the U.S. Department of Energy, Office of Science,
467 and Office of Biological & Environmental Research under contract number DE-AC02-
468 05CH11231, the Office of the Vice President for Research at the University of Oklahoma (J.Z.),
469 and the National Program on Key Basic Research Project of China (973 Program, No.
470 2015CB150505 to S.C.); the subsequent integration work (October 2018 to March 2020) of this
471 study, including supplementary experiments, data analysis and manuscript writing, was mainly
472 supported by the National Natural Science Foundation of China (Grant numbers 31770539 and
473 91951207 to Z.H., grant number 31672262 to Q.Y.), and the Special Funds for Scientific and
474 Technological Innovation in Guangdong Province (Grant number 2018A030310302 to B.W.)

475

476 **SUPPLEMENTARY MATERIAL**

477 Supplementary information for this article is available at The ISME Journal's website.

References

1. Zhou JZ, He Q, Hemme CL, Mukhopadhyay A, Hillesland K, Zhou AF et al. How sulphate-reducing microorganisms cope with stress: lessons from systems biology. Nat Rev Microbiol. 2011; **9:** 452-466.
2. Bowles MW, Mogollon JM, Kasten S, Zabel M, Hinrichs KU. Global rates of marine sulfate reduction and implications for sub-sea-floor metabolic activities. Science. 2014; **344:** 889-891.
3. Plugge CM, Zhang W, Scholten JC, Stams AJ. Metabolic flexibility of sulfate-reducing bacteria. Front Microbiol. 2011; **2:** 81.
4. He Q, He Z, Joyner DC, Joachimiak M, Price MN, Yang ZK et al. Impact of elevated nitrate on sulfate-reducing bacteria: a comparative study of *Desulfovibrio vulgaris*. ISME J. 2010; **4:** 1386-1397.
5. Kamp A, Hogslund S, Risgaard-Petersen N, Stief P. Nitrate Storage and Dissimilatory Nitrate Reduction by Eukaryotic Microbes. Front Microbiol. 2015; **6:** 1492.
6. Hubert C, Voordouw G. Oil field souring control by nitrate-reducing *Sulfurospirillum* spp. that outcompete sulfate-reducing bacteria for organic electron donors. Appl Environ Microbiol. 2007; **73:** 2644-2652.
7. Canter LW. *Nitrates in Groundwater*, Routledge, 2019.
8. Ascott MJ, Gooddy DC, Wang L, Stuart ME, Lewis MA, Ward RS et al. Global patterns of nitrate storage in the vadose zone. Nat Commun. 2017; **8:** 1416.
9. Green SJ, Prakash O, Jasrotia P, Overholt WA, Cardenas E, Hubbard D et al. Denitrifying Bacteria from the Genus *Rhodanobacter* Dominate Bacterial Communities in the Highly Contaminated Subsurface of a Nuclear Legacy Waste Site. Appl Environ Microbiol. 2012; **78:** 1039-1047.
10. Hubert C. Microbial Ecology of Oil Reservoir Souring and its Control by Nitrate Injection. Springer:Berlin, Germany, 2010, pp 2753-2766.

11. Suri N, Voordouw J, Voordouw G. The Effectiveness of Nitrate-Mediated Control of the Oil Field Sulfur Cycle Depends on the Toluene Content of the Oil. Front Microbiol. 2017; **8:** 956.
12. Lucassen ECHET, Smolders AJP, Van der Salm AL, Roelofs JGM. High groundwater nitrate concentrations inhibit eutrophication of sulphate-rich freshwater wetlands. Biogeochemistry. 2004; **67:** 249-267.
13. Marietou A. Nitrate reduction in sulfate-reducing bacteria. FEMS Microbiol Lett. 2016; **363:** fnw155.
14. Cadby IT, Ibrahim SA, Faulkner M, Lee DJ, Browning D, Busby SJ et al. Regulation, sensory domains and roles of two *Desulfovibrio desulfuricans* ATCC27774 Crp family transcription factors, HcpR1 and HcpR2, in response to nitrosative stress. Mol Microbiol. 2016; **102:** 1120-1137.
15. Haveman SA, Greene EA, Voordouw G. Gene expression analysis of the mechanism of inhibition of *Desulfovibrio vulgaris* Hildenborough by nitrate-reducing, sulfide-oxidizing bacteria. Environ Microbiol. 2005; **7:** 1461-1465.
16. Callbeck CM, Agrawal A, Voordouw G. Acetate production from oil under sulfate-reducing conditions in bioreactors injected with sulfate and nitrate. Appl Environ Microbiol. 2013; **79:** 5059-5068.
17. Redding AM, Mukhopadhyay A, Joyner DC, Hazen TC, Keasling JD. Study of nitrate stress in *Desulfovibrio vulgaris* Hildenborough using iTRAQ proteomics. Brief Funct Genomic Proteomic. 2006; **5:** 133-143.
18. Korte HL, Fels SR, Christensen GA, Price MN, Kuehl JV, Zane GM et al. Genetic basis for nitrate resistance in *Desulfovibrio* strains. Front Microbiol. 2014; **5:** 153.
19. Korte HL, Saini A, Trotter VV, Butland GP, Arkin AP, Wall JD. Independence of nitrate and nitrite inhibition of *Desulfovibrio vulgaris* Hildenborough and use of nitrite as a substrate for growth. Environ Sci Technol. 2014; **49:** 924-931.
20. Elena SF, Lenski RE. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. Nat Rev Genet. 2003; **4:** 457-469.

21. Lenski RE. Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations. ISME J. 2017; **11:** 2181.

22. Barrick JE, Lenski RE. Genome dynamics during experimental evolution. Nat Rev Genet. 2013; **14:** 827-839.

23. Zhou A, Hillesland KL, He Z, Schackwitz W, Tu Q, Zane GM et al. Rapid selective sweep of pre-existing polymorphisms and slow fixation of new mutations in experimental evolution of *Desulfovibrio vulgaris*. ISME J. 2015; **9:** 2360-2372.

24. Adams J, Rosenzweig F. Experimental microbial evolution: history and conceptual underpinnings. Genomics. 2014; **104:** 393-398.

25. van Ditmarsch D, Boyle KE, Sakhtah H, Oyler JE, Nadell CD, Deziel E et al. Convergent Evolution of Hyperswarming Leads to Impaired Biofilm Formation in Pathogenic Bacteria. Cell Rep. 2013; **4:** 697-708.

26. Gerstein AC, Chun HJE, Grant A, Otto SP. Genomic convergence toward diploidy in *Saccharomyces cerevisiae*. PLoS Genet. 2006; **2:** 1396-1401.

27. Kerr B, Neuhauser C, Bohannan BJ, Dean AM. Local migration promotes competitive restraint in a host-pathogen 'tragedy of the commons'. Nature. 2006; **442:** 75-78.

28. Burch CL, Chao L. Evolution by Small Steps and Rugged Landscapes in the RNA Virus ϕ6. Genetics. 1999; **151:** 921-927.

29. Levy SF, Blundell JR, Venkataram S, Petrov DA, Fisher DS, Sherlock G. Quantitative evolutionary dynamics using high-resolution lineage tracking. Nature. 2015; **519:** 181.

30. Bell G. Evolutionary rescue of a green alga kept in the dark. Biol Lett. 2013; **9:** 20120823.

31. Stern DL. The genetic causes of convergent evolution. Nat Rev Genet. 2013; **14:** 751-764.

32. Tenaillon O, Barrick JE, Ribeck N, Deatherage DE, Blanchard JL, Dasgupta A et al. Tempo and mode of genome evolution in a 50,000-generation experiment. Nature. 2016; **536:** 165.

33. Tenaillon O, Rodriguez-Verdugo A, Gaut RL, McDonald P, Bennett AF, Long AD et al. The Molecular Diversity of Adaptive Convergence. Science. 2012; **335:** 457-461.

554 34. Vogwill T, Phillips RL, Gifford DR, MacLean RC. Divergent evolution peaks under intermediate
555 population bottlenecks during bacterial experimental evolution. P Roy Soc B-Biol Sci. 2016; **283**.
556 35. Cooper VS, Schneider D, Blot M, Lenski RE. Mechanisms causing rapid and parallel losses of
557 ribose catabolism in evolving populations of *Escherichia coli* B. J Bacteriol. 2001; **183:** 2834-2841.
558 36. Mehta-Kolte MG, Stoeva MK, Mehra A, Redford SA, Youngblut MD, Zane G et al. Adaptation of
559 *Desulfovibrio alaskensis* G20 to perchlorate, a specific inhibitor of sulfate reduction. Environ
560 Microbiol. 2019; **21:** 1395-1406.
561 37. Zhou A, Baidoo E, He Z, Mukhopadhyay A, Baumohl JK, Benke P et al. Characterization of NaCl
562 tolerance in *Desulfovibrio vulgaris* Hildenborough through experimental evolution. ISME J. 2013;
563 **7:** 1790-1802.
564 38. Mukhopadhyay A, He Z, Alm EJ, Arkin AP, Baidoo EE, Borglin SC et al. Salt stress in
565 *Desulfovibrio vulgaris* Hildenborough: an integrated genomics approach. J Bacteriol. 2006; **188:**
566 4068-4078.
567 39. Kong Y. Btrim: a fast, lightweight adapter and quality trimming program for next-generation
568 sequencing technologies. Genomics. 2011; **98:** 152-153.
569 40. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012; **9:** 357.
570 41. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N et al. The Sequence Alignment/Map
571 format and SAMtools. Bioinformatics. 2009; **25:** 2078-2079.
572 42. Narasimhan V, Danecek P, Scally A, Xue YL, Tyler-Smith C, Durbin R. BCFtools/RoH: a hidden
573 Markov model approach for detecting autozygosity from next-generation sequencing data.
574 Bioinformatics. 2016; **32:** 1749-1751.
575 43. Deatherage DE, Barrick JE. Identification of Mutations in Laboratory-Evolved Microbes from
576 Next-Generation Sequencing Data Using breseq. In: Sun L, Shou W (eds). *Engineering and*
577 *Analyzing Multicellular Systems: Methods and Protocols*. Springer New York: New York, NY,
578 2014, pp 165-188.

44. Good BH, McDonald MJ, Barrick JE, Lenski RE, Desai MM. The dynamics of molecular evolution over 60,000 generations. Nature. 2017; **551:** 45-50.

45. Heidelberg JF, Seshadri R, Haveman SA, Hemme CL, Paulsen IT, Kolonay JF et al. The genome sequence of the anaerobic, sulfate-reducing bacterium *Desulfovibrio vulgaris* Hildenborough. Nat Biotechnol. 2004; **22:** 554-559.

46. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M et al. The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. Nucleic Acids Res. 2016; gkw937.

47. Pelley JW. Protein Synthesis and Degradation. In: Pelley JW (ed). *Elsevier's Integrated Review Biochemistry (Second Edition)*. W.B. Saunders, Philadelphia:2012, pp 149-160.

48. Kelley LA, Mezulis S, Yates CM, Wass MN. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc. 2015; **10:** 845-858.

49. Keller KL, Bender KS, Wall JD. Development of a markerless genetic exchange system for *Desulfovibrio vulgaris* Hildenborough and its use in generating a strain with increased transformation efficiency. Appl Environ Microbiol. 2009; **75:** 7682-7691.

50. Aragao D, Macedo S, Mitchell EP, Romao CV, Liu MY, Frazao C et al. Reduced hybrid cluster proteins (HCP) from *Desulfovibrio desulfuricans* ATCC 27774 and *Desulfovibrio vulgaris* (Hildenborough): X-ray structures at high resolution using synchrotron radiation. J Biol Inorg Chem. 2003; **8:** 540-548.

51. Price MN, Huang KH, Alm EJ, Arkin AP. A novel method for accurate operon predictions in all sequenced prokaryotes. Nucleic Acids Res. 2005; **33:** 880-892.

52. Rajeev L, Luning EG, Dehal PS, Price MN, Arkin AP, Mukhopadhyay A. Systematic mapping of two component response regulators to gene targets in a model sulfate reducing bacterium. Genome Biol. 2011; **12:** R99.

603 53. Bender KS, Yen HC, Hemme CL, Yang Z, He Z, He Q et al. Analysis of a ferric uptake regulator
604 (Fur) mutant of *Desulfovibrio vulgaris* Hildenborough. Appl Environ Microbiol. 2007; **73:** 5389-
605 5400.
606 54. Zhou A, Chen YI, Zane GM, He Z, Hemme CL, Joachimiak MP et al. Functional characterization
607 of Crp/Fnr-type global transcriptional regulators in *Desulfovibrio vulgaris* Hildenborough. Appl
608 Environ Microbiol. 2012; **78:** 1168-1177.
609 55. He Q, Huang KH, He Z, Alm EJ, Fields MW, Hazen TC et al. Energetic consequences of nitrite
610 stress in *Desulfovibrio vulgaris* Hildenborough, inferred from global transcriptional analysis. Appl
611 Environ Microbiol. 2006; **72:** 4370-4381.
612 56. Rodionov DA, Dubchak I, Arkin A, Alm E, Gelfand MS. Reconstruction of regulatory and
613 metabolic pathways in metal-reducing delta-proteobacteria. Genome Biol. 2004; **5:** R90.
614 57. Cadby IT, Busby SJ, Cole JA. An HcpR homologue from *Desulfovibrio desulfuricans* and its
615 possible role in nitrate reduction and nitrosative stress. Biochem Soc Trans. 2011; **39:** 224-229.
616 58. Haveman SA, Brunelle V, Voordouw JK, Voordouw G, Heidelberg JF, Rabus R. Gene expression
617 analysis of energy metabolism mutants of *Desulfovibrio vulgaris* Hildenborough indicates an
618 important role for alcohol dehydrogenase. J Bacteriol. 2003; **185:** 4345-4353.
619 59. Walian PJ, Allen S, Shatsky M, Zeng L, Szakal ED, Liu H et al. High-throughput Isolation and
620 Characterization of Untagged Membrane Protein Complexes: Outer Membrane Complexes of
621 *Desulfovibrio vulgaris*. J Proteome Res. 2012; **11:** 5720-5735.
622 60. Los DA, Murata N. Membrane fluidity and its roles in the perception of environmental signals.
623 Biochim Biophys Acta. 2004; **1666:** 142-157.
624 61. Rowlett VW, Mallampalli V, Karlstaedt A, Dowhan W, Taegtmeyer H, Margolin W et al. Impact
625 of Membrane Phospholipid Alterations in Escherichia coli on Cellular Function and Bacterial
626 Stress Adaptation. J Bacteriol. 2017; **199**.

62. Zhou A, Lau R, Baran R, Ma J, von Netzer F, Shi W et al. Key Metabolites and Mechanistic Changes for Salt Tolerance in an Experimentally Evolved Sulfate-Reducing Bacterium, *Desulfovibrio vulgaris*. mBio. 2017; **8:** e01780-01717.

63. Haveman SA, Greene EA, Stilwell CP, Voordouw JK, Voordouw G. Physiological and gene expression analysis of inhibition of *Desulfovibrio vulgaris* Hildenborough by nitrite. J Bacteriol. 2004; **186:** 7944-7950.

64. Finkel SE. Long-term survival during stationary phase: evolution and the GASP phenotype. Nat Rev Microbiol. 2006; **4:** 113-120.

65. Farrell MJ, Finkel SE. The growth advantage in stationary-phase phenotype conferred by rpoS mutations is dependent on the pH and nutrient environment. J Bacteriol. 2003; **185:** 7044-7052.

66. Clark ME, He Q, He Z, Huang KH, Alm EJ, Wan XF et al. Temporal transcriptomic analysis as *Desulfovibrio vulgaris* hildenborough transitions into stationary phase during electron donor depletion. Appl Environ Microbiol. 2006; **72:** 5578-5588.

67. Galhardo RS, Hastings PJ, Rosenberg SM. Mutation as a stress response and the regulation of evolvability. Crit Rev Biochem Mol Biol. 2007; **42:** 399-435.

68. Rolfe MD, Rice CJ, Lucchini S, Pin C, Thompson A, Cameron ADS et al. Lag Phase Is a Distinct Growth Phase That Prepares Bacteria for Exponential Growth and Involves Transient Metal Accumulation. J Bacteriol. 2012; **194:** 686-701.

Figure titles and legends

Figure 1 Increased $NaNO_3$ tolerance in EN populations compared to AN populations. The lag phase (a), growth rate (b), and biomass yield (c) of AN and EN populations in LS4D, LS4D + 30 mM $NaNO_3$, and LS4D + 100 mM $NaNO_3$ for testing nitrate tolerance. Data are shown as mean ± standard deviation, and statistical significance was based on the pairwise ANOVA (*: $p < 0.05$). No growth was detected for AN populations within 70 h under 100 mM $NaNO_3$.

Figure 2 Growth of DvH parental strain (JZ001) and marker exchange mutants (ME2543, ME2547 and ME2548) of nitrosative stress response genes (NSR) under (a) no stress (LS4D), (b) nitrite stress (LS4D + 0.15 mM $NaNO_2$), (c) nitrate stress (LS4D + 30 mM $NaNO_3$) and (d) the growth phenotype of complemented mutants with extra 30 mM $NaNO_3$ in the medium. JZ001P: JZ001 with empty vector pMO719; ME2543(C): the complemented strain of ME2543; ME2547(C): the complemented strain of ME2547. The complemented strain for ME2548 was not obtained.

Figure 3 Growth of DvH parental strain (JZ001) and marker exchange mutants (ME2394, ME2395, ME2396 and ME2405) of nitrogen regulatory protein family C genes (NRC) under (a) no stress (LS4D), (b) salt stress (LS4D + 100 mM NaCl), (c) nitrate stress (LS4D + 30 mM $NaNO_3$) and (d) the growth phenotype of complemented mutants with extra 30 mM $NaNO_3$ in the medium. JZ001P: JZ001 with empty vector pMO719; ME2394(C): the complemented strain of ME2394; ME2395(C): the complemented strain of ME2395; ME2396(C): the complemented strain of ME2396; ME2405(C): the complemented strain of ME2405.

669

670 **Figure 4** A schematic representation of possible nitrate tolerance mechanisms in DvH. Frequently
671 mutated genes were clustered into six groups (DVU#), including (a) nitrosative stress response
672 (NSR) genes, (b) nitrogen regulation protein C family (NRC) genes, (c) nitrate cluster, (d) fatty
673 acid synthesis genes, (e) iron regulatory genes, and (f) two component system *LytR*/*LytS*. Nonsense
674 and frameshift mutations were identified in the first three gene groups (a, b, and c), indicating that
675 a functional loss of some of those genes/groups conferred nitrate tolerance in DvH. Possible
676 mechanisms for increased nitrate tolerance may include a shift of energy metabolism through
677 derepression of sulfate reduction (a), lacking alcohol oxidation pathways catalyzed by alcohol
678 dehydrogenase (b), barring entry of nitrate into cell by (c) [18], altering overall membrane
679 characteristics (d) [23, 62], and conferring growth advantages by regulating iron homeostasis (e)
680 and carbon metabolism (carbon starvation genes *CstA* and *CstB*) (f) at the stationary phase [66].
681 Dashed boxes represent possible nitrate tolerance mechanisms. Key genes were confirmed by
682 mutagenesis analysis for their contribution to nitrate tolerance (a, b), and arrows represent
683 processes demonstrated by previous studies [18, 23, 52, 56, 66]. ADH: alcohol dehydrogenase;
684 HK: sensor histidine kinase; RR: response regulator; Ⓟ: phosphoryl group; gene$^{(R)}$: global
685 transcriptional regulator.

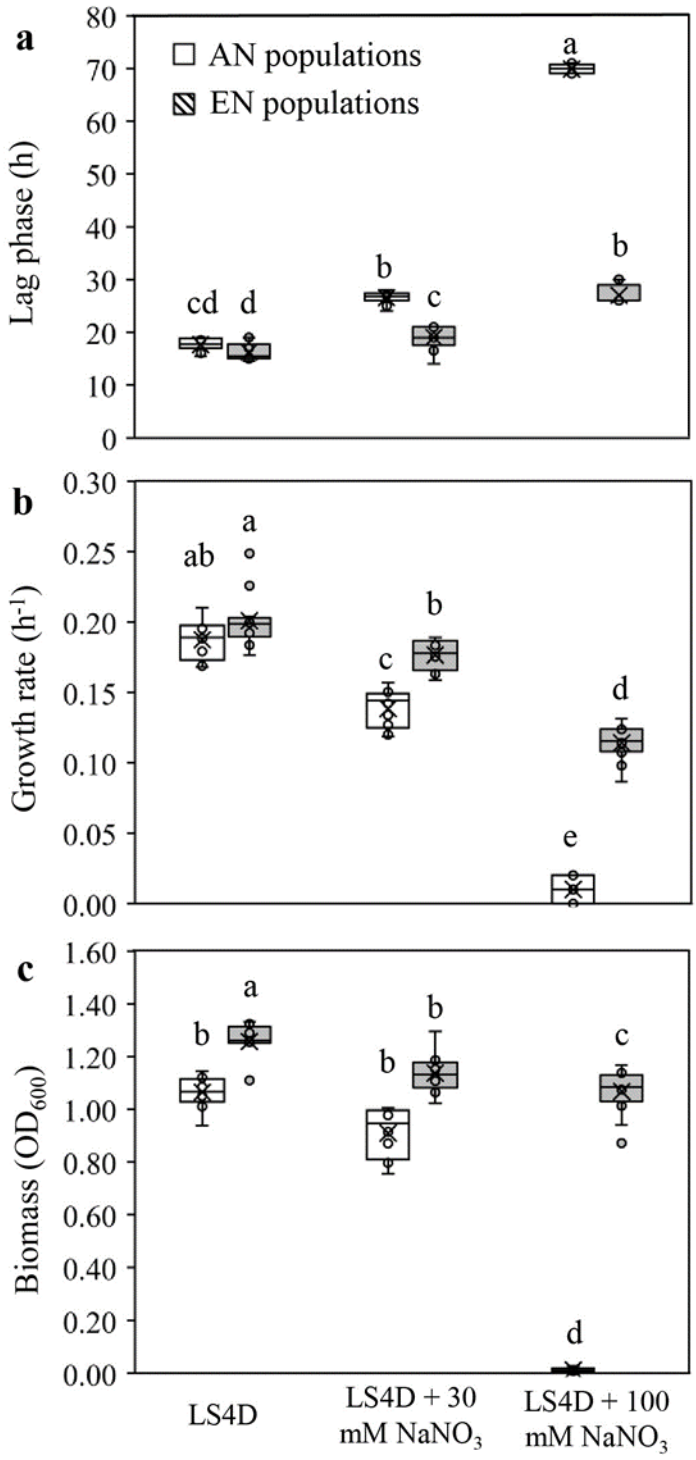
a
AN populations
EN populations
Lag phase (h)
80
70
60
50
40
30
20
10
0
cd
d
b
c
a
b
b
Growth rate (h-1)
0.30
0.25
0.20
0.15
0.10
0.05
0.00
ab
a
c
b
e
d
c
Biomass (OD600)
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
b
a
b
b
d
c
LS4D
LS4D + 30 mM NaNO3
LS4D + 100 mM NaNO3

Figure 1

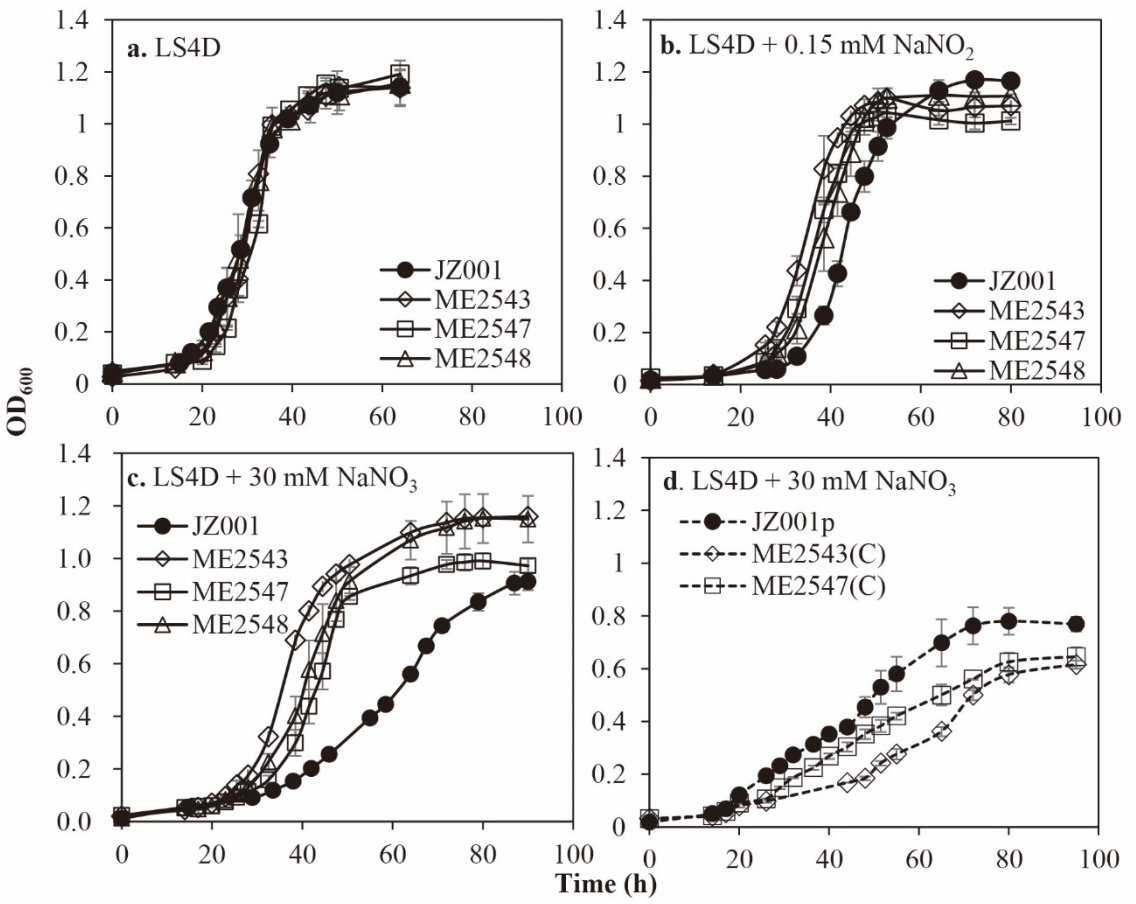
a. LS4D
b. LS4D + 0.15 mM NaNO2
c. LS4D + 30 mM NaNO3
d. LS4D + 30 mM NaNO3
JZ001
ME2543
ME2547
ME2548
JZ001p
ME2543(C)
ME2547(C)
OD600
Time (h)

Figure 2

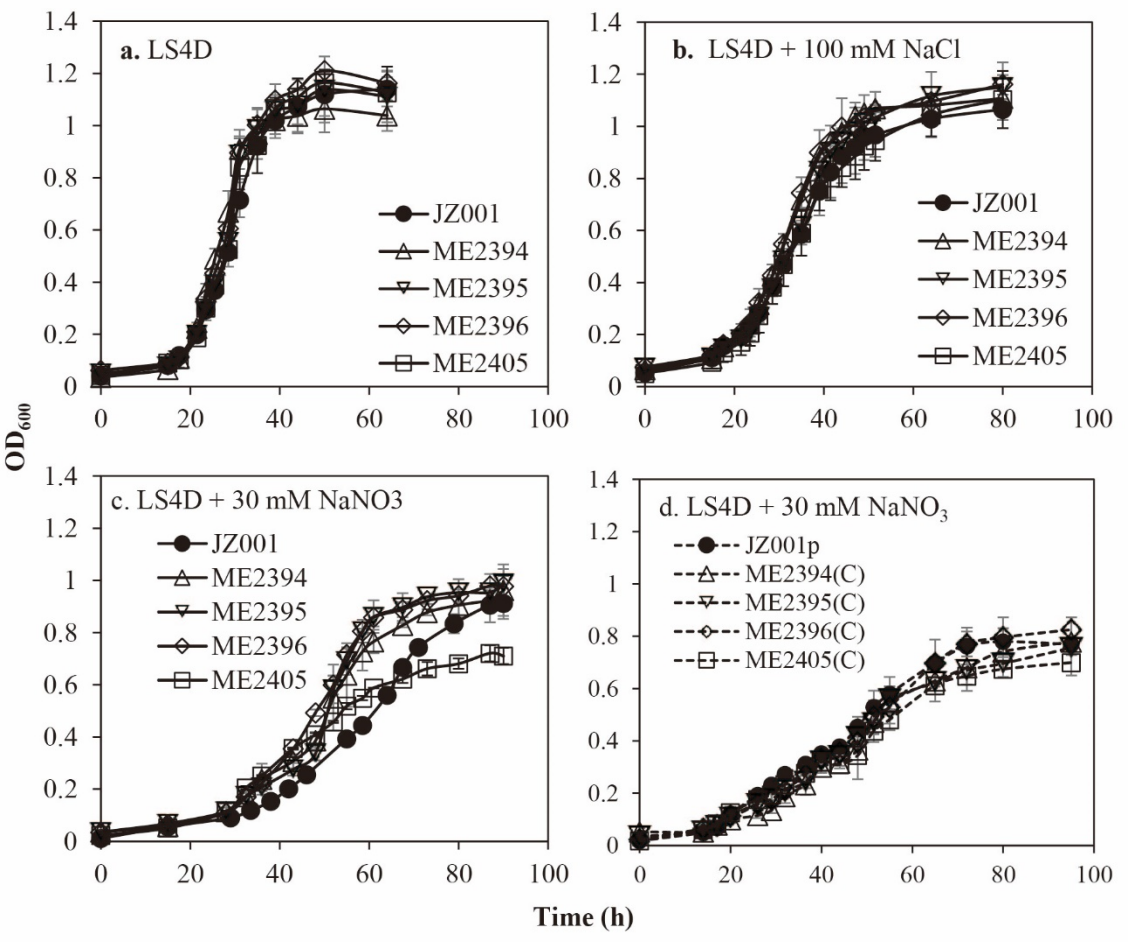
a. LS4D
b. LS4D + 100 mM NaCl
c. LS4D + 30 mM NaNO3
d. LS4D + 30 mM NaNO3
JZ001
ME2394
ME2395
ME2396
ME2405
JZ001p
ME2394(C)
ME2395(C)
ME2396(C)
ME2405(C)
OD600
Time (h)

Figure 3

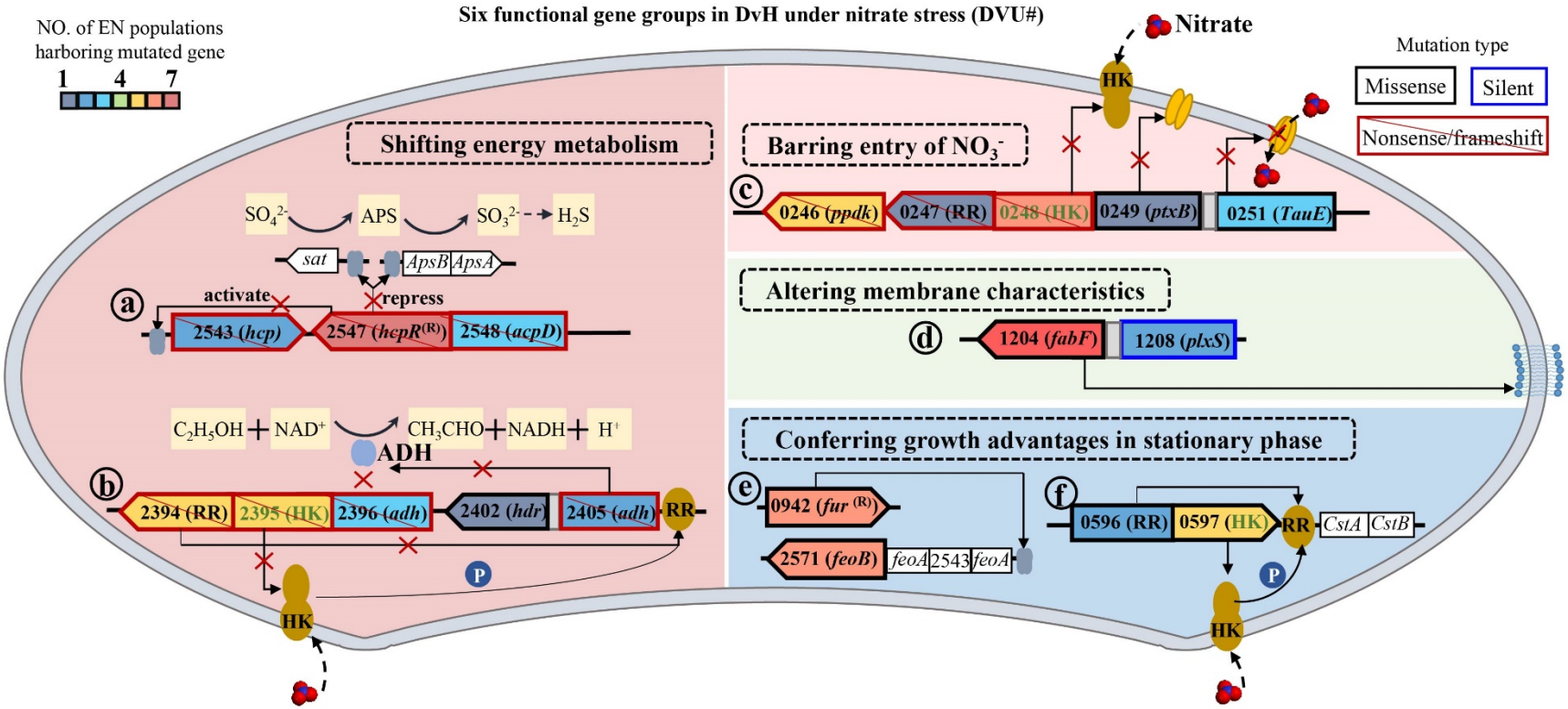
NO. of EN populations harboring mutated gene
1 4 7
Six functional gene groups in DvH under nitrate stress (DVU#)
Nitrate
Mutation type
Missense
Silent
Nonsense/frameshift
Shifting energy metabolism
Barring entry of NO3-
Altering membrane characteristics
Conferring growth advantages in stationary phase
SO4 2- APS SO3 2- H2S
sat
ApsB ApsA
activate
repress
2543 (hcp)
2547 (hcpR(R))
2548 (acpD)
C2H5OH + NAD+ CH3CHO + NADH + H+
ADH
2394 (RR)
2395 (HK)
2396 (adh)
2402 (hdr)
2405 (adh)
RR
P
HK
0246 (ppdk)
0247 (RR)
0248 (HK)
0249 (ptxB)
0251 (TauE)
1204 (fabF)
1208 (plxS)
0942 (fur (R))
2571 (feoB)
feoA 2543 feoA
0596 (RR)
0597 (HK)
CstA CstB

Figure 4

Table 1 Strains and plasmids used in this study.

Strain/plasmid	Genotype or relevant characteristics	Source or reference
Escherichia coli α-select (silver efficiency)	Φ80d/*lac*ZΔM15, Δ（*lacZYA-argF*）U169, *recA*1, *endA*1, *hsdR*17（r_K^-, m_K^+）, *phoA*, *supE*44, λ^-, *thi*-1, *gyrA*96, *relA*1	Bioline
Desulfovibrio vulgaris		
JZ001	Wild type (pDV1) Δ*upp*; 5-FU^r (Parent strain for marker exchange mutants)	This study
ME2394	JZ001 ΔDVU2394 :: (*npt upp*); Km^r, 5-FU^s	This study
ME2395	JZ001 ΔDVU2395 :: (*npt upp*); Km^r, 5-FU^s	This study
ME2396	JZ001 ΔDVU2396 :: (*npt upp*); Km^r, 5-FU^s	This study
ME2405	JZ001 ΔDVU2405 :: (*npt upp*); Km^r, 5-FU^s	This study
ME2543	JZ001 ΔDVU2543 :: (*npt upp*); Km^r, 5-FU^s	This study
ME2547	JZ001 ΔDVU2547 :: (*npt upp*); Km^r, 5-FU^s	This study
ME2548	JZ001 ΔDVU2548 :: (*npt upp*); Km^r, 5-FU^s	This study
pMO719	pCR8/GW/TOPO containing SRB replicon (pBG1); Sp^r	[1]
pMO746	pCR4-TOPO containing P_{npt}-*npt-upp*;Km^r;	[1]
pMO719-DVU2394	pMO719 containing DNA fragment 2500764-2500995::DVU2394: Sp^r	This study
pMO719-DVU2395	pMO719 containing DNA fragment 2500764-2500995::DVU2395: Sp^r	This study
pMO719-DVU2396	pMO719 containing DNA fragment 2500764-2500995::DVU2396: Sp^r	This study
pMO719-DVU2405	pMO719 containing DNA fragment 2511308-2511461::DVU2405: Sp^r	This study
pMO719-DVU2543	pMO719 containing DNA fragment 2653710-2653863::DVU2543; Sp^r	This study
pMO719-DVU2547	pMO719 containing DNA fragment 2661298-2661384::DVU2547; Sp^r	This study

Km: Kanamycin; Sp: spectinomycin; 5-Fu: 5-fluorouracil; superscript “r”: resistance; superscript “s”: sensitivity.

Table 2 The number of newly acquired mutations in 19 genes of six gene groups in the EN populations.

Gene groups	Gene	COGs	Products	No. of EN populations bearing mutations	
				in gene	in gene group
Nitrosative stress response (NSR) genes	DVU2543	C	Hydroxylamine reductase (*hcp*)	2	12
	DVU2547	T	Transcriptional regulator (*hcpR*)	7	
	DVU2548	I	Acyl carrier protein phosphodiesterase (*acpD*)	3	
Nitrogen regulatory protein family C (NRC) genes	DVU2394	T	Sigma-54 dependent transcriptional regulator/response regulator (RR)	5	12
	DVU2395	T	Sensor histidine kinase (HK)	5	
	DVU2396	C	Alcohol dehydrogenase, iron-containing (*adh*)	3	
	DVU2402	C	Heterodisulfide reductase, A subunit (*hdr*)	1	
	DVU2405	C	Alcohol dehydrogenase, iron-containing (*adh*)	2	
Nitrate cluster	DVU0246	S	Pyruvate phosphate dikinase PEP/pyruvate binding subunit (*ppdk*)	5	12
	DVU0247	T	CheY-like superfamily, response regulator (RR)	1	
	DVU0248	T	Signal transduction histidine kinase (HK)	6	
	DVU0249	P	Probable phosphite transport system-binding protein, putative (*ptxB*)	1	
	DVU0251	R	Transmembrane protein TauE like	3	
Fatty acid synthesis genes	DVU1204	IQ	3-oxoacyl-(acyl-carrier-protein) synthase II (*fab*)	7	9
	DVU1208	I	Putative glycerol-3-phosphate acyltransferase (*plsX*)	2	
Iron regulatory genes	DVU0942	P	FUR family transcriptional regulator (*fur*)	6	12
	DVU2571	P	Ferrous iron transport protein B (*feoB*)	6	
LytR/LytS two component system	DVU0596	KT	DNA-binding response regulator (*lytR*)	1	6
	DVU0597	T	Histidine kinase (*lytS*)	5	

Table 3 Newly acquired mutations in nitrosative stress response (NSR) genes in the EN populations.

Gene	Populations	Mutation type	Allele frequency (%)	Position	Nucleotide change	Amino-acid change
DVU2543 *hcp*	EN4	SNP	99	2654808	G → A	**Trp → STOP**
	EN5	SNP	79	2654808	G → A	**Trp → STOP**
DVU2547 *hcpR*	EN1	SNP	74	2660114	G → T	Gly → Val
		Deletion	13	2660166	-4: GGCG	27 amino acids changed
		SNP	17	2660388	C → T	Pro → Ser
	EN3	SNP	66	2660235	C → T	Leu → Phe
		SNP	36	2660404	C →A	Phe → Leu
	EN6	SNP	98	2660273	C → T	Ser → Leu
	EN7	SNP	100	2660403	C → T	**Gln → STOP**
	EN8	SNP	99	2660403	C → T	**Gln → STOP**
	EN9	SNP	94	2660477	C → T	Ser → Phe
	EN10	Deletion	72	2660450	-15:GGCGCGAGGCCATCA	five amino acids loss
DVU2548 *acpD*	EN2	Deletion	93	2661113	GA → A	**Frameshift mutation**
		SNP	46	2661114	A → C	Ser → Pro
	EN8	Deletion	21	2661089	-57 bp	**Frameshift mutation**
	EN11	Insertion	71	2661301	+4: CACT	three bases upstream of the start codon
			62	2661319	+12: CTTGATATACAG	21 bases upstream of the start codon
	EN12	Deletion	100	2661182	CGG → CG	**Frameshift mutation**

Table 4 Newly acquired mutations in nitrogen regulatory protein C family (NRC) genes in the EN populations.

Gene	Populations	Mutation type	Allele frequency (%)	Mutation position	Nucleotide change	Amino-acid change
DVU2394	EN3	Deletion	16	2497229	-1: G	**2** amino acids changed
		SNP	59	2497736	C → T	Glu → Lys
	EN5	SNP	16	2496979	T → C	Tyr → Cys
	EN6	Deletion	49	2497411	-1: G	**Frameshift mutation**
	EN7	Deletion	100	2497258	-18: GCAATTCACGCACGTTGC	6 amino acids loss
	EN9	Deletion	36	2497708	-19:GAACAGTTCGCTACGGGCC	**Frameshift mutation**
		Insertion	11	2497728	+12: GACCGCGGACAT	4 amino acids inserted
DVU2395	EN4	Deletion	100	2499096	-8: CGGCGAGC	**Frameshift mutation**
	EN8	Insertion	46	2498544	+2: CG	**4** amino acids changed
			26	2499078	+14:CCTCATGCGCCACC	**Frameshift mutation**
	EN9	Deletion	54	2499384	-1: T	**Frameshift mutation**
	EN10	SNP	15	2498363	T →C	Silent mutation
		SNP	23	2498737	T →C	Lys → Glu
		SNP	45	2498935	G → A	**Gln → STOP**
	EN12	SNP	100	2499244	G → A	**Gln → STOP**
DVU2396	EN1	SNP	20	2499926	G → A	Pro → Ser
		SNP	72	2500012	G → T	**Ser → STOP**
	EN2	Deletion	98	2499984	-1: C	**Frameshift mutation**
	EN6	Deletion	42	2499735	-1: C	**Frameshift mutation**
DVU2402	EN2	SNP	96	2507986	G → A	Arg → Cys
DVU2405	EN5	SNP	81	2510695	C → A	**Glu → STOP**
	EN11	Deletion	100	2510695	-1: C	**Frameshift mutation**
		SNP	100	2510695	C → T	**Glu → Lys**

Table 5 Newly acquired mutations in nitrate cluster in the EN populations.

Gene	Populations	Mutation type	Allele frequency (%)	Mutation position	Nucleotide change	Amino-acid change
DVU0246	EN1	Deletion	77	279997	AGG → AG	**Frameshift mutation**
	EN2	Deletion	15	279123	CG → G	**Frameshift mutation**
		Deletion	53	279997	AGG → AG	**Frameshift mutation**
	EN7	Deletion	48	279074	-234 bp	**Frameshift mutation**
	EN8	Deletion	21	280374	-11: TGTTGCCGTCG	**Frameshift mutation**
	EN9	SNP	59	279507	C → A	**Gly → STOP**
	EN11	Deletion	48	279702	GT → G	**Frameshift mutation**
DVU0247	EN6	Deletion	34	281155	AT → A	**Frameshift mutation**
		Insertion	50	281286	+4:ACGA	**Frameshift mutation**
DVU0248	EN3	Deletion	14	281561	AGG → AG	**Frameshift mutation**
	EN4	SNP	63	283154	C → T	Asp → Asn
	EN5	Deletion	18	281703	GT → G	**Nonsense mutation**
	EN8	SNP	65	282788	C → T	Ala → Thr
	EN9	Deletion	14	293167	-63 bp	**Frameshift mutation**
	EN10	SNP	18	283136	C → T	Gly → Arg
	EN12	SNP	47	283204	G → T	Ser → Tyr
DVU0249	EN3	SNP	17	283782	C → T	Gly → Asp
DVU0251	EN3	SNP	13	285479	G → A	Thr → Ile
		SNP	19	285947	C → G	Gly → Ala
		SNP	13	286439	C → T	109 bases upstream of the start codon
	EN7	SNP	33	285948	C → T	Gly → Ser
	EN12	SNP	23	285948	C → T	Gly → Ser
DVU0246-0251	EN5	Deletion	28	279540	-9920 bp	△DVU0246-0251
	EN10		33	279032	-15883 bp	
	EN12		32	272887	-38514 bp	
DVU0250-0251	EN5	Deletion	27	284766	-5814 bp	△DVU0250-0251

1. Keller KL, Bender KS, Wall JD. Development of a markerless genetic exchange system for Desulfovibrio vulgaris Hildenborough and its use in generating a strain with increased transformation efficiency. *Appl Environ Microbiol* 2009;**75:**7682-7691.

Supporting Information

Experimental evolution reveals nitrate tolerance mechanisms in *Desulfovibrio vulgaris*

Bo Wu[1, 2], Feifei Liu[2, 3], Aifen Zhou[2], Juan Li[4], Longfei Shu[1, 2], Megan L. Kempher[2], Xueqin Yang[1, 2], Daliang Ning[2], Feiyan Pan[2], Grant M. Zane[5], Judy D. Wall[5], Joy D. Van Nostrand[2], [6]Philippe Juneau; Shouwen Chen [7,8 *], Qingyun Yan[1, 2 *], Jizhong Zhou[2, 9, 10,*], and Zhili He[1, 2, 4*]

[1]Environmental Microbiomics Research Center, School of Environmental Science and Engineering, Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Sun Yat-sen University, Guangzhou 510006, China ; [2]Institute for Environmental Genomics and Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK 73019; [3]Guangdong Provincial Key Lab of Microbial Culture Collection and Application, Guangdong Institute of Microbiology and State Key Laboratory of Applied Microbiology Southern China, Guangzhou 510070, China; [4]College of Agronomy, Hunan Agricultural University, Changsha 410128, China; [5]Departments of Biochemistry and Molecular Microbiology & Immunology, University of Missouri-Columbia, Columbia, MO 65211; [6]Département des Sciences Biologiques, TOXEN, Ecotoxicology of Aquatic Microorganisms Laboratory, Université du Québec à Montréal, Montréal, Canada;[7]State Key Laboratory of Agricultural Microbiology, Huazhong Agricultural University, Wuhan 430070, China; [8]Hubei Collaborative Innovation Center for Green Transformation of Bio-Resources, College of Life Sciences, Hubei University, Wuhan 430062, China; [9]Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA; [10]School of Environment, Tsinghua University, Beijing 100084, China

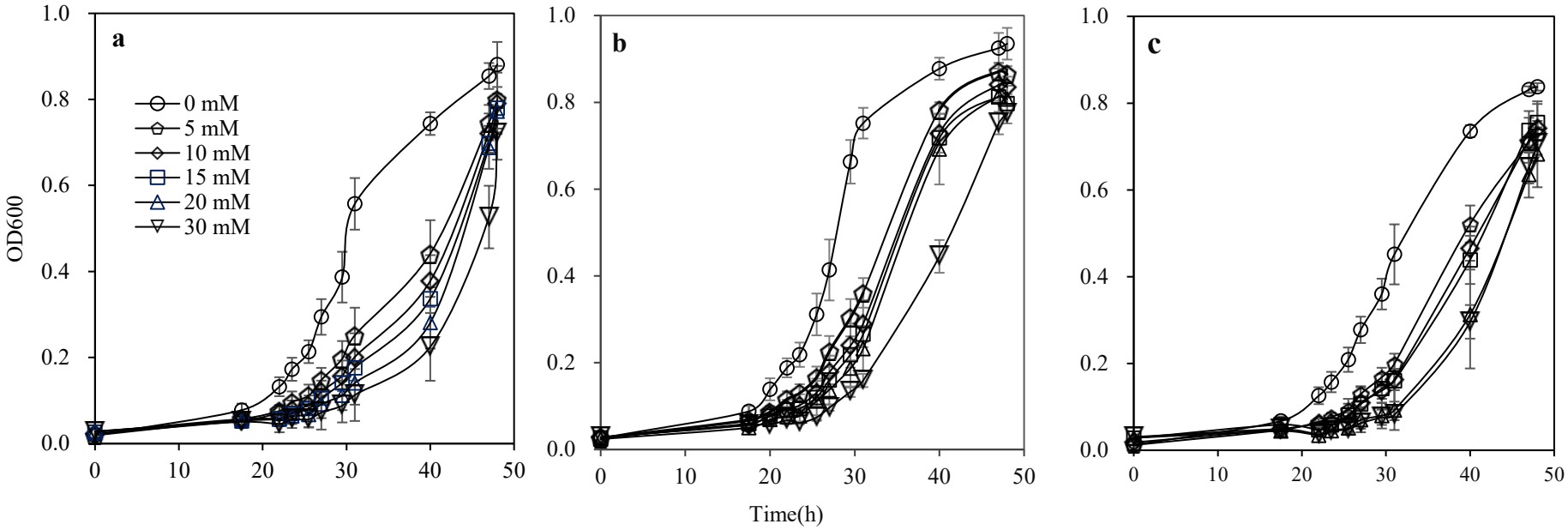

Figure S1 Determination of $NaNO_3$ concentrations for a long-term experimental evolution with ancestral populations AN2 (a), AN8 (b), and AN11 (c).

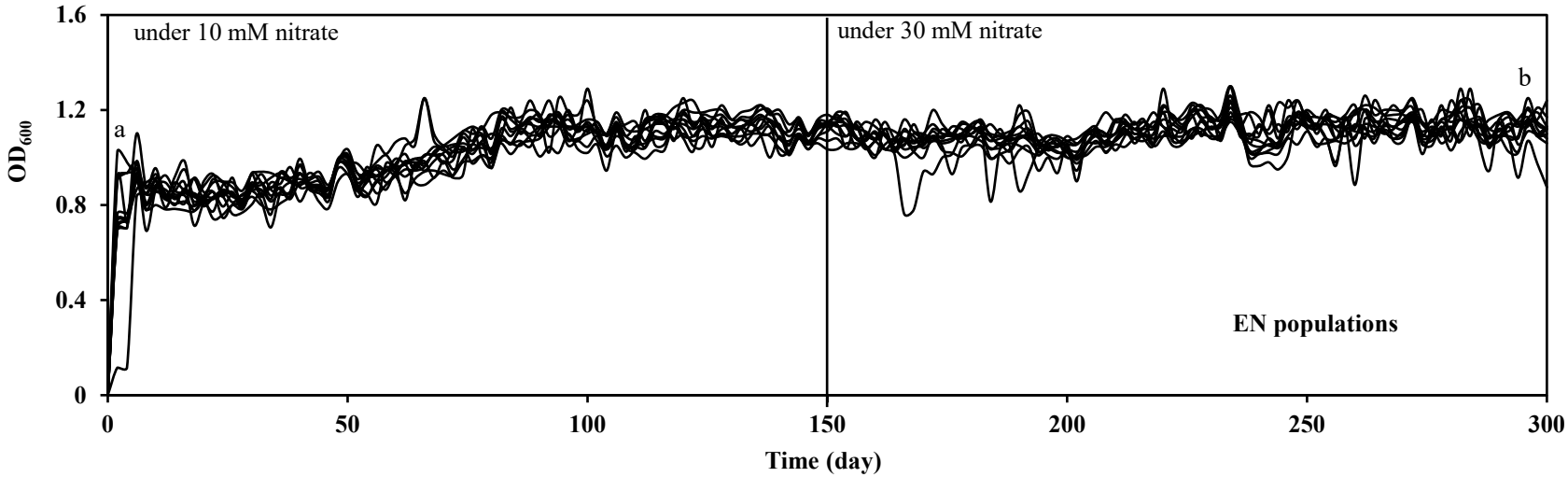

Figure S2 Growth records of 12 ancestral (AN) populations under elevated nitrate. Nitrate-evolved (EN) populations were obtained by transferring 12 ancestral DvH populations under 10 mM nitrate for 500 generations, then under 30 mM for 500 generations. Statistical significance was evaluated between biomass of first transfer and last transfer based on the Student's t-test ($p < 0.001$).

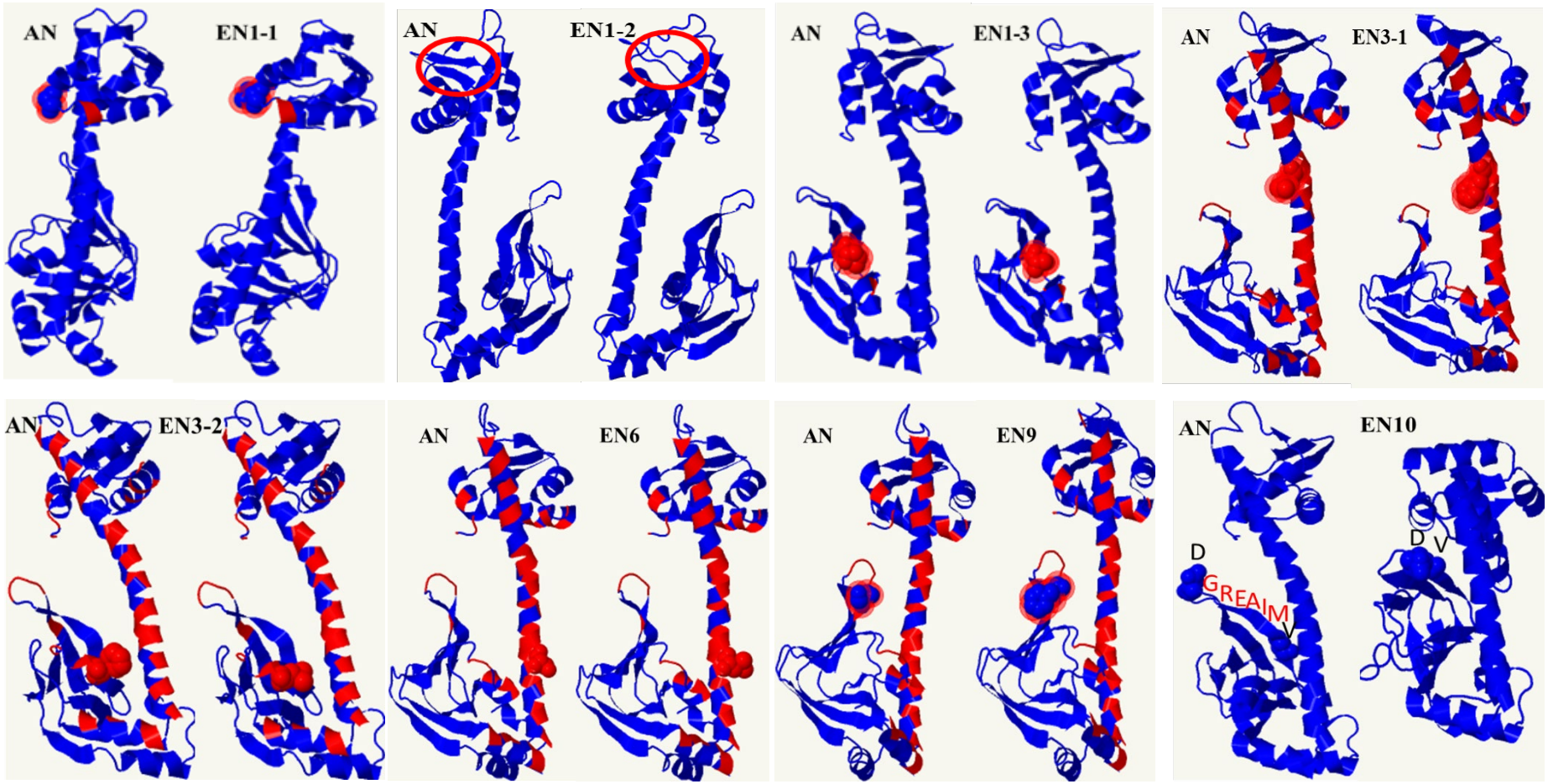

Figure S3 Changes in the tertiary structure of HcpR protein in the nitrate-evolved populations predicted by Phyre2 using Jmol [1]. The site surrounded by red lines presents the changes in the tertiary structure in HcpR.

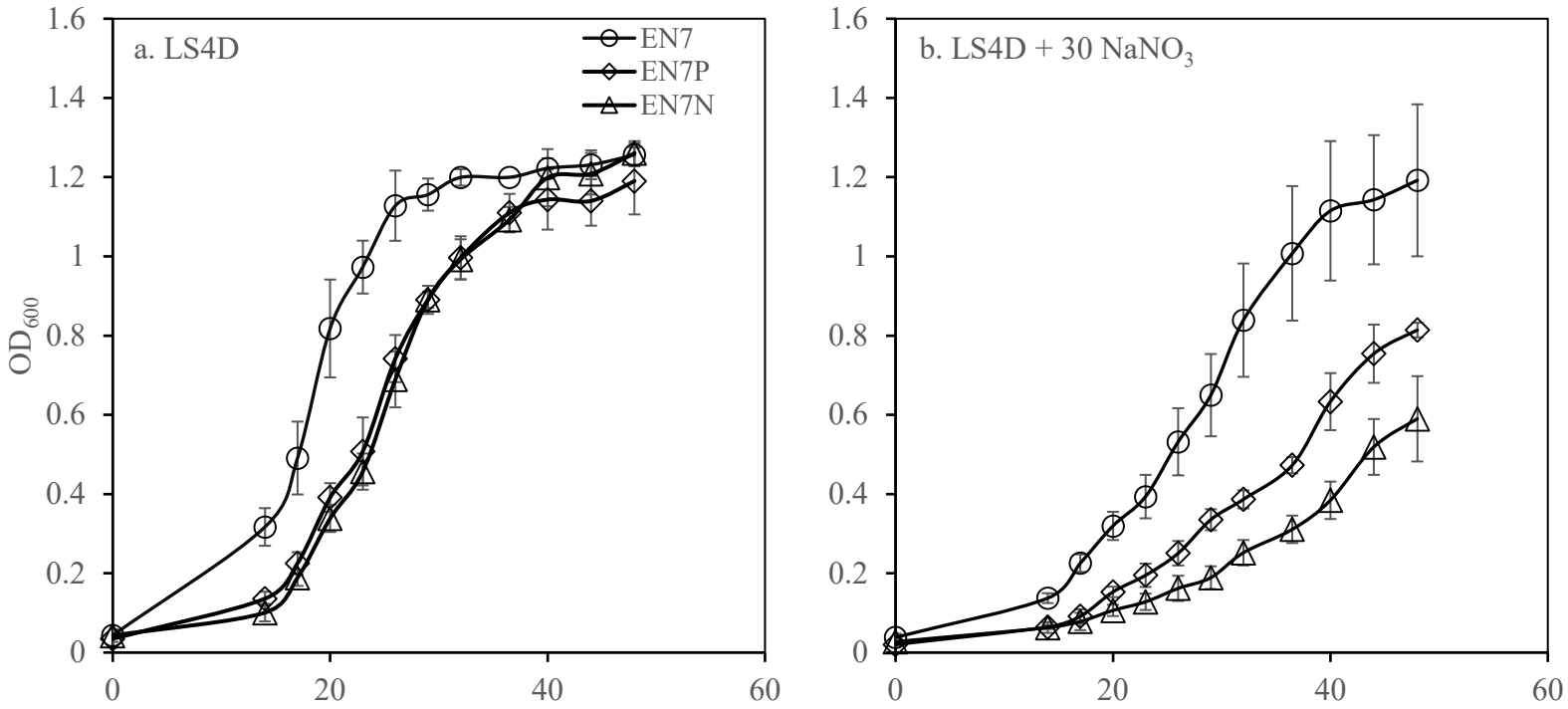

Figure S4 Growth of EN7, EN7P (EN7 with empty vector pMO719), and EN7N (EN7 with DVU2547 complement plasmid) under (a) no stress (LS4D) and (b) nitrate stress (LS4D + 30 mM $NaNO_3$).

Table S1 Primers used to validate the randomly chosen mutations.

Primers	Sequences	Length	Populations (filter score/allele frequency in the population)					
0246L	CCTCTCGACGACCTTCTCC	576	EN1	EN2	EN8	EN11		
0246R	GGTGTGGATGGCGATGAC		(27.5/77.32%)	(29.5/53.39%)	(149/21.43%)	(217/59.00%)		
0248L	CAAGATCGTCTCCGACCTG	580	EN3	EN5				
0248R	GGGGTTGATGGGTTCGAG		(74.5/14.29%)	(12.7/10.75%)				
0251L	GTGCTCGGCCTCTACCTCTT	587	EN3					
0251R	TAGCCGAAGATGCGGTAGTT		(44/13.33%)					
1260L	CGAGGGCTTTGCACTCTATG	601	EN1	EN4				
1260R	CTGGTCGTTGAGGCGATAAT		(217/72.73%)	(215/61.84%)				
2394L	ATCCAGTCGCTGGGCACT	598	EN6	EN7	EN9			
2394R	GTAACGCATGTTCGGTGATG		(33.5/49.44%)	(214/100%)	(107/35.63%)			
2395L	CGAGCCTCGACGATCTTATC	559	EN4	EN8	EN12			
2395R	AAGCCGTTGATGGCTGTC		(214/100%)	(217/46.21%)	(222/100%)			
2547L	GTGCAGCGTATCCTTGACCT	686	EN1	EN3	EN6	EN7	EN8	EN9
2547R	TAGAGTTGCTCCGTCCCTTC		(53.5/12.5%)	(225/66.32%)	(222/98.06%)	(222/100%)	(222/98.94%)	(222/94.32%)
2571L	GTGGCCTGTTCATCCACAC	604	EN1	EN2	EN3	EN6	EN8	EN12
2571R	GCCAGTACCGTGTTGAACG		(222/100%)	(214/100%)	(222/100%)	(222/98.96%)	(222/100%)	(222/100%)
2287L	ACTCGACCCGGTGATGCT	350	AN2	AN7	EN2	EN6	EN9	
2287R	GGTAGAAGGTGAACGCCTGA		(225/72.22%)	(225/67.07%)	(222/100%)	(222/100%)	(222/100%)	

Table S2 Mutations identified in 12 AN populations (MP*: the number of populations with mutations).

Gene	COG	Predicted function	Position	Reference	Alter	Type	AN1	AN2	AN3	AN4	AN5	AN6	AN7	AN8	AN9	AN10	AN11	AN12	MP*
DVU0129	T	Sensory box protein	165841	C	A	Gene							0.15						1
DVU0162	EF	Carbamoyl-phosphate synthase, large subunit	201043	C	G	Gene								0.50			0.23		2
DVU0281	S	Exopolysaccharide biosynthesis protein, putative	326403	G	A	Gene	0.28	0.33	0.77		0.16	0.23	0.39		0.38		0.16	0.27	9
DVU0448	M	GDP-mannose 4,6-dehydratase	508909	G	C	Gene		0.13											1
DVU0467	E	Anthranilate phosphoribosyltransferase	535249	G	C	Gene	0.26	0.34	0.57	0.15	0.15	0.25	0.22		0.35		0.14	0.31	10
DVU0564	S	ISDvu4, transposase, truncation	629934	G	GCTAC	Intergenic					0.16								1
DVU0564	S	ISDvu4, transposase, truncation	629934	GA	AGGGCTAC	Intergenic										0.32			1
DVU0700	NT	Methyl-accepting chemotaxis protein	778393	A	G	Gene	0.17												1
DVU0788	M	Rod shape-determining protein MreC	874923	G	C	Gene		0.38											1
DVU0796	E	Histidinol dehydrogenase	882512	C	G	Intergenic	0.38	0.25	0.50	0.44				0.60	0.20	0.40		0.75	8
DVU1087	S	Hypothetical protein DVU1087	1191167	C	T	Gene			0.50	0.50		0.50				0.40	1.00	0.33	6
DVU1104	S	Baseplate assembly protein, putative	1210118	C	G	Gene		0.26											1
DVU1204	IQ	3-oxoacyl-(acyl-carrier-protein) synthase II	1295951	G	C	Gene	0.14												1
DVU1238	ET	Amino acid ABC transporter, periplasmic amino acid-binding protein	1327518	G	C	Gene							0.13						1
DVU1272	NU	General secretion pathway protein E, putative	1363430	C	A	Gene	0.25			0.49				0.79	0.36		0.52	0.18	6
DVU1349	H	Geranylgeranyl diphosphate synthase	1426830	T	G	Gene	0.49	0.66	0.50	0.38	0.86	0.80	0.62	0.24	0.28	0.99	0.30	0.58	12
DVU1360	M	UDP-glucose 4-epimerase	1439686	CATA	CA	Gene												0.27	1
DVU1368	R	Rhodanese-like domain-containing protein	1447894	G	A	Gene								0.30			0.26		2
DVU1371	R	HAD family hydrolase	1449911	C	G	Gene							0.28						1
DVU1530	J	Metallo-beta-lactamase family protein	1599469	C	T	Gene	0.35	0.37	0.61	0.14		0.21	0.37		0.43		0.17	0.28	9
DVU1698	S	Hypothetical protein DVU1698	1773342	CGG	CG	Gene												1.00	1
DVU1698	S	Hypothetical protein DVU1698	1773345	A	G	Gene		0.64	0.33	1.00			0.44		0.50			1.00	6
DVU1722	S	Hypothetical protein DVU1722	1796665	C	A	Gene							0.14						1
DVU1769	R	Periplasmic hydA	1833244	C	A	Gene												0.25	1
DVU1819	U	Preprotein translocase subunit SecD	1884277	G	C	Gene									0.35				1
DVU1988	R	Hypothetical protein DVU1988	2068188	T	A	Gene	0.25			0.45				0.80	0.34		0.55		5
DVU2023	S	Hypothetical protein DVU2023	2104739	G	C	Gene	0.35	0.64	0.35	0.28	0.84	0.76	0.74	0.15	0.29	0.99	0.44	0.60	12
DVU2063	L	Hypothetical protein DVU2063	2142086	C	G	Gene	0.18			0.55				0.74	0.34		0.57		5
DVU2078	NT	Protein-glutamate methylesterase CheB	2162235	C	A	Gene					0.17						0.49		2
DVU2078	NT	Protein-glutamate methylesterase CheB	2163270	T	C	Gene	0.18	0.13		0.52				0.73	0.26				5
DVU2287	C	Hydrogenase, CooK subunit, selenocysteine-containing, putative	2381876	T	G	Gene	0.24	0.28	0.64	0.13		0.17	0.33		0.41			0.39	8
DVU2287	C	Hydrogenase, CooK subunit, selenocysteine-containing, putative	2381877	G	C	Gene	0.76	0.72	0.35	0.87	0.92	0.83	0.68	0.94	0.59	0.99	0.89	0.62	12
DVU2379	R	M16 family peptidase putative	2477501	G	T	Gene					0.15								1
DVU2498	S	Hypothetical protein DVU2498	2609413	G	A	Gene												0.14	1
DVU2505	D	Cell cycle protein FtsW	2616993	A	T	Gene							0.27						1
DVU2557	H	BirA bifunctional protein	2669556	A	C	Gene	0.27	0.14		0.53				0.72	0.48		0.64	0.16	7
DVU2617	P	Sodium/calcium exchanger family protein	2738682	C	G	Gene									0.16				1
DVU2664	P	Phosphate ABC transporter, ATP-binding protein, putative	2775672	C	G	Gene	0.45	0.64	0.37	0.29	0.86	0.78	0.79	0.15	0.31	0.99	0.32	0.73	12
DVU2802	K	GntR family transcriptional regulator	2905021	T	G	Gene				0.48								0.11	2
DVU2802	K	GntR family transcriptional regulator	2905143	T	G	Gene	0.17												1
DVU2802	K	GntR family transcriptional regulator	2905203	G	A	Gene	0.24	0.39	0.66			0.18	0.30		0.39			0.28	7
DVU2802	K	GntR family transcriptional regulator	2905275	G	C	Gene								0.40	0.45		0.29		3
DVU2959	S	Hypothetical protein DVU2959	3066925	CACACA	GCAC	Intergenic		1.00						1.00					2
DVU3022	T	Sensory box histidine kinase/response regulator	3140578	A	T	Gene	1.00	1.00	1.00	1.00	1.00			1.00	1.00		1.00	1.00	9
DVU3022	T	Sensory box histidine kinase/response regulator	3140578	A	G	Gene		1.00	1.00	1.00	1.00			1.00	1.00			1.00	7
DVU3022	T	Sensory box histidine kinase/response regulator	3140597	T	TAT	Gene	0.83				0.86								2
DVU3045	T	Sensory box histidine kinase/response regulator	3169435	G	C	Gene	0.29	0.24	0.62		0.12	0.23	0.27		0.35			0.26	8
DVU3106	T	GGDEF domain-containing protein	3254670	G	C	Intergenic								0.25					1
DVU3129	S	Hypothetical protein DVU3129	3276860	G	C	Gene					0.17						0.13		2
intergenic mutations could not be assigned to a gene		Sits between DVU0645 (Methyl-accepting chemotaxis protein) and DVU0646 (Precorrin-2 C20-methyltransferase) that face opposite ways	716874	T	A	Intergenic	0.42	0.31	0.18	0.38	0.61	0.68	0.72	0.13	0.25		0.32	0.36	11
		+ 216 bp of DVU0978（ABC transporter, periplasmic substrate-binding protein, putative)	1073326	C	T	Intergenic	0.22			0.63				0.83	0.36		0.56		5
		Sits between DVU1176 (Hypothetical protein DVU1176) and DVU1177 (Hypothetical protein DVU1177) that face opposite ways	1268170	G	C	Intergenic								1.00			0.18		2
		Sits between DVU1999 (Sulfate transporter family protein) and DVU2000 (Hypothetical protein DVU2000) that face opposite ways	2082600	T	A	Intergenic				0.19				0.20					2
		Sits between DVU2112 (Hypothetical protein DVU2112) and DVU2113 (Xanthine/uracil permease family protein) that face opposite ways	2207959	G	C	Intergenic		0.28	0.52			0.24						0.29	4
		+ 198 bp of DVU2397（Hypothetical protein DVU2397)	2502193	G	C	Intergenic	0.43	0.65	0.30	0.40	0.86	0.84	0.70	0.16	0.22	1.00	0.31	0.56	12
		Sits between DVU2683 (L-lactate permease family protein) and DVU2684 (Hypothetical protein DVU2684) that face opposite ways	2798285	C	G	Intergenic		0.22											1
		+ 141 bp of DVU2885（Alcohol dehydrogenase, iron-containing)	2982786	C	CTG	Intergenic						0.33							1
		+ 138 bp of DVU2885（Alcohol dehydrogenase, iron-containing)	2982789	C	CCA	Intergenic												1.00	1
No. of mutations in each population							**24**	**24**	**18**	**22**	**16**	**15**	**18**	**22**	**25**	**8**	**22**	**26**	

Table S3 Mutations identified in 12 EN populations (MP*: the number of populations with mutations)

Gene	COG	Predicted function	Position	Reference	Alter	Type	EN1	EN2	EN3	EN4	EN5	EN6	EN7	EN8	EN9	EN10	EN11	EN12	MP*
DVU0126	CP	ABC transporter, ATP-binding protein	162600	C	G	Gene				1.00	0.82								2
DVU0225	S	Hypothetical protein DVU0225	262421	C	T	Gene						0.43							1
DVU0246	S	Pyruvate phosphate dikinase PEP/pyruvate binding subunit	279123	CG	C	Gene		0.15											1
DVU0246	S	Pyruvate phosphate dikinase PEP/pyruvate binding subunit	279507	C	A	Gene									0.59				1
DVU0246	S	Pyruvate phosphate dikinase PEP/pyruvate binding subunit	279702	GT	G	Gene											0.47		1
DVU0246	S	Pyruvate phosphate dikinase PEP/pyruvate binding subunit	279997	AGGGGGGG	AGGGGGG	Gene	0.77	0.53											2
DVU0246	S	Pyruvate phosphate dikinase PEP/pyruvate binding subunit	280373	GTTGCCGT	C	Gene								0.21					1
DVU0247	T	Response regulator	281155	AT	A	Gene						0.34							1
DVU0247	T	Response regulator	281286	TACGA	TACGAACGA	Gene						0.50							1
DVU0248	T	Signal transduction histidine kinase	281561	AGGGG	AGG	Gene			0.14										1
DVU0248	T	Signal transduction histidine kinase	281703	GTT	G	Gene					0.18								1
DVU0248	T	Signal transduction histidine kinase	282788	C	T	Gene								0.65					1
DVU0248	T	Signal transduction histidine kinase	283136	C	T	Gene										0.18			1
DVU0248	T	Signal transduction histidine kinase	283154	C	T	Gene				0.63									1
DVU0248	T	Signal transduction histidine kinase	283204	G	T	Gene												0.47	1
DVU0249	P	PtxB, putative	283782	C	T	Gene			0.17										1
DVU0251	R	Transmembrane protein TauE like	285479	G	A	Gene			0.13										1
DVU0251	R	Transmembrane protein TauE like	285947	C	G	Gene			0.19										1
DVU0251	R	Transmembrane protein TauE like	285948	C	T	Gene							0.34					0.23	2
DVU0281	S	Exopolysaccharide biosynthesis protein, putative	325822	GT	G	Gene												1.00	1
DVU0281	S	Exopolysaccharide biosynthesis protein, putative	326403	G	A	Gene	1.00		1.00	1.00	0.76			1.00					5
DVU0408	M	EAL domain-containing protein	456487	A	G	Intergenic					0.19								1
DVU0467	E	Anthranilate phosphoribosyltransferase	535249	G	C	Gene	1.00		1.00	1.00	0.89			1.00					5
DVU0564	S	ISDvu4, transposase, truncation	629931	C	A	Intergenic							0.48					0.22	2
DVU0564	S	ISDvu4, transposase, truncation	629934	GA	GAGGGCTAC	Intergenic		0.21					0.73						2
DVU0564	S	ISDvu4, transposase, truncation	629934	G	GCTAC	Intergenic									0.29				1
DVU0596	KT	DNA-binding response regulator LytR	664509	C	T	Gene	0.11												1
DVU0597	T	regulatory protein LytS	665366	C	T	Intergenic			0.12										1
DVU0597	T	Regulatory protein LytS	665986	C	T	Gene						0.42							1
DVU0597	T	Regulatory protein LytS	666383	C	T	Gene				0.11			0.12						2
DVU0597	T	Regulatory protein LytS	666479	A	C	Gene											0.25		1
DVU0796	E	Histidinol dehydrogenase	882512	C	G	Intergenic	0.80	0.44		1.00	0.29	0.38		0.40	1.00		0.40		8
DVU0797	S	Hypothetical protein DVU0797	883424	C	T	Gene								0.38					1
DVU0799	S	Hypothetical protein DVU0799	885159	G	GTCGGAACCGTTGAGGTTCTTGCCG	Gene								0.55					1
DVU799	S	Hypothetical protein DVU799	885499	G	A	Gene								0.48					1
DVU0847	C	Adenylylsulfate reductase	934997	C	T	Gene							0.99						1
DVU0942	P	FUR family transcriptional regulator	1034641	A	C	Gene										1.00			1
DVU0942	P	FUR family transcriptional regulator	1034711	A	T	Gene				1.00	0.88				1.00				3
DVU0942	P	FUR family transcriptional regulator	1034799	A	T	Gene											1.00		1
DVU0942	P	FUR family transcriptional regulator	1034881	G	A	Gene							1.00						1
DVU2571	P	Ferrous iron transport protein B	2685074	A	C	Gene												1.00	1
DVU2571	P	Ferrous iron transport protein B	2685298	T	C	Gene			1.00										1
DVU2571	P	Ferrous iron transport protein B	2685402	CTCAAGGGCGGCCTCGGCCCGTTCGGCATCAAGGGCGGC	CTCAAGGGCGGC	Gene		1.00											1
DVU2571	P	Ferrous iron transport protein B	2685709	G	A	Gene	1.00												1
DVU2571	P	Ferrous iron transport protein B	2685734	G	A	Gene								1.00					1
DVU2571	P	Ferrous iron transport protein B	2685977	A	T	Gene						0.99							1
DVU1087	S	Hypothetical protein DVU1087	1191167	C	T	Gene			0.40		0.50				0.75		0.57	0.33	5
DVU1204	IQ	3-oxoacyl-(acyl-carrier-protein) synthase II	1295686	C	T	Gene						1.00							1
DVU1204	IQ	3-oxoacyl-(acyl-carrier-protein) synthase II	1295951	G	C	Gene	1.00												1
DVU1204	IQ	3-oxoacyl-(acyl-carrier-protein) synthase II	1296057	C	G	Gene				1.00	0.92								2
DVU1204	IQ	3-oxoacyl-(acyl-carrier-protein) synthase II	1296081	T	C	Gene									0.99				1
DVU1204	IQ	3-oxoacyl-(acyl-carrier-protein) synthase II	1296268	C	T	Gene		1.00											1
DVU1204	IQ	3-oxoacyl-(acyl-carrier-protein) synthase II	1296277	C	T	Gene											1.00		1
DVU1204	IQ	3-oxoacyl-(acyl-carrier-protein) synthase II	1296453	T	A	Gene			0.96										1
DVU1208	I	Putative glycerol-3-phosphate acyltransferase PlsX	1299241	T	G	Gene								0.43		1.00			2
DVU1260	I	Outer membrane protein P1, putative	1346776	TGG	TG	Gene				0.16									1
DVU1260	I	Outer membrane protein P1, putative	1347079	CTATAATGTCGGCATCCA	C	Gene				0.62									1
DVU1260	I	Outer membrane protein P1, putative	1347740	CT	C	Gene	0.73												1
DVU1272	NU	General secretion pathway protein E, putative	1363430	C	A	Gene											1.00		1
DVU1349	H	Geranylgeranyl diphosphate synthase	1426830	T	G	Gene		1.00			0.15	1.00	1.00		1.00	1.00		1.00	7
DVU1371	R	HAD family hydrolase	1449911	C	G	Gene								1.00					1
DVU1408	S	Hypothetical protein DVU1408	1477702	A	G	Gene	0.18												1
DVU1469	J	30S ribosomal protein S1	1551144	C	T	Gene								0.56					1
DVU1469	J	30S ribosomal protein S1	1551147	C	T	Gene						0.98						1.00	2
DVU1530	J	Metallo-beta-lactamase family protein	1599469	C	T	Gene	1.00		0.99	1.00	0.83			1.00					5
DVU1545	S	Hemolysin-type calcium-binding repeat/calx-beta domain-containing protein	1620196	C	T	Gene						0.52							1
DVU1576	I	4-diphosphocytidyl-2C-methyl-D-erythritol kinase	1658379	CGTATG	CG	Gene											0.94		1
DVU1588	F	Hypoxanthine phosphoribosyltransferase	1671284	CGGGGGG	CGGGGGG	Gene						0.59							1
DVU1634	S	Hypothetical protein DVU1634	1717112	C	T	Gene					0.13								1
DVU1698	S	Hypothetical protein DVU1698	1773345	A	G	Gene	1.00				1.00		1.00	1.00	1.00		0.75	1.00	7
DVU1785	U	MarC membrane protein	1848460	A	C	Gene						0.97							1
DVU1785	U	MarC membrane protein	1848475	G	A	Gene											0.94		1
DVU1862	T	GGDEF domain-containing protein	1929597	CA	C	Gene					0.29								1
DVU1862	T	GGDEF domain-containing protein	1930141	C	T	Gene			0.57	0.16									2
DVU1862	T	GGDEF domain-containing protein	1930239	CA	C	Gene								0.42					1
DVU1862	T	GGDEF domain-containing protein	1930258	A	T	Gene			0.31										1
DVU1988	R	Hypothetical protein DVU1988	2068188	T	A	Gene											1.00		1
DVU2023	S	Hypothetical protein DVU2023	2104739	G	C	Gene		1.00			0.24	0.99	1.00		1.00	1.00		1.00	7
DVU2063	L	Hypothetical protein DVU2063	2142086	C	G	Gene											1.00		1
DVU2078	NT	Protein-glutamate methylesterase CheB	2162235	C	A	Gene									1.00				1
DVU2078	NT	Protein-glutamate methylesterase CheB	2163270	T	C	Gene											1.00		1
DVU2284	S	Hypothetical protein DVU2284	2374803	A	C	Gene						0.20							1
DVU2287	C	Hydrogenase, CooK subunit, selenocysteine-containing, putative	2381876	T	G	Gene	1.00		1.00	1.00	0.82			0.99					5
DVU2287	C	Hydrogenase, CooK subunit, selenocysteine-containing, putative	2381877	G	C	Gene		1.00			0.18	1.00	1.00		1.00	1.00	1.00	1.00	8
DVU2394	T	Sigma-54 dependent transcriptional regulator/response regulator	2496979	T	C	Gene					0.16								1
DVU2394	T	Sigma-54 dependent transcriptional regulator/response regulator	2497229	TG	T	Gene			0.16										1
DVU2394	T	Sigma-54 dependent transcriptional regulator/response regulator	2497258	AGCACGTTGCGCAATTCACGCACGTTGC	AGCACGTTGC	Gene							1.00						1
DVU2394	T	Sigma-54 dependent transcriptional regulator/response regulator	2497411	AGGGGGGG	AGGGGGGGG	Gene						0.49							1
DVU2394	T	Sigma-54 dependent transcriptional regulator/response regulator	2497708	TGGCCGAACAGTTCGCTACGGGCC	TGGCC	Gene									0.36				1

Table S3 (Continued)

DVU2394	T	Sigma-54 dependent transcriptional regulator/response regulator	2497728	GG	GACCGCG GACATGG	Gene								0.11				1
DVU2394	T	Sigma-54 dependent transcriptional regulator/response regulator	2497736	C	T	Gene			0.59									1
DVU2395	T	Sensor histidine kinase	2498363	T	C	Gene									0.16			1
DVU2395	T	Sensor histidine kinase	2498544	TGCG	TGCGCG	Gene							0.46					1
DVU2395	T	Sensor histidine kinase	2498737	T	C	Gene									0.23			1
DVU2395	T	Sensor histidine kinase	2498935	G	A	Gene									0.45			1
DVU2395	T	Sensor histidine kinase	2498936	TGA	CTCATGCG	Gene							0.26					1
DVU2395	T	Sensor histidine kinase	2499096	GCGGCGAG	G	Gene				1.00								1
DVU2395	T	Sensor histidine kinase	2499244	G	A	Gene											1.00	1
DVU2395	T	Sensor histidine kinase	2499383	CTTT	CTT	Gene								0.54				1
DVU2396	C	Alcohol dehydrogenase, iron-containing	2499736	TC	T	Gene						0.42						1
DVU2396	C	Alcohol dehydrogenase, iron-containing	2499926	G	A	Gene	0.20											1
DVU2396	C	Alcohol dehydrogenase, iron-containing	2499983	GCCC	GCC	Gene		0.98										1
DVU2396	C	Alcohol dehydrogenase, iron-containing	2500012	G	T	Gene	0.72											1
DVU2402	C	Heterodisulfide reductase, A subunit	2507986	G	A	Gene		0.96										1
DVU2405	C	Alcohol dehydrogenase, iron-containing	2510694	TCC	TC	Gene										1.00		1
DVU2405	C	Alcohol dehydrogenase, iron-containing	2510695	C	A	Gene					0.86							1
DVU2405	C	Alcohol dehydrogenase, iron-containing	2510695	C	T	Gene										1.00		1
DVU2406	S	Hypothetical protein DVU2406	2511389	AGC	A	Gene							0.29					1
DVU2411	S	EF hand domain-containing protein	2515114	GCCCCCCC	GCCCCCCC	Gene						0.42						1
DVU2505	D	Cell cycle protein FtsW	2616993	A	T	Gene							1.00					1
DVU2543	C	Hydroxylamine reductase	2654808	G	A	Gene				0.99	0.79							2
DVU2547	T	Transcriptional regulator, putative	2660114	C	A	Gene	0.74											1
DVU2547	T	Transcriptional regulator, putative	2660166	CCCGCCCGC	CCCCGCC	Gene	0.13											1
DVU2547	T	Transcriptional regulator, putative	2660235	G	A	Gene			0.66									1
DVU2547	T	Transcriptional regulator, putative	2660273	G	A	Gene						0.99						1
DVU2547	T	Transcriptional regulator, putative	2660388	G	A	Gene	0.17											1
DVU2547	T	Transcriptional regulator, putative	2660403	G	A	Gene							1.00	0.99				2
DVU2547	T	Transcriptional regulator, putative	2660404	G	T	Gene			0.36									1
DVU2547	T	Transcriptional regulator, putative	2660450	ATGATGG CCTCGCGC C	A	Gene									0.72			1
DVU2547	T	Transcriptional regulator, putative	2660477	G	A	Gene								0.94				1
DVU2548	I	Acyl carrier protein phosphodiesterase	2661113	GA	G	Gene		0.93										1
DVU2548	I	Acyl carrier protein phosphodiesterase	2661114	A	G	Gene		0.46										1
DVU2548	I	Acyl carrier protein phosphodiesterase	2661182	CGG	CG	Gene											1.00	1
DVU2548	I	Acyl carrier protein phosphodiesterase	2661301	A	ACACT	Intergenic										0.71		1
DVU2548	I	Acyl carrier protein phosphodiesterase	2661319	GGT	GCCTTGAT ATACAGT	Intergenic										0.62		1
DVU2557	H	BirA bifunctional protein	2669556	A	C	Gene										1.00		1
DVU2664	P	Phosphate ABC transporter, ATP-binding protein, putative	2775672	C	G	Gene		1.00			0.22	1.00	1.00		1.00	1.00	1.00	7
DVU2747	S	Hypothetical protein DVU2747	2853618	G	A	Gene		0.57										1
DVU2802	K	GntR family transcriptional regulator	2905203	G	A	Gene	1.00		1.00	1.00	0.83			1.00				5
DVU2802	K	GntR family transcriptional regulator	2905275	G	C	Gene										1.00		1
DVU2894	T	Sigma-54 dependent transcriptional regulator	2991001	A	G	Gene	0.14											1
DVU2950	S	Sensory box protein/GGDEF domain protein, authentic frameshift (TIGR)	3056210	GGA	GGAGA	Gene		1.00										1
DVU2959	S	Hypothetical protein DVU2959	3066925	CACACAC	GCAC	Intergenic	1.00	1.00						1.00	1.00		1.00	5
DVU3022	T	Sensory box histidine kinase/response regulator	3140578	A	T	Gene	1.00	1.00	1.00		1.00	1.00	1.00	1.00	1.00		1.00	10
DVU3022	T	Sensory box histidine kinase/response regulator	3140578	A	G	Gene	1.00	1.00	1.00					1.00	1.00		1.00	6
DVU3022	T	Sensory box histidine kinase/response regulator	3140595	GAT	GATAT	Gene									1.00			2
DVU3022	T	Sensory box histidine kinase/response regulator	3140597	T	TAT	Gene			1.00	1.00						1.00		3
DVU3045	T	Sensory box histidine kinase/response regulator	3169435	G	C	Gene	1.00		1.00	1.00	0.78			0.99				5
DVU3129	S	Hypothetical protein DVU3129	3276860	G	C	Gene										1.00		1
DVU3190	S	Hypothetical protein DVU3190	3350162	G	A	Gene										0.96		1
DVU3264	C	Fumarate hydratase	3440138	C	T	Gene					0.19							1
DVU3304	S	Hypothetical protein DVU3282	3481721	G	A	Gene		0.14										1
DVU3304	S	Hypothetical protein DVU3283	3482319	G	A	Gene	0.23											1
U_tRNA-A	E	tRNA-Arg	3484653	C	T	Intergenic					0.20							1
intergenic mutations could not be assigned to a gene		+ 110 bp of DVU0251（Transmembrane protein TauE like)	286439	C	T	Intergenic			0.13									1
		+ 309 bp of DVU0353（Alcohol dehydrogenase, iron-containing)	402212	A	G	Intergenic					0.29							1
		+ 503 bp of DVU0360（Acetolactate synthase catalytic subunit)	408888	G	C	Intergenic			0.14		0.16							2
		Sits between DVU0645 (Methyl-accepting chemotaxis protein) and DVU0646 (Precorrin-2 C20-methyltransferase) that face opposite ways	716874	T	A	Intergenic		1.00				1.00	1.00		1.00		1.00	5
		+ 216 bp of DVU0978（ABC transporter, periplasmic substrate-binding protein, putative)	1073326	C	T	Intergenic										1.00		1
		+ 133 bp of DVU1562（HAMP domain-containing protein)	1644297	C	T	Intergenic		0.90										1
		Sits between DVU1999 (Sulfate transporter family protein) and DVU2000 (Hypothetical protein DVU2000) that face opposite ways	2082582	A	G	Intergenic										0.30		1
		Sits between DVU1999 (Sulfate transporter family protein) and DVU2000 (Hypothetical protein DVU2000) that face opposite ways	2082589	G	A	Intergenic										0.38		1
		Sits between DVU1999 (Sulfate transporter family protein) and DVU2000 (Hypothetical protein DVU2000) that face opposite ways	2082600	T	A	Intergenic										0.53		1
		Sits between DVU1999 (Sulfate transporter family protein) and DVU2000 (Hypothetical protein DVU2000) that face opposite ways	2082602	TA	T	Intergenic										0.46		1
		Sits between DVU2112 (Hypothetical protein DVU2112) and DVU2113 (Xanthine/uracil permease family protein) that face opposite ways	2207959	G	C	Intergenic			1.00	1.00	0.78							3
		+ 198 bp of DVU2397（Hypothetical protein DVU2397)	2502193	G	C	Intergenic		1.00			0.17	1.00	1.00		1.00	1.00	1.00	7
		+ 138 bp of DVU2885（Alcohol dehydrogenase, iron-containing)	2982789	C	CCA	Intergenic									1.00			1
		+ 144 bp of DVU3047（Aminotransferase, class IV)	3174094	CGGGGGG	CGGGGGG	Intergenic						0.54						1
Gene		**Description**	**Position**		**Junction**													
DVU0246-0253		Functional loss of DVU0246-0253	279540		△9920 bp						0.28							
DVU0250-0255		Functional loss of DVU0250-0255	284766		△5814 bp						0.27							
DVU0246		Frameshift mutation	279074		△234 bp								0.48					
DVU0248		Frameshift mutation	283167		△63 bp									0.14				
DVU0246-0256		Functional loss of DVU0246-0256	279032		△15883 bp										0.33			
DVU0240-0270		Functional loss of DVU0240-0270	272887		△38514 bp												0.32	
DVU1862		Frameshift mutation	1930123		△146 bp				0.68									
DVU1862		86 amino acid loss	1929985		△258 bp						0.25							
DVU2548		Frameshift mutation	2661089		△57 bp											0.21		
No. of mutations in each populations							24	23	27	19	33	23	18	26	24	13	31	20	

Mutations in green background are identified using breseq.

Table S4 Newly acquired mutations in 12 EN populations (MP*: the number of populations with mutations)

Gene	COG	Predicted function	Position	Reference	Alter	Type	EN1	EN2	EN3	EN4	EN5	EN6	EN7	EN8	EN9	EN10	EN11	EN12	MP*
DVU0126	CP	ABC transporter, ATP-binding protein	162600	C	G	Gene				1.00	0.82								2
DVU0225	S	Hypothetical protein DVU0225	262421	C	T	Gene						0.43							1
DVU0246	S	Pyruvate phosphate dikinase PEP/pyruvate binding subunit	279123	CG	C	Gene		0.15											1
DVU0246	S	Pyruvate phosphate dikinase PEP/pyruvate binding subunit	279507	C	A	Gene									0.59				1
DVU0246	S	Pyruvate phosphate dikinase PEP/pyruvate binding subunit	279702	GT	G	Gene											0.47		1
DVU0246	S	Pyruvate phosphate dikinase PEP/pyruvate binding subunit	279997	AGGGGGG	AGGGGGG	Gene	0.77	0.53											2
DVU0246	S	Pyruvate phosphate dikinase PEP/pyruvate binding subunit	280373	CTGTTGCCGTCG	C	Gene								0.21					1
DVU0247	T	Response regulator	281155	AT	A	Gene						0.34							1
DVU0247	T	Response regulator	281286	TACGA	TACGAACGA	Gene						0.50							1
DVU0248	T	Signal transduction histidine kinase	281561	AGGGG	AGG	Gene			0.14										1
DVU0248	T	Signal transduction histidine kinase	281703	GTT	G	Gene					0.18								1
DVU0248	T	Signal transduction histidine kinase	282788	C	T	Gene								0.65					1
DVU0248	T	Signal transduction histidine kinase	283136	C	T	Gene										0.18			1
DVU0248	T	Signal transduction histidine kinase	283154	C	T	Gene				0.63									1
DVU0248	T	Signal transduction histidine kinase	283204	G	T	Gene												0.47	1
DVU0249	P	PtxB, putative	283782	C	T	Gene			0.17										1
DVU0251	R	Transmembrane protein TauE like	285479	G	A	Gene			0.13										1
DVU0251	R	Transmembrane protein TauE like	285947	C	G	Gene			0.19										1
DVU0251	R	Transmembrane protein TauE like	285948	C	T	Gene							0.34					0.23	2
DVU0281	S	Exopolysaccharide biosynthesis protein, putative	325822	GT	G	Gene												1.00	1
DVU0596	KT	DNA-binding response regulator LytR	664509	C	T	Gene	0.11												1
DVU0597	T	regulatory protein LytS	665366	C	T	Intergenic			0.12[a]										1
DVU0597	T	Regulatory protein LytS	665986	C	T	Gene						0.42							1
DVU0597	T	Regulatory protein LytS	666383	C	T	Gene				0.11			0.12						2
DVU0597	T	Regulatory protein LytS	666479	A	C	Gene											0.25		1
DVU0797	S	Hypothetical protein DVU0797	883424	C	T	Gene								0.38					1
DVU0799	S	Hypothetical protein DVU0799	885158	CCG	CCGTCGGAACCGT TGAGGTTCTTGCCG	Gene								0.55					1
DVU0847	C	Adenylylsulfate reductase	934997	C	T	Gene							0.99						1
DVU1204	IQ	3-oxoacyl-(acyl-carrier-protein) synthase II	1295686	C	T	Gene							1.00						1
DVU1204	IQ	3-oxoacyl-(acyl-carrier-protein) synthase II	1296057	C	G	Gene				1.00	0.92								2
DVU1204	IQ	3-oxoacyl-(acyl-carrier-protein) synthase II	1296081	T	C	Gene									0.99				1
DVU1204	IQ	3-oxoacyl-(acyl-carrier-protein) synthase II	1296268	C	T	Gene		1.00											1
DVU1204	IQ	3-oxoacyl-(acyl-carrier-protein) synthase II	1296277	C	T	Gene											1.00		1
DVU1204	IQ	3-oxoacyl-(acyl-carrier-protein) synthase II	1296453	T	A	Gene			0.96										1
DVU1208	I	Putative glycerol-3-phosphate acyltransferase PlsX	1299241	T	G	Gene								0.43		1.00			2
DVU1260	I	Outer membrane protein P1, putative	1346776	TGG	TG	Gene				0.16									1
DVU1260	I	Outer membrane protein P1, putative	1347079	TATAATGTCGGCATCC	C	Gene				0.62									1
DVU1260	I	Outer membrane protein P1, putative	1347740	CT	C	Gene	0.73												1
DVU1408	S	Hypothetical protein DVU1408	1477702	A	G	Gene	0.18												1
DVU1469	J	30S ribosomal protein S1	1551144	C	T	Gene								0.56					1
DVU1469	J	30S ribosomal protein S1	1551147	C	T	Gene						0.98						1.00	2
DVU1545	S	Hemolysin-type calcium-binding repeat	1620196	C	T	Gene						0.52							1
DVU1576	I	4-diphosphocytidyl-2C-methyl-D-erythritol kinase	1658379	CGTATG	CG	Gene											0.94		1
DVU1588	F	Hypoxanthine phosphoribosyltransferase	1671284	CGGGGGG	CGGGGGG	Gene						0.59							1
DVU1634	S	Hypothetical protein DVU1634	1717112	C	T	Gene					0.13								1
DVU1785	U	MarC membrane protein	1848460	A	C	Gene						0.97							1
DVU1785	U	MarC membrane protein	1848475	G	A	Gene											0.94		1
DVU1862	T	GGDEF domain-containing protein	1929597	CA	C	Gene					0.29								1
DVU1862	T	GGDEF domain-containing protein	1930141	C	T	Gene			0.57	0.16									2
DVU1862	T	GGDEF domain-containing protein	1930239	CA	C	Gene								0.42					1
DVU1862	T	GGDEF domain-containing protein	1930258	A	T	Gene			0.31										1
DVU2284	S	Hypothetical protein DVU2284	2374803	A	C	Gene						0.20							1
DVU2394	T	Sigma-54 dependent transcriptional regulator	2496979	T	C	Gene					0.16								1
DVU2394	T	Sigma-54 dependent transcriptional regulator	2497229	TG	T	Gene			0.16										1
DVU2394	T	Sigma-54 dependent transcriptional regulator	2497258	AGCACGTTGCGCAA TTCACGCACGTTGC	AGCACGTTGC	Gene							1.00						1
DVU2394	T	Sigma-54 dependent transcriptional regulator	2497411	AGGGGGGG	AGGGGGGGG	Gene						0.49							1
DVU2394	T	Sigma-54 dependent transcriptional regulator	2497708	TGGCCGAACAGT TCGCTACGGGCC	TGGCC	Gene									0.36				1
DVU2394	T	Sigma-54 dependent transcriptional regulator	2497728	GG	GACCGCGGACATGG	Gene									0.11				1
DVU2394	T	Sigma-54 dependent transcriptional regulator	2497736	C	T	Gene			0.59										1
DVU2395	T	Sensor histidine kinase	2498363	T	C	Gene										0.16			1
DVU2395	T	Sensor histidine kinase	2498544	TGCG	TGCGCG	Gene								0.46					1
DVU2395	T	Sensor histidine kinase	2498737	T	C	Gene										0.23			1
DVU2395	T	Sensor histidine kinase	2498935	G	A	Gene										0.45			1
DVU2395	T	Sensor histidine kinase	2498936	TGA	TGACCTCATGCGCCACC	Gene								0.26					1
DVU2395	T	Sensor histidine kinase	2499096	GCGGCGAGC	G	Gene				1.00									1
DVU2395	T	Sensor histidine kinase	2499244	G	A	Gene												1.00	1
DVU2395	T	Sensor histidine kinase	2499383	CTTT	CTT	Gene									0.54				1
DVU2396	C	Alcohol dehydrogenase, iron-containing	2499736	TC	T	Gene						0.42							1
DVU2396	C	Alcohol dehydrogenase, iron-containing	2499926	G	A	Gene	0.20												1
DVU2396	C	Alcohol dehydrogenase, iron-containing	2499983	GCCC	GCC	Gene		0.98											1
DVU2396	C	Alcohol dehydrogenase, iron-containing	2500012	G	T	Gene	0.72												1
DVU2402	C	Heterodisulfide reductase, A subunit	2507986	G	A	Gene		0.96											1
DVU2405	C	Alcohol dehydrogenase, iron-containing	2510694	TCC	TC	Gene											1.00		1
DVU2405	C	Alcohol dehydrogenase, iron-containing	2510695	C	A	Gene					0.86								1
DVU2405	C	Alcohol dehydrogenase, iron-containing	2510695	C	T	Gene											1.00		1
DVU2406	S	Hypothetical protein DVU2406	2511389	AGC	A	Gene								0.29					1
DVU2411	S	EF hand domain-containing protein	2515114	GCCCCCCC	GCCCCCCCC	Gene						0.42							1
DVU2543	C	Hydroxylamine reductase	2654808	G	A	Gene				0.99	0.79								2
DVU2547	T	Transcriptional regulator, putative	2660114	C	A	Gene	0.74												1
DVU2547	T	Transcriptional regulator, putative	2660166	CCCCGCCCGCC	CCCCGCC	Gene	0.13												1
DVU2547	T	Transcriptional regulator, putative	2660235	G	A	Gene			0.66										1
DVU2547	T	Transcriptional regulator, putative	2660273	G	A	Gene						0.99							1
DVU2547	T	Transcriptional regulator, putative	2660388	G	A	Gene	0.17												1
DVU2547	T	Transcriptional regulator, putative	2660403	G	A	Gene							1.00	0.99					2
DVU2547	T	Transcriptional regulator, putative	2660404	G	T	Gene			0.36										1
DVU2547	T	Transcriptional regulator, putative	2660450	ATGATGGCCTCGCGCC	A	Gene										0.72			1
DVU2547	T	Transcriptional regulator, putative	2660477	G	A	Gene									0.94				1
DVU2548	I	Acyl carrier protein phosphodiesterase	2661113	GA	G	Gene		0.93											1
DVU2548	I	Acyl carrier protein phosphodiesterase	2661114	A	G	Gene		0.46											1
DVU2548	I	Acyl carrier protein phosphodiesterase	2661182	CGG	CG	Gene												1.00	1
DVU2548	I	Acyl carrier protein phosphodiesterase	2661301	A	ACACT	Intergenic											0.71[a]		1
DVU2548	I	Acyl carrier protein phosphodiesterase	2661319	GGT	GCCTTGATATACAGT	Intergenic											0.62[a]		1

Table S4 (Continued)

DVU0942	P	FUR family transcriptional regulator	1034641	A	C	Gene									1.00			1
DVU0942	P	FUR family transcriptional regulator	1034711	A	T	Gene				1.00	0.88				1.00			3
DVU0942	P	FUR family transcriptional regulator	1034799	A	T	Gene										1.00		1
DVU0942	P	FUR family transcriptional regulator	1034881	G	A	Gene							1.00					1
DVU2571	P	Ferrous iron transport protein B	2685074	A	C	Gene											1.00	1
DVU2571	P	Ferrous iron transport protein B	2685298	T	C	Gene			1.00									1
DVU2571	P	Ferrous iron transport protein B	2685402	CTCAAGGGCGGCCT CGGCCCGTTCGGC ATCAAGGGCGGC	CTCAAGGGCGGC	Gene		1.00										1
DVU2571	P	Ferrous iron transport protein B	2685709	G	A	Gene	1.00											1
DVU2571	P	Ferrous iron transport protein B	2685734	G	A	Gene								1.00				1
DVU2571	P	Ferrous iron transport protein B	2685977	A	T	Gene						0.99						1
DVU2747	S	Hypothetical protein DVU2747	2853618	G	A	Gene		0.57										1
DVU2894	T	Sigma-54 dependent transcriptional regulator	2991001	A	G	Gene	0.14											1
DVU2950	S	Sensory box protein/GGDEF domain protein	3056210	GGA	GGAGA	Gene		1.00										1
DVU3022	T	Sensory box histidine kinase/response regulator	3140595	GAT	GATAT	Gene									1.00		1.00	2
DVU3190	S	Hypothetical protein DVU3190	3350162	G	A	Gene											0.96	1
DVU3264	C	Fumarate hydratase	3440138	C	T	Gene					0.19							1
DVU3304	S	Hypothetical protein DVU3282	3481721	G	A	Gene		0.14										1
DVU3304	S	Hypothetical protein DVU3283	3482319	G	A	Gene	0.23											1
Gene		**Description**	**Position**	**Junction**														
DVU0246-0253		Functional loss of DVU0246-0253	279540	△9920 bp						0.28								1
DVU0250-0255		Functional loss of DVU0250-0255	284766	△5814 bp						0.27								1
DVU0246		Frameshift mutation	279074	△234 bp								0.48						1
DVU0248		Frameshift mutation	283167	△63 bp										0.14				1
DVU0246-0256		Functional loss of DVU0246-0256	279032	△15883 bp											0.33			1
DVU0240-0270		Functional loss of DVU0240-0270	272887	△38514 bp													0.32	1
DVU1862		Frameshift mutation	1930123	△146 bp				0.68										1
DVU1862		86 amino acid loss	1929985	△258 bp						0.25								1
DVU2548		Frameshift mutation	2661089	△57 bp												0.21		1
No. of mutations in each populations						12	11	13	10	13	14	8	12	9	8	11	8	

[a] intergenic mutation identified within 100 bp of the upstream of start codon of the gene
Mutations in green background are identified using breseq.

Table S5 Mutated functional genes with newly acquired mutations in EN populations (MP*: the number of populations with mutations)

Gene	COG	Predicted function	MP*	EN1	EN2	EN3	EN4	EN5	EN6	EN7	EN8	EN9	EN10	EN11	EN12
DVU0126	CP	ABC transporter, ATP-binding protein	2				1.00	0.82							
DVU0225	S	Hypothetical protein DVU0225	1						0.43						
DVU0246	S	Pyruvate phosphate dikinase PEP/pyruvate binding subunit	5	0.77	0.53					0.48	0.21	0.59		0.47	
DVU0247	T	Response regulator	1						0.50						
DVU0248	T	Signal transduction histidine kinase	6			0.14	0.63	0.18			0.65	0.14	0.18		0.47
DVU0249	P	PtxB, putative	1			0.17									
DVU0251	R	Transmembrane protein TauE like	3			0.19				0.34					0.23
△DVU0246-0253 △DVU0246-0256 △DVU0240-0270		Functional loss of DVU0246-0251	3					0.28					0.33		0.32
DVU0281	S	Exopolysaccharide biosynthesis protein, putative	1												1.00
DVU0596	KT	DNA-binding response regulator LytR	1	0.11											
DVU0597	T	regulatory protein LytS	5			0.12[a]	0.11		0.42	0.12				0.25	
DVU0797	S	Hypothetical protein DVU0797	1								0.38				
DVU0799	S	Hypothetical protein DVU0799	1								0.55				
DVU0847	C	Adenylylsulfate reductase	1							0.99					
DVU0942	P	FUR family transcriptional regulator	6				1.00	0.88		1.00		1.00	1.00	1.00	
DVU2571	P	Ferrous iron transport protein B	6	1.00	1.00	1.00			0.99		1.00				1.00
DVU1204	IQ	3-oxoacyl-(acyl-carrier-protein) synthase II	7		1.00	0.96	1.00	0.92		1.00		0.99		1.00	
DVU1208	I	Putative glycerol-3-phosphate acyltransferase PlsX	2								0.43		1.00		
DVU1260	I	Outer membrane protein P1, putative	2	0.73			0.62								
DVU1408	S	Hypothetical protein DVU1408	1	0.18											
DVU1469	J	30S ribosomal protein S1	3						0.98		0.56				1.00
DVU1545	S	Hemolysin-type calcium-binding repeat	1						0.52						
DVU1576	I	4-diphosphocytidyl-2C-methyl-D-erythritol kinase	1											0.94	
DVU1588	F	Hypoxanthine phosphoribosyltransferase	1						0.59						
DVU1634	S	Hypothetical protein DVU1634	1					0.13							
DVU1785	U	MarC membrane protein	2						0.97					0.94	
DVU1862	T	GGDEF domain-containing protein	4			0.68	0.16	0.29			0.42				
DVU2284	S	Hypothetical protein DVU2284	1						0.20						
DVU2394	T	Sigma-54 dependent transcriptional regulator	5			0.59		0.16	0.49	1.00		0.36			
DVU2395	T	Sensor histidine kinase	5				1.00				0.46	0.54	0.45		1.00
DVU2396	C	Alcohol dehydrogenase, iron-containing	3	0.72	0.98				0.42						
DVU2402	C	Heterodisulfide reductase, A subunit	1		0.96										
DVU2405	C	Alcohol dehydrogenase, iron-containing	2					0.86						1.00	
DVU2406	S	Hypothetical protein DVU2406	1								0.29				
DVU2411	S	EF hand domain-containing protein	1						0.42						
DVU2543	C	Hydroxylamine reductase	2				0.99	0.79							
DVU2547	T	Transcriptional regulator, putative	7	0.74		0.66			0.99	1.00	0.99	0.94	0.72		
DVU2548	I	Acyl carrier protein phosphodiesterase	3		0.93									0.71[a]	1.00
DVU2747	S	Hypothetical protein DVU2747	1		0.57										
DVU2894	T	Sigma-54 dependent transcriptional regulator	1	0.14											
DVU2950	S	Sensory box protein/GGDEF domain protein	1		1.00										
DVU3022	T	Sensory box histidine kinase/response regulator	2									1.00		1.00	
DVU3190	S	Hypothetical protein DVU3190	1											0.96	
DVU3264	C	Fumarate hydratase	1					0.19							
DVU3304	S	Hypothetical protein DVU3282	2	0.23	0.14										

[a] intergenic mutation identified within 100 bp of the upstream of start codon of the gene

Table S6 Newly arising mutations occurred more than once in the EN populations in addition to those in six gene groups.

Gene	Products	COG	Mutation position	Nucleotide change	Allele frequency	Populations
DVU0126	ABC transporter, ATP-binding protein	CP	**162600**	**C → G**	**100**	**EN4**
					100	**EN5**
DVU1260	outer membrane protein P1, putative	I	1347740	CT → C	72.73	EN1
			1346776	TGG → TG	16.28	EN4
			1347080	-17:TATAATGTCGGCATCCA	61.84	
DVU1469	30S ribosomal protein S1	J	1551144	C → T	50.56	EN8
			1551147	**C → T**	**98.02**	**EN6**
					100	**EN12**
DVU1785	MarC membrane protein	U	1848460	A → C	96.83	EN6
			1848475	G → A	93.67	EN11
DVU1862	GGDEF domain-containing protein	T	1929985	△258 bp	25.10	EN5
			1930123	△146 bp	67.50	EN3
			1930141	**C → T**	**56.67**	**EN3**
					16.36	**EN4**
			1930258	A → T	31.11	EN3
			1929597	CA → C	28.99	EN5
			1930239	CA → C	42.11	EN8
DVU3022	response regulator	T	**3140595**	**T → G**	**81.82**	**EN6**
					99.04	**EN11**
DVU3304	Signal transduction histidine kinase	T	3482319	G → A	23.47	EN1
			3481721	G → A	13.08	EN2

Table S7 Newly arising mutations (single nucleotide polymorphisms) in DVU1204 (*fabF*) and DVU1208 (*plsX*) of EN populations.

Gene	Populations	Allele frequency (%)	Mutation position	Nucleotide change	Amino-acid change
DVU1204 *fabF*	EN2	100	1296268	C → T	Ala → Thr
	EN3	96	1296453	T → A	Gln → Leu
	EN4	100	1296057	C → G	Gly → Ala
	EN5	90	1296057	C → G	Gly → Ala
	EN7	100	1295686	C → T	Gly → Ser
	EN9	99	1296081	T → C	His → Arg
	EN11	100	1296277	C → T	Ala → Thr
DVU1208 *plsX*	EN8	43	1299241	T → G	Silent mutation
	EN10	100	1299241	T → G	Silent mutation

Table S8 Newly arising mutations in DVU0924 (*fur*) and DVU2571 (*feoB*) of EN populations.

Gene	Populations	Mutation type	Allele frequency (%)	Mutation position	Nucleotide change	Amino-acid change
DVU0924 *fur*	EN4	SNP	100	1034711	A → T	Glu → Asp
	EN5	SNP	88	1034711	A → T	Glu → Asp
	EN7	SNP	100	1034881	G → A	Arg → Gln
	EN9	SNP	100	1034711	A → T	Glu → Asp
	EN10	SNP	100	1034641	A → C	Asp → Ala
	EN11	SNP	100	1034799	A → T	Asn → Tyr
DVU2571 *feoB*	EN1	SNP	100	2685709	G → A	Ser → Phe
	EN2	Deletion	100	2685414	-27: CTCGGCCCGTTCGGCATCAAGGGCGGC	nine amino acids loss
	EN3	SNP	100	2685298	T → C	Lys → Arg
	EN6	SNP	99	2685977	A → T	Phe → Ile
	EN8	SNP	100	2685734	G → A	Arg → Cys
	EN12	SNP	100	2685074	A → C	Phe → Val

Table S9 Newly arising mutations in two component system *LytR*/*LytS* (DVU0596/DVU0597) of EN populations.

Gene	Populations	Allele frequency (%)	Mutation position	Nucleotide change	Amino-acid change
DVU0596 *lytR*	EN1	11	664509	C → T	Pro → Leu
DVU0597 *LytS*	EN3	12	665366	C → T	four bases upstream of the start codon
	EN4	11	666383	C → T	Ala → Val
	EN6	40	665986	C → T	Leu → Phe
	EN7	12	666383	C → T	Ala → Val
	EN11	25	666479	A → C	Asn → Thr

References

1. Kelley LA, Mezulis S, Yates CM, Wass MN. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc. 2015;**10:**845-858.