Supplementary Information for

Rapid selective sweep of pre-existing polymorphisms and slow fixation of new mutations in experimental evolution of *Desulfovibrio vulgaris*

Aifen Zhou[1], Kristina L. Hillesland[2], Zhili He[1], Wendy Schackwitz[3], Qichao Tu[1], Grant M. Zane[4], Qiao Ma[1,5], Yuanyuan Qu[1,5], David A. Stahl[6], Judy D. Wall[4], Terry C. Hazen[7,8], Matthew W. Fields[9], Adam P. Arkin[10], and Jizhong Zhou[1,11,12 *]

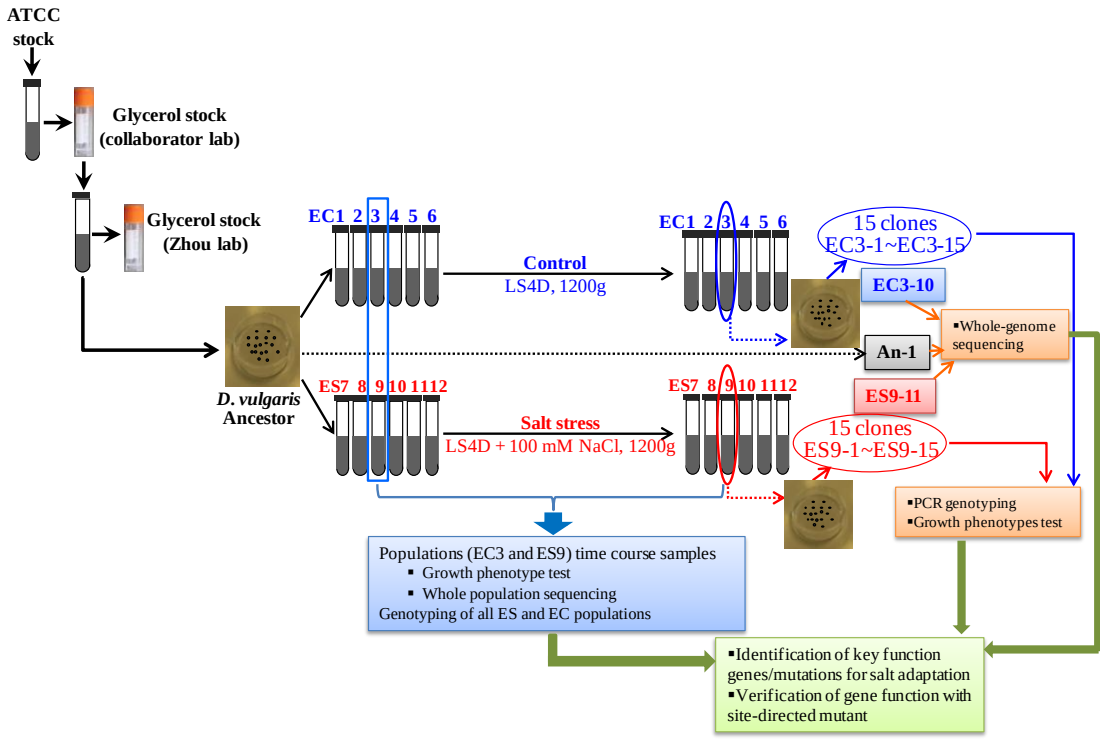
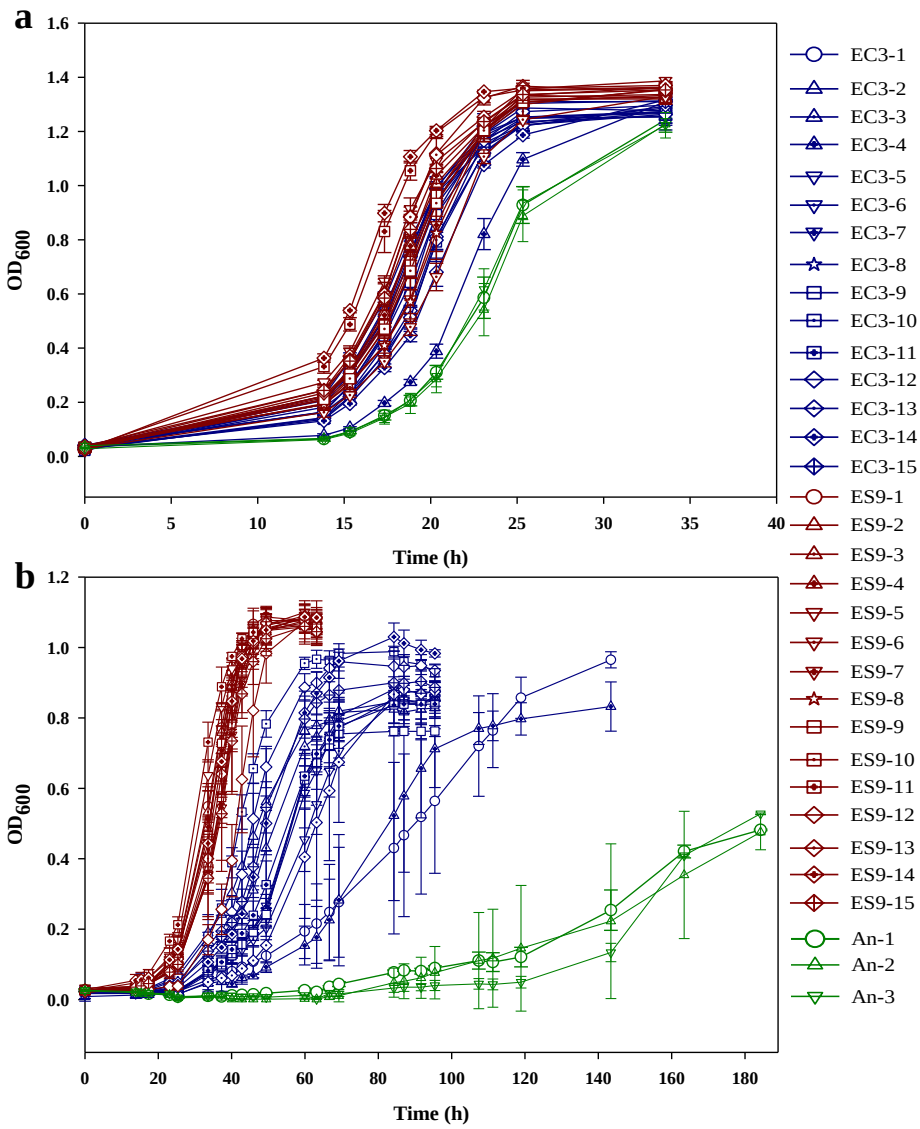


Figure S1 Overview of the experiment. Experimental evolution was founded with individual clones from pure culture of *D. vulgaris* Hildenborough (DvH). Populations EC1~EC6 were evolved in non-stress environment with defined medium LS4D, Populations ES7~ES12 were evolved in mild salt stress environment with LS4D + 100 mM NaCl, for 1200 generations. Genomic DNA samples of single-colony based strains (EC3-10, ES9-11) isolated from evolved populations EC3 or ES9 were sequenced. The dynamics of salt tolerance phenotypes or mutation frequencies during evolution were investigated with representative populations EC3 and ES9. Population genotyping was conducted to verify whether the mutation selection was common in independently evolved ES or EC populations. The gene function was demonstrated with site-directed mutants.



Growth parameter differences among ES9, EC3, and ancestral strains (P<0.05, t-test)

		LS4D		LS4D+250 mM NaCl	
Growth parameter	Strains	Average	Stdev	Average	Stdev
growth rate (hr^{-1})	ES9	0.24^{b}	0.01	0.12^{a}	0.01
	EC3	0.25^{a}	0.01	0.06^{b}	0.01
	An	0.23^{b}	0.00	0.03^{c}	0.01
Biomass (OD_{600})	ES9	1.35^{a}	0.02	1.08^{a}	0.01
	EC3	1.29^{b}	0.02	0.90^{b}	0.06
	An	1.28^{b}	0.01	0.50^{c}	0.03
Lag phase (hr)	ES9	9.53^{c}	0.88	23.98^{c}	2.11
	EC3	11.16^{b}	1.34	36.96^{b}	7.06
	An	15.84^{a}	0.15	119.93^{a}	11.94

Figure S2 Growth phenotypes of single-colony based strains isolated from 1200-generation populations ES9 or EC3 and the ancestor (An) under non-stress (LS4D, a) or salt stress (LS4D + 250 mM NaCl, b) conditions. Error bar: standard deviation.

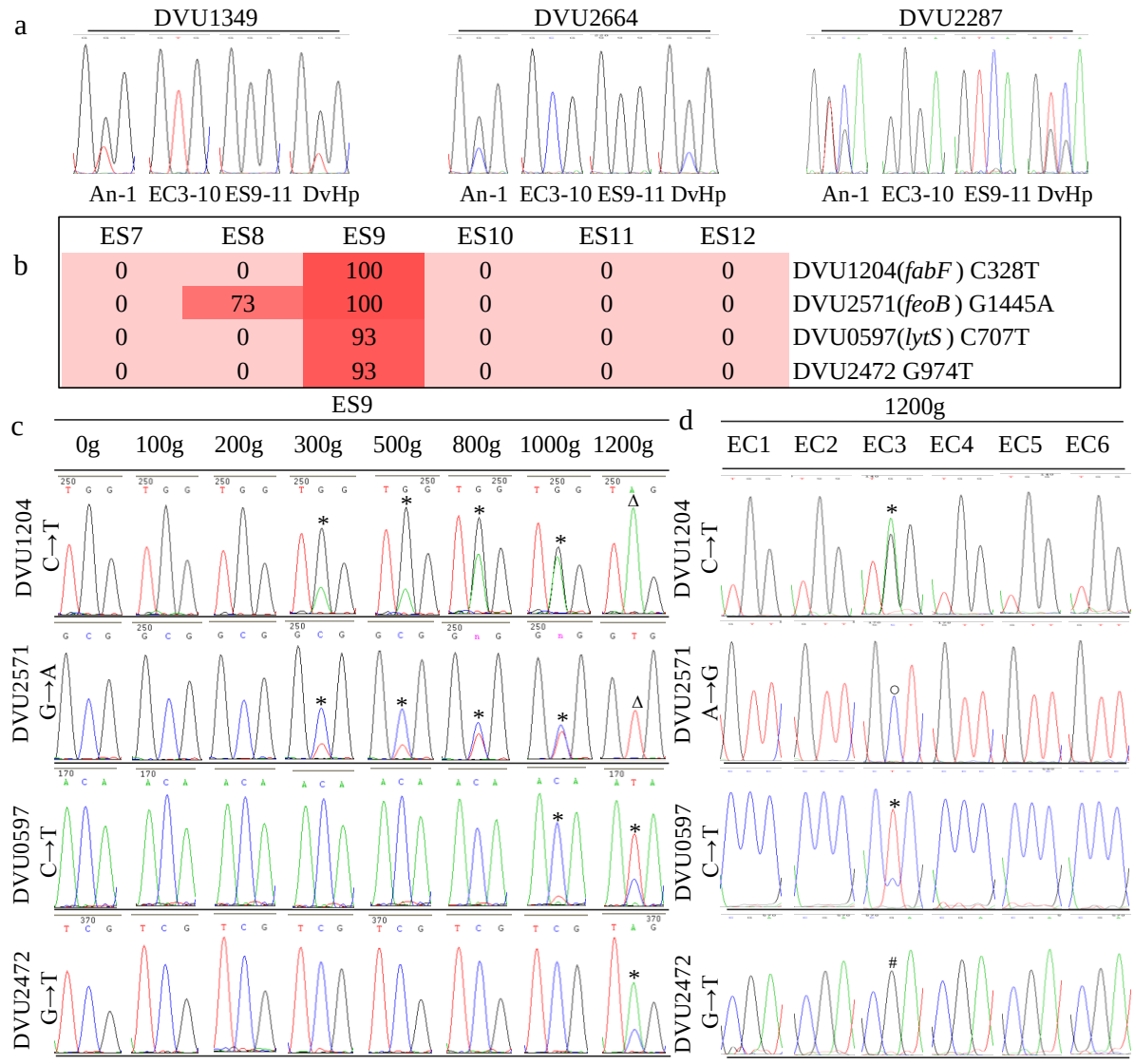


Figure S3 Sanger sequencing confirmation of mutations in colony isolates or populations. a. mutations in colony isolates (ES9-11 and EC3-10), ancestral colony isolate (An-1), or ancestral population (DvHp). b. PCR genotyping of new mutations in six stress-evolved ES populations (1200g). c. mutation frequencies of four new mutations in ES9 over evolution. d. mutation frequencies in six EC populations (1200g). Chromatograms of the sequences of the nucleotide(s) with mutations and the adjacent nucleotides are shown. *: polymorphisms; Δ: ES-type mutation; ○: EC-type mutation; #: no mutation.

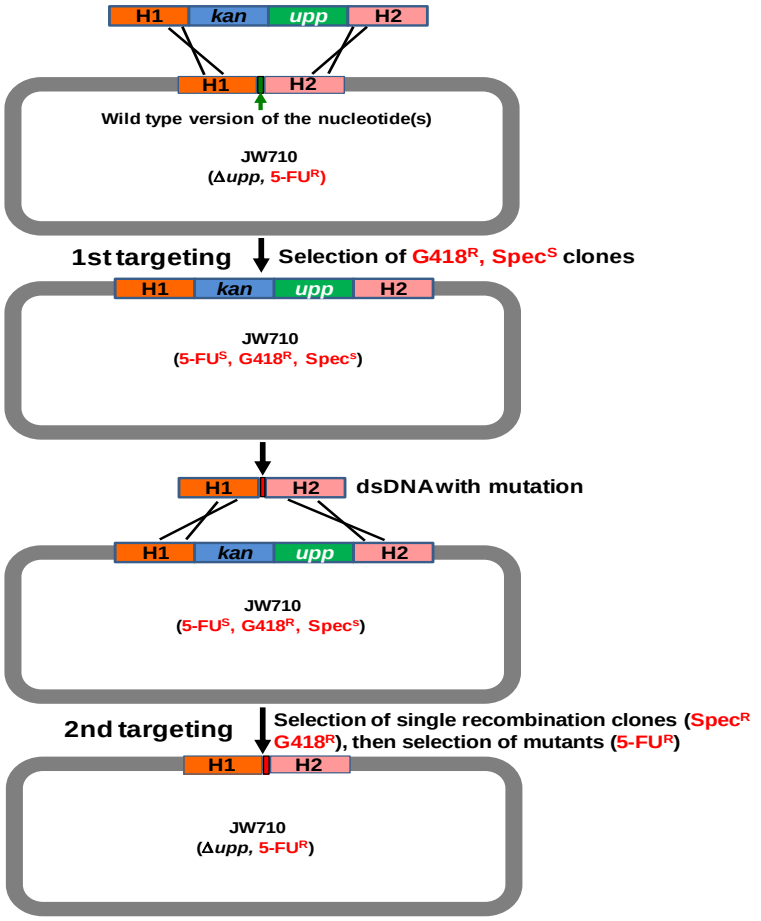


Figure S4 Overview of site-directed mutant generation in DvH. i). construction of 1st targeting marker replacement suicide vector (only the insert is shown, the vector backbone is not shown). Four DNA fragments (1, the upstream homology region (H1, ~ 700 bp) immediately adjacent to the point mutation; 2, kanamycin selection marker and *upp* gene driven by kanamycin promoter (KAN confers G418 resistance); 3, down-stream homology region (H2, ~ 700 bp) immediately adjacent to the point mutation; 4, backbone region, spectinomycin resistance and pUC origin) were assembled into a marker replacement suicide vector with the SLIC procedure. H1 and H2 were PCR amplified with genomic DNA template. Fragments 2 and 4 were amplified from plasmids pMO746 or pMO719, respectively. After electroporation of the marker replacement suicide vector into strain JW710, the wild-type version of the single nucleotide was replaced by active kanamycin selection marker and *upp* gene with homologous recombination (grey circle: DvH genome). The derived mutant was selected as $G418^R$ $Spec^S$. ii). construction of 2nd targeting suicide vector. The suicide vector was assembled with three fragments H1, H2, the backbone region 4 containing spectinomycin resistance and pUC origin with the SLIC procedure. The desired SNP/mutation was included in H1 and H2. After electroporation of the 2nd targeting suicide vector into the $G418^R$ $Spec^S$ mutant, recombination events resulted in the replacement of kanamycin selection marker and *upp* gene by the SNP/mutation. The resulted site-directed mutant (SDM) was selected by 5-FU resistance and confirmed by Sanger sequencing of the PCR fragments harboring the SNP/mutation with genomic DNA as PCR template.

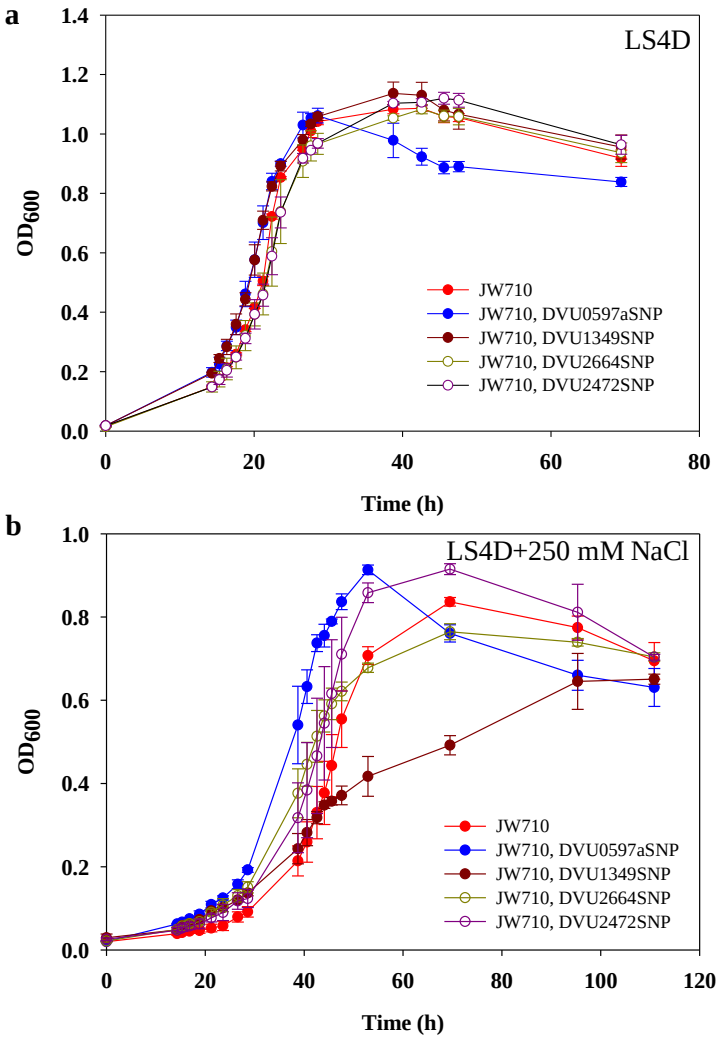


Figure S5 Growth curves of site-directed mutants carrying individual SNPs under non-stress (LS4D, a) or salt stress (LS4D +250 mM NaCl, b) conditions. Error bars indicate standard deviations.

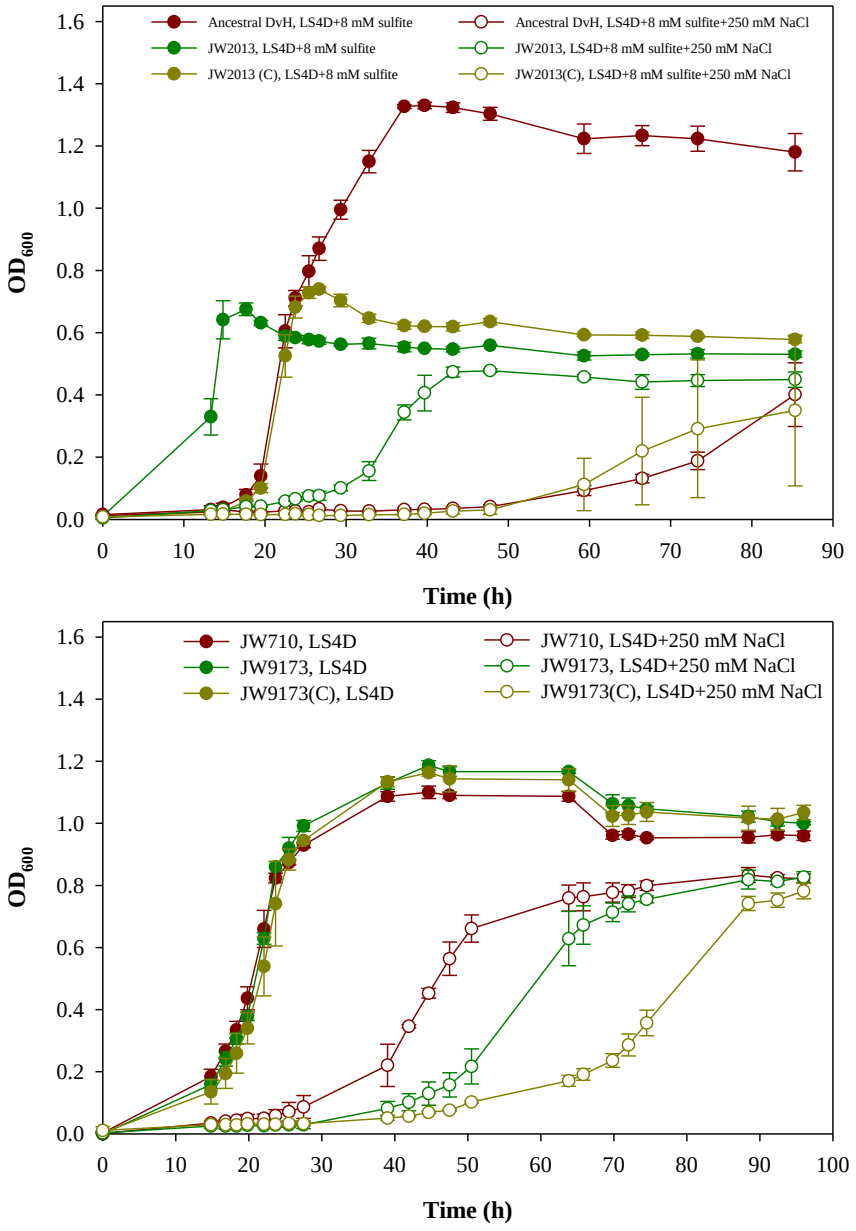


Figure S6 Growth curves of deletion mutants, complemented mutants and respective wild type controls under non-stress (LS4D) and salt stress (LS4D + 250 mM NaCl) conditions. JW2013: deletion mutant of DVU0597. JW2013(C): JW2013 containing plasmid carrying the wild type version of operon DVU0596-DVU0597 and the upstream and downstream flanking region. JW9173: deletion mutant of DVU2472. JW9173(C): JW2013 containing plasmid carrying the wild type version of DVU2472 driven by the constitutive Kan^R promoter. JW710: *D. vulgaris* strain with deletion of *upp* which encodes the pyrimidine salvage enzyme uracil phosphoribosyl transferase, this strain was used to generate deletion mutant of DVU2472.

Table S1 Mutations in single-colony based ancestral *D. vulgaris* and evolved strains ES9-11 and EC3-10

Mutation type	Location of mutation	Mutation position	Nucleotide change	Amino acid change	Affected gene(s)
SNP	Gene	1773256	G → C	Gln→Glu	DVU1698, hypothetical protein
SNP	Gene	1913197	T→G	Asp→Ala	DVU1842, lipoprotein, putative
SNP	Gene	3064064	C→G	Ala→Gly	DVU2955, hypothetical protein
SNP	Gene	3064065	G→C	Ala→Gly	DVU2955, hypothetical protein
SNP	Gene	3142198	A→G	Asp→Gly	DVU3023, *atoC*, sigma-54 dependent DNA-binding response regulator
Ins	Gene	1897096	1:T	147 more amino acids	DVU1831, transporter, putative, authentic frameshift, pseudogene become real gene
Ins	Gene	2083308	1:C	243 more amino acids	DVU2001, site-specific recombinase, phage integrase family, authentic frameshift, pseudogene become real gene
Ins	Gene	3056212	1:G	400 amino acids	DVU2950, sensory box protein/GGDEF domain protein, authentic frameshift, pseudogene become real gene
Ins	Gene	3455727	1:T	228 amino acids	DVU3280, peptide ABC transporter, ATP-binding protein, authentic frameshift, pseudogene become real gene
Ins	Gene	3140596	2:GA	6 amino acid changes	DVU3022, sensory box histidine kinase/response regulator
Ins	Gene	3276864	1:C	22 more amino acids	DVU3129, hypothetical protein
Ins	Intergenic	42868	1:G	No	+57 of DVU0036 (hypothetical protein) and -127 of DVU0037 (hypothetical protein)
Ins	Intergenic	211389	1:A	No	-397of DVU0169 (oligopeptide/dipeptide ABC transporter, periplasmic oligopeptide/dipeptide-binding protein), -39 of DVU0170 (methyl-accepting chemotaxis protein)
Ins	Intergenic	1144620	1:C	No	+76 of DVU1042 (twin-arginine translocation protein TatB), -394 of DVU1041(Sec-independent protein translocase TatC)
Ins	Intergenic	1313342	1:C	No	+127 of DVU1222 (hypothetical protein), +43 of DVU1223 (hypothetical protein)
Ins	Intergenic	2982788	2:CC	No	+31 of DVU2884 (peptidase, M18 family), -140 of DVU2885 (dhaT, alcohol dehydrogenase, iron-containing)
Ins	Intergenic	3066926	3:ACG	No	+54 of DVU2959 (hypothetical protein), -33 of DVU2960 (sigma-54 dependent transcriptional regulator)
Ins	Intergenic	3457155	3:CGC	No	+165 of DVU3281 (hypothetical protein), +70 of DVU3282 (ADP-ribosylglycohydrolase family protein)
Del	Intergenic	882517	-1:C	No	-101 of DVU0796 (hisD, histidinol dehydrogenase), +29 of DVU0797 (hypothetical protein)
Del	Gene	1191169	-1:C	73 more amino acids	DVU1087, conserved hypothetical protein
Del	Gene	1773346	-1:A	truncated protein	DVU1698, hypothetical protein
Del	Gene	3140577	-2:CA	6 amino acid changes	DVU3022, sensory box histidine kinase/response regulator

Table S2 Polymorphic loci in ancestral *D. vulgaris* and selected by evolution environment (confirmed by Sanger sequencing)

Location of mutation	Mutation position	Affected gene	Sanger sequencing			Pyro-sequencing			Amino acid change		
			An	EC3-10	ES9-11	An	EC3-10	ES9-11	An	EC3-10	ES9-11
Gene	2775672	DVU2664, *pstB-2* , phosphate ABC transporter, ATP-binding protein	G/C	C	G	Het	C	G	mix	No	Ala → Pro
Gene	2381876, 2381877	DVU2287, *CooK* , hydrogenase, CooK subunit, selenocysteine-containing, putative	T/GC/G	GG	TC	Het	G(2381876)	C(2381877)	mix	stop → Gly	stop → Ser
Gene	1426830	DVU1349, *SelGGPS* , geranylgeranyl diphosphate synthase	G/T	T	G	Het	T	G	mix	No	Val → Gly
Gene	2104739	DVU2023, hypothetical protein	C/G	G	C	Het	G	C	mix	No	Val = Val
Intergenic	2502193	DVU2397, hypothetical protein	C/G	G	C	Het	G	C	mix	No	non-coding (-199)
Gene	2905203	DVU2802, transcriptional regulator, GntR family	G/A	A	G	Het	A	G	mix	Ala → Thr	No
Gene	3169435	DVU3045, *fexB* , sensory box histidine kinase/response regulator	G/C	C	G	Het	C	G	mix	Gly → Arg	No
Gene	1599469	DVU1530, metallo-beta-lactamase family protein	C/T	T	C	Het	T	C	mix	Asp = Asp	No
Gene	326403	DVU0281, exopolysaccharide biosynthesis protein, putative	G/A	A	G	Het	A	G	mix	Asp = Asp	No
Gene	535249	DVU0467, *trpD* , anthranilate phosphoribosyltransferase	C/G	C	G	Het	C	G	mix	Leu=Leu	No

Smaller font indicates minor peak in chromatograms of Sanger sequencing. Het: heterozygous. Mix: mixture.
Pink: mutations selected in NaCl-evolved ES9-11. Light blue: mutations selected in control-evolved EC3-10.

Table S3 Primers used for amplification of DNA fragments harboring mutations

DVU #	Primer sequence	Product size (bp)
DVU0597	Forward: CCCATGATATTGCTCAATGG	760
	Reverse: GAATCTGCTCCACCTCTTCG	
DVU2571	Forward: ACATGCTCGACAGGGTGTTC	584
	Reverse: CTCAAGTTCGGTGGTGTCCT	
DVU1204	Forward: TGGAAATGACATCGAAACCA	501
	Reverse: TCGATACGTCCCAGCTTGAT	
DVU2472	Forward: CGAGACCTACGGACACCATT	579
	Reverse: CATGCACAACTGGCGAAG	
DVU2664	Forward: TTGCAACGTCTCAACGAATG	594
	Reverse: TTCGAGTTCAGCCACATCAG	
DVU2287	Forward: ATGATGCTGGCCTACGAGAT	500
	Reverse: GTGGGCAGCTTGAGGTAGAA	
DVU1349	Forward: GAAGGCTACCTTGCCACTTG	510
	Reverse: GAGAGCCCCGGTCTTCAT	
DVU2023	Forward: GACGACCCGTCGATTTATGA	518
	Reverse: TCGCCTCCTCAAGCTTCTTA	
DVU2397	Forward: ACTCACCGCCATGCTTGT	552
	Reverse: GTATTCCATCCGTGCCTCAT	
DVU0942	Forward: CATCGCCGTATTTCAGGATT	402
	Reverse: GCGCAGATGCCATAGAGATA	
DVU2395	Forward: GATGCTCTCGAAGGATTTGC	592
	Reverse: CACCACGTTGAGGTTCACAG	
DVU0797	Forward: AGGTCCGTCGTTCGTACATC	543
	Reverse: CTTCCTTGGCACCGTAAGAG	
DVU0799	Forward: CCTGAAGGTCCGTATGGGTA	558
	Reverse: GGGGTCATCATGCTCATCTT	
DVU2802	Forward: CAGATCATGAAGCTCGTCCA	522
	Reverse: GGCGAAGAACGTCTGAAGG	
DVU3045	Forward: GACTGGCATGAGAGTCATCG	534
	Reverse: TCGTACTGCCAGATGGTGTC	
DVU1530	Forward: ATCGTCTTCGTGGGCTATCA	514
	Reverse: ACTGCGAAAGGAAATGCAAC	
DVU0467	Forward: ATAGCGACCATCCTCGAGAC	532
	Reverse: CCAGCAGGTTGAAGAGGGTA	
DVU0281	Forward: TCGTAAGTCTCGTGGCAGTG	531
	Reverse: CCATGACGGGTAGAATCAGG	
DVU0467	Forward: CGTACAGTTCCATCCCGAAT	671
	Reverse: AGCAGGTTGAAGAGGGTACG	
DVU2112	Forward: ATGGGAAAAGGATGGCTTCT	507
	Reverse: TATGCCTTCGTCAAGACGTG	
DVU1698	Forward: CGATGTCATCCGCTACACAA	529
	Reverse: GGTATGCGGCGTGTAGTCTC	
DVU1842	Forward: AGGCTATTCGACCCCAGAAG	572
	Reverse: TGTAGCGGGTAACCCAGAAG	
DVU2955	Forward: CGTACCCTGCGCTACAAGTT	507
	Reverse: GCACTGGATAGACGGGCTTA	
DVU3023	Forward: GTTCAGGTCTGCGAAACCAT	595
	Reverse: AGAGGGTGCTTTCCACGAG	
DVU2001	Forward: AAACATCAGGCTTGGACGAC	590
	Reverse: GTTCCCGAGGTGAAGTTGAG	
DVU2950	Forward: CTCTTCATCGAACCGGAGTC	568
	Reverse: ATGCGGTACTCTGCCTCGTA	
7.8kb deletion (PCR1)	UP-F: AGGTCGACATGCTGTTCTCC	550
	UP-R: ATTTCTCGACGGACGTATGG	
7.8kb deletion (PCR2)	DOWN-F: CAGAAAGTCACGGCTCAACA	524
	DOWN-R: TGCGTGATGTTACGTCCTTC	
7.8kb deletion (PCR3)	UP-F	207
	DOWN-R	
DVU1862	Forward: TCAAGTCGGTCAACGACAAC	592
	Reverse: GTCGCAGTCCATGAGCAGTA	
DVU2349	Forward: GGGGTGCTCCTGTGTAAAGA	556
	Reverse: CGAATAGATGGGAAGGCAGA	

Table S4. Specificity of mutation within evolved populations

	Gene	DVU0597	DVU2571	DVU1204	DVU2472	DVU0597	DVU2571	DVU1204
Illumina sequencing results (whole genome sequencing)	Mutation position	666077	2685757	1296562	2581002	666481	2685839	1296677
	Nucleotide change	C→T	C→T	G→A	G→T	C→T	T→C	G→A
	Identified in strain	ES9-11	ES9-11	ES9-11	ES9-11	EC3-10	EC3-10	EC3-10
Sanger sequencing results (PCR fragment)	An-1	C	C	G	G	C	T	G
	An-2	C	C	G	G	C	T	G
	An-4	C	C	G	G	C	T	G
	An-5	C	C	G	G	C	T	G
	An-6	C	C	G	G	C	T	G
	Mutation frequecny in An (%)	0	0	0	0	0	0	0
	EC3-1	C	C	G	G	T	C	A
	EC3-2	C	C	G	G	T	C	A
	EC3-3	C	C	G	G	T	C	A
	EC3-4	C	C	G	G	T/c Het	C	A
	EC3-5	C	C	G	G	T	C	A
	EC3-6	C	C	G	G	T	C	A
	EC3-7	C	C	G	G	T	C	A
	EC3-8	C	C	G	G	T	C	A
	EC3-9	C	C	G	G	T	C	A
	EC3-10	C	C	G	G	T	C	A
	EC3-11	C	C	G	G	T	C	A
	EC3-12	C	C	G	G	T	C	A
	EC3-13	C	C	G	G	T	C	A
	EC3-14	C	C	G	G	T	C	A
	EC3-15	C	C	G	G	T	C	A
	Mutation frequecny in EC3 (%)	0	0	0	0	100	100	100
	ES9-1	T	T	A	T	C	T	G
	ES9-2	T	T	A	T	C	T	G
	ES9-3	T	T	A	T	C	T	G
	ES9-4	T	T	A	T	C	T	G
	ES9-5	T	T	A	T	C	T	G
	ES9-6	T	T	A	T	C	T	G
	ES9-7	T/c Het	T	A	T	C	T	G
	ES9-8	T	T	A	T	C	T	G
	ES9-9	T	T	A	T	C	T	G
	ES9-10	T	T	A	T	C	T	G
	ES9-11	T	T	A	T	C	T	G
	ES9-12	C	T	A	G	C	T	G
	ES9-13	T	T	A	T	C	T	G
	ES9-14	T	T	A	T	C	T	G
	ES9-15	T	T	A	T	C	T	G
	Mutation frequecny in ES9 (%)	93	100	100	93	0	0	0

An-1 to An-6: colony isolates from ancetoral population (An); EC3-1 to EC3-15: clones isolated from 1200-generation non-stress-evolved population EC3; ES9-1 to ES9-15: clones isolated from 1200-generation stress-evolved population ES9; Het: polymorphic loci; yellow highlight: no mutation identified.

Table S5 PCR Primers used for generation of site-directed mutants

Primer name	Sequence(5'-3')	Description
1349-1	Forward: gccttttgctggccttttgctcacatAGCGACTTCGGGTTACGTC	amplification of upstream fragment of DVU1349 SNP
1349-2	Reverse: GCGACAAGATATTCGGCACCAAGTAAGCAGGCCGAAGGCGTTTGCCA	
1349-3	Forward:GCGCCCCAGCTGGCAATTCCGGGCTCTGCCTCTCGACGGCT	amplification of downstream fragment of DVU1349 SNP
1349-4	Reverse: gtCGAGGCATTTCTGTCCTGGCTGGCGAAGGAAGTCTGCATCCTC	
1349-5	Forward: CGTCGAGAGGCAGAGCCCAGGCCGAAGGCGTTTGCCA	amplification of fragment containing DVU1349 SNP
1349-6	Reverse: CAAACGCCTTCGGCCTGGGCTCTGCCTCTCGACGGCT	
2664-1	Forward: gccttttgctggccttttgctcacatGCGAGTATCAGGGGAGGATA	amplification of upstream fragment of DVU2664 SNP
2664-2	Reverse: GCGACAAGATATTCGGCACCAAGTAAGAGACGCGATGTCCAGAGAC	
2664-3	Forward: GCGCCCCAGCTGGCAATTCCGGCACGGGCGATCGAGGAGCGCA	amplification of downstream fragment of DVU2664 SNP
2664-4	Reverse: gtCGAGGCATTTCTGTCCTGGCTGGCCTTCTCTTCAAGGCCAACC	
2664-5	Forward:TCCTCGATCGCCCGTGGAGACGCGATGTCCAGAGA	amplification of fragment containing DVU2664 SNP
2664-6	Reverse: TCTGGACATCGCGTCTCCACGGGCGATCGAGGAGCGCAT	
0597-1	Forward: gccttttgctggccttttgctcacatGCTTCGTGGTTTCATCGTTT	amplification of upstream fragment of DVU0597 SNP
0597-2	Reverse: GCGACAAGATATTCGGCACCAAGTAAGTCTCGTCGAGGATGATCTCG	
0597-3	Forward: GCGCCCCAGCTGGCAATTCCGGACGGGTGGCTGCCGTGGCT	amplification of downstream fragment of DVU0597 SNP
0597-4	Reverse: gtCGAGGCATTTCTGTCCTGGCTGGTTCCATGCTCGTGGCAAG	
0597-5	Forward: ACGGCAGCCACCCGTATCTCGTCGAGGATGATCT	amplification of fragment containing DVU0597 SNP
0597-6	Reverse: AGATCATCCTCGACGAGATACGGGTGGCTGCCGT	
2472-1	Forward: gccttttgctggccttttgctcacatGCGCTGGATGGACAGACAC	amplification of upstream fragment of DVU2472 SNP
2472-2	Reverse: GCGACAAGATATTCGGCACCAAGTAAGGCTCGACGCCGTGGGCGCAA	
2472-3	Forward: GCGCCCCAGCTGGCAATTCCGGATGGCGTAGCGACTGCGGCT	amplification of downstream fragment of DVU2472 SNP
2472-4	Reverse: gtCGAGGCATTTCTGTCCTGGCTGGGCGAAGTGCACCGTCTCT	
2472-5	Forward: GTCGCTACGCCATAGCTCGACGCCGTGGGCGCA	amplification of fragment containing DVU2472 SNP
2472-6	Reverse: TGCGCCCACGGCGTCGAGCTATGGCGTAGCGACTG	