The rapid evolution of flagellar ion selectivity in experimental populations of *E. coli*

Pietro Ridone¹, Tsubasa Ishida²,³, Angela Lin¹, David T. Humphreys⁴,⁵, Eleni Giannoulatou⁴, Yoshiyuki Sowa²,³, Matthew A. B. Baker¹,⁶

Determining which cellular processes facilitate adaptation requires a tractable experimental model where an environmental cue can generate variants that rescue function. The bacterial flagellar motor (BFM) is an excellent candidate—an ancient and highly conserved molecular complex for bacterial propulsion toward favorable environments. Motor rotation is often powered by $\text{H}^+$ or $\text{Na}^+$ ion transit through the torque-generating stator subunit of the motor complex, and ion selectivity has adapted over evolutionary time scales. Here, we used CRISPR engineering to replace the native *Escherichia coli* $\text{H}^+$-powered stator with $\text{Na}^+$-powered stator genes and report the spontaneous reversion of our edit in a low-sodium environment. We followed the evolution of the stators during their reversion to $\text{H}^+$-powered motility and used both whole-genome and RNA sequencing to identify genes involved in the cell’s adaptation. Our transplant of an unfit protein and the cells’ rapid response to this edit demonstrate the adaptability of the stator subunit and highlight the hierarchical modularity of the flagellar motor.

**INTRODUCTION**

Bacterial motility via the flagellar motor represents one of the earliest forms of locomotion (1). This rotary motility imparts such a substantial selective advantage (2, 3) that resources are allocated to chemotaxis even in the absence of nutrient gradients (4, 5). The evolutionary origins and subsequent adaptation of the motor are of significant scientific and public interest (6), because the bacterial flagellar motor (BFM) holds prominence as an ancient and large molecular complex of high sophistication. Furthermore, the BFM is an ideal model for studies in molecular evolution because it demonstrates modularity (7, 8) and single-nucleotide variants, which result in changes in motility that are easily experimentally selected (9).

The torque that drives the BFM is supplied by motor-associated transmembrane protein complexes known as stators. The stator complex, an asymmetric heteroheptamer (in *Escherichia coli*: MotA₂MotA₂) most likely acts itself as a miniature rotating nanomachine coupling ion transit to rotation (10, 11). The stators are essential for motility, as they drive rotation, and are accessible for studies in experimental evolution due to their unambiguous role in connecting a specific environmental cue (the presence of the coupling ion) to an easily discernible phenotype (cell swimming). Furthermore, the stators have been subjected to protein engineering approaches for many years, particularly the synthesis of chimeric stator constructs that enable the motor of *E. coli*, natively proton-driven, to be powered by sodium ion flow (12–15). The majority of stators are proton-driven, but many that are sodium-driven can be found in nature (16), and this divergence is presumed to have occurred in the distant past (7, 17, 18). Past reports have argued that $\text{H}^+$-coupled motility diverged from $\text{Na}^+$-coupled machinery in ancestral times (19), but the molecular basis for this adaptation and the evolutionary landscape that constrains stator adaptation remain unclear.

To simulate the effects of natural evolution on stator adaptation, we designed an experiment where an *E. coli* strain, expressing only a sodium-powered stator, would be introduced to a nonlethal environment (soft agar swim plate), which lacked the power source for the stator ($\text{Na}^+$). Our hypothesis was that the population would undergo selection for upmotile variants, adapting its stators to function in the new environment.

We used genomic editing techniques [no-SCAR (scarless Cas9 assisted recombineering) CRISPR-Cas9 (20) and λ-Red (21)] to replace the native *motA motB* stator genes of the *E. coli* BFM with chimeric sodium-powered *pomA pomB* (henceforth Pots) stator genes derived from *Vibrio alginolyticus* (12). We transplanted the *pomA pomB* stator genes at the same location and orientation of the native *motAmotB* locus to preserve the native genomic context of the motile RP437 *E. coli* strain. We then examined which genetic changes occurred during growth on soft agar in depleted sodium, that is, under selective pressure for proton-driven motility. We performed our directed evolution experiments of our Pots *E. coli* strain in the absence of antibiotics to avoid additional, undesired selective pressures (22).

**RESULTS**

**Directed evolution of Pots on low-sodium swim plates**

The RP437 strain was edited to carry the Pots stator genes in place of the native *E. coli* *motAmotB* genes via the no-SCAR method (20) and the traditional λ-Red recombineering, respectively (figs. S1 and S2). A single no-SCAR Pots clone was selected and tested on swim plates (fig. S3) after verification of successful editing by colony polymerase chain reaction (PCR) and Sanger sequencing (fig. S4). The edited strain was able to swim on sodium-rich (Na⁺LB: ~100 mM NaCl) soft agar plates but not on potassium-rich sodium-poor swim plates (K⁺LB: 67 mM KCl, ~15 mM Na⁺) (fig. S3A). This edited strain exhibited the same swimming behavior as the control statorless strain with motility restored via an inducible plasmid vector that could express the Pots construct (RP6894 ΔmotAmotB + pSHU1234, hereby pPots).

¹School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, Australia. ²Department of Frontier Bioscience, Hosei University, Tokyo, Japan. ³Research Center for Micro-Nano Technology, Hosei University, Tokyo, Japan. ⁴Victor Chang Cardiac Research Institute, Sydney, Australia. ⁵School of Clinical Medicine, Faculty of Medicine and Health, UNSW Sydney, Australia. ⁶ARC Centre of Excellence in Synthetic Biology, University of New South Wales, Sydney, Australia. ⁷Corresponding author. Email: matthew.baker@unsw.edu.au

Copyright © 2022, the authors. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (Creative Commons Attribution 4.0 International License).
We next challenged this Pots strain to survive on K⁺-based soft agar (K⁺LB) for prolonged periods (fig. S3B). Motile subpopulations arose spontaneously from inoculated colonies within a few days. Cells from the edge of these motile flares were passaged onto fresh swim agar for up to five passages at 3- and 4-day intervals (fig. S5). When multiple flares occurred in a single swim ring, each was individually passaged (fig. S3B) and recapitulated (Fig. 1C and fig. S3C). Directed evolution consistently generated swimming flares when Pots clones were cultured on agar containing yeast extract and tryptone (K⁺LB swim plate: ~15 mM [Na⁺]) but not on minimal media (K⁺MM swim plate: ~1 mM total [Na⁺]) or when the Pots construct was encoded on a plasmid (fig. S6). One Pots strain generated using λ-Red methods (21), which carried the native V. alginolyticus Shine-Dalgarno sequence, also successfully produced flares (fig. S7).

**Whole-genome sequencing of evolved lineages**

Lineages were selected for whole-genome sequencing (WGS) after a preliminary screening for mutations in the stator genes by Sanger sequencing PCR amplicons spanning the genomic pomA/potB locus (Fig. 1). Variant calling to the MG1655 reference genome was used to compare single-nucleotide polymorphisms (SNPs) between members of the same lineage. Our intended pomA/potB edit was the only difference between the RP437 and Pots genomes, indicating that neither no-SCAR nor λ-Red editing had resulted in off-target edits. One hundred fifty-three SNPs were called as variants between our experimental parent RP437 and the MG1655 reference that were shared across all lineage members (table S1).

Several lineages whose descendants could swim in reduced sodium had mutations at the pomA/potB locus (L3.3–4–5: potB G20V; L3.3–4–5: pomA L183F; L3.3–4–5: potB G20W). In contrast, lineages passaged only on ~100 mM Na⁺ agar (L1 and L2) accumulated mutations not in stators but in the flagellar components. Lineages passaged on ~15 mM sodium-poor agar whose descendants could not swim (L4) had no mutations on any flagellar genes.

**Differential expression across upmotile lineages**

**G20V and L183F**

To determine which genes may be involved in adaptation, we performed RNA sequencing (RNA-seq) experiments to measure transcript levels for two lineages that evolved different stator mutations over different lengths of time (Fig. 2 and table S2). These were pomA L183F (Na⁺-powered phenotype, 14 days) and potB G20V (H⁺-powered phenotype, 7 days). From these, we selected the common ancestor to the two lineages (Pots), the last lineage member before the mutation occurred (L5.3 and L3.2, the prefixation sample) and the first available member carrying that lineage mutation on the chromosome (L6.4 and L3.3, the fixed variant). Notably, L5.3 displayed improved swimming in K⁺LB despite being isogenic to its parent strain, the Pots common ancestor. Member L3.2 of the G20V lineage, the immediate predecessor to L3.3, in contrast, was non-motile. We closely examined the transcripts mapped to the pomA/potB locus in all samples and saw no enrichment of mutant transcripts above the background noise in the prefixation samples (fig. S8). This confirmed that a chromosomal mutation was responsible for the observed stator variants. To identify the common processes which could lead to variant fixation, we calculated which differentially expressed genes were present in both lineages and performed pathway analysis to identify which biological processes were relevant to the shared genes (Fig. 2).

Both lineages increased expression of lrrhA at the prefixation point, which acts as an inhibitor of the flhDC flagellar master regulator. Expression response was dominated by pathways under control of RpoD (σ70)/FliA (σ28) in L183F and RpoD (σ70) in potB G20V (fig. S9). The prominence of FliA regulation in pomA L183F was reflected by up-regulation of motility genes at the fixation stage. Those same genes were not significantly affected in the G20V lineage, which instead promoted primarily adhesion/biofilm behavior in response to the changed environment.

**Phenotypic characterization of evolved strains in the presence and absence of sodium**

We characterized rotational and free-swimming phenotypes of the parent and evolved strains in the presence and absence of sodium using a tethered cell assay with particular focus on the upmotile G20V variant, L3.3 (Fig. 3 and fig. S10). The potB G20V fixed variant was clearly distinguishable from Pots in this assay and maintained rotation both in the absence of sodium and in the presence of phenamil, an amidorlide derivative and sodium channel blocker (Fig. 3A) (23). We confirmed the potB G20V H⁺-powered phenotype by introducing the same point mutation (GGG to GTG) on our plasmid vector (pPots) and testing motility in tethered cell assay when the protein was expressed in the statorless ΔmotAB RP6894 (Fig. 3B). Measurement of tethered rotational speed versus lithium and versus sodium showed that the potB G20V variant was motile at 0 mM Na⁺ and 0 mM Li⁺, with a dependence on the concentration of ion, in contrast to motA/ motB, whose rotational speed was independent of sodium or lithium concentration (fig. S11, A to C). All stator types had cessation of rotation at 50 μM carbonyl cyanide 3-chlorophenylhydrazone (CCCP) and showed no measurable effect as pH was varied between 6.0 and 8.0 (fig. S11, D and E).

We further tracked single-cell rotation in the tethered cell assay with sequential exchange of buffers, including characterization of pomA P177Q and pomA L183F (fig. S12). There, potB G20V (L3.3) actively rotated in K⁺MBEDTA and in Na⁺ MB (motility buffer) + 100 μM phenamil. K⁺MB indicates sodium-independent rotation. We further synthesized all possible variants at site G20 (24), and of these, only G20V was able to rotate in the absence of sodium (fig. S13).

**Natural prevalence of motB G20 and motB V20**

We examined the natural prevalence of glycine and valine at site 20 across a phylogeny from a collection of 82 motB sequences. Of these sequences, G20 was present only in the clade corresponding to Vibrio sp. pomB (7 sequences), whereas V20 was distributed across the phylogeny more broadly (41 sequences) (Fig. 1D and fig. S14). Ancestral reconstruction across all nodes predicted that the ancestral phenotype of this phylogeny was V20.

**Reproducibility of stator mutagenesis and capacity for reversion**

To examine mutation reproducibility in the stators, we subjected 55 Pots colonies to directed evolution in K⁺LB swim plates. These yielded a total of 42 flares within the first 3 days of incubation, which were then passaged four more times at 3-day intervals. At the end of this experiment, we selected the 20 terminal lineage members that produced the largest swim rings and Sanger-sequenced them at the pomA/potB locus. For these, we observed a total of five mutations in pomA (S25C, D31N, P177A, P177Q, and L183F) and one more new mutation in potB (L36Q) (Fig. 3C). We mapped monomer models
for PomA and PotB generated using Alphafold (25) to a model PomA\textsubscript{5}PotB\textsubscript{2} complex using the published high-resolution Bacillus subtilis MotA\textsubscript{5}B\textsubscript{2} structure (Fig. 2, C and D) (11). All stator mutations accumulated at sites proximal to or within the predicted ion transport pore, at the interface between the PotB transmembrane domain and the third and fourth transmembrane domains of PomA (Fig. 2D).

Over the course of these experiments, we found that three stator residues underwent mutation at the same site twice [PomA L183F (2×), PomA P177A and P177Q, and PotB G20V and G20W]. WGS

---

Fig. 1. Directed evolution of the flagellar motor. (A) Sodium-swimming strain was repeatedly passaged on either Na\textsuperscript{+}LB (~100 mM [Na\textsuperscript{+}]) or K\textsuperscript{+}LB (~15 mM [Na\textsuperscript{+}]) plates. Flares, indicating a potentially upmotile variant, were passaged and sent for sequencing. (B) Edited \textit{E. coli} strains Pots and Pots\textsuperscript{λ}, obtained from \textit{E. coli} RP437 via no-SCAR and \textsuperscript{λ}-Red recombineering, respectively, were passaged on soft agar (colored background, yellow: Na\textsuperscript{+}; blue: K\textsuperscript{+}) over an 18-day period, and eight lineages (L1 to L8) were selected for further investigation, each composed of five members (i.e., L1.3 indicates the first lineage and the third passage). Swimming ability in the presence of high or low sodium is displayed by a yellow or blue ring, respectively, corresponding to swim size on swim plate (C). Lack of motility on K\textsuperscript{+} soft agar is represented by a blue dot indicating colony growth only. Colonies that were nonmotile in K\textsuperscript{+} plates were confirmed with further incubation (fig. S3C). Red and boxed lineage member labels indicate WGS and RNA-seq data availability (including Pots). SNPs identified relative to the Pots reference genome are annotated next to each respective lineage member and in table S1. Highlighted genes other than \textit{pomA} and \textit{potB}: \textit{pitA} (metal phosphate:H\textsuperscript{+} symporter), \textit{flgL} (flagellar hook-filament junction protein 2), \textit{fliM} (flagellar motor switch protein), \textit{cdsA} (cardiolipin-diglyceride synthase), \textit{icd} (isocitrate dehydrogenase), and \textit{rrsG} (16S ribosomal RNA). Scale bar, 10 mm. (C) Recapitulation of the directed evolution experiment. Na\textsuperscript{+} (left) and K\textsuperscript{+} (right) soft agar plates inoculated with a 1-μl aliquot of glycerol stock of each strain indicated in (A) (except RP437) and arranged in the same order as (A). (D) Phylogeny of \textit{motB} across 82 species with ancestral reconstruction at the G20 site. G20 is conserved in the \textit{Vibrio} spp. clade. Full phylogeny is shown in fig. S14.
Fig. 2. Differential expression analysis of RNA-seq data. Differentially expressed genes (DEGs; adjusted P value < 0.01) measured across selected members of the L183F lineage (A) or G20V lineage (B). rRNA, ribosomal RNA; TCA, tricarboxylic acid; CoA, coenzyme A. The average read counts (n = 3) for each gene were normalized by z-score and displayed as clustered heatmaps, flanked by their respective dendrograms. DEG clusters are labeled by their representative biological process (PANTHER). Genes up-regulated with respect to the Pots strain are labeled in red, down-regulated ones are in blue. The ring diagrams above each heatmap are taken from Fig. 1 and indicate the motility phenotype of each lineage member. (C) Venn diagram indicating the number of DEGs in each dataset and the DEGs in common, shown in the heatmap on the right. AA, amino acid.

Ridone et al. Sci. Adv. 8, eabq2492 (2022) 23 November 2022
examined all stator mutants that swam on K⁺ LB swim plates in conscious media when starting from a more favorable vantage point. We tested whether evolution could be more easily directed on minimal media. Mutations in a subsequent generation are underlined. (A) View from the extracellular side of the transmembrane portion of PomA5PotB2 stator complex (see the "Structural modeling" section). One PomA subunit (blue) and the transmembrane domains (TM) of the PotB (yellow) are highlighted. Mutant sites are labeled in green, and the catalytic aspartate residue essential for function is highlighted in red. The red circle indicates the predicted location of the ion transport pore. (B) Side view of the area highlighted by the dashed box in (D). Residues P151 (PomA) and D24 (PotB) are also highlighted in magenta and red, respectively. The arrow at the interface between (A) TM3-4 and (B) TM indicates the predicted location of the ion transport pore. The inset highlights the change in the pore region due to the G20V substitution in PotB (green).

Fig. 3. Functional characterization of evolved stators. (A) Tethered single-cell speed measurements (Hz, revolutions/s). Blue bar indicates speed in 67 mM K⁺ motility buffer, red bar indicates 85 mM Na⁺ motility buffer, and red patterned bar indicates 85 mM NaCl + 100 μM phenamil (Phen) motility buffer. Number of cells analyzed per condition (from left to right): RP437: 27, 51, and 25; Pots: n.a., 78, and n.a.; Pots L3.3 potB G20V: 45, 36, and 39 (n.a. indicates no visible rotating cell). Error bars indicate SD. (B) Tethered single-cell speed measurements of RP6894 ΔmotAB strain coexpressing pomA and potB G20V from pPots plasmid, color coded as in (A). Number of cells analyzed per condition: ΔmotAB + pPots: n.a., 32, and n.a.; ΔmotAB + pPots potB G20V: 40, 63, and 48. Error bars indicate SD. Single-cell tracked data shown in fig. S12. (C) Graphical summary of stator gene mutations detected across all directed evolution experiments and the growth conditions under which these mutations arose. MM indicates agar in minimal media. Mutations in a subsequent generation are underlined. (D) revealed that the pitA gene had mutated in three separate lineages (L2.5, L3.5, and L8.5) with one of the mutations occurring twice (pitA W112*Stop).

To test for the capacity for reversion, we took sequenced lineages that swam in K⁺ LB swim plates (L3.3-4-5, L6.4-5, and L8.3-4-5) and provided in Fig. 3C. We also verified the capacity of strains at the premutation stage to replicate the mutation of their lineage and found that L3.2 (pre-G20V) consistently replicated the potB G20V mutation over two independent experiments (six of six flares) and in one instance incorporated a potB L36Q mutation in addition to G20V (fig. S16).

Comparison of fitness and motility in competition
Last, we compared fitness and competitive motility directly between the L3.3 (potB G20V) and the Pots ancestor. The potB G20V variant conferred a clear and discernible motility advantage in a mixed population on K⁺ LB swim plates while displaying similar growth in both Na⁺ and K⁺ LB liquid media (fig. S17).
adaptation are not so well understood (26, 27). Epistasis and functional redundancy in biochemical pathways impede the accurate forecasting of mutagenic events responsible for rescuing a phenotype, which could be the result of loss-of-function mutations in negative regulators or gain-of-function mutations in positive regulators (28). The flagellar system and its regulatory elements have been known to be under selective pressure due to their associated energetic and fitness costs, not only always resulting in positive selection but also often resulting in gene deletion or inactivation by insertion sequences (29–34). By targeting the stator subunit of the flagellar motor, we have been able to study the molecular events leading to the spontaneous adaptation of a unique module within a highly conserved molecular complex (35).

Previous reports have shown that bacterial motility can adapt (9) and be rescued (36) via remodeling of the flagellar regulatory network. Ni et al. (9) observed that evolutionary adaptation of motility occurs via remodeling of the checkpoint regulating flagellar gene expression. Their experiments tracked adaptive changes in swim plates, matching our experiments, but their only selection criteria were for improved swimming in an unhindered swimming population. In agreement with our results (fliM A161V), they found fliM (M67I and T192N) to be among the first genes to mutate in the improved swimmer population, but they did not report any changes in fliL (ΔA57-Q58), nor, significantly, did they see any mutations in any stator genes. Flagellum-mediated motility also appears to be naturally robust to the loss of regulatory factors, such as the enhancer-binding protein FleQ in *Pseudomonas fluorescens*, which function can be substituted by distantly related homologous proteins (35).

In contrast, our *E. coli* PotB strain faced selective pressure from ion scarcity. Our scenario is reminiscent of previous semisolid agar experimental evolution studies on the adaptation of antibiotic resistance and produced similar results. In the Microbial Evolution and Growth Arena plate experiments of Baym et al. (22), they similarly saw that the phosphate transporter *pitA* was repeatedly mutated, often to a frameshifted or a nonsense variant. In similar experiments, the isocitrate dehydrogenase *icd* was also seen to mutate often (1, 37).

Differential expression analysis of our RNA-seq datasets revealed that our two chosen lineages displayed different transcriptomic signatures during the adaptation of the stator genes. While the L183F lineage displayed many up-regulated genes involved in motility, the G20V lineage was found to regulate genes involved in biofilm formation.

The improvement in low-sodium motility observed in L5.3 (pre-L183F mutation stage) in comparison with its PotB ancestor may be explained by the up-regulation of flagellar and chemotaxis genes in L5.3. Notably, L6.4 (pomA L183F) did not swim in the total absence of sodium (fig. S12D); however, the swim ring size of both L5.3 and L6.4 was greater relative to Pots on low-sodium K+ LB plates. The genes involved have also been measured as up-regulated in a similar study on the adaptation of *E. coli* to swimming in soft agar (8), and it could be the case that up-regulation of several flagellar components, including the chimeric sodium stators themselves (PomA and PotB), improves motility in these ~15 mM [Na+] plates. In contrast, the expression profile of the Na+-dependent swimmer L3.2 (pre-G20V mutation stage) was characterized by the regulation of genes involved in metabolic pathways indicative of nitrogen starvation, fermentation of products of catabolism such as amino acids and nucleotides, and the transition to a biofilm lifestyle. Roughly 10% of genes showed significant changes in transcription in both lineage trios (10.8 and 13.2% in L183F and G20V, respectively), and all of these were up-regulated during adaptation.

After the G20V mutation was fixed, L3.3 (G20V) was found to up-regulate chemotaxis and motility genes and other markers of adaptation to soft agar (cheA, trp, tar, and tap) (9). Both lineages up-regulated the *flhDC* regulator *LrhA* at the premutation stage, hinting that down-regulation of the flagellar biosynthetic cascade is a shared trait in the early stages of adaptation to a sodium-poor environment. Similarly, both lineages up-regulated nucleotide catabolic processes and salvage pathways, a feature also observed in antibiotic response, and which can affect mutation rates by disturbing the nucleotide triphosphate balance in the cellular pool (38–40). This might suggest that the mutations were a product of stress-induced mutagenesis, a known facilitator of evolution (41), which has been proposed to involve RpoS-mediated up-regulation of the DinB error-prone polymerase (42). In our RNA-seq data, we saw minimal involvement of RpoS signaling (fig. S9) and no evidence of up-regulation of known error-prone polymerases (table S2). This suggests that the molecular events leading to fixation are not resolvable from transcriptomic analysis of only two lineages or that alternate mechanisms could have facilitated mutagenesis. These may include mutagenesis via redox events on DNA, as seen in antibiotic resistance (43), or via transcription-dependent mechanisms (44, 45). The biological pathways leading to mutation remain to be elucidated.

We observed a convergence of mutations on the stator genes and even to the very same nucleotide [GAG (L) to GAA (F)] in two separate pomA L183F lineages. Stator genes were the first to mutate in all of our WGS lineages under pressure from sodium scarcity. Given that this was from a clonal population under identical environmental constraints, it suggests that adaptation of the stators provides a strong selective benefit in changing environments.

Mutations in stators are known to affect ion usage and may confer dual ion-coupling capacity (13–15, 46). For example, the substrate preference of the *Bacillus alcalophilus* MotPS stator (Na+/K+ and Rb+) was changed with the single-mutation M33L in motS, causing the loss of both K+ and Rb+ -coupling motility in *E. coli* (47). Similarly, a bifunctional *Bacillus clausii* MotAB stator (Na+/H+ ) triple mutant (motB V37L, A40S, and G42S) was selective only for sodium ions, while the combination of mutations G42S, Q43S, and Q46A made MotB selective only for H+ (48). The previously reported S25C (49) and D31N (50) amino acid substitutions in PomA have been shown to reduce motility and, in the case of D31N, affect ion usage. Furthermore, single-point mutations in stator genes of *Vibrio* spp. (e.g., *pomB* G20V/G20R/P16S) have been shown to impart phenamil resistance both in *V. alginolyticus* and in our Pots strain (51–53).

The PotB-G20V variant directly evolved here is capable of rotation in the absence of sodium, but its swimming speed does increase with increased availability of either sodium or lithium, with an energization profile more akin to the parental Pots strain than the proton-powered *E. coli* wild type (fig. S11). This perhaps suggests that PotB-G20V is an opportunistic stator adaptation that enables promiscuity, allowing the passage of protons under sodium-scarce conditions.

The key difference in this work compared with previous efforts for directed evolution of the stators via mutagenesis (54, 55) is that, here, we edited the stators directly onto the *E. coli* genome to direct stator evolution in vivo in the native *E. coli* genomic context. This would not be possible in *Vibrio* sp. because *Vibrio* cells do not survive at low sodium. Conversely, in our system, it is difficult to use
directed evolution to revert the ion selectivity of the stator (H\(^+\) to Na\(^+\)) because, in *E. coli*, it is not possible to drastically reduce the proton concentration without affecting essential systems. Nevertheless, the observation of no revertant agrees with previous work suggesting that requirements for Na\(^+\) binding are stricter than for H\(^+\) binding and that mutations that convert a Na\(^+\) motor to an H\(^+\) are more accessible than the reverse (56, 57).

Mutation of pore-proximal residues into hydrophobic residues (e.g., G20V) hinted at a mechanism for varying constrictions in the pore to alter the efficiency of ion binding. However, in contrast, none of the bulkier, hydrophobic amino acid replacements at potB G20 (e.g., P and W) resulted in a similar G20V-like H\(^+\)-powered phenotype (fig. S13C). This suggests a selectivity mechanism enabled by G20V that is not driven simply by size. We propose an alternate mechanism whereby selectivity is maintained through perturbing the electrostatic environment in the vicinity of PotB D24 and the conserved P151 of PomA (*E. coli* MotA P173) (58).

Upon examination of the phylogenetic record with specific focus on the G20V locus, we observed that valine was more prevalent and distributed more broadly across microbial strains. This contemporary prevalence and ancestral sequence reconstruction across our phylogeny implied that the ancestral state of MotB was more to be likely V20. While G20V point mutations arose spontaneously in our experiments within a few days, these transitions do not appear to have occurred in the evolutionary record. This may indicate constraints on the adaptation of the sodium-powered stator units when considered in their native sodium-dependent hosts.

Caution is required when applying learnings from directed evolution to natural evolution because selection pressures in the wild are not typically general (59). In this study, we have leveraged our system and our experimental design to obtain large, measurable phenotypic change through a single mutation at the G20 locus. The stators of the flagellar motor appear ready to evolve in our experiments: They do not require transit through additional cryptic or neutral mutations and, thus, are a model system for exploring molecular adaptation of ion selectivity. Alternative candidates such as sodium porters and pumps often have built in redundancy (60), and many marine microbes require sodium to be viable and thus evolutionary pressure cannot be applied with such specificity to a single protein complex. In our idealized system, we are able to examine isogenic bacterial colonies without competitive effects in a homogeneous medium without local niches. Nevertheless, it still remains difficult to quantify rates of adaptation (61). For these reasons, our system is optimized to produce stator variants, but it may be that this ease for adaptation via single-point mutation is accessible precisely because cryptic mutations have been accumulated because of exposure to changing environments in the distant past (62, 63).

Motility confers a fitness advantage that is worth significant energetic investment despite the high cost of synthesizing the flagellar machinery (5, 34, 64). This advantage can only be seized if the correct ions are available for stator conversion into torque. We have shown here that the flagellar apparatus is capable of single-site mutation to adapt the stator genes to harness other available ions. While the modularity of the overall flagellar motor is now well shown (65), here, we have observed further modularity and adaptability within the stator complex. Chimeric functional stators cannot only be engineered from stator components in various species, but they are also subsequently capable of rapid mutation to use ions more promiscuously. Motility provides many benefits to an organism, but the exact evolutionary event that resulted in the first flagellar motor is not known (1). Nevertheless, our work shows that, following emergence, subsequent environmental adaptation can occur rapidly.

**MATERIALS AND METHODS**

*E. coli* strains, plasmids, and culture media

*E. coli* strain RP437 was used as the parent strain for genomic editing experiments (66, 67). The pSHU1234 (pPots) plasmid encoding *pomA* and *potB* (51) was used as the template to generate the double-stranded donor DNA. This was used to replace the *motA* and *motB* gene on the RP437 chromosome. Point mutations in plasmids were generated using the QuikChange technique, while saturation mutagenesis of the *potB* G20 site was performed using the “22-c trick” technique (24). Liquid cell culturing was done using LB broth (NaCl or KCl, 0.5% yeast extract, and 1% Bacto tryptone). Cells were cultured on agar plates composed of LB broth and 1% Bacto agar (BD Biosciences, USA). Swim plate cultures were performed on the same substrates adjusted for agar content (0.3% Bacto agar). MM was used to replace yeast extract and tryptone in soft agar swim plates. MM composition is as follows: 10 mM KH\(_2\)PO\(_4\) (KPi), 1 mM (NH\(_4\))\(_2\)SO\(_4\), 1 mM MgSO\(_4\), thiamine (1 μg/ml), and 0.1 mM each of the amino acids Thr, Leu, His, Met, and Ser. Inhibition of Na\(^+\)-dependent motility was performed using phenamil methanesulfonate (P203, Sigma-Aldrich) at 50 and 100 μM concentrations, while H\(^+\)-dependent motility was inhibited using the protonophore CCCP (C2759, Sigma-Aldrich).

**Editing *E. coli* with Cas9-assisted recombineering**

This procedure was adapted from the no-SCAR method (20). The target strain to be edited (*E. coli* RP437) was sequentially transformed first with the pKD-sgRNA-3′MotA (Sm\(^+\)) plasmid, encoding a single guide RNA (sgRNA) sequence directed at the 3′ end of *motA*, and then with the pCas9cr4 (Cm\(^+\)) plasmid to yield a parent strain harboring both plasmids. The overlap extension PCR technique (68) was used to assemble linear double-stranded DNA (dsDNA) molecules using three starting dsDNA fragments. The resulting donor DNA was electroporated in the plasmid-bearing host, and the successfully edited clones were selected via colony PCR and Sanger sequencing of the *motAB* locus. A list of primers and PCR protocols used in this work is provided in fig. S18.

**Construction of Pots by λ-Red recombineering**

Chromosomal replacement from *motAmotB* to *pomApotB* was achieved using a λ-Red recombination system, with plasmid pKD46 encoding the Red system and positive selection for the recovery of swimming ability (69). Motile clones were selected by isolating motile flares on swim plates.

**Measurement of sodium concentration in solutions using atomic absorption spectroscopy**

The amount of residual sodium in motility buffers and culture media was measured using atomic absorption spectroscopy (AAS) (ANA-182, Tokyo Photo Electric Co. Ltd., Japan). AAS measurements are displayed in table S3.

**Tethered cell assay preparation and analysis**

The tethered cell assay with anti-FliC antibody (69, 70) was performed as previously described (23). Briefly, 1 ml of cells [optical density at 600 nm (OD\(_{600}\)) = 0.5] grown in K\(^{+}\)TB (tryptone buffer)
(tryptone, 85 mM KCl) was sheared by passing the cell suspension through a 26-gauge syringe needle 30 times. These cells were then washed three times in 1 ml of motility buffer [K’MB: 85 mM KCl and 10 mM KPi (pH 7.0)] and lastly resuspended in 500 μl of motility buffer. Then, 20 μl of suspension was loaded into a tunnel slide pre-filled with motility buffer that had previously been incubated with anti-FliC antibodies for 15 min at room temperature (1:300 dilution in water). The unbound cells were then removed from the tunnel slide by washing with a total of 200 μl of motility buffer (~10 times the tunnel slide volume). The tethered cell time-lapse videos were recorded at ×40 magnification on a phase contrast microscope (Nikon). Time-lapse videos were collected using a camera (Chameleon3 CM3, Point Grey Research) recording 20-s-long videos at 20 frames/s. Experiments involving single-cell tracking during buffer exchange were recorded at 60 frames/s, with cells washed and resuspended in K’MB + 0.1 mM EDTA-2 K (K’MBEDTA). Free-swimming cells were grown overnight in K’+TB at 30°C to OD_{600} of ~0.5 and then washed three times in 1 ml K’MB before resuspension in 500 μl of K’MB and imaging in a tunnel slide. A custom LabVIEW software (18, 23, 51) was used as previously reported to estimate specific rotational parameters of the tethered cells such as rotation frequency (speed), clockwise and counterclockwise bias, switching frequency, and speed of swimming cells. FliC-sticky *E. coli* RP437 cells (ΔmotAB ΔcheY ΔpilA fiC<sup>Δ</sup>) (71) were used to collect data presented in fig. S11. Visualization of the data was performed using GraphPad Prism 8.

**Fitness comparison assay between Pots and L3.3 (potB G20v)**

Overnight cultures of the two strains grown in K’LB were adjusted to equal OD_{600} and then mixed at 1:1, 1:10, and 1:100 (L3.3: Pots) ratios in a 100-μl volume. Ten microliters of each mixture was then used to inoculate 2 ml of K’LB liquid culture at 30°C for 24 hours. The resulting dense culture was then diluted to OD_{600} of 0.25 and then further diluted 10<sup>6</sup>-fold in K’LB before streaking onto K’LB swim plates (20-μl streaks) and incubating for 48 hours at 30°C.

**SNP analysis**

WGS of 22 *E. coli* strains was performed using a MiSeq 2 × 150–base pair (bp) chip on an Illumina sequencing platform. Sequencing was carried out at the Ramaciotti Centre for Genomics, Kensington and delivered as demultiplexed fastq files (quality control: >80% bases higher than Q30 at 2 × 150 bp). The SNP calling and analysis were performed using Snippy (72, 73). The short reads from sequencing were aligned to the MG1655 reference *E. coli* genome (GenBank: U00096.2) and to a synthetic genome based on MG1655, edited to contain the Pots stator sequences from pPots (pomA/potB) at the motAB locus.

**Transcriptomics**

RNA was extracted from bacterial cultures inoculated with glycerol stocks of the relevant strains and grown in K’LB broth at 30°C until OD_{600} of 0.5. Total RNA was extracted from a 0.5-ml aliquot of the culture using the RNAeasy Protect Bacteria Mini Kit (74524, QIAGEN) with on-column deoxyribonuclease digestion, as indicated in the manufacturer protocol. RNA quality was assessed using a TapeStation System (RNA ScreenTape, Agilent). All RNA samples selected for sequencing had an RNA integrity number > 8. Library preparation and sequencing were performed at the Ramaciotti Centre using the NextSeq 500 platform (Illumina) running for 150 cycles using a mid-output flowcell in paired-end read mode (2 × 75 bp). Fastq files containing the RNA-seq reads underwent quality control using FastQC (74) and then processed with FASTP (version 0.20.1) (75), to remove low-quality reads and trim adaptor sequences. Reads were aligned to the Pots reference genome using HISAT2 (76), transcripts were assembled and quantified using STRINGTIE (73), and differential expression analysis was carried out using DESeq2 (77). Heatmap dendograms were generated using the Heatmap2 function from the R gplots package (Heatmap2). Complete clustering was performed using the Euclidean Distance method. All the analysis tools described above were run on the Galaxy webserver (https://usegalaxy.org/). Nucleotide variations that were present in motB and pomA were quantified from RNA-seq data using the Rsamtools pileup function (78). This involved the writing of custom R scripts (available at https://github.com/VCCRI) that compared Rsamtools output to the genome reference. We then performed pathway analysis to identify which biological processes were relevant to the shared genes using the EcoCyc database (79).

**Structural modeling**

The PomA2PotB<sub>2</sub> model was assembled by modeling each monomer using the Colabfold pipeline (25) and by aligning the resulting monomers to each subunit of the *B. subtilis* MotA<sub>2</sub>St<sub>2</sub> structure (PDB:6YSL) (11).

**Phylogenetics and ancestral reconstruction**

Phylogeny was generated with RAxML-HPC version 8 on XSEDE (80) through the CIPRES Science Gateway (81). The phylogeny was calculated using the PROTGAMMA protein substitution model, LG protein substitution matrix, and a parsimony seed value of 12,345. Ancestral sequences were calculated using CodeML, a maximum likelihood program from the PAML package, using the LG rate file with the Empirical+F model, and using eight categories in dG of NSites models (82). Ancestral sequence reconstructions at each node were used to determine G20/V20 identity at each node. Genomic context for the stators was pulled from the Kyoto Encyclopedia of Genes and Genomes Database (83).

**SUPPLEMENTARY MATERIALS**

Supplementary material for this article is available at https://science.org/doi/10.1126/sciadv.abq2492

View/request a protocol for this paper from Bio-protocol.

**REFERENCES AND NOTES**

1. M. Miyata, R. C. Robinson, T. Q. P. Uyeda, Y. Fukumori, S. I. Fukushima, S. Haruta, M. Homa, K. Inaba, M. Ito, C. Kaito, K. Kató, K. Kenri, Y. Kinosita, S. Kojima, T. Minamino, H. Mori, S. Nakamura, D. Nakane, K. Nakayama, M. Nishiyama, S. Shibata, K. Shimabukuro, M. Yamakoshi, A. Taoka, Y. Tashiro, I. Tulum, H. Wada, K. I. Wakabayashi, Tree of motility – A proposed history of motility systems in the tree of life. Genes Cells 25, 6–21 (2020).
2. D. T. Freahe, B. Mickalide, D. Schnitkey, J. Merritt, T. E. Kuhlman, S. Kuehn, Environment determines evolutionary trajectory in a constrained phenotypic space. eLife 6, e24669 (2017).
3. D. B. Roszak, R. R. Colwell, Survival strategies of bacteria in the natural environment. Microbiol. Rev. 51, 365–379 (1987).
4. S. Gude, E. Pinçe, K. M. Taute, A. B. Seinen, T. S. Shimizu, S. J. Tans, Bacterial coexistence driven by motility and spatial competition. Nature 578, 588–592 (2020).
5. B. Ni, R. Colin, H. Link, R. G. Endres, V. Sourjik, Growth rate dependent resource investment in bacterial motile behavior quantitatively follows potential benefit of chemotaxis. Proc. Natl. Acad. Sci. U.S.A. 117, 595–601 (2020).
6. M. J. Pallen, N. J. Matzke, From the origin of species to the origin of bacterial flagella. Nat. Rev. Microbiol. 4, 784–790 (2006).
9. B. Ni, B. Ghosh, S. F. Palady, R. Colin, T. Heimerl, V. Sourjik, Evolutionary remodeling of bacterial motility checkpoint control. Cell Rep. 18, 866–877 (2017).

10. M. Santivere, A. Roa-Eguiara, C. Kühne, N. Wadhwa, H. C. Berg, M. Erhardt, N. M. I. Taylor, J. C. Deme, S. Johnson, O. Vickery, A. Aron, H. Monkhouse, T. Griffiths, R. H. James, M. Baym, T. D. Lieberman, E. D. Kelsic, R. Chait, R. Gross, I. Yelin, R. Kishony, M. Mirdita, S. Ovchinnikov, M. Steinegger, ColabFold - Making protein folding accessible.

11. J. Takekawa, K. Imada, M. Homma, Structure and energy-conversion mechanism of the bacterial Na+−driven flagellar motor. Trends Microbiol. 28, 719–730 (2020).

12. N. Takekawa, M. Isumi, H. Terashima, S. Zhi, Y. Nishino, M. Sakuma, S. Kojima, M. Homma, K. Imada, Structure of vibrio FliL, a new stomatin-like protein that assists the bacterial flagellar motor function. mbio 10, e00292-19 (2019).

13. M. I. Islam, A. Lin, Y. W. Lai, N. J. Matzke, M. A. B. Baker, Ancestral sequence reconstructions of MotF are proton-motive and require mota for motility. Front. Microbiol. 11, 625837 (2020).

14. N. Takekawa, M. Nishiyama, T. Kanekis, T. Kanai, H. Aomi, S. Kojima, M. Homma, Sodium-driven energy conversion for flagellar rotation of the earliest divergent hyperthermophilic bacterium. Sci. Rep. 5, 12711 (2015).

15. C. R. Reisch, K. L. Prather, The no-SCAR (scarless Cas9 assisted recombinase) system for genome editing in Escherichia coli. Sci. Rep. 5, 15096 (2015).

16. K. A. Dansekon, B. L. Wanner, One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc. Natl. Acad. Sci. U.S.A. 97, 6640–6645 (2000).

17. M. Baym, T. D. Lieberman, E. D. Kelsic, R. Chait, R. Gross, I. Yelin, R. Kishony, M. Mirdita, S. Ovchinnikov, M. Steinegger, ColabFold - Making protein folding accessible.

18. N. J. Matzke, A. Lin, G. Y. Dong, M. Dong, K. Taghizadeh, M. R. Sullivan, C. E. Elmquist, R. P. Cunningham, P. C. Dedon, Defects in purine nucleotide metabolism lead to substantial incorporation of xanthine and hypoxanthine into DNA and RNA. Proc. Natl. Acad. Sci. U.S.A. 109, 2319–2324 (2012).

19. A. J. Lopatin, J. H. Yang, Digital insights into nucleotide metabolism and antibiotic treatment failure. Front Digit Health 3, 583468 (2021).

20. P. L. Foster, H. Lee, E. Popodi, J. P. Townes, H. Tang, Determinants of spontaneous mutation in the bacterium Escherichia coli as revealed by whole-genome sequencing. Proc. Natl. Acad. Sci. U.S.A. 112, E5990–E5999 (2015).

21. R. S. Galhardo, P. J. Hastings, S. M. Rosenberg, Mutation as a stress response and the regulation of evolvability. Crit. Rev. Biochem. Mol. Biol. 42, 399–435 (2007).

22. C. Shee, J. L. Gibson, M. C. Darrow, C. Gonzalez, S. M. Rosenberg, Impact of a stress-inducible switch to mutagenic repair of DNA breaks in mutation in Escherichia coli. Proc. Natl. Acad. Sci. U.S.A. 108, 13659–13664 (2011).

23. M. A. Kohanski, M. A. DePristo, J. C. Collins, Sublethal antibiotic treatment leads to multidrug resistance via radical-induced mutagenesis. Mol. Cell 37, 313–320 (2010).

24. C. P. Selby, Mfd protein and transcription-repair coupling in Escherichia coli. Photobiol. Chem. 93, 280–295 (2017).

25. H. Merrick, Y. Zhang, A. D. Grossman, J. D. Wang, Replication-transcription conflicts in bacteria. Nat. Rev. Microbiol. 10, 449–458 (2012).

26. Y. Onoue, M. Ivaki, A. Shinobu, Y. Nishihara, H. Iwatsuki, H. Terashima, A. Kitao, H. Kendori, M. Homma, Essential ion binding residues for Na+ flow in stator complex of the Vibrio flagellar motor. Sci. Rep. 9, 11216 (2019).

27. N. Terahara, M. Sano, M. Ito, A bacillus flagellar motor that can use both Na+ and K+ as a coupling ion is converted by a single mutation to use only Na+. PLoS ONE 7, e46248 (2012).

28. N. Terahara, T. A. Kruilwish, M. Ito, Mutations after the sodium versus potassium use of a Bacillus clausii flagellar motor and confer dual ion use on Bacillus subtilis motors. Proc. Natl. Acad. Sci. U.S.A. 105, 14359–14364 (2008).

29. Y. Asai, T. Shoji, I. Kawagishi, M. Homma, Cysteine-scanning mutagenesis of the periplasmic loop regions of PomA, a putative channel component of the sodium-driven flagellar motor of Vibrio alginolyticus. J. Bacteriol. 182, 1001–1007 (2000).

30. S. Kojima, T. Shoji, Y. Asai, I. Kawagishi, M. Homma, A slow-motility phenotype caused by substitutions at residue Asp31 in the PomA channel component of a sodium-driven flagellar motor. J. Bacteriol. 182, 3314–3318 (2000).

31. T. Ishida, R. Ito, J. Clark, N. J. Matzke, Y. Sowa, M. A. B. Baker, Sodium-powered stators of the bacterial flagellar motor can generate torque in the presence of phenamil with mutations near the peptidoglycan-binding region. Mol. Microbiol. 111, 1689–1699 (2019).

32. S. Jacques, Y. K. Kim, L. L. McCarter, Mutations conferring resistance to phenamil and amiloride, inhibitors of sodium-driven motility of Vibrio parahaemolyticus. Proc. Natl. Acad. Sci. U.S.A. 96, 5740–5745 (1999).

33. S. Kojima, Y. Asai, T. Atsumi, I. Kawagishi, M. Homma, Na+−driven flagellar motor resistant to phenamil, an amiloride analog, caused by mutations in putative channel components. J. Mol. Biol. 285, 1537–1547 (1999).

34. A. G. Garza, L. W. Harris-Haller, R. A. Stoebner, M. D. Manson, Motility protein interactions of the bacterial flagellar motor can generate torque in the presence of phenamil with mutations near the peptidoglycan-binding region. Mol. Microbiol. 111, 1689–1699 (2019).

35. S. Jaques, Y. K. Kim, L. L. McCarter, Mutations conferring resistance to phenamil and amiloride, inhibitors of sodium-driven motility of Vibrio parahaemolyticus. Proc. Natl. Acad. Sci. U.S.A. 96, 5740–5745 (1999).

36. S. Kojima, Y. Asai, T. Atsumi, I. Kawagishi, M. Homma, Na+−driven flagellar motor resistant to phenamil, an amiloride analog, caused by mutations in putative channel components. J. Mol. Biol. 285, 1537–1547 (1999).
59. J. D. Bloom, F. H. Arnold, In the light of directed evolution: Pathways of adaptive protein evolution. Proc. Natl. Acad. Sci. U.S.A. 106 Suppl 1, 9995–10000 (2009).

60. J. Atia, C. McCluskey, A. S. Shmygol, D. A. Rand, H. A. van den Berg, A. M. Blanks, Reconstruction of cell surface densities of ion pumps, exchangers, and channels from mRNA expression, conductance kinetics, whole-cell calcium, and current-clamp voltage recordings, with an application to human uterine smooth muscle cells. PLoS Comput. Biol. 12, e1004828 (2016).

61. M. J. McDonald, Microbial experimental evolution—A proving ground for evolutionary theory and a tool for discovery. EMBO Rep. 20, e66992 (2019).

62. J. A. Draghi, J. B. Plotkin, Molecular evolution: Hidden diversity sparks adaptation. Nature 474, 45–46 (2011).

63. E. J. Hayden, E. Ferrada, A. Wagner, Cryptic genetic variation promotes rapid evolutionary adaptation in an RNA enzyme. Nature 474, 92–95 (2011).

64. R. Colin, V. Sourij, Emergent properties of bacterial chemotaxis pathway. Curr. Opin. Microbiol. 39, 24–33 (2017).

65. M. Kaplan, C. M. Okonomou, C. R. Wood, G. Chreifi, P. Subramanian, D. R. Ortega, Y.-W. Chang, M. Beeby, C. L. Shaffer, G. J. Jensen, Novel transient cytoplasmic rings stabilize assembling bacterial flagellar motors. EMBO J. 41, e109523 (2022).

66. J. S. Parkinson, Complementation analysis and deletion mapping of Escherichia coli mutants defective in chemotaxis. J. Bacteriol. 135, 45–53 (1978).

67. S. M. Block, D. F. Blair, H. C. Berg. Compliance of bacterial flagella measured with optical tweezers. Nature 338, 514–518 (1989).

68. R. Higuchi, B. Krummel, R. K. Saiki, A general method of in vitro preparation and specific mutagenesis of DNA fragments: Study of protein and DNA interactions. Nucleic Acids Res. 16, 7351–7367 (1988).

69. Y. Kinosita, T. Ishida, M. Yoshida, R. Ito, Y. V. Morimoto, K. Goto, R. M. Berry, T. Nishizaka, Y. Sowa, Distinct chemotactic behavior in the original Escherichia coli K-12 depending on forward-and-backward swimming, not on run-tumble movements. Science 290, 5887 (2020).

70. M. Nishiyama, S. Kojima, Bacterial motility measured by a miniature chamber for high-pressure microscopy. Int. J. Mol. Sci. 13, 9225–9239 (2012).

71. Y. Sowa, A. D. Rowe, M. C. Leake, T. Yakushi, M. Homma, A. Ishijima, R. M. Berry, Direct observation of steps in rotation of the bacterial flagellar motor. Nature 437, 916–919 (2005).

72. S. J. Bush, D. Foster, D. W. Eyre, E. L. Clark, N. de Maio, L. P. Shaw, N. Stoesser, T. E. A. Peto, D. W. Crook, A. S. Walker, Genomic diversity affects the accuracy of bacterial single-nucleotide polymorphism-calling pipelines. Gigascience 9, giaa007 (2020).

73. M. Pertea, G. M. Pertea, C. M. Antonescu, T. C. Chang, J. T. Mendell, S. L. Salzberg, StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat. Biotechnol. 33, 290–295 (2015).

74. S. Andrews, FastQC: A Quality Control Tool for High Throughput Sequence Data. Online http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (2010).

75. S. Chen, Y. Zhou, Y. Chen, J. Gu, fastp: An ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34, 884–890 (2018).

76. D. Kim, B. Langmead, S. L. Salzberg, HISAT: A fast spliced aligner with low memory requirements. Nat. Methods 12, 357–360 (2015).

77. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550 (2014).

78. P. H. Morgan M, Obenchain V, Hayden N, Binary alignment (BAM), FASTA, variant call (BCF), and tabix file import (2021); https://bioconductor.org/packages/Rsamtools.

79. P. D. Karp, M. Riley, M. Saier, I. T. Paulsen, J. Collado-Vides, S. M. Paley, A. Pellegrini-Toole, C. Bonavides, S. Gama-Castro, The EcoCyc database. Nucleic Acids Res. 30, 56–58 (2002).

80. A. Stamatakis, RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313 (2014).

81. M. A. Miller, W. Pfeiffer, T. Schwartz, in Proceedings of the 2011 TeraGrid Conference: Extreme Digital Discovery. (2011), pp. 1–8.

82. Z. Yang, PALM: A program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. 13, 555–556 (1997).

83. M. Kanehisa, S. Goto, KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 28, 27–30 (2000).

Acknowledgments: We would like to acknowledge M. Yoshida and R. Ito for technical assistance. Funding: D.T.H. and E.G. acknowledge The Victor Chang Cardiac Research Institute Innovation Centre (funded by the New South Wales Government Ministry of Health), Y.S. was supported by JSPS KAKENHI (JP19H02475 and JP20K06564), MEXT KAKENHI (JP19H05404, JP 21H00410, and JP21H05892), and Takeda Science Foundation. M.A.B.B. was supported by a UNSW Scientia Research Fellowship, a CSIRO Future Science Platform in Synthetic Biology Project Grant, an ARC Discovery Project Grant DP190100497, and an HFSP Young Investigator Project Grant RGY0072/21. Author contributions: P.R. and M.A.B.B. designed the experiments and executed the experiments in strain editing, molecular biology, and microbiology. T.J and Y.S executed the experiments in strain editing and microbiology. M.A.B.B. executed bioinformatics surrounding variant calling. A.L and M.A.B.B. executed bioinformatics surrounding phyllogenetics. P.R., D.T.H., E.G., and M.B. supervised the design, execution, and writing of the project. All authors contributed to writing and revision of the manuscript. Competing interests: The authors declare that they have no competing interests. Data and materials availability: All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. All WGS and RNA-seq data are deposited in BioProject, accession: PRJNA729860. Pots genome is deposited in GenBank: CP083410.1.

Submitted 27 March 2022
Accepted 6 October 2022
Published 23 November 2022

10.1126/sciadv.abq2492