Spontaneous adaptation of ion selectivity in a bacterial flagellar motor

Pietro Ridone[1], Tsubasa Ishida[2], Angela Lin[1], David T Humphreys[3], Eleni Giannoulatou[3], Yoshiyuki Sowa[2,4], Matthew A. B. Baker[1,5*]

[1]School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, Australia.

[2]Department of Frontier Bioscience, Hosei University, Tokyo, Japan.

[3]Victor Chang Cardiac Research Institute, Sydney, Australia

[4]Research Center for Micro-Nano Technology, Hosei University, Tokyo, Japan.

[5]CSIRO Synthetic Biology Future Science Platform, Brisbane, Australia.

*correspondence: matthew.baker@unsw.edu.au

ABSTRACT

Motility provides a selective advantage to many bacterial species and is often achieved by rotation of flagella that propel the cell towards more favourable conditions. In most species, the rotation of the flagellum, driven by the Bacterial Flagellar Motor (BFM), is powered by H^+ or Na^+ ion transit through the torque-generating stator subunit of the motor complex. The ionic requirements for motility appear to have adapted to environmental changes throughout history but the molecular basis of this adaptation remains unknown. Here we used CRISPR engineering to replace the native *Escherichia coli* H^+-powered stator with Na^+-powered stator genes and report the rapid and spontaneous reversion to H^+-power in a low sodium environment. We follow the evolution of the stators during their reversion to H^+-powered motility and use whole genome sequencing to identify both flagellar- and non-flagellar-associated genes involved in the cell's adaptation to new power sources. This work highlights the utility of the flagellar stator system for studying the molecular mechanisms underlying adaptation and demonstrates how environmental change can rapidly alter the function of an ion transporter.

INTRODUCTION

Bacterial motility via the flagellar motor represents one of the earliest forms of locomotion (Miyata et al., 2020). This rotary motility imparts such significant selective advantage (Roszak and Colwell 1987, Fraebel, Mickalide et al. 2017) that resources are allocated to chemotaxis even in the absence of nutrient gradients (Gude, Pince et al. 2020, Ni, Colin et al. 2020). Furthermore, the evolutionary origins, and subsequent adaptation of the motor are of significant scientific and public interest (Pallen and Matzke 2006), since the BFM holds prominence as an ancient, large, molecular complex of high sophistication.

The torque that drives the BFM is supplied by motor-associated transmembrane protein-complexes known as stators. The stator complex, an asymmetric heteroheptamer (in *E. coli:* $MotA_5MotA_2$) most likely acts itself as a miniature rotating nanomachine coupling ion transit to rotation (Deme, Johnson et al. 2020, Santiveri, Roa-Eguiara et al. 2020). 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 (presence of the coupling ion) to an easily discernible phenotype (cell swimming). Furthermore, the stators have been subject to protein engineering approaches for many years, in particular the synthesis of chimeric stator constructs that enable the motor of *E. coli*, natively proton-driven, to be powered by sodium ion flow (Asai, Yakushi et al. 2003).The majority of stators are proton driven, but many that are sodium-driven can be found in nature, and this divergence is presumed to have occurred in the distant past (Takekawa, Isumi et al. 2019, Islam, Lin et al. 2020, Lai, Ridone et al. 2020). Past reports have argued that H^+-coupled motility diverged from Na^+-coupled machinery in ancestral times (Takekawa, Nishiyama et al. 2015), but the molecular basis for this adaptation, and the evolutionary landscape that constrains stator adaptation remains unclear.

In order 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 non-lethal environment (soft agar swim plate) which lacked the power source for the stator (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 CRISPR/Cas9 (Reisch and Prather 2015) and λ-Red (Datsenko and Wanner 2000)) to replace the native *motA motB* stator genes of the *E. coli* BFM with chimeric sodium-powered *pomA potB* (henceforth *pots*) stator genes derived from *Vibrio alginolyticus* (Asai, Yakushi et al. 2003). We transplanted the *pots* stator genes at the same location and orientation of the native *motA motB* 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 (Baym, Lieberman et al. 2016).

RESULTS

The RP437 strain was edited to carry the *pots* stator genes in place of the native *E. coli motAmotB* genes (Fig. 1A) via the no-SCAR Method (Reisch and Prather 2015) and traditional λ-Red recombineering respectively (Supplementary Fig. 1 & 2). A single no-SCAR Pots clone was selected and tested on swim plates (Supplementary Fig. 3) after verification of successful editing by colony PCR and Sanger Sequencing (Supplementary Fig. 4). The edited strain was able to swim on sodium-rich (85 mM NaCl) soft agar plates but not on potassium-rich sodium-poor (67 mM KCl, ~8 mM [Na^+]) plates (Supplementary Fig. 3A). This edited strain exhibited the same swimming behaviour as the control stator-less strain with motility restored via an inducible plasmid vector that could express the *Pots* construct (RP6894 Δ*motAmotB* + pSHU1234, hereby pPots).

We next challenged this *pots* strains to survive on K^+ based soft agar for prolonged periods (Supplementary Fig. 3B). 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 5 passages at 3-4 days intervals (Supplementary Fig. 5). When multiple flares occurred in a single swim ring, each was individually passaged (Supplementary Fig. 3B), and could be recapitulated (Fig. 1). Directed evolution consistently generated swimming flares when *pots* clones were cultured on agar containing yeast extract and tryptone (~8 mM [Na^+]), but not on minimal media (~ 1mM total [Na^+]) or when the *pots* construct was encoded on a plasmid (Supplementary Fig. 6). One *pots* strain generated using λ-Red methods

95 (Datsenko and Wanner 2000), which carried the native *V. alginolyticus pomA*
96 promoter, also successfully produced flares (Supplementary Fig. 7).

97 Lineages were selected for whole-genome sequencing (WGS) after a preliminary
98 screening for mutations in the stator genes by Sanger-sequencing PCR amplicons
99 spanning the genomic *pomApotB* locus (Fig. 1). Variant calling to the MG1655
100 reference genome was used to compare single nucleotide polymorphisms (SNPs)
101 between members of the same lineage. Our intended *pomApotB* edit was the only
102 difference between the RP437 and *Pots* genomes, indicating that neither no-SCAR
103 nor λ-Red editing had resulted in off-target edits. 153 SNPs were called as variants
104 between our experimental parent RP437 and the MG1655 reference which were
105 shared across all lineage members (Supplementary Table 1).

106 Several lineages whose descendants could swim in reduced sodium had mutations at
107 the *pots* locus (L3.3-4-5: PotB G20V; L6.4-5: PomA L183F; L8.3-4-5: PotB G20W). In
108 contrast, lineages passaged only on 85 mM Na^+ agar (L1 & L2) accumulated mutations
109 not in stators but in the flagellar components. Lineages passaged on ~8 mM sodium-
110 poor agar whose descendants could not swim (L4) had no mutations on any flagellar
111 genes.

112 To determine which genes may be involved in the adaptation process, we performed
113 RNAseq experiments to measure transcript levels for two lineages which evolved
114 different stator mutations over different lengths of time (Fig. 2, Supplementary Table
115 2). These were: PomA L183F (Na+ powered phenotype, 14 days) and PotB G20V (H+
116 phenotype, 7 days). From these, we selected the common ancestor to the two lineages
117 (Pots), the last lineage member before the mutation occurred (L5.3 and L3.2, the pre-
118 fixation sample) and the first available member carrying that lineage mutation on the
119 chromosome (L6.4 and L3.3, the fixed variant). Of note, L5.3 displayed improved
120 swimming despite being isogenic to its parent strain, the Pots common ancestor.
121 Member L3.2 of the G20V lineage, the immediate predecessor to L3.3, in contrast,
122 was non-motile. We closely examined the transcripts mapped to the PomA/PotB locus
123 in all samples and saw no significant enrichment of mutant transcripts above the
124 background noise in the pre-fixation samples (Supplementary Fig. 8). This confirmed
125 that a chromosomal mutation, rather than a RNA editing event, was responsible for
126 the observed stator variants. To identify common processes which led to variant
127 fixation in these two lineages we focused first on the differentially expressed genes
128 present in both lineages (Fig. 2). We then performed pathway analysis to identify which

129 biological processes were relevant to the shared genes and analysed which regulatory
130 elements controlled their expression.

131 Both lineages increased expression of *lrhA* at the pre-fixation point, which acts as an
132 inhibitor of the *flhDC* flagellar master regulator. Expression response was dominated
133 by pathways under control of RpoD (σ70) /FliA (σ28) in L183F, and RpoD (σ70) in
134 G20V (Supplementary Fig. 9). The prominence of FliA regulation in L183F was
135 reflected by upregulation of motility genes at the fixation stage. Those same genes
136 were not significantly affected in the G20V lineage, which instead promoted primarily
137 adhesion/biofilm behavior in response to the changed environment.

138 We characterized rotational phenotypes of the parent and evolved strains in the
139 presence and absence of sodium using a tethered cell assay with particular focus on
140 the upmotile G20V variant, L3.3 (Fig. 3, Supplementary Fig. 10). The *potB* G20V fixed
141 variant was clearly distinguishable from pots in this assay, and maintained rotation in
142 both the absence of sodium and in the presence of the sodium blocker phenamil (Fig.
143 3A). We confirmed the PotB G20V H^+-powered phenotype by introducing the same
144 point mutation (GGG to GTG) on our plasmid vector (pPots) and testing motility in
145 tethered cell assay when the protein was expressed in the stator-less Δ*motAB* RP6894
146 (Fig. 3B). Measurement of rotational speed vs lithium and sodium showed that the
147 PotB-G20V variant was motile at 0 mM Na+ and 0 mM Li+, but showed a dependence
148 on the concentration of ion presence, similar to pots, and in contrast to MotA/MotB
149 (Supplementary Fig. 11A-C). All stator types had cessation of rotation at 50 μM CCCP,
150 and showed no measurable effect as pH was varied between 6.0 and 8.0
151 (Supplementary Fig. 11DE).

152 We further tracked single cell rotation in the tethered cell assay with sequential
153 exchange of buffers (Supplementary Fig. 12). There, *potB* G20V (L3.3) actively rotated
154 in sodium-free buffers and in presence of 100 μM Phenamil – indicating sodium-free
155 rotation and a disrupted phenamil binding site. We further synthesised all possible
156 variants at site G20, and of these, only G20V was able to rotate in the absence of
157 sodium (Supplementary Fig. 13).

158 We examined the natural prevalence of glycine and valine at site 20 across a
159 phylogeny from a collection of 82 MotB sequences. Of these sequences, G20 was
160 present only in the clade corresponding to *Vibrio sp.* PomB (7 sequences), whereas
161 V20 was distributed across the phylogeny more broadly (41 sequences)

162 (Supplementary Fig. 14). Ancestral reconstruction across every node of this phylogeny
163 predicted that the ancestral phenotype was G20.

164 To examine mutation reproducibility in the stators at higher throughput, we subjected
165 55 *pots* colonies to directed evolution in 8 mM [Na^+] agar. These yielded a total of 42
166 flares within the first 3 days of incubation, which were then passaged four more times
167 at 3-day intervals. At the end of this experiment we selected the 20 terminal lineage
168 members which produced the largest swim rings and Sanger-sequenced them at the
169 *pots* locus. For these we observed a total of 5 mutations in *pomA* (S25C, D31N,
170 P177A, P177Q and L183F) and one more new mutation in *potB* (L36Q) (Fig. 2B). We
171 mapped monomer models for PomA and PotB generated using Alphafold (Mirdita,
172 Schütze et al. 2021) to a model $PomA_5PotB_2$ complex using the recently published
173 high-resolution *B. subtilis* $MotA_5B_2$ structure (Deme, Johnson et al. 2020) (Fig. 2C &
174 2D). All stator mutations accumulated at sites proximal to or within the predicted ion-
175 transport pore, at the interface between the PotB transmembrane domain and the third
176 and fourth transmembrane domains of PomA (Fig. 2D).

177 Over the course of these experiments, we found that three stator residues underwent
178 mutation at the same site twice (PomA L183F (2x), PomA P177A & P177Q, PotB
179 G20V & G20W). WGS revealed that the *pitA* gene had mutated in three separate
180 lineages (L2.5, L3.5, L8.5) with one of the mutations occurring twice (PitA W112*Stop).

181 To test for the capacity for reversion, we took sequenced lineages that swam in 8 mM
182 [Na^+] agar (L3.3-4-5, L6.4-5, L8.3-4-5) and reintroduced them to an environment with
183 85 mM [Na^+] (Supplementary Fig. 15A). After 10 rounds of daily passaging, no
184 reversion in the mutants that had enabled low-sodium swimming was observed, with
185 only a single additional mutation gained: a *potB* T21A mutation in the terminal
186 descendant of *potB* G20W *pitA* W112*stop (L8.5).

187 Finally, we tested whether evolution could be more easily directed on minimal media
188 when starting from a more favourable vantage point. We examined all stator mutants
189 that swam on 8 mM [Na^+] plates in conditions of further sodium scarcity (minimal
190 media: 1 mM [Na^+]) (Supplementary Fig. 15B). Initially, only the strains with *potB* G20V
191 and *pomA* P177Q mutations could swim, and this capacity was maintained following
192 five passages over 12 days. Sanger sequencing revealed that the terminal descendant
193 of p*otB* G20V *pitA* G432R mutant (L3.5) gained a further mutation in *pomA* (M20L). A

194 summary of all stator gene mutations obtained from all directed evolution experiments
195 is provided in Fig. 2B.
196

DISCUSSION

Previous reports have shown that bacterial motility can adapt (Ni, Ghosh et al. 2017) and be rescued (Taylor, Mulley et al. 2015) via remodelling of the flagellar regulatory network. Ni et al. observed that evolutionary adaptation of motility occurs via remodeling of the checkpoint regulating flagellar gene expression (Ni, Ghosh et al. 2017). Their experiments tracked adaptive changes in swim plates, matching our experiments, however 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 the amongst the first genes to mutate in the improved swimmer population but they did not report any changes in *flgL* (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 *P. fluorescens*, which function can be substituted by distantly related homologous proteins (Taylor, Mulley et al. 2015).

In contrast, our *E. coli pots* strain faced selective pressure from ion scarcity. Our scenario is reminiscent of previous semi-solid agar experimental evolution studies on the adaptation of antibiotic resistance and recapitulates similar results. In the MEGA plate experiments of Baym et al., they similarly saw that the phosphate transporter *pitA* was repeatedly mutated, often to a frameshifted or nonsense variant (Baym, Lieberman et al. 2016). In similar experiments the isocitrate dehydrogenase *icd* was also seen to mutate often (Fraebel, Mickalide et al. 2017, Anand, Olson et al. 2019).

Differential expression analysis of our RNAseq datasets revealed that our two chosen lineages underwent different transcriptomic pathways before fixing a mutation in the stator genes. While the L183F lineage displayed many upregulated genes involved in motility, the G20V lineage was found to regulate genes involved in biofilm formation.

The motile behaviour of the L5.3 member (pre-L183F mutation stage) was reflected by the upregulation of flagellar and chemotaxis genes. The genes involved have also been measured as upregulated in a similar study on the adaptation of *E. coli* to swimming in soft agar (Ni, Ghosh et al. 2017) which may explain how the strain can swim in very low Na+ through upregulating several flagellar components including the chimeric sodium stators themselves (PomA and PotB). In contrast, the expression profile of the Na+-only 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

231 nucleotides and the transition to a biofilm lifestyle. Roughly 10% of genes showed
232 significant changes in transcription in both lineage trios (10.8% and 13.2% in L183F
233 and G20V respectively) and all of these were upregulated during adaptation.

234 When the G20V mutation was fixed, L3.3 (G20V) was found to upregulate chemotaxis
235 and motility genes, and other markers of adaptation to soft agar (cheA, trp, tar, tap)(Ni,
236 Ghosh et al. 2017). Both lineages upregulated the flhDC regulator LrhA at the pre-
237 mutation stage, suggesting that downregulation of the flagellar biosynthetic cascade
238 is a shared trait contributing to stator adaptation. Similarly, both lineages upregulated
239 nucleotide catabolic processes and salvage pathways, a feature also observed also in
240 antibiotic response and which can affect mutation rates by disturbing the NTP balance
241 in the cellular pool (Pang, McFaline et al. 2012, Foster, Lee et al. 2015, Lopatkin and
242 Yang 2021).

243 Mutations in stators are known to affect ion usage and may confer dual-ion coupling
244 capacity. For example, the substate preference of the *B. alcalophilus* MotPS stator
245 (Na^+/K^+ and Rb^+) was changed with the single mutation M33L in MotS, causing the
246 loss of both K^+- and Rb^+-coupling motility in *E. coli* (Terahara, Sano et al. 2012).
247 Similarly a bi-functional *B. clausii* MotAB stator (Na^+/H^+) triple mutant (MotB V37L,
248 A40S and G42S) was selective only for sodium ions while the combination of
249 mutations G42S, Q43S and Q46A made MotB selective only for H^+ (Terahara,
250 Krulwich et al. 2008). The previously reported S25C (Asai, Shoji et al. 2000) and D31N
251 (Kojima, Shoji et al. 2000) amino acid substitutions in PomA have been shown to
252 reduce motility and, in the case of D31N, affect ion usage. Furthermore, single point
253 mutations in stator genes of *Vibrio* spp. (eg. *pomB* G20V / G20R / P16S) have been
254 shown to impart phenamil resistance both in *Vibrio alginolyticus* and in our pots strain
255 (Jaques, Kim et al. 1999, Kojima, Asai et al. 1999, Ishida, Ito et al. 2019).

256 The PotB-G20V variant directly evolved here is capable of rotation in the total absence
257 of sodium, however its swimming speed does increase with increased availability of
258 either sodium or lithium, with an energisation profile more akin to the parental pots
259 strain than the proton powered *E. coli* wild-type (Supplementary Fig. 11). This perhaps
260 suggests that PotB-G20V is an opportunistic stator adaptation that enables
261 promiscuity, allowing the passage of protons in sodium-scarce conditions.

262 The key difference in this work compared with previous efforts for directed evolution
263 of the stators via mutagenesis (Garza, Harris-Haller et al. 1995, Kojima, Kuroda et al.

1999) 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 not possible in *Vibrio sp.* since *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 confounding, essential systems. Nevertheless, the observation of no revertants agrees with previous work suggesting that requirements for Na^+ binding are more strict than for H^+ binding, and that mutations that convert a Na^+ motor to an H^+ are more accessible than the reverse (Hase, Fedorova et al. 2001, Mulkidjanian, Galperin et al. 2008).

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 *pom*A L183F lineages). Stator genes were the first to mutate in all of our WGS lineages under pressure for H^+ motility. Given this was from a clonal population under identical environmental constraints, it suggests that adaptation of the stators is prioritized in changing environments. Exactly how this may occur remains unclear. Mutation of pore-proximal residues into hydrophobic residues (eg. 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 (eg. F,W) resulted in a similar G20V-like H^+-powered phenotype (Supplementary Fig. 13C). 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) (Mandadapu, Nirody et al. 2015).

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 broader distribution, 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.

As noted by pioneers in directed evolution (Bloom and Arnold 2009), caution is required when applying learnings from directed evolution to natural evolution since

298 selection pressures in the wild are not typically general. In this study we have
299 leveraged our system, and our experimental design, to obtain large, measureable
300 phenotypic change through a single mutation at the G20 locus. The stators of the
301 flagellar motor appear ready to evolve in our experiments: they do not require transit
302 through additional cryptic or neutral mutations, and thus are a model system for
303 exploring molecular adaptation of ion selectivity. Alternative candidates such as
304 sodium porters and pumps often have built in redundancy (Atia, McCloskey et al.
305 2016), and many marine microbes require sodium to be viable and thus evolutionary
306 pressure cannot be applied with such specificity to a single protein complex. In our
307 idealised system, we are able to examine isogenic bacterial colonies without
308 competitive effects, in a homogeneous medium without local niches. Nevertheless, it
309 still remains difficult to quantify rates of adaptation (McDonald 2019). For these
310 reasons, our system is optimised to produce stator variants, but it may be that this
311 ease for adaptation via single-point mutation is accessible precisely because cryptic
312 mutations have been accumulated due to exposure to changing envinronments in the
313 distant past (Draghi and Plotkin 2011, Hayden, Ferrada et al. 2011).

314 Overall, motility confers a fitness advantage that is worth significant energetic
315 investment despite the high cost of synthesizing the flagellar machinery (Colin and
316 Sourjik 2017, Ni, Colin et al. 2020). This advantage can only be seized if the correct
317 ions are available for stator-conversion into torque, otherwise the cell appears to
318 prioritize the adaptation of the stator genes to harness other available ions to restore
319 function. CRISPR/Cas9 has become a widespread method used for precise genomic
320 edits, yet the reversion of such edits appears rapid and gene targeted when focused
321 selective pressure is applied. Here, ion scarcity supplied a strong selective pressure
322 that allowed us to identify novel mutations which were correlated with altered ion-
323 specificity. Our transplant of an unfit protein and the cells' rapid reversion of this edit
324 demonstrates tight evolutionary regulation of the stator subunit in an ancient molecular
325 complex.

326 **MATERIALS AND METHODS**

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

328 *E. coli* strain RP437 was used as the parent strain for genomic editing experiments
329 (Parkinson 1978, Block, Blair et al. 1989). The pSHU1234 (pPots) plasmid encoding
330 *pomA* and *potB* (Ishida, Ito et al. 2019) was used as the template to generate the
331 double stranded donor DNA. This was used to replace the *motA* and *motB* gene on

332 the RP437 chromosome. Liquid cell culturing was done using LB broth (NaCl or KCl,
333 Yeast Extract, Bacto Tryptone). Cells were cultured on agar plates composed of LB
334 Broth and 1% Bacto Agar (BD, U.S.A.). Swim plate cultures were performed on the
335 same substrates adjusted for agar content (0.3% Bacto Agar). Minimal Media (MM)
336 was used to replace Yeast Extract and Tryptone in soft agar swim plates. MM
337 composition: 10 mM KPO_4 , 1mM $(NH_4)_2SO_4$, 1mM $MgSO_4$, 1 µg/ml Thiamine, 0.1 mM
338 of each of the amino acids Thr, Leu, His, Met and Ser.

339 **Editing *E. coli* with Cas9-assisted Recombineering.**

340 This procedure was adapted from the no-SCAR method (Reisch and Prather 2015) .
341 The target strain to be edited (*E. coli* RP437) was sequentially transformed first with
342 the pKD-sgRNA-3'MotA (Sm^+) plasmid, encoding a sgRNA sequence directed at the
343 3' end of *motA*, and then with the pCas9cr4 (Cm^+) plasmid to yield a parent strain
344 harboring both plasmids. The overlap extension PCR technique (Higuchi, Krummel et
345 al. 1988) was employed to assemble linear double stranded DNA molecules (dsDNA)
346 using 3 starting dsDNA fragments. The resulting donor DNA was electroporated in the
347 plasmid-bearing host and the successfully edited clones were selected via colony PCR
348 and Sanger sequencing of the *motAB* locus. A list of primers and PCR protocols used
349 in this work is provided in Supplementary Fig. 16.

350 **Construction of Pots by λ-Red Recombineering**

351 Chromosomal replacement from *motAmotB* to *pomApotB* was achieved by using a λ
352 Red recombination system, with plasmid pKD46 encoding the Red system and
353 positive selection for the recovery of swimming ability (Kinosita, Ishida et al. 2020).
354 Motile clones were selected by isolating motile flares on swim plates.

355 **Tethered cell assay preparation and analysis**

356 The tethered cell assay with anti-FliC-antibody was performed as previously described
357 (Nishiyama and Kojima 2012, Kinosita, Ishida et al. 2020). The tethered cells time
358 lapse videos were recorded at 40x magnification on a phase contrast microscope
359 (Nikon). Time lapse videos were collected using a camera (Chameleon3 CM3, Point
360 Grey Research) recording 20s-long videos at 20 frames per second. Experiments
361 involving single cell tracking during buffer exchange were recorded at 60 frames per
362 second. A custom LabView software (Ishida, Ito et al. 2019, Islam, Lin et al. 2020) was
363 employed as previously reported to estimate specific rotational parameters of the
364 tethered cells such as rotation frequency (speed), clockwise and counterclockwise

365 bias and switching frequency. Visualization of the data was performed using Graph
366 Pad Prism 8.

367 **Single Nucleotide Polymorphism (SNP) analysis**

368 Whole genome sequencing of 22 *E. coli* strains was performed using a MiSeq 2 x
369 150bp chip on an Illumina sequencing platform. Sequencing was carried out at the
370 Ramaciotti Centre for Genomics, Kensington and delivered as demultiplexed fastQ
371 files (Quality Control: >80% bases higher than Q30 at 2×150 bp). The SNP calling and
372 analysis was performed using Snippy (Pertea, Pertea et al. 2015, Bush, Foster et al.
373 2020). The short reads from sequencing were aligned to the MG1655 reference *E. coli*
374 genome (GenBank: U00096.2) and to a synthetic genome based on MG1655, edited
375 to contain the Pots stator sequences from pPots (*pomA*/*potB*) at the *motAB* locus.

376 **Transcriptomics**

377 RNA was extracted from bacterial cultures inoculated with glycerol stocks of the
378 relevant strains and grown in K+ TB broth at 30°C until OD_{600}=0.5. Total RNA was
379 extracted from a 0.5 ml aliquot of the culture using the RNAeasy Protect Bacteria Mini
380 Kit (74524, QIAGEN) with on-column DNAse digestion, as indicated in the
381 manufacturer protocol. RNA quality was assessed using a TapeStation System (RNA
382 ScreenTape, Agilent). All RNA samples selected for sequncing had an RNA Integrity
383 Number (RIN) > 8. Library preparation and sequencing were performed at the
384 Ramaciotti Centre using the NextSeq 500 platform (Illumina) running for 150 cycle
385 using a MID flowcell in paired-end read mode (2x75bp). Fastq files containing the
386 RNAseq reads underwent quality control using FastQC (Andrews 2010) and then
387 processed with FASTP (Version 0.20.1)(Chen, Zhou et al. 2018), to remove low quality
388 reads and trim adaptor sequences. Reads were aligned to the Pots reference genome
389 using HISAT2(Kim, Langmead et al. 2015), transcripts were assembled and quantified
390 using STRINGTIE (Pertea, Pertea et al. 2015) and differential expression analysis was
391 carried out using DESeq2 (Love, Huber et al. 2014). Heatmaps dendrograms were
392 generated using the Heatmap2 function from the R gplots package (Heatmap2).
393 Complete clustering was performed using the Euclidean Distance method. All the
394 analysis tools described above were run on the Galaxy webserver
395 (https://usegalaxy.org/). Nucleotide variations that were present in motB and pomA
396 were quantified from RNA-Seq data using the Rsamtools pileup function (Morgan M
397 2021). This involved the writing of custom R scripts (available upon request) that
398 compared Rsamtools output to the genome reference.

399 We then performed pathway analysis to identify which biological processes were
400 relevant to the shared genes using the EcoCyc database (Karp, Riley et al. 2002).
401

402 **Structural Modelling**

403 The $PomA_5PotB_2$ model was assembled by modelling each monomer using the
404 Colabfold pipeline (Mirdita, Schütze et al. 2021), and by aligning the resulting
405 monomers to each subunit of the *B. subtilis* $MotA_5B_2$ structure (PDB:6YSL)(Deme,
406 Johnson et al. 2020).
407

408 **Phylogenetics and Ancestral Reconstruction**

409 Phylogeny was generated with RAxML-HPC v.8 on XSEDE (Stamatakis 2014) through
410 the CIPRES Science Gateway (Miller, Pfeiffer et al. 2011). The phylogeny was
411 calculated using the PROTGAMMA protein substitution model, LG protein substitution
412 matrix, and a parsimony seed value of 12345. Ancestral sequences were calculated
413 using CodeML, a maximum likelihood program from the PAML package, using the LG
414 rate file with the Empirical+F model and using 8 categories in dG of NSsites models
415 (Yang 1997). Ancestral Sequence Reconstructions at each node were used to
416 determine G20/V20 identity at each node. Genomic context for the stators was was
417 pulled from the KEGG Database (Kanehisa and Goto 2000).
418

419 **AUTHOR CONTRIBUTIONS**

420 PR and MABB designed experiments and executed experiments in strain editing,
421 molecular biology and microbiology. TI and YS executed experiments in strain
422 editing. MABB executed bioinformatics surrounding variant calling. AL and MABB
423 executed bioinformatics surrounding phylogenetics. PR, DTH, EG and MB executed
424 bioinformatics surrounding the transcriptomics. MB supervised the design, execution
425 and writing of the project. All authors contributed to writing and revision of the
426 manuscript.

427 **ACKNOWLEDGEMENTS**

428 We would like to acknowledge Myu Yoshida and Rie Ito for technical assistance.

429 **FUNDING**

430 YS was supported by JSPS KAKENHI (JP18H02475 and JP20K06564), MEXT
431 KAKENHI (JP19H05404 and JP 21H00410) and Takeda Science Foundation. MABB
432 was supported by a UNSW Scientia Research Fellowship, a CSIRO Synthetic

Biology Future Science Platform 2018 Project Grant, and ARC Discovery Project DP190100497.

436 **REFERENCES**

437 Anand, A., C. A. Olson, L. Yang, A. V. Sastry, E. Catoiu, K. S. Choudhary, P. V. Phaneuf, T. E. Sandberg,
438 S. Xu, Y. Hefner, R. Szubin, A. M. Feist and B. O. Palsson (2019). "Pseudogene repair driven by
439 selection pressure applied in experimental evolution." Nat Microbiol **4**(3): 386-389.
440 Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online].
441 Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
442 Asai, Y., T. Shoji, I. Kawagishi and M. Homma (2000). "Cysteine-scanning mutagenesis of the
443 periplasmic loop regions of PomA, a putative channel component of the sodium-driven flagellar
444 motor in Vibrio alginolyticus." J Bacteriol **182**(4): 1001-1007.
445 Asai, Y., T. Yakushi, I. Kawagishi and M. Homma (2003). "Ion-coupling determinants of Na+-driven
446 and H+-driven flagellar motors." J Mol Biol **327**(2): 453-463.
447 Atia, J., C. McCloskey, A. S. Shmygol, D. A. Rand, H. A. van den Berg and A. M. Blanks (2016).
448 "Reconstruction of Cell Surface Densities of Ion Pumps, Exchangers, and Channels from mRNA
449 Expression, Conductance Kinetics, Whole-Cell Calcium, and Current-Clamp Voltage Recordings, with
450 an Application to Human Uterine Smooth Muscle Cells." PLoS Comput Biol **12**(4): e1004828.
451 Baym, M., T. D. Lieberman, E. D. Kelsic, R. Chait, R. Gross, I. Yelin and R. Kishony (2016).
452 "Spatiotemporal microbial evolution on antibiotic landscapes." Science **353**(6304): 1147-1151.
453 Block, S. M., D. F. Blair and H. C. Berg (1989). "Compliance of bacterial flagella measured with optical
454 tweezers." Nature **338**(6215): 514-518.
455 Bloom, J. D. and F. H. Arnold (2009). "In the light of directed evolution: pathways of adaptive protein
456 evolution." Proc Natl Acad Sci U S A **106 Suppl 1**: 9995-10000.
457 Bush, S. J., D. Foster, D. W. Eyre, E. L. Clark, N. De Maio, L. P. Shaw, N. Stoesser, T. E. A. Peto, D. W.
458 Crook and A. S. Walker (2020). "Genomic diversity affects the accuracy of bacterial single-nucleotide
459 polymorphism-calling pipelines." Gigascience **9**(2).
460 Chen, S., Y. Zhou, Y. Chen and J. Gu (2018). "fastp: an ultra-fast all-in-one FASTQ preprocessor."
461 Bioinformatics **34**(17): i884-i890.
462 Colin, R. and V. Sourjik (2017). "Emergent properties of bacterial chemotaxis pathway." Curr Opin
463 Microbiol **39**: 24-33.
464 Datsenko, K. A. and B. L. Wanner (2000). "One-step inactivation of chromosomal genes in Escherichia
465 coli K-12 using PCR products." Proc Natl Acad Sci U S A **97**(12): 6640-6645.
466 Deme, J. C., S. Johnson, O. Vickery, A. Aron, H. Monkhouse, T. Griffiths, R. H. James, B. C. Berks, J. W.
467 Coulton, P. J. Stansfeld and S. M. Lea (2020). "Structures of the stator complex that drives rotation of
468 the bacterial flagellum." Nat Microbiol **5**(12): 1553-1564.
469 Deme, J. C., S. Johnson, O. Vickery, A. Muellbauer, H. Monkhouse, T. Griffiths, R. H. James, B. C.
470 Berks, J. W. Coulton, P. J. Stansfeld and S. M. Lea (2020). "Structures of the stator complex that
471 drives rotation of the bacterial flagellum." Nature Microbiology **5**(12): 1553-+.
472 Draghi, J. A. and J. B. Plotkin (2011). "Molecular evolution: Hidden diversity sparks adaptation."
473 Nature **474**(7349): 45-46.
474 Foster, P. L., H. Lee, E. Popodi, J. P. Townes and H. Tang (2015). "Determinants of spontaneous
475 mutation in the bacterium Escherichia coli as revealed by whole-genome sequencing." Proc Natl
476 Acad Sci U S A **112**(44): E5990-5999.
477 Fraebel, D. T., H. Mickalide, D. Schnitkey, J. Merritt, T. E. Kuhlman and S. Kuehn (2017).
478 "Environment determines evolutionary trajectory in a constrained phenotypic space." Elife **6**.
479 Garza, A. G., L. W. Harris-Haller, R. A. Stoebner and M. D. Manson (1995). "Motility protein
480 interactions in the bacterial flagellar motor." Proc Natl Acad Sci U S A **92**(6): 1970-1974.
481 Gude, S., E. Pince, K. M. Taute, A. B. Seinen, T. S. Shimizu and S. J. Tans (2020). "Bacterial coexistence
482 driven by motility and spatial competition." Nature **578**(7796): 588-592.
483 Hase, C. C., N. D. Fedorova, M. Y. Galperin and P. A. Dibrov (2001). "Sodium ion cycle in bacterial
484 pathogens: evidence from cross-genome comparisons." Microbiol Mol Biol Rev **65**(3): 353-370, table
485 of contents.
486 Hayden, E. J., E. Ferrada and A. Wagner (2011). "Cryptic genetic variation promotes rapid
487 evolutionary adaptation in an RNA enzyme." Nature **474**(7349): 92-95.

488 Higuchi, R., B. Krummel and R. K. Saiki (1988). "A general method of in vitro preparation and specific
489 mutagenesis of DNA fragments: study of protein and DNA interactions." Nucleic Acids Res **16**(15):
490 7351-7367.
491 Ishida, T., R. Ito, J. Clark, N. J. Matzke, Y. Sowa and M. A. B. Baker (2019). "Sodium-powered stators
492 of the bacterial flagellar motor can generate torque in the presence of phenamil with mutations
493 near the peptidoglycan-binding region." Mol Microbiol **111**(6): 1689-1699.
494 Islam, M. I., A. Lin, Y. W. Lai, N. J. Matzke and M. A. B. Baker (2020). "Ancestral Sequence
495 Reconstructions of MotB Are Proton-Motile and Require MotA for Motility." Front Microbiol **11**:
496 625837.
497 Jaques, S., Y. K. Kim and L. L. McCarter (1999). "Mutations conferring resistance to phenamil and
498 amiloride, inhibitors of sodium-driven motility of Vibrio parahaemolyticus." Proc Natl Acad Sci U S A
499 **96**(10): 5740-5745.
500 Kanehisa, M. and S. Goto (2000). "KEGG: kyoto encyclopedia of genes and genomes." Nucleic Acids
501 Res **28**(1): 27-30.
502 Karp, P. D., M. Riley, M. Saier, I. T. Paulsen, J. Collado-Vides, S. M. Paley, A. Pellegrini-Toole, C.
503 Bonavides and S. Gama-Castro (2002). "The EcoCyc Database." Nucleic Acids Res **30**(1): 56-58.
504 Kim, D., B. Langmead and S. L. Salzberg (2015). "HISAT: a fast spliced aligner with low memory
505 requirements." Nat Methods **12**(4): 357-360.
506 Kinosita, Y., T. Ishida, M. Yoshida, R. Ito, Y. V. Morimoto, K. Goto, R. M. Berry, T. Nishizaka and Y.
507 Sowa (2020). "Distinct chemotactic behavior in the original Escherichia coli K-12 depending on
508 forward-and-backward swimming, not on run-tumble movements." Sci Rep **10**(1): 15887.
509 Kojima, S., Y. Asai, T. Atsumi, I. Kawagishi and M. Homma (1999). "Na+-driven flagellar motor
510 resistant to phenamil, an amiloride analog, caused by mutations in putative channel components." J
511 Mol Biol **285**(4): 1537-1547.
512 Kojima, S., M. Kuroda, I. Kawagishi and M. Homma (1999). "Random mutagenesis of the pomA gene
513 encoding a putative channel component of the Na(+)-driven polar flagellar motor of Vibrio
514 alginolyticus." Microbiology (Reading) **145 (Pt 7)**: 1759-1767.
515 Kojima, S., T. Shoji, Y. Asai, I. Kawagishi and M. Homma (2000). "A slow-motility phenotype caused
516 by substitutions at residue Asp31 in the PomA channel component of a sodium-driven flagellar
517 motor." J Bacteriol **182**(11): 3314-3318.
518 Lai, Y. W., P. Ridone, G. Peralta, M. M. Tanaka and M. A. B. Baker (2020). "Evolution of the Stator
519 Elements of Rotary Prokaryote Motors." J Bacteriol **202**(3).
520 Lopatkin, A. J. and J. H. Yang (2021). "Digital Insights Into Nucleotide Metabolism and Antibiotic
521 Treatment Failure." Front Digit Health **3**.
522 Love, M. I., W. Huber and S. Anders (2014). "Moderated estimation of fold change and dispersion for
523 RNA-seq data with DESeq2." Genome Biol **15**(12): 550.
524 Mandadapu, K. K., J. A. Nirody, R. M. Berry and G. Oster (2015). "Mechanics of torque generation in
525 the bacterial flagellar motor." Proc Natl Acad Sci U S A **112**(32): E4381-4389.
526 McDonald, M. J. (2019). "Microbial Experimental Evolution - a proving ground for evolutionary
527 theory and a tool for discovery." EMBO Rep **20**(8): e46992.
528 Miller, M. A., W. Pfeiffer and T. Schwartz (2011). The CIPRES science gateway: a community resource
529 for phylogenetic analyses. Proceedings of the 2011 TeraGrid Conference: extreme digital discovery.
530 Mirdita, M., K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov and M. Steinegger (2021). "ColabFold -
531 Making protein folding accessible to all." 2021.2008.2015.456425.
532 Morgan M, P. H., Obenchain V, Hayden N. (2021). " Rsamtools: Binary alignment (BAM), FASTA,
533 variant call (BCF), and tabix file import. R package version 2.10.0."
534 https://bioconductor.org/packages/Rsamtools. .
535 Mulkidjanian, A. Y., M. Y. Galperin, K. S. Makarova, Y. I. Wolf and E. V. Koonin (2008). "Evolutionary
536 primacy of sodium bioenergetics." Biol Direct **3**: 13.
537 Ni, B., R. Colin, H. Link, R. G. Endres and V. Sourjik (2020). "Growth-rate dependent resource
538 investment in bacterial motile behavior quantitatively follows potential benefit of chemotaxis." Proc
539 Natl Acad Sci U S A **117**(1): 595-601.
540 Ni, B., B. Ghosh, F. S. Paldy, R. Colin, T. Heimerl and V. Sourjik (2017). "Evolutionary Remodeling of
541 Bacterial Motility Checkpoint Control." Cell Rep **18**(4): 866-877.

542 Nishino, Y., Y. Onoue, S. Kojima and M. Homma (2015). "Functional chimeras of flagellar stator
543 proteins between E-coli MotB and Vibrio PomB at the periplasmic region in Vibrio or E-coli."
544 Microbiologyopen **4**(2): 323-331.
545 Nishiyama, M. and S. Kojima (2012). "Bacterial motility measured by a miniature chamber for high-
546 pressure microscopy." Int J Mol Sci **13**(7): 9225-9239.
547 Pallen, M. J. and N. J. Matzke (2006). "From The Origin of Species to the origin of bacterial flagella."
548 Nat Rev Microbiol **4**(10): 784-790.
549 Pang, B., J. L. McFaline, N. E. Burgis, M. Dong, K. Taghizadeh, M. R. Sullivan, C. E. Elmquist, R. P.
550 Cunningham and P. C. Dedon (2012). "Defects in purine nucleotide metabolism lead to substantial
551 incorporation of xanthine and hypoxanthine into DNA and RNA." Proc Natl Acad Sci U S A **109**(7):
552 2319-2324.
553 Parkinson, J. S. (1978). "Complementation analysis and deletion mapping of Escherichia coli mutants
554 defective in chemotaxis." J Bacteriol **135**(1): 45-53.
555 Pertea, M., G. M. Pertea, C. M. Antonescu, T. C. Chang, J. T. Mendell and S. L. Salzberg (2015).
556 "StringTie enables improved reconstruction of a transcriptome from RNA-seq reads." Nat Biotechnol
557 **33**(3): 290-295.
558 Reisch, C. R. and K. L. Prather (2015). "The no-SCAR (Scarless Cas9 Assisted Recombineering) system
559 for genome editing in Escherichia coli." Sci Rep **5**: 15096.
560 Roszak, D. B. and R. R. Colwell (1987). "Survival strategies of bacteria in the natural environment."
561 Microbiol Rev **51**(3): 365-379.
562 Santiveri, M., A. Roa-Eguiara, C. Kühne, N. Wadhwa, H. C. Berg, M. Erhardt and N. M. I. Taylor (2020).
563 "Structure and function of stator units of the bacterial flagellar motor." 2020.2005.2015.096610.
564 Stamatakis, A. (2014). "RAxML version 8: a tool for phylogenetic analysis and post-analysis of large
565 phylogenies." Bioinformatics (Oxford, England) **30**(9): 1312-1313.
566 Sudo, Y., H. Terashima, R. Abe-Yoshizumi, S. Kojima and M. Homma (2009). "Comparative study of
567 the ion flux pathway in stator units of proton- and sodium-driven flagellar motors." Biophysics
568 (Nagoya-shi) **5**: 45-52.
569 Takekawa, N., M. Isumi, H. Terashima, S. Zhu, Y. Nishino, M. Sakuma, S. Kojima, M. Homma and K.
570 Imada (2019). "Structure of Vibrio FliL, a New Stomatin-like Protein That Assists the Bacterial
571 Flagellar Motor Function." mBio **10**(2).
572 Takekawa, N., M. Nishiyama, T. Kaneseki, T. Kanai, H. Atomi, S. Kojima and M. Homma (2015).
573 "Sodium-driven energy conversion for flagellar rotation of the earliest divergent hyperthermophilic
574 bacterium." Sci Rep **5**: 12711.
575 Taylor, T. B., G. Mulley, A. H. Dills, A. S. Alsohim, L. J. McGuffin, D. J. Studholme, M. W. Silby, M. A.
576 Brockhurst, L. J. Johnson and R. W. Jackson (2015). "Evolution. Evolutionary resurrection of flagellar
577 motility via rewiring of the nitrogen regulation system." Science **347**(6225): 1014-1017.
578 Terahara, N., T. A. Krulwich and M. Ito (2008). "Mutations alter the sodium versus proton use of a
579 Bacillus clausii flagellar motor and confer dual ion use on Bacillus subtilis motors." Proceedings of
580 the National Academy of Sciences of the United States of America **105**(38): 14359-14364.
581 Terahara, N., M. Sano and M. Ito (2012). "A Bacillus Flagellar Motor That Can Use Both Na+ and K+ as
582 a Coupling Ion Is Converted by a Single Mutation to Use Only Na+." Plos One **7**(9).
583 Yang, Z. (1997). "PAML: a program package for phylogenetic analysis by maximum likelihood."
584 Computer applications in the biosciences **13**(5): 555-556.

585

586

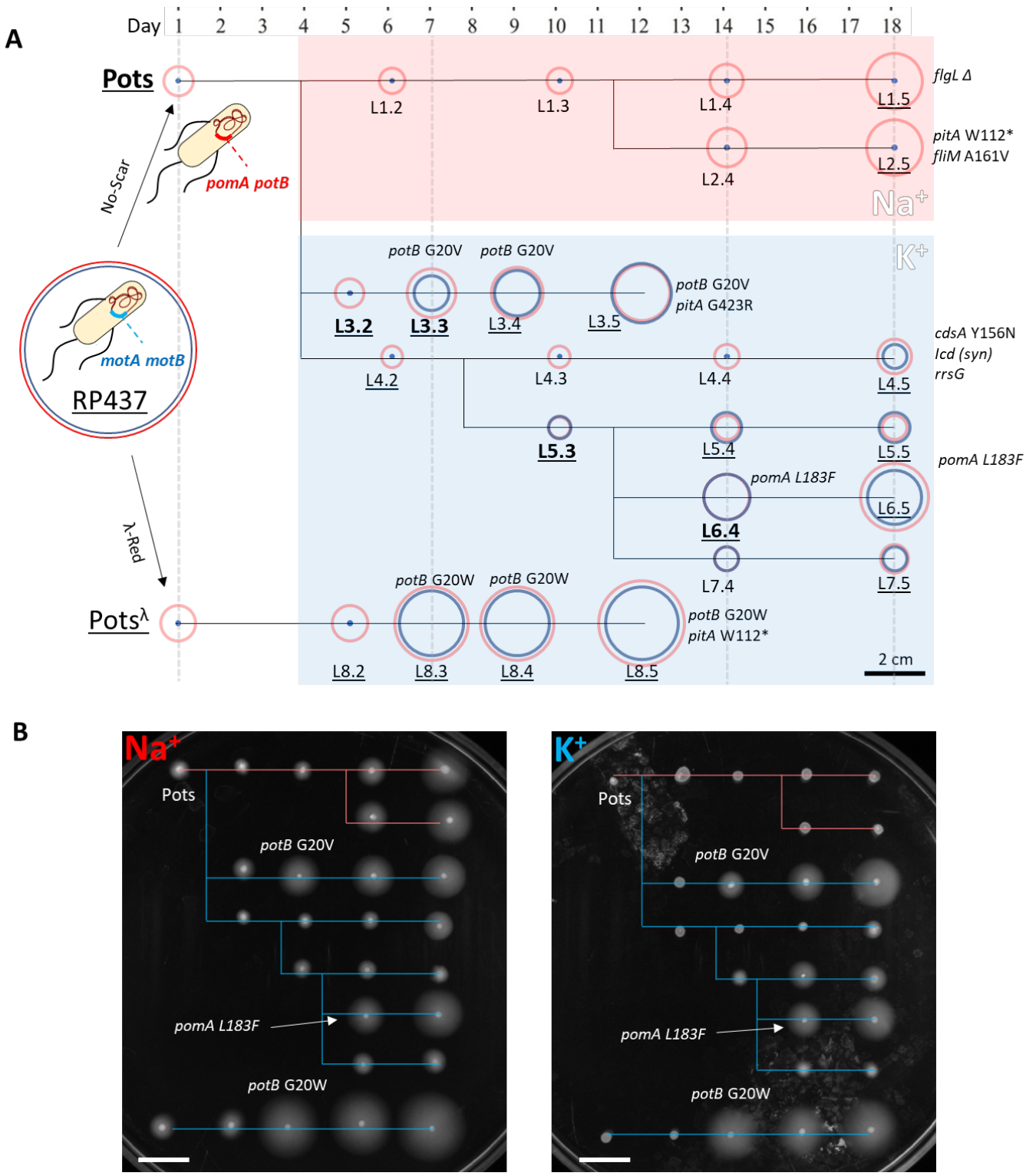

Figure 1. Directed Evolution of the Flagellar Motor

A) Summary of the directed evolution experiments. The edited *E. coli* strains Pots and Pots$^{\lambda}$, obtained from *E. coli* RP437 via No-Scar and λ-Red recombineering respectively, were passaged on soft agar (colored background, red: Na+; blue: K+) over an 18-day period. The summary comprises a total of 8 lineages (L1-L8) selected for further investigation, each comprised of 5 members (ie. L1.3 indicates the first lineage and the third passage of a motile flare). The day of collection and re-inoculation is indicated above each lineage member. The ability of each lineage

member to swim in the presence of either Na+ or K+ is displayed by a red or blue ring, respectively, corresponding to swim size on a swim plate. Lack of motility after 24 hr of incubation on K+ soft agar is represented by a blue dot. Ring sizes were measured from pair of large swim plates presented in panel B. Member labels are underscored to indicate WGS data availability and bolded to indicate RNAseq data availability. All strains were locally Sanger sequenced at the *pomApotB* locus, but 21 strains (underscored) underwent whole genome sequencing (WGS). The full list of SNPs identified via variant calling relative to the Pots reference genome are annotated next to each respective lineage member. Highlighted genes other than *pomA* and *potB*: *pitA* (metal phosphate:H^+ symporter), *flgL* (flagellar hook-filament junction protein 2), *fliM* (flagellar motor switch protein), *cdsA* (cardiolipin-diglyceride synthase), *icd* (isocitrate dehydrogenase), *rrsG* (16S ribosomal RNA). Black scale bar = 2 cm. B) Recapitulation of the directed evolution experiment. Na+ (left) and K+(right) soft agar plates inoculated with a 1µl aliquot of glycerol stock of each strain indicated in A (except RP437 [Supplementary Fig. 5]) and arranged in the same order as (A) on a large petri dish (ø = 30 cm). The colored lines indicate the type of soft agar each lineage was propagated in and isolated from. Arrows and labels indicate the first member in each lineage to display a mutation in the stator genes PomA or PotB. Each plate was incubated for 24 hr at 30°C . White scale bar = 2 cm.

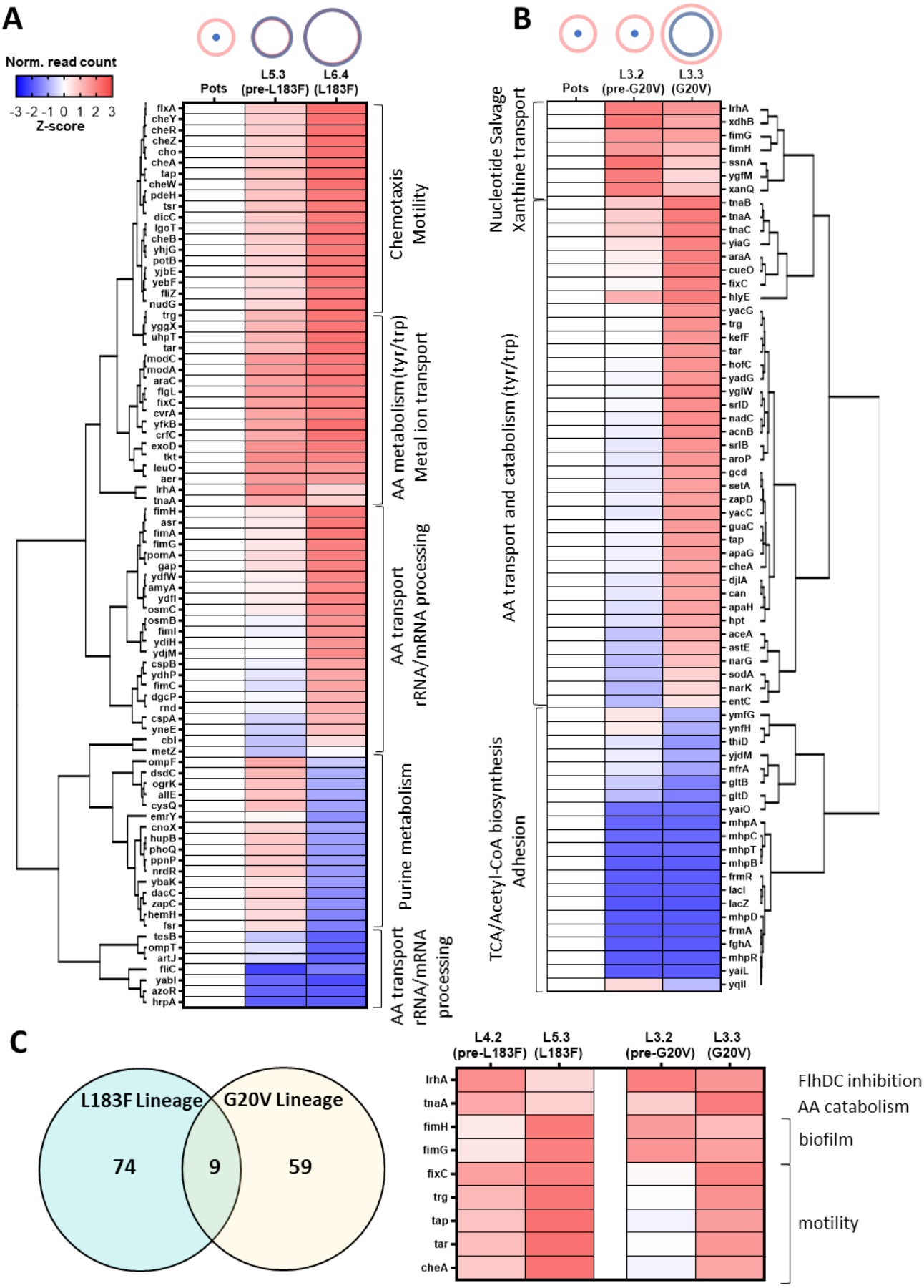

Fig.2. Differential expression analysis of RNAseq data. Differentially expressed genes (DEGs, adjusted p-value < 0.01) measured across selected members of the

L183F lineage (A) or G20V lineage (B). 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 labelled by their representative biological process (PANTHER) . Genes upregulated with respect to the Pots strain are labelled in red, downregulated ones 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.

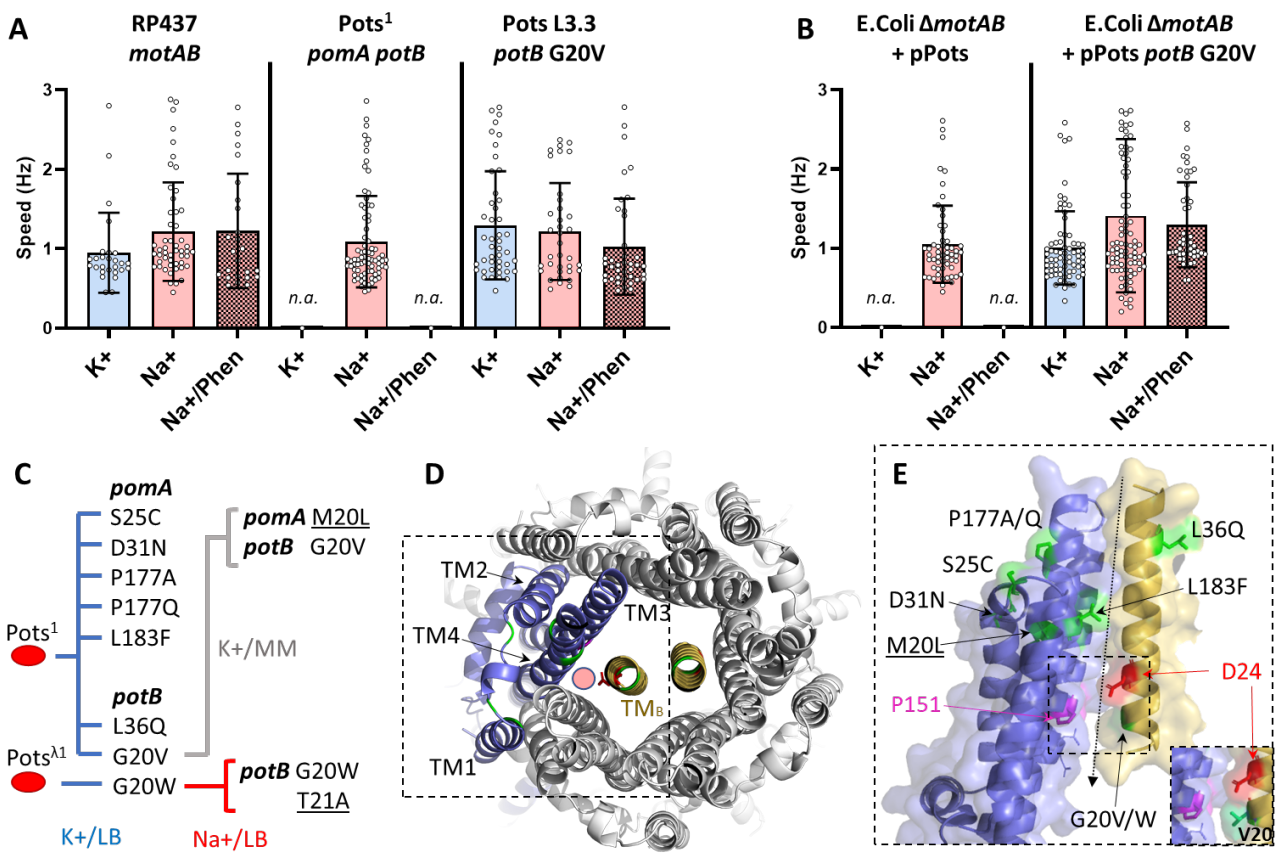

Figure 3. Functional characterization of evolved stators. A) Single cell speed measurements using the tethered cell assay measured in Hz (revolutions/s). Blue bar indicates speed in 67 mM KCl motility Buffer, red bar: 85 mM NaCl motility Buffer; red patterned bar: 85 mM NaCL + 100 µM phenamil motility buffer. Number of cells analysed per condition (from left to right): RP437: 27, 51, 25; pots: n.a, 78, n.a; pots *potB* G20V 45, 36, 39 (n.a. indicates no visible rotating cell). Error bars indicate Standard Deviation (S.D.). B) Single cell speed measurements using the tethered cell assay in RP6894 Δ*motAB* strain co-expressing *pomA* and *potB* G20V from pPots plasmid. Blue bar indicates speed in 67 mM KCl motility buffer, red bar: 85 mM NaCl motility buffer; red patterned bar: 85 mM NaCl + 100 µM phenamil motility buffer. Number of cells analysed per condition (from left to right): (Δ*motAB* + pPots: n.a., 32, n.a; Δ*motAB* + pPots *potB* G20V: 40, 63, 48). Error bars indicate S.D. Single cell tracked data shown in Supplementary Fig. 12. C) Graphical summary of stator gene mutations detected across all directed evolution experiments and the growth conditions under which these mutations arose. LB indicates agar containing yeast extract and tryptone. MM indicates agar in minimal media. Mutations in a subsequent generation are underlined. D) View from the extracellular side of the transmembrane portion of $PomA_5PotB_2$ stator complex (see Methods: Structural Modelling). One monomer of subunit A is coloured in blue and the TM domains of the B subunits are

649 coloured in yellow. Mutant sites obtained in the directed evolution experiments are
650 labelled in green, the catalytic aspartate residue essential for function is highlighted in
651 red. The red circle indicates the predicted location of the ion transport pore (inward
652 conduction). E) side view of the area highlighted by the dashed box in D. Residues
653 P151 (PomA) and D24 (PotB) are also highlighted in magenta and red respectively.
654 Primary mutation sites are indicated in black, while secondary mutation sites are
655 underscored. The arrow at the interface between (A)TM3-4 and (B)TM indicates the
656 predicted location of the ion transport pore. The inset highlights the change in the pore
657 region due to the G20V substitution in PotB (green).

658