Evolutionary innovation through transcription factor promiscuity in microbes is constrained by pre-existing gene regulatory network architecture

Authors: Matthew J. Shepherd[1], Aidan P. Pierce[1], Tiffany B. Taylor[1]*

1 - Milner Centre for Evolution, Department of Life Sciences, University of Bath, Claverton Down, Bath BA2 7AY, UK

* - Corresponding author: Tiffany B. Taylor, Milner Centre for Evolution, Department of Life Sciences, University of Bath, Claverton Down, Bath BA2 7AY, UK, +44 (0)1225 384398, t.b.taylor@bath.ac.uk

The contributions to this work were as follows:

Experimental work performed by MJS. Rhamnose expression plasmids built by APP. Manuscript written by MJS. Manuscript editing by all authors. Project supervised by TBT. Funding secured by TBT.

Acknowledgements:

This work was funded by a Royal Society Research Fellows Grant (RG160491; awarded to TBT) supporting MJS; Royal Society Research Fellows Enhancement Award (RGF\EA\201057; awarded to TBT) supporting APP; a Royal Society Dorothy Hodgkin Research Fellowship (DH150169; awarded to and supporting TBT). Bioinformatics analyses for the article was carried out using the Medical Research Council's (MRC) Cloud Infrastructure for Microbial Bioinformatics (CLIMB), Illumina Whole-Genome Sequencing conducted by MicrobesNG, Birmingham, UK (http://www.microbesng.com) which was supported by the BBSRC (grant number BB/L024209/1), and Illumina RNA sequencing with ribodepletion was conducted by Oxford Genomics Centre, Oxford UK. We would like to thank Prof. Laurence Hurst for insightful discussion of earlier manuscript versions, Dr. Jonathan Nzakizwanayo and Dr. Iain MacArthur for providing us with *E.coli* strains SP50 pRK2073, S17 pTNS2 and ST18, Dr. Ellie Harrison for providing the R script that Fig.5A is based upon, and Dr. James Horton, Louise Flanagan and Josie Elliott for helpful discussion during the course of the study.

Data availability:

All raw data for this study is available on the Open Science Framework (OSF), and can be accessed at https://osf.io/pcdhx/

Abstract

The survival of a population during environmental shifts depends on whether the rate of phenotypic adaptation keeps up with the rate of changing conditions. A common way to achieve this is via change to gene regulatory network (GRN) connections – known as rewiring – that facilitate novel interactions and innovation of transcription factors. To understand the success of rapidly adapting organisms, we therefore need to determine the rules that create and constrain opportunities for GRN rewiring. Here, using an experimental microbial model system we reveal a hierarchy among transcription factors that are rewired to rescue lost function, with alternative rewiring pathways only unmasked after the preferred pathway is eliminated. We identify three key properties - high activation, high expression, and pre-existing low-level affinity for novel target genes – that facilitate transcription factor innovation via gain of functional promiscuity. Ease of acquiring these properties is constrained by pre-existing GRN architecture, which was overcome in our experimental system by both targeted and global network alterations. This work reveals the key properties that determine transcription factor evolvability, and as such, the evolution of GRNs.

Introduction

During conditions of environmental upheaval and niche transition events, rapid phenotypic adaptation is essential for evolutionary success[1–3]. To understand patterns of diversification in novel environments we need to understand why some evolutionary transitions occur more rapidly than others, and what allows some organisms to succeed where others fail. Key to this is understanding the evolution of gene regulatory networks (GRNs), control circuits common throughout the domains of life that determine the magnitude and timing of gene expression[4] in response to environmental and internal signals[5,6]. GRNs are frequent sites of adaptive mutation driving phenotypic evolution[7–9], and often underscore adaptation to - and survival in - new and changeable environments[10–13]. Determining how these networks evolve and identifying rules governing when and how a regulatory circuit adapts will allow us to better understand their role in determining the evolutionary success of an organism.

Within GRNs, a key mechanism of innovation is transcription factor rewiring, in which transcription factors can gain or lose regulatory connections to target genes creating new network architectures and opportunity for phenotypic innovation[14]. Rewiring

events can have dramatic effects on the transcriptome[15] that can drive phenotypic diversification. However, the majority of past studies on transcription factor rewiring involve retrospective experimental dissection of networks that have already diverged[16–21], which allows inference of past rewiring events but does not address the evolutionary factors driving the rewiring process. For rewiring to occur, a transcription factor must first have the potential for promiscuous activity (interaction with non-cognate regulatory targets) allowing regulation of a new gene[22], as is the case for innovation in other proteins[23,24]. Promiscuity is also referred to as cross-talk, however this term can describe cognate interactions where transcription factors have overlapping regulons, so we avoid it in this study. These non-cognate interactions can become meaningful and drive adaptation if favoured by natural selection[25,26]. However, promiscuity cannot be commonplace within GRNs as this will likely result in dysregulation of gene expression and fitness costs for an organism in the environment to which it is adapted[27,28].

Whilst there exists some empirical evidence of the importance of transcription factor promiscuity in driving evolutionary innovation[21,29,30], in each case it is unknown why the transcription factor in question became promiscuous as opposed to any of the other regulators within their protein family; what role – if any – GRN structure played in the process; and whether the rewired transcription factors were unique in their ability to become promiscuous within each study system. To address these questions, we set out to investigate the evolution of transcription factor promiscuity using rescue of flagellar motility via rewiring of NtrC documented by Taylor et al. (2015b). In this model system, *Pseudomonas fluorescens* SBW25 is engineered to be non-motile via deletion of the master regulator for flagellar synthesis (*fleQ*) and abolishment of biosurfactant production. Under strong selection for motility, the bacteria rapidly and reliably evolve new regulatory network wiring to rescue flagellar motility (a schematic of these experiments is laid out in Fig.1A)[31] with the same transcription factor (*ntrC*) repurposed to rescue flagellar driven motility each time, to the exclusion of all other homologous transcription factors within the protein family[32]. However, we do not know the factors that determine this evolutionary preference. To test this, we constructed a double *fleQ ntrC* knockout and placed this mutant under strong selection for swimming motility. This forces evolutionary utilisation of an alternative transcription factor in order to rescue motility, and is an established

method for unveiling hidden evolutionary pathways[33]. We reveal a hierarchy among transcription factors that are rewired to rescue lost flagellar function, with alternative rewiring pathways only revealed after the preferred co-opted transcription factor, NtrC, is eliminated. Identification of additional transcription factors capable of gaining promiscuity within the same GRN background allows us to investigate rules governing when and where a transcription factor becomes promiscuous within its GRN, and is important for understanding how these regulatory systems innovate during adaptation to environmental challenges.

Results

Evolutionary rescue of flagellar motility can occur in the absence of *ntrC*, through *de-novo* mutation to an alternative two-component system

Is the reason we only see NtrC co-opted to rescue FleQ function because it is the only transcription factor capable of doing this? It is known that within families of homologous transcription factors there is variation in the ability to bind non-cognate sites[34], so it is possible that NtrC is unique in its ability to become promiscuous and rewire. FleQ and NtrC are part of a family of 22 structurally related transcription factors called RpoN-dependent enhancer binding proteins (RpoN-EBPs), many of which are predicted to be more structurally similar to FleQ than NtrC[32]. To identify if any other RpoN-EBPs were capable of promiscuity, the double knockout non-motile *P. fluorescens* (Δ*fleQ*Δ*ntrC*) was challenged to rescue flagellar motility in 0.25% agar LB plates. Motile zones re-emerged in a two-step manner (a slow-swimming variant, then a faster-swimming variant) as typical of previous studies[31]. Motile isolates were sampled and whole genome re-sequenced at each step. Motility-granting mutations were identified in the gene *PFLU1131* for all first-step motile isolates (n=15), and in 13/15 cases this was the only mutant gene (Supplementary excel file E1). *PFLU1131* is unstudied and encodes a putative sensor histidine kinase. The gene is situated in an operon between two other genes, *PFLU1130* and *PFLU1132* which encode a hypothetical GNAT-acetyltransferase, and a putative RpoN-dependent transcriptional regulator respectively. PFLU1132 is a known FleQ-homolog[32] and together with PFLU1131 forms a putative two-component system. The most frequent mutation in *PFLU1131* was an identical 15bp deletion (Fig.2A) in the histidine-kinase phospho-acceptor domain (73%). This mutation results in loss of 5 amino acids (368-GEVAM-372) in the protein product (henceforth referred to as PFLU1131-del15). Other

137 mutations included a highly similar 15bp deletion resulting in loss of 5 amino acids a
138 few amino acids downstream (369-EVAMG-373), as well as a SNP resulting in
139 A375V. All of these mutations (86% of the first-step mutations) cluster to the same
140 26bp of the 1770bp *PFLU1131* ORF, and result in amino acid changes in a site
141 directly adjacent to the catalytically active H-box[35] (amino acids 376-382, Fig.2A)
142 suggesting a significant effect on the catalytic function of the putative kinase. We
143 additionally constructed a *PFLU1132* knockout in the *ΔfleQΔntrC* background to
144 identify any further RpoN-EBPs capable of promiscuity. *ΔfleQΔntrCΔPFLU1132* failed
145 to rescue flagellar motility within the 6-week assay cut-off (Supplementary excel file
146 E2, n= 192), suggesting evolutionary gain of promiscuity in other transcription factors
147 may not be easily achieved under the conditions we tested.
148 Together, these results suggest that in the absence of FleQ and NtrC, flagellar
149 motility can be rescued via rewiring of an alternative transcription factor in the same
150 protein family, PFLU1132. To confirm that observed motility phenotypes were
151 dependent on the PFLU1131/2 two-component system, the *PFLU1132* gene was
152 deleted and complemented with and without the presence of the most common first-
153 step kinase mutation (*PFLU1131*-del15). Knockout of *PFLU1132* abolished flagellar
154 motility, and complementation restored motility only in the presence of the kinase
155 mutation (Fig.2B). This experiment was repeated for *ntrBC* and produced the same
156 result (Supplementary Fig.S1). The *PFLU1131*-del15 mutation also grants flagellar
157 motility when *ntrC* is present (i.e., in a Δ*fleQ* background), so does not depend on
158 the absence of *ntrC* to function (Supplementary Fig.S2). Transcriptomic analysis of
159 the *PFLU1131*-del15 mutant by RNA sequencing indicates an alteration to the
160 activity of the PFLU1131/2 two-component system that results in a net up-regulatory
161 effect on the transcriptome (Fig.2C). These regulatory changes include upregulation
162 of the flagellar genes (Supplementary Fig.S4) and the *PFLU1130/1/2* operon. In
163 sum, in the "commonly rewired route", repeatable mutations in the NtrBC (Fig.1B)
164 two-component system were found to rescue flagellar motility. In the absence of both
165 FleQ and NtrC, we identified the "unmasked rewiring route", via PFLU1131/2
166 (Fig.1C), that was also capable of rescuing flagellar motility.
167
168 **First-step mutations in the unmasked rewiring pathway grant promiscuous**
169 **activity to a FleQ homologous transcription factor**

170 In the common rewiring pathway, NtrC, a homolog of FleQ, was recruited to recover
171 lost FleQ function. This occurred as a two-step process; an initial motility-granting
172 mutation to the gene *ntrB* encoding NtrC's cognate kinase and a secondary motility-
173 enhancing mutation to the helix-turn-helix (HTH) DNA binding domain of the NtrC
174 transcription factor. Similarly, In the Δ*fleQ*Δ*ntrC* background strain, first-step
175 mutations restored slow motility, showing that alternative rewiring routes were
176 possible, but not utilised in the presence of *ntrC*. To determine the mechanism of this
177 unmasked rewiring route we performed whole genome sequencing and
178 transcriptome analysis on ancestral and evolved isolates. In 13/15 first-step mutants,
179 only a single mutation to *PFLU1131* was identified. To confirm that no other
180 mutations were needed for flagellar motility, the *PFLU1131*-del15 mutation was
181 introduced into the ancestor (Δ*fleQ*Δ*ntrC*). The resulting engineered strain was motile
182 (Supplementary Fig.S2). As no mutational change to the DNA binding domain of
183 PFLU1132 is needed, this suggests a non-specific mechanism through which this
184 transcription factor induces the flagellar genes.
185 To investigate this, our transcriptomic data was used to assess the impact of the
186 *PFLU1131*-del15 mutation on the expression of a list of genes[36] controlled by all 22
187 homologous transcription factors (part of the RpoN-EBP family) present in *P.*
188 *fluorescens* SBW25[37]. Regulatory changes predictably include upregulation of the
189 flagellar genes (Supplementary Fig.S3) and the *PFLU1130/1/2* operon, but also
190 include upregulation of many other genes. In particular, we saw significant
191 upregulation for 54% of all RpoN-EBP controlled genes in the unmasked rewiring
192 pathway (Fig.3). For comparison, 70% of RpoN-EBP controlled genes were
193 upregulated in the common rewiring pathway (via NtrC)[34]. Flagellar genes account
194 for 12% of all RpoN-EBP controlled genes, and whilst the native regulatory targets of
195 PFLU1132 are unknown it is unlikely that it natively regulates 54% of these.
196 Similarly, the first step motile mutant reported in Taylor et al., (2015b) leads to
197 upregulation of genes known to be involved in nitrogen assimilation (as expected as
198 mutations are located in the Ntr pathway), however it also results in upregulation of
199 many genes with no known regulatory link to NtrC. The large percentages of the
200 RpoN regulon induced by these specific kinase mutations, in both the common and
201 unmasked rewiring pathways, suggest first-step mutations confer a state of

202 promiscuous regulatory activity to PFLU1132 resulting in non-cognate interactions
203 across many RpoN-EBP controlled genes including the flagellar genes.
204
205 **Evolutionary rescue of motility through unmasked rewiring pathway is**
206 **significantly constrained**
207 If alternative rewiring routes are available to natural selection – why have we not seen
208 them utilised in the presence of the commonly rewired route? In the unmasked rewiring
209 pathway, we considered the time taken for swimming mutants to evolve and the
210 strength of the evolved phenotype in comparison to the common rewiring route. Time
211 to emergence (the length of time taken for motility to evolve on the soft agar plate) was
212 recorded for all replicate experimental lines of the Δ*fleQ*Δ*ntrC* ancestor, as well as
213 Δ*fleQ* as a comparison. Strikingly, only 9.4% of replicate experimental lines (n=160) of
214 Δ*fleQ*Δ*ntrC* evolved within the 6-week assay cut-off compared to 100% for the Δ*fleQ*
215 background (n=22). Initial first-step motile Δ*fleQ*Δ*ntrC* strains also evolved within an
216 average of 18.7 days from the assay start (Fig.4A), significantly longer than Δ*fleQ* took
217 to rescue motility (4.2 days, P = 0.002, Dunn test). After first-step PFLU1131
218 mutations, second-step mutants evolved rapidly in Δ*fleQ*Δ*ntrC* lines within an average
219 of 2.2 days (Fig.4A). This was faster than the second-step mutants in the Δ*fleQ*
220 background which took 3.3 days to emerge (*P* = 0.0017, Dunn test).
221 The greatly increased time, and low frequency for rescue motility in an Δ*fleQ*Δ*ntrC*
222 background may be reflective of a small pool of accessible PFLU1131 mutations that
223 can grant promiscuity to PFLU1132, perhaps due to a small mutational target size.
224 One replicate line gained a *de novo* mutation in the DNA mismatch-repair gene *mutS*
225 (Supplementary excel file E1), resulting in a frame shift and probable loss of
226 function[38]. This strain, along with strains derived from it possessed large numbers of
227 additional SNPs (>80) including a SNP in PFLU1131. Loss of function to *mutS* is
228 known to enhance mutability in *Pseudomonads*[39], and its occurrence may increase
229 access to motility-rescuing mutations in PFLU1131.
230 From a phenotypic perspective, the first-step mutations in the *PFLU1131* pathway in
231 a Δ*fleQ*Δ*ntrC* strain, provide a far poorer motility phenotype than the analogous first-
232 step mutations in a Δ*fleQ* background (Fig.4B). In a race assay, *PFLU1131*-del15
233 mutants swam 0.31 mm for every 1 mm swam by the most common first step mutant
234 in the common rewiring route (P=0.000356, Wilcox test). One possible explanation

235 for this poor motility is a lack of significant upregulation for the flagellar filament
236 subunit FliC (Supplementary Fig.S4) in the *PFLU1131*-del15 mutant.
237 The *PFLU1131*-del15 mutant additionally does not confer a significant defect to
238 growth in shaking LB broth compared to its ancestral strain (Fig.4C, average relative
239 area under the growth curve (rAUGC) of 1.55 and 1.48 respectively, P=0.1132, Dunn
240 test), and grew significantly better than the most common first step mutant in the
241 common rewiring route (rAUGC of 1, P=0.0081, Dunn test). The poor motility
242 phenotype and the lack of significant fitness cost associated with *PFLU1131*-del15
243 likely indicates that the PFLU1131/2 system is a weaker activator of the flagellar
244 motility that incurs less sever pleiotropy compared to NtrBC.
245
246 These findings indicate that the unmasked pathway is the far poorer option for
247 gaining promiscuous regulation of the flagellar genes compared to the common
248 NtrBC pathway - motility rescue in the Δ*fleQ*Δ*ntrC* background took significantly
249 longer to occur, was significantly less frequent, and provides a significantly poorer
250 flagellar motility phenotype with lower pleiotropic fitness cost. However, it is not clear
251 why this should be the case – both NtrC and PFLU1132 are FleQ-homologs, and are
252 gaining mutations to their cognate kinases that unlock promiscuity. This suggests
253 that there are other factors that constrain the evolution of promiscuity in PFLU1132.
254
255 **Motility enhancing second-step mutations suggest alterations to global gene**
256 **regulatory network can facilitate promiscuity in unmasked rewiring pathway**
257
258 Following the first-step mutations in PFLU1131 that unlock flagellar motility, our
259 motile isolates develop second-step mutations that act to boost motility speed
260 (Supplementary Fig.S5A), often accompanied by significant pleiotropic fitness costs
261 for growth in LB broth (Fig. S5B). These second-step mutations offer clues to the
262 nature of constraining factors limiting innovation through promiscuity of PFLU1132,
263 as they represent evolutionary solutions to the poor motility provided to the first-step
264 mutations.
265
266 Whole genome resequencing identified a diverse set of motility enhancing second-
267 step mutations occurring at both a local (PFLU1131/2 locus) and global regulatory

268 scale (Fig.5A, full details of all mutations provided in supplementary excel file E1,
269 second step n = 18 (several first step lineages generates multiple second step
270 zones). In contrast to the frequent second-step mutations observed in the HTH DNA-
271 binding domain of NtrC[31], no analogous mutations were observed in the same
272 domain of PFLU1132. Mutations that did occur in the *PFLU1132* gene were parallel
273 and identified in only 2 lines, impacting the receiver domain of the transcription
274 factor, which may further increase activity of the PFLU1132 regulator by modifying
275 the interactions with its kinase. One line gained a secondary *PFLU1131* mutation, a
276 SNP in the HATPase domain along with the first-step SNP A547D already present
277 (Fig.2B), which may further boost kinase interactions. A mutation to the promoter
278 region of the PFLU1131/2 operon was also observed.
279 Other secondary mutations can be grouped into 3 other broad categories. (i) The first
280 are mutations in the operon *PFLU1583*/*4*, accounting for 24% of secondary
281 mutations (Fig.5A, Purple). This gene pair encodes a putative anti-sigma factor
282 PP2C-like phosphatase and a putative STAS family anti-sigma factor antagonist
283 (*PFLU1583* and *PFLU1584* respectively). Mutations in *PFLU1583* were generally
284 loss of function suggesting loss of repression on an unknown sigma factor. This may
285 act on RpoN the partner sigma factor of PFLU1132 in regulating gene expression,
286 however this cannot be tested easily through *rpoN* knockout, and *P. fluorescens*
287 SBW25 encodes ~31 other putative sigma factors[37] which may instead be the
288 mechanistic targets of this mutation. Transcriptomic analysis of a *PFLU1583* mutant
289 (PFLU1583Δ48-74) identified sigma factor RpoE as upregulated, however *rpoE*
290 knockout did not negate the motility enhancing effect of the mutation (Supplementary
291 Fig.S6A). To test that PFLU1583/4 are not acting to enhance motility through
292 another RpoN-EBP, we engineered *PFLU1583*Δ48-74 in a Δ*fleQ*Δ*ntrC* background in
293 the absence of a *PFLU1131* mutation. This strain was immotile, indicating that these
294 anti-sigma factor mutations depend on the function of the PFLU1131/2 system to
295 confer a swimming phenotype (Supplementary Fig.S6B). (ii) The second category of
296 mutations impact core gene expression components that will affect global GRN
297 function (Fig.5A, Green). These included mutations to *rho*, *rpsK* and *rpoC* that
298 encode core gene expression machinery and will impact expression of most genes in
299 the genome. (iii) The final category of motility enhancing mutations fit a metabolic
300 theme (Fig.5A, blue). These mutations occur in genes involved in central carbon

301 metabolism, which will likely have a significant impact on global gene expression
302 through alterations to central carbon flux and may act on RpoN indirectly – for
303 example via core carbon metabolic regulator CbrB (an RpoN-EBP) – or signal to the
304 PFLU1131 sensor-kinase via internal metabolic flux.
305 The diversity and nature of these second-step mutations suggest global regulatory
306 strategies to facilitate flagellar gene expression through the unmasked rewiring
307 pathway. To understand their regulatory impact, transcriptome analysis was
308 performed on a pair of representative mutants: the most common second-step
309 mutant (anti-sigma factor mutant, *PFLU1583* Δ48-74) which represented a mutation
310 with global regulatory effects, and the *PFLU1131/2* promoter mutation as a
311 representative of mutation with predicted "local" regulatory effects. Together these
312 mutations represent the two loci that constitute 56% of identified second-step
313 mutations.
314 Principle component analysis indicates that the transcriptomic profile of the anti-
315 sigma factor mutant differs significantly from the profile of the *PFLU1131/2* promoter
316 mutation. Both mutations resulted in similar variation across PC1 relative to the first-
317 step *PFLU1131*-del15 mutant, but opposite variation in PC2 (Fig.5B). Both the
318 *PFLU1131/2* promoter and anti-sigma factor mutations result in net up-regulatory
319 effects on the transcriptome reflected by positive skewed volcano plots (Fig.5C and
320 Fig.5D) albeit with differing expression patterns. Both mutations have the effect of
321 further upregulating RpoN-EBP controlled genes with 22% and 60% being
322 upregulated for the anti-sigma factor mutant and the *PFLU1131/2* promoter mutant
323 respectively (Fig.5E).
324 In general, for the unmasked rewiring pathway, genomic and transcriptomic analysis
325 of second-step mutations indicate that second-step mutations are increasing
326 promiscuous activity of the rewired transcription factor (PFLU1132), albeit through
327 differing (local and global) mechanisms. These results highlight that diverse
328 mutations with both targeted and global regulatory effects can influence promiscuity
329 to facilitate rewiring and innovation in a transcription factor.
330
331 **Second-step promoter capture mutation suggests importance of transcription**
332 **factor expression in gaining promiscuity**

333 Although second-step mutations were diverse in the unmasked rewiring pathway,
334 one promoter capture event in particular provides evidence for the role of increased
335 expression of the rewired transcription factor in strengthening the motility phenotype.
336 For promiscuous activation of non-cognate genes, the PFLU1132 transcription factor
337 will need to first saturate its native regulatory interactions before it can engage in
338 low-affinity promiscuous interactions with FleQ-controlled genes. High expression of
339 active transcription factor can provide these conditions, elevating the concentration
340 of transcription factor in the cell to permit more frequent promiscuous action[40,41].
341 Whilst many of the second-step mutations detailed above may impact *PFLU1131/2*
342 expression indirectly (RpoC, Rho), the *PFLU1131/2* promoter mutation directly
343 effects expression of the two-component system. This mutation is a 1.59Kbp deletion
344 resulting in total loss of the *PFLU1130* gene, and in *PFLU1131/2* becoming part of
345 the *PFLU1127/8/9* operon (Fig.6A). Upstream of this new combined operon sits a
346 predicted RpoN binding site[36], so this deletion positions *PFLU1132* in a new operon
347 under the control of RpoN which may create a positive feedback loop where this
348 RpoN-dependent regulator drives its own expression – either through promiscuous
349 regulation or native control of this promoter. This genetic rearrangement significantly
350 upregulates both *PFLU1131* and *PFLU1132* with 4.26 and 3.79 Log2Fold increases
351 in expression compared to the non-motile Δ*fleQ* strain respectively (Fig.6B, raw
352 values for this and other tested strains provided in supplementary excel file E5).
353 Upregulation of this two-component system corresponds with the 60% increase in
354 expression of RpoN-EBP controlled genes discussed above (Fig.5E), indicating an
355 increase in promiscuous transcription factor activity. The 1.59Kbp deletion results in
356 loss of an intergenic region that typically separates the two combined operons. This
357 region is predicted to contain a rho-dependent terminator (Supplementary Excel file
358 E3) which may typically prevent transcriptional readthrough into *PFLU1130/1/2*.
359 Such a terminator will restrict *PFLU1131/2* expression via readthrough and constrain
360 its ability to achieve high concentrations that may facilitate promiscuity. The *rho*
361 mutation observed in another second-step motile strain may also act to increase
362 readthrough at this site. In comparison to this promoter mutation, the other second-
363 step mutation tested (anti-sigma factor *PFLU1583* Δ48-74) does not significantly
364 impact *PFLU1131/2* expression (Fig.6B), so this mutation likely influences the
365 promiscuity of the *PFLU1132* transcription factor through a separate mechanism.

366 In summary, expression of *PFLU113/2* appears to be a major factor constraining
367 promiscuous regulatory activity and motility rescue through this regulatory system,
368 and overcoming this constraint facilitates evolutionary innovation through this
369 unmasked rewiring route.
370
371 **Increasing transcription factor expression in motile mutants results in faster**
372 **motility phenotypes**
373 To test the importance of transcription factor gene expression in gain of promiscuous
374 activity, we made use of the complementation strains presented in Fig.2C and
375 supplementary Fig.S1. These experiments produced strains where the common
376 (*ntrC)* and unmasked (*PFLU1132)* rewired transcription factor genes are deleted
377 from their native loci and reintroduced on an L-rhamnose titratable promoter system.
378 This was done in the presence of their respective kinase mutations that conferred
379 first-step slow spreading motility in each rewiring pathway. Concentration of L-
380 rhamnose added to the media positively correlated with distance moved (Fig.7A) for
381 both the common rewiring pathway (i.e., NtrC expression system; ρ = 0.984,
382 Spearman's test) and the unmasked rewiring pathway (i.e., PFLU1132 expression
383 system; ρ = 0.865, Spearman's test). We confirmed increasing L-rhamnose
384 concentration increased expression from the L-rhamnose titratable system (RhaSR-
385 PrhaBAD) in our bacterial strain backgrounds by testing LacZ activity with increasing
386 L-rhamnose concentration for *lacZ* under control of the L-rhamnose titratable
387 promoter construct in the Δ*fleQ* and Δ*fleQ*Δ*ntrC* backgrounds. In both cases LacZ
388 activity positively correlated with L-rhamnose concentration (ρ = 0.981 and ρ = 0.955
389 respectively, Spearman's tests), indicating that increasing L-rhamnose concentration
390 results in increasing expression of the gene introduced in the *rhaSR-PrhaBAD*
391 expression system (Supplementary Fig.S7).
392 This allows us to conclude that increasing expression of *ntrC* and *PFLU1132* in the
393 presence of their first-step kinase mutations granted stronger motility phenotypes,
394 likely through stronger promiscuous action. The expression level of a transcription
395 factor can therefore significantly impact its propensity for promiscuous regulatory
396 activity, and more highly expressed transcription factor may be expected to engage
397 in non-canonical regulatory interactions more readily.
398

399 **Discussion**

400 Essential to understanding why some organisms adapt faster than others in the face
401 of environmental challenges is knowledge of how GRNs and regulatory systems
402 evolve. Evolutionary innovation of transcription factors within these networks
403 depends on gain of functional promiscuity as a first step, however the rules of when
404 and where such properties develop within the context of the GRN is currently
405 unknown. Using our microbial model system, we have identified two available
406 pathways for rescuing flagellar motility in a maladapted Δ*fleQ* network; through
407 mutations to the NtrBC and PFLU1131/2 two component systems. However, one
408 pathway – the NtrBC pathway – is always used to the exclusion of the other when
409 both are present. By comparing these available evolutionary pathways, we can
410 resolve why NtrC is more accessible for gaining promiscuity and rewiring due to its
411 position within its GRN. In addition, we can identify three key properties that facilitate
412 transcription factor promiscuity and opportunities for evolutionary innovation.
413 **Firstly, our results confirm the importance of shared structural homology**
414 **between the absent and recruited transcription factor.** In our system, the shared
415 structural homology of NtrC and PFLU1132 with FleQ grants a basal affinity for
416 regulating flagellar genes that evolution can act upon. This likely explains why
417 structurally unrelated transcription factors do not rewire in our model, and indicates
418 that similarity to FleQ facilitates the gain of promiscuity observed in NtrC and
419 PFLU1132. However it is not the case that the regulators that are most structurally
420 homologous to FleQ are the first to be repurposed[32], so homology is not the deciding
421 factor in which RpoN-EBPs become promiscuous to rescue flagellar motility. It is
422 possible that the RpoN-EBPs represent a hyper-evolvable family of transcription
423 factors through the nature of their connectivity and mechanism of transcriptional
424 regulation via RpoN. To investigate this evolutionary rescue of a non-RpoN
425 dependent trait would need to be tested using a different maladapted GRN rescue
426 model. **Secondly, hyperactivation of the co-opted transcription factor benefits**
427 **promiscuous regulatory activity**. In our model system, first step mutations that
428 granted a slow motile phenotype were always achieved through mutation of a kinase
429 – resulting in hyper-activation of their cognate transcription factors. Most
430 transcription factors are subject to some form of post-translational control, most
431 commonly through phosphorylation of a REC receiver domain in bacteria[42].

432 Promiscuous binding may occur transiently or only briefly, so ensuring that the
433 transcription factor is constantly activated increases the likelihood of these non-
434 cognate regulatory events leading to altered gene expression. **Finally - and**
435 **perhaps most importantly - high gene expression of the transcription factor**
436 **boosts promiscuous activity**. High expression of a transcription factor can lead to
437 saturation of native binding sites. Any unbound, active transcription factor is then
438 free to engage in promiscuous activity. These low affinity interactions will occur more
439 readily if there is a higher concentration of transcription factor present, following laws
440 of mass action kinetics[43]. Typically, transcription factors are expressed at low levels
441 in bacteria[44], due to only a small protein copy number being needed for a small
442 number of binding sites in most cases[45]. Transcription factor expression level may
443 well be optimised at a low level to prevent mis-regulation. Indeed, highly expressed
444 and connected transcription factors were found to result in greater network
445 perturbations when synthetically rewired[15], which may reflect a greater ability to act
446 promiscuously. Many molecular biology studies use overexpression/overactivation of
447 proteins of interest as a means to determining their biological functions[46]. Our
448 findings also suggest that this approach should perhaps be used more cautiously,
449 certainly when studying transcription factor function.
450
451 Aside from mutations directly targeting activation and expression of *PFLU1131/2*, we
452 observed several distinct strategies that enhance PFLU1132 promiscuity through
453 global regulator changes. These strategies likely impact the wider GRN and
454 transcriptome significantly, with mutations to RpoC, RpsK and Rho being particularly
455 striking. RpoC and RpsK mutations may act to increase expression of most genes
456 including *PFLU1131/2* and the flagellar genes. Mutations to the *PFLU1583/4* anti-
457 sigma factor system also indicated a strategy to boost promiscuity that would have a
458 significant impact of the wider GRN. Global regulators are often sites of mutation and
459 innovation[47,48], with mutations affecting sigma factor function in particular having
460 large impacts on global gene expression which can drive phenotypic innovation[49].
461 Changes to sigma factors have been shown to facilitate GRN remodelling, as was
462 the case for a study of *P. aeruginosa* during adaptation to the CF lung[13]. Our study
463 highlights for the first time that such mutational strategies can also achieve increases
464 in promiscuity for specific transcription factors to facilitate evolutionary innovation.
465 Each of these mutational strategies that enhanced functional promiscuity in

466 PFLU1132 acted to alter the pre-existing GRN architecture around the transcription
467 factor, whether at a local level by altering the *PFLU1131/2* promoter or at a global
468 regulatory level.
469
470 Our findings suggest that pre-existing GRN architecture can be a key constraining
471 factor in gaining functional promiscuity and evolutionary innovation by a transcription
472 factor. For example, the ease by which a transcription factor can gain high
473 expression depends on the GRN structure it sits within. Many transcription factors
474 promote their own expression[50] (positive autoregulation) and can generate high self-
475 expression through runaway positive feedback. Such positive feedback loops are
476 typically contained by a negative repressor[51–53], including in the case of NtrBC
477 (Fig.1B). Mutational loss of repression can therefore be an evolutionarily accessible
478 mechanism to achieve promiscuity-enhancing high expression levels of a
479 transcription factor. In our model system, GlnK prevents runaway *ntrBC*
480 autoregulation by inhibiting NtrB[54] (Fig.1B). All previously observed NtrBC pathway
481 mutations were single *de-novo* mutations that act to remove repression through
482 GlnK, both directly or indirectly[31,55,56]. These have a dual effect, both hyperactivating
483 the NtrBC system, and strongly upregulating *ntrBC* expression through their positive
484 autoregulatory loop. In contrast, we have identified in this study that to achieve the
485 same effect of high activity and high expression in PFLU1131/2, multiple mutations
486 were required (Fig.1C). This highlights the clear evolutionary advantage NtrBC has
487 in becoming promiscuous compared to PFLU1131/2.
488
489 From the evidence outlined, pre-existing GRN motifs can be predicted to significantly
490 bias the evolvability of the transcription factors under their control. This is strongly
491 supported by the literature: mutational loss of a negative repressor leading to
492 overactivity and overexpression of a positively autoregulated transcription factor has
493 been observed to drive adaptation in several experimental and clinical settings[57–59].
494 Considering the alternative scenario, a transcription factor that represses its own
495 expression (negative autoregulation) will hinder the gain of a high expression level,
496 due to the nature of self-repression. Interestingly, the hierarchy favouring NtrBC over
497 PFLU1132 matches previous studies of evolutionary pathway heirarchies[33], where
498 mutational target size predicts that loss of repressors will be favoured over other
499 precise mutations. Our study highlights that in the context of GRN structures,

connectivity to other network components can constrain innovation through gain of promiscuity, alongside other effects including mutational target size.

Our study provides empirical evidence for mechanisms and evolutionary drivers of regulatory evolution. GRNs are key to the adaptability of organisms, and our results reveal the genetic regulatory factors that generate evolvable control systems. We find that feedback loops that can generate high expression of regulators after a small number of mutations can facilitate rapid innovation and adaptation. Previously, we have hypothesised that some transcription factors may be 'primed' for evolutionary innovation by virtue of the GRN architecture they exist within[26]. Our data support this by suggesting that promiscuity can be more easily evolved in transcription factors with positive feedback structures. These findings can also be utilised in the design of genetic regulatory systems to have reduced evolvability, a key future goal for synthetic biology projects where the evolutionary change of a constructed system is detrimental[60,61]. Designing systems that avoid easily accessible routes to hyperactivation and hyperexpression of their components could help prevent rewiring of engineered GRNs. Pre-existing GRN structure can be expected to significantly bias patterns of diversification in regulatory systems, and subsequently shape genomic evolution by influencing the speed in which genetic circuitry can adapt to changing conditions. This is highly important for populations adapting during sudden changes such as niche transitions including the emergence of new pathogens. Our data moves us closer to the eventual goal of empirically building a set of principles by which GRNs adapt and offers unique insights into how these systems can function and influence the evolution of their components.

Materials and Methods

Strains and culture conditions

All ancestral lines in this study are derived from *Pseudomonas fluorescens* SBW25Δ*fleQ* IS-ΩKm-hah: PFLU2552 (referred to throughout as Δ*fleQ*). Removal of the flagellum master regulator FleQ and transposon-insertional disruption of the gene *viscB* (PFLU2552), rendering this strain immotile as detailed previously[31,62]. All strains were cultured on lysogeny broth (LB; Miller) media at 27°C. *Escherichia coli*

532 strains for cloning were cultured on LB media at 37°C. Strains, plasmids, media
533 supplements and primers are detailed in Supplementary Excel file E4.

534 **Knockout mutant construction by two-step allelic exchange**

535 Knockouts of ntrC and PFLU1132 were achieved using two-step allelic exchange
536 following the protocol of Hmelo et al., 2015 with some alterations. In brief, 400bp
537 flanking regions up- and down-stream of the target gene were amplified and joined
538 by strand overlap extension (SOE) PCR to create a knockout allele. This was
539 inserted into the allelic-exchange suicide vector pTS1 (Containing *sacB* and *tetR*) by
540 SOE-cloning[64] to create the knockout plasmid, and transformed into the conjugal *E.*
541 *coli* strain ST18 by chemical-competence heat-shock. Knockout plasmids were
542 transferred from *E. coli* ST18 to Δ*fleQ* by two-parent puddle-mating and merodiploids
543 selected for on LB supplemented with kanamycin sulphate and tetracycline
544 hydrochloride but lacking 5-ALA supplement required for growth of the *E. coli* ST18
545 auxotrophic Δ*hemA* mutant. Merodiploids were cultured overnight in LB broth lacking
546 tetracycline selection and diluted before spread plating onto NSLB media
547 supplemented with 15% w/v sucrose for *sacB* levansucrase-mediated
548 counterselection of the pTS1 plasmid backbone. Sucrose-resistant colonies were
549 isolated and screened for tetracycline sensitivity. Chromosomal presence of the
550 knockout allele and absence of target gene coding sequences was confirmed by
551 colony PCR. Knockout of ntrC left an ORF encoding an 11 amino acid truncated
552 protein, and knockout of PFLU1132 left an ORF encoding an 8 amino acid truncated
553 protein. All other genomic features including operon structure and terminator regions
554 were left intact.

555 **Motility evolution experiments**

556 Δ*fleQ*Δ*ntrC* and Δ*fleQ*Δ*ntrC*Δ*PFLU1132* were challenged to rescue motility in the
557 absence of the FleQ master flagellar regulator on soft agar, as described
558 previously[31,56]. Pure colonies were picked and inoculated into 0.25% agar LB plates
559 made as described in Alsohim et al., 2014, and incubated at 27°C. Plates were
560 checked a minimum of twice daily for motility, recording time to emergence. Motile
561 zones were sampled immediately and always from the leading edge. Motile isolates
562 were streaked on LB agar, and a pure colony picked and stored at -80°C as glycerol
563 stocks of LB overnight cultures. All subsequent analysis was conducted on these

564 pure motile isolates. Experiment was run for six weeks and any replicates without
565 motility after this cut-off recorded as having not evolved.

566 **Bacterial growth and motility fitness assays**

567 The fitness of the motility phenotype of evolved Δ*fleQ*Δ*ntrC* strains was tested by
568 measuring distance moved over 24 hrs of incubation in 0.25% agar LB plates. Six
569 biological replicates of each strain were grown as separate overnight cultures.
570 Cultures were adjusted to an OD595 = 1 and resuspended in PBS. Soft agar plates
571 were inoculated with 1 µL of these suspensions, by piercing the surface of the plate
572 with the pipette tip, and then effusing the sample into the gap left by the tip. Plates
573 were incubated for 24 h at 27°C, and photographs taken of motile zones. Surface
574 area moved was then calculated from the radius of the concentric motile zone
575 measured from these images ($A = \pi r^2$). Values were square root transformed before
576 plotting.

577 Growth in shaking LB broth was measured by inoculating 99 µL of sterile LB broth
578 with 1 µL of the OD595=1 PBS cell suspensions for each replicate in a 96-well plate.
579 Plates were incubated at 27°C with 180 rpm shaking in a plate reader, recording
580 OD595 every hour for 24 hours. Area under the bacterial growth curve was
581 calculated using the growthcurver package in R and plotted[65]. Area under the
582 bacterial growth curve provides a metric for fitness, as it accounts for the
583 characteristics of lag- and log-phase growth, as well as the final carrying capacity of
584 the population. All growth curves were plotted and inspected prior to calculation of
585 AUGC to ensure similar AUGC values corresponded to similar curve shapes
586 (Supplementary Fig. S3)

587 **Mutation identification by whole genome resequencing and PCR amplicon**
588 **sanger sequencing**

589 To identify motility rescuing mutations, genomic DNA was extracted from motile
590 strains and their ancestral strain using the Thermo Scientific GeneJET Genomic
591 DNA Purification Kit. Genomic DNA was quality checked using BR dsDNA Qubit
592 spectrophotometry to determine concentration and nanodrop spectrophotometry to
593 determine purity. Illumina HiSeq sequencing was provided by MicrobesNG
594 (Birmingham, UK), with a minimum 30x coverage and a sample average of 114x
595 coverage. Returned paired-end reads were aligned to the *P. fluorescens* SBW25
596 reference genome[37] using the Galaxy platform[66]. INDELs were identified using the

597 integrated genomics viewer[67]. For SNP identification, the variant calling software
598 SNIPPY was used with default parameters (Seemann, 2015). Protein domains
599 affected by mutations detailed in this article were predicted from amino acid
600 sequences using pfam, SMART and BLASTp. For a subset of additional motile
601 mutants, PCR amplification and subsequent sanger sequencing of *PFLU1131* was
602 performed as this gene was mutated in all previously sequenced strains. This was
603 done using the service provided by Eurofins Genomics. PCR amplicons were
604 purified using the Monarch® PCR & DNA Cleanup Kit (New England Biolabs).
605 Mutations were identified by alignment of the returned PFLU1131 sequence against
606 the *P. fluorescens* SBW25 reference genome using NCBI BLAST.
607 Three second-step isolates derived from line 24 were sampled due to observed
608 colony morphology variation. After whole genome re-sequencing it became apparent
609 that this variation was due to the presence of a *mutS* mutation and not due to diverse
610 motility strategies. These isolates all have the mutation *PFLU1132*-D111G
611 (Supplementary excel file E1, isolate ID's 24-S1, 24-LCV, and 24-SCV), however this
612 is likely not due to parallel evolution, but instead a single instance of a *PFLU1132*-
613 D111G mutant that then diversified due to *mutS* mutation. For all subsequent
614 analysis these three isolates were treated as a single instance of *PFLU1132*-D111G.
615 **RNA-sequencing**
616 Whole-cell RNA was extracted from 20 OD units of *P. fluorescens* cultures in mid-log
617 phase growth (OD_{595} ~1.5). Cultures were incubated in LB broth at 27°C and 180
618 rpm shaking. Extractions were performed for biological triplicates of each strain of
619 interest. Upon reaching the desired OD, growth and RNA expression was halted by
620 addition of a ½ culture volume of ice-cold killing buffer (20 mM NaN_3, 20 mM Tris-
621 HCl, 5 mM $MgCl_2$). Cells were pelleted, and the killing buffer removed. A lysis buffer
622 of β-mercaptoethanol in buffer RLT from the Qiagen RNeasy extraction kit was used
623 to resuspend pellets, which were then lysed by bead-milling at 4500 rpm for 45 s
624 with lysing matrix B. Lysates were spun through columns from the Qiagen RNeasy
625 extraction kit, and the extraction completed following the RNeasy kit protocol. A
626 DNase I treatment step was included between washes with buffer RW1, by adding
627 DNase I directly to the column from the RNase-free DNase kit (Qiagen) following kit
628 protocol. Samples were eluted in nuclease-free water, and subsequently treated with
629 TURBO DNase from the Turbo DNA-free kit (Invitrogen) following kit protocols.

630 Purified RNA concentration was measured by Qubit RNA BR assay (Thermo-
631 scientific), RNA quality by nanodrop spectrophotometry, and RNA integrity by
632 agarose gel electrophoresis. RNA sequencing was provided by Oxford Genomics
633 (Oxford UK). Samples were ribodepleted, and the mRNA fraction converted to cDNA
634 with dUTP incorporated during second strand synthesis. The cDNA was end-
635 repaired, A-tailed and adapter-ligated. Prior to amplification, samples underwent
636 uridine digestion. The prepared libraries were size selected, multiplexed and QC'ed
637 before paired end sequencing over one unit of a flow cell in a NovaSeq 6000.
638 Returned data was quality checked before it was returned in fastq format.
639 Reference-based transcript assembly was performed using the Rockhopper RNA
640 sequencing analysis software[68], to generate transcript read counts. Differential gene
641 expression analysis was performed using the DESeq2 package of the R statistical
642 coding language to create lists of differentially expressed genes and corresponding
643 Log2Fold changes in gene expression.

644 **Rhamnose-inducible expression system constructs**

645 Complementation of *ntrC* and *PFLU1132* knockouts, as well as inducible expression
646 experiments were performed by single-copy chromosomal re-introduction of the
647 RpoN-EBP genes under control of a *rhaSR-PrhaBAD* L-rhamnose inducible
648 expression system[69]. The RpoN-EBP ORF was amplified and a strong ribosome
649 binding site (stRBS) introduced upstream after a 7bp short spacer between the site
650 and the start codon. This construct was placed downstream of the *PrhaBAD*
651 promoter on the miniTn7 suicide-vector pJM220 (obtained from the Addgene plasmid
652 repository, plasmid #110559) by restriction-ligation and transferred to *E. coli* DH5α
653 by chemical-competence heat-shock. The miniTn7 transposon containing the *rhaSR*
654 genes and the PrhaBAD-stRBS-RpoN-EBP construct were transferred to the *P.*
655 *fluorescens* chromosome by transposonal insertion downstream of the *glmS* gene
656 via four-parent puddle-mating conjugation[70]. The relevant *E. coli* DH5α pJM220-
657 derived plasmid donor was combined with recipient *P. fluorescens* Δ*fleQ* strains,
658 transposition helper *E. coli* SM10 λpir pTNS2 and conjugation helper *E. coli* SP50
659 pRK2073, and Gentamicin resistant *Pseudomonas* selected for on LB supplemented
660 with Gentamicin sulphate and Kanamycin sulphate. Chromosomal insertion of the
661 correct miniTn7 transposon was confirmed by colony PCR.

662 **β-galactosidase activity assay**

663 To validate the expression dynamics of the rhaSR-PrhaBAD system in the genetic
664 backgrounds used in this work, the same system described above was introduced
665 into *P. fluorescens* Δ*fleQ* and Δ*fleQ*Δ*ntrC* with *lacZ* under the control of PrhaBAD.
666 This was obtained from the pJM230 plasmid[69] (obtained from the Addgene plasmid
667 repository, plasmid #110560). Overnight cultures of each strain were set up in
668 biological triplicate, and used to inoculate 9 mL cultures of LB broth supplemented
669 with a range of L-rhamnose concentrations at a bacterial OD_{600} = 0.05. Cultures
670 were incubated shaking 180 r.p.m. at 27°C until they reached and OD_{600} = 0.5, 2 mL
671 was spun down and frozen at -20°C until needed. Cell pellets were resuspended in
672 Z-buffer (60mM Na_2HPO_4*7H_2O, 40mM NaH_2PO4*H_2O, 10mM KCl, 1mM
673 $MgSO_4$*7H_2O), OD_{600} recorded, and 400µL cells added to 600µL fresh Z-buffer. Cell
674 suspension was lysed with addition of 40 µL Chloroform, 20 µL 0.1% SDS and
675 vortexing followed by 10 minutes incubation at 30°C. β-galactosidase activity assay
676 was begun with addition of 200µL o-nitrophenyl-β-D-galactopyraniside (ONPG) and
677 incubated at 30°C for 20 minutes. If reactions become visibly yellow they were halted
678 immediately by addition of 500 µL Na_2CO_3 and the time recorded. All other reactions
679 were halted in this manner after 20 minutes. Reactions were centrifuged to remove
680 debris and 900 µL added to a cuvette and A_{420} measured. LacZ reporter activity
681 (MU) was calculated by: MU = (A_{420} * 1000) / (Time(mins) * 0.4 * OD_{600} of initial Z-
682 buffer cell suspension).

683 **Statistical analyses and data handling**

684 All statistical analysis and data handling was performed using R core statistical
685 packages and the Dunn.test package. Shapiro-Wilks normality tests were performed
686 to confirm non-normality of datasets. To compare group medians for more than two
687 groups, a Kruskal-Wallis test with post-hoc Dunn test and Benjamini-Hochberg
688 correction was performed, with a $P \leq 0.025$ taken to indicate significance. For
689 comparisons of two groups only, a two-sample Wilcox test was used. Spearman's
690 correlation was used to test significance of correlations between motility rate and
691 rhamnose concentration, using the R core statistical function cor.test with the
692 spearman method specified.
693 For analysis of transcriptomic data, RNA sequencing output was run through
694 Rockhopper for reference-based transcript assembly. Transcript count data were
695 then analysed using DESeq2 package in R, for differential gene expression analysis,

696 as well as generation of PCA and Volcano plots. A gene was taken as being
697 differentially expressed if it differed by a Log2Fold change of ≥1 or ≤ -1 between two
698 samples (representing a doubling or halving of expression level respectively) with an
699 adjusted P-value of ≤0.001.
700 To identify whether Rho dependent terminators were present downstream of genes
701 of interest, the RhoTermPredict algorithm[71] was run on the *P. fluorescens* SBW25
702 reference genome in Python.
703 **References**

704 1. Sheppard SK, Guttman DS, Fitzgerald JR. Population genomics of bacterial
705 host adaptation. *Nat Rev Genet*. 2018;19(9):549-565. doi:10.1038/s41576-
706 018-0032-z
707 2. Cohan FM, Koeppel AF. The Origins of Ecological Diversity in Prokaryotes.
708 *Curr Biol*. 2008;18(21):1024-1034. doi:10.1016/j.cub.2008.09.014
709 3. Martin CH, Richards EJ. The Paradox behind the Pattern of Rapid Adaptive
710 Radiation: How Can the Speciation Process Sustain Itself Through an Early
711 Burst? *Annu Rev Ecol Evol Syst*. 2019;50:569-593. doi:10.1146/annurev-
712 ecolsys-110617-062443
713 4. Fang X, Sastry A, Mih N, et al. Global transcriptional regulatory network for
714 Escherichia coli robustly connects gene expression to transcription factor
715 activities. *Proc Natl Acad Sci U S A*. 2017;114(38):10286-10291.
716 doi:10.1073/pnas.1702581114
717 5. López-Maury L, Marguerat S, Bähler J. Tuning gene expression to changing
718 environments: From rapid responses to evolutionary adaptation. *Nat Rev*
719 *Genet*. 2009;10(1):68. doi:10.1038/nrg2500
720 6. Martínez-Antonio A, Janga SC, Salgado H, Collado-Vides J. Internal-sensing
721 machinery directs the activity of the regulatory network in Escherichia coli.
722 *Trends Microbiol*. 2006;14(1):22-27. doi:10.1016/j.tim.2005.11.002
723 7. Mcdonald JMC, Reed RD. Patterns of selection across gene regulatory
724 networks. *Semin Cell Dev Biol*. 2022;(xxxx).
725 doi:10.1016/j.semcdb.2022.03.029
726 8. Isalan M, Lemerle C, Michalodimitrakis K, et al. Evolvability and hierarchy in
727 rewired bacterial gene networks. *Nature*. 2008;452(7189):840-845.
728 doi:10.1038/nature06847
729 9. Mack KL, Ballinger MA, Phifer-Rixey M, Nachman MW. Gene regulation

730 underlies environmental adaptation in house mice. *Genome Res*.
731 2018;28(11):1636-1645. doi:10.1101/gr.238998.118
732 10. Luo X, Song R, Acar M. Multi-component gene network design as a survival
733 strategy in diverse environments. *BMC Syst Biol*. 2018;12(1):85.
734 doi:10.1186/s12918-018-0609-3
735 11. Ali F, Seshasayee ASN. Dynamics of genetic variation in transcription factors
736 and its implications for the evolution of regulatory networks in Bacteria. *Nucleic*
737 *Acids Res*. 2021;48(8):4100-4114. doi:10.1093/NAR/GKAA162
738 12. Turkarslan S, Reiss DJ, Gibbins G, et al. Niche adaptation by expansion and
739 reprogramming of general transcription factors. *Mol Syst Biol*. 2011;7(554):1-
740 16. doi:10.1038/msb.2011.87
741 13. Damkiær S, Yang L, Molin S, Jelsbak L. Evolutionary remodeling of global
742 regulatory networks during long-term bacterial adaptation to human hosts.
743 *Proc Natl Acad Sci U S A*. 2013;110(19):7766-7771.
744 doi:10.1073/pnas.1221466110
745 14. Johnson AD. The rewiring of transcription circuits in evolution. *Curr Opin*
746 *Genet Dev*. 2017;47:121-127. doi:10.1016/j.gde.2017.09.004
747 15. Baumstark R, Hänzelmann S, Tsuru S, et al. The propagation of perturbations
748 in rewired bacterial gene networks. *Nat Commun*. 2015;6:1-5.
749 doi:10.1038/ncomms10105
750 16. Kruzel EK, Giles SS, Hull CM. Analysis of cryptococcus neoformans sexual
751 development reveals rewiring of the pheromone-response network by a
752 change in transcription factor identity. *Genetics*. 2012;191(2):435-449.
753 doi:10.1534/genetics.112.138958
754 17. Reece-Hoyes JS, Pons C, Diallo A, et al. Extensive rewiring and complex
755 evolutionary dynamics in a C.elegans multiparameter transcription factor
756 network. *Mol Cell*. 2013;51(1):116-127. doi:10.1016/j.molcel.2013.05.018
757 18. Tebung WA, Choudhury BI, Tebbji F, Morschhäuser J. Rewiring of the Ppr1
758 Zinc Cluster Transcription Factor from Purine Catabolism to Pyrimidine
759 Biogenesis in the Saccharomycetaceae. *Curr Biol*. 2016;26(13):1677-1687.
760 doi:10.1016/j.cub.2016.04.064
761 19. Adhikari S, Erill I, Curtis PD. Transcriptional rewiring of the GcrA/CcrM
762 bacterial epigenetic regulatory system in closely related bacteria. *PLoS Genet*.
763 2021;17(3):1-30. doi:10.1371/journal.pgen.1009433

764 20. Martchenko M, Levitin A, Hogues H, Nantel A, Whiteway M. Transcriptional
765 Rewiring of Fungal Galactose-Metabolism Circuitry. *Curr Biol*.
766 2007;17(12):1007-1013. doi:10.1016/j.cub.2007.05.017
767 21. Hsu PC, Lu TC, Hung PH, et al. Plastic Rewiring of Sef1 Transcriptional
768 Networks and the Potential of Nonfunctional Transcription Factor Binding in
769 Facilitating Adaptive Evolution. *Mol Biol Evol*. 2021;38(11):4732-4747.
770 doi:10.1093/molbev/msab192
771 22. Voordeckers K, Pougach K, Verstrepen KJ. How do regulatory networks
772 evolve and expand throughout evolution? *Curr Opin Biotechnol*. 2015;34:180-
773 188. doi:10.1016/j.copbio.2015.02.001
774 23. Copley SD. An evolutionary biochemist's perspective on promiscuity. *Trends*
775 *Biochem Sci*. 2015;40(2):72-78. doi:10.1016/j.tibs.2014.12.004
776 24. Guzmán GI, Sandberg TE, LaCroix RA, et al. Enzyme promiscuity shapes
777 adaptation to novel growth substrates. *Mol Syst Biol*. 2019;15(4):1-14.
778 doi:10.15252/msb.20188462
779 25. Aharoni A, Gaidukov L, Khersonsky O, Gould SM, Roodveldt C, Tawfik DS.
780 The "evolvability" of promiscuous protein functions. *Nat Genet*. 2005;37(1):73-
781 76. doi:10.1038/ng1482
782 26. Taylor TB, Shepherd MJ, Jackson RW, Silby MW. Natural selection on
783 crosstalk between gene regulatory networks facilitates bacterial adaptation to
784 novel environments. *Curr Opin Microbiol*. 2022;67:102140.
785 doi:10.1016/j.mib.2022.02.002
786 27. Capra EJ, Perchuk BS, Skerker JM, Laub MT. Adaptive mutations that prevent
787 crosstalk enable the expansion of paralogous signaling protein families. *Cell*.
788 2012;150(1):222-232. doi:10.1016/j.cell.2012.05.033
789 28. Zarrinpar A, Park SH, Lim WA. Optimization of specificity in a cellular protein
790 interaction network by negative selection. *Nature*. 2003;426(6967):676-680.
791 doi:10.1038/nature02178
792 29. Pougach K, Voet A, Kondrashov FA, et al. Duplication of a promiscuous
793 transcription factor drives the emergence of a new regulatory network. *Nat*
794 *Commun*. 2014;5:1-11. doi:10.1038/ncomms5868
795 30. Brunkard JO, Runkel AM, Zambryski PC. A promiscuous intermediate
796 underlies the evolution of LEAFY DNA binding specificity. *Science (80-)*.
797 2015;347(6222):645-649. doi:10.1126/science.1255437

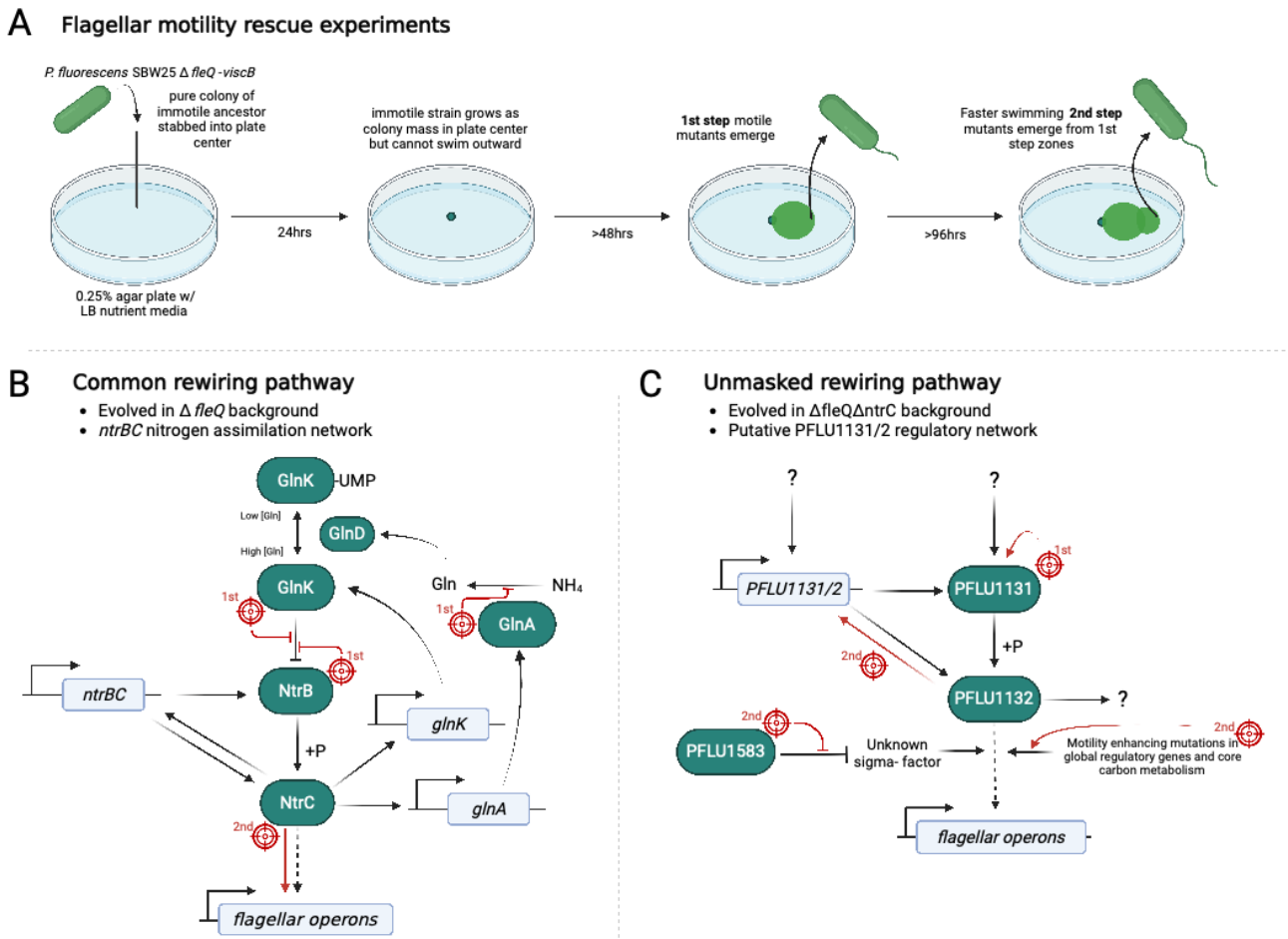
798 31. Taylor TB, Mulley G, Dills AH, et al. Evolutionary resurrection of flagellar
799 motility via rewiring of the nitrogen regulation system. *Science (80-)*.
800 2015;347(6225):1014-1017. doi:10.1126/science.1259145
801 32. Taylor T, Mulley G, McGuffin L, et al. Evolutionary rewiring of bacterial
802 regulatory networks. *Microb Cell*. 2015;2(7):256-258.
803 doi:10.15698/mic2015.07.215
804 33. Lind PA, Farr AD, Rainey PB. Experimental evolution reveals hidden diversity
805 in evolutionary pathways. *Elife*. 2015;4:1-17. doi:10.7554/elife.07074
806 34. Igler C, Lagator M, Tkačik G, Bollback JP, Guet CC. Evolutionary potential of
807 transcription factors for gene regulatory rewiring. *Nat Ecol Evol*.
808 2018;2(10):1633-1643. doi:10.1038/s41559-018-0651-y
809 35. Kim DJ, Forst S. Genomic analysis of the histidine kinase family in bacteria
810 and archaea. *Microbiology*. 2001;147(5):1197-1212. doi:10.1099/00221287-
811 147-5-1197
812 36. Jones J, Studholme DJ, Knight CG, Preston GM. Integrated bioinformatic and
813 phenotypic analysis of RpoN-dependent traits in the plant growth-promoting
814 bacterium Pseudomonas fluorescens SBW25. *Environ Microbiol*.
815 2007;9(12):3046-3064. doi:10.1111/j.1462-2920.2007.01416.x
816 37. Silby MW, Cerdeño-Tárraga AM, Vernikos GS, et al. Genomic and genetic
817 analyses of diversity and plant interactions of Pseudomonas fluorescens.
818 *Genome Biol*. 2009;10(5):1-16. doi:10.1186/gb-2009-10-5-r51
819 38. Oliver A, Baquero F, Blázquez J. The mismatch repair system (mutS, mutL
820 and uvrD genes) in Pseudomonas aeruginosa: Molecular characterization of
821 naturally occurring mutants. *Mol Microbiol*. 2002;43(6):1641-1650.
822 doi:10.1046/j.1365-2958.2002.02855.x
823 39. Luján AM, Maciá MD, Yang L, Molin S, Oliver A, Smania AM. Evolution and
824 adaptation in pseudomonas aeruginosa biofilms driven by mismatch repair
825 system-deficient mutators. *PLoS One*. 2011;6(11).
826 doi:10.1371/journal.pone.0027842
827 40. Gutierrez PS, Monteoliva D, Diambra L. Cooperative Binding of Transcription
828 Factors Promotes Bimodal Gene Expression Response. *PLoS One*. 2012;7(9).
829 doi:10.1371/journal.pone.0044812
830 41. Kribelbauer JF, Rastogi C, Bussemaker HJ, Mann RS. Low-affinity binding
831 sites and the transcription factor specificity paradox in eukaryotes. *Annu Rev*

832 *Cell Dev Biol*. 2019;35:357-379. doi:10.1146/annurev-cellbio-100617-062719
833 42. Galperin MY. Structural classification of bacterial response regulators:
834 Diversity of output domains and domain combinations. *J Bacteriol*.
835 2006;188(12):4169-4182. doi:10.1128/JB.01887-05
836 43. Kim JK, Tyson JJ. Misuse of the Michaelis–Menten rate law for protein
837 interaction networks and its remedy. *PLoS Comput Biol*. 2020;16(10):1-21.
838 doi:10.1371/journal.pcbi.1008258
839 44. Ishihama Y, Schmidt T, Rappsilber J, et al. Protein abundance profiling of the
840 Escherichia coli cytosol. *BMC Genomics*. 2008;9:1-17. doi:10.1186/1471-
841 2164-9-102
842 45. Alkema WBL, Lenhard B, Wasserman WW. Regulog analysis: Detection of
843 conserved regulatory networks across bacteria: Application to Staphylococcus
844 aureus. *Genome Res*. 2004;14(7):1362-1373. doi:10.1101/gr.2242604
845 46. Prelich G. Gene overexpression: Uses, mechanisms, and interpretation.
846 *Genetics*. 2012;190(3):841-854. doi:10.1534/genetics.111.136911
847 47. Philippe N, Crozat E, Lenski RE, Schneider D. Evolution of global regulatory
848 networks during a long-term experiment with Escherichia coli. *BioEssays*.
849 2007;29(9):846-860. doi:10.1002/bies.20629
850 48. Harrison E, Guymer D, Spiers AJ, Paterson S, Brockhurst MA. Parallel
851 Compensatory Evolution Stabilizes Plasmids across the Parasitism-Mutualism
852 Continuum. *Curr Biol*. 2015;25(15):2034-2039. doi:10.1016/j.cub.2015.06.024
853 49. Pletnev P, Pupov D, Pshanichnaya L, et al. Rewiring of growth-dependent
854 transcription regulation by a point mutation in region 1.1 of the housekeeping σ
855 factor. *Nucleic Acids Res*. 2020;48(19):10802-10819. doi:10.1093/nar/gkaa798
856 50. Macneil LT, Walhout AJM. Gene regulatory networks and the role of
857 robustness and stochasticity in the control of gene expression. *Genome Res*.
858 2011;21:645-657. doi:10.1101/gr.097378.109.21
859 51. Rao SD, Igoshin OA. Overlaid positive and negative feedback loops shape
860 dynamical properties of PhoPQ two-component system. *PLoS Comput Biol*.
861 2021;17(1 December):1-18. doi:10.1371/journal.pcbi.1008130
862 52. Weyder M, Prudhomme M, Bergé M, Polard P, Fichant G. Dynamic modeling
863 of Streptococcus pneumoniae competence provides regulatory mechanistic
864 insights into its tight temporal regulation. *Front Microbiol*. 2018;9(JUL):1-25.
865 doi:10.3389/fmicb.2018.01637

866 53. diCenzo GC, Sharthiya H, Nanda A, Zamani M, Finan TM. PhoU allows rapid
867 adaptation to high phosphate concentrations by modulating PstSCAB transport
868 rate in Sinorhizobium meliloti. *J Bacteriol*. 2017;199(18):1-20.
869 doi:10.1128/JB.00143-17
870 54. Hervás AB, Canosa I, Santero E. Regulation of glutamate dehydrogenase
871 expression in Pseudomonas putida results from its direct repression by NtrC
872 under nitrogen-limiting conditions. *Mol Microbiol*. 2010;78(2):305-319.
873 doi:10.1111/j.1365-2958.2010.07329.x
874 55. Shepherd MJ, Horton JS, Taylor TB. A near-deterministic mutational hotspot in
875 Pseudomonas fluorescens is constructed by multiple interacting genomic
876 features. *Mol Biol Evol*. 2022;39(6):1-7. doi:10.1093/molbev/msac132
877 56. Horton JS, Flanagan LM, Jackson RW, Priest NK, Taylor TB. A mutational
878 hotspot that determines highly repeatable evolution can be built and broken by
879 silent genetic changes. *Nat Commun*. 2021;12(1). doi:10.1038/s41467-021-
880 26286-9
881 57. Bretl DJ, Ladd KM, Atkinson SN, Müller S, Kirby JR. Suppressor mutations
882 reveal an NtrC-like response regulator, NmpR, for modulation of Type-IV Pili-
883 dependent motility in Myxococcus xanthus. *PLoS Genet*. 2018;14(10):1-28.
884 doi:10.1371/journal.pgen.1007714
885 58. Patel V, Matange N. Adaptation and compensation in a bacterial gene
886 regulatory network evolving under antibiotic selection. *Elife*. 2021;10:1-27.
887 doi:10.7554/eLife.70931
888 59. Schofield MC, Rodriguez DQ, Kidman AA, et al. The anti-sigma factor MucA is
889 required for viability in Pseudomonas aeruginosa. *Mol Microbiol*.
890 2021;116(2):550-563. doi:10.1111/mmi.14732
891 60. Gallup O, Ming H, Ellis T. Ten future challenges for synthetic biology. *Eng Biol*.
892 2021;5(3):51-59. doi:10.1049/enb2.12011
893 61. Chalovich JM, Eisenberg E. Engineering reduced evolutionary potential for
894 synthetic biology. *Biophys Chem*. 2005;257(5):2432-2437.
895 doi:10.1039/c3mb70606k.Engineering
896 62. Alsohim AS, Taylor TB, Barrett GA, et al. The biosurfactant viscosin produced
897 by Pseudomonas fluorescensSBW25 aids spreading motility and plant growth
898 promotion. *Environ Microbiol*. 2014;16(7):2267-2281. doi:10.1111/1462-
899 2920.12469

63. Hmelo LR, Borlee BR, Almblad H, et al. Precision-engineering the Pseudomonas aeruginosa genome with two-step allelic exchange. *Nat Protoc*. 2015;10(11):1820-1841. doi:10.1038/nprot.2015.115
64. Bryksin A V., Matsumura I. Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids. *Biotechniques*. 2010;48(6):463-465. doi:10.2144/000113418.Overlap
65. Sprouffske K, Wagner A. Growthcurver: An R package for obtaining interpretable metrics from microbial growth curves. *BMC Bioinformatics*. 2016;17(1):17-20. doi:10.1186/s12859-016-1016-7
66. Afgan E, Baker D, Batut B, et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res*. 2018;46(W1):W537-W544. doi:10.1093/nar/gky379
67. Robinson JT, Thorvaldsdóttir H, Winckler W, et al. Integrative genomics viewer. *Nat Biotechnol*. 2011;29(1):24-26. doi:10.1038/nbt.1754
68. Brian Tjaden. A Computational System for Identifying Operons Based on RNA-Seq Data. *Methods*. 2020;176:60-72. doi:10.1016/j.ymeth.2019.03.026.A
69. Meisner J, Goldberg JB. The Escherichia coli rhaSR-PrhaBAD inducible promoter system allows tightly controlled gene expression over a wide range in Pseudomonas aeruginosa. *Appl Environ Microbiol*. 2016;82(22):6715-6727. doi:10.1128/AEM.02041-16
70. Choi KH, Schweizer HP. mini-Tn7 insertion in bacteria with single attTn7 sites: Example Pseudomonas aeruginosa. *Nat Protoc*. 2006;1(1):153-161. doi:10.1038/nprot.2006.24
71. Di Salvo M, Puccio S, Peano C, Lacour S, Alifano P. RhoTermPredict: An algorithm for predicting Rho-dependent transcription terminators based on Escherichia coli, Bacillus subtilis and Salmonella enterica databases. *BMC Bioinformatics*. 2019;20(1):1-11. doi:10.1186/s12859-019-2704-x

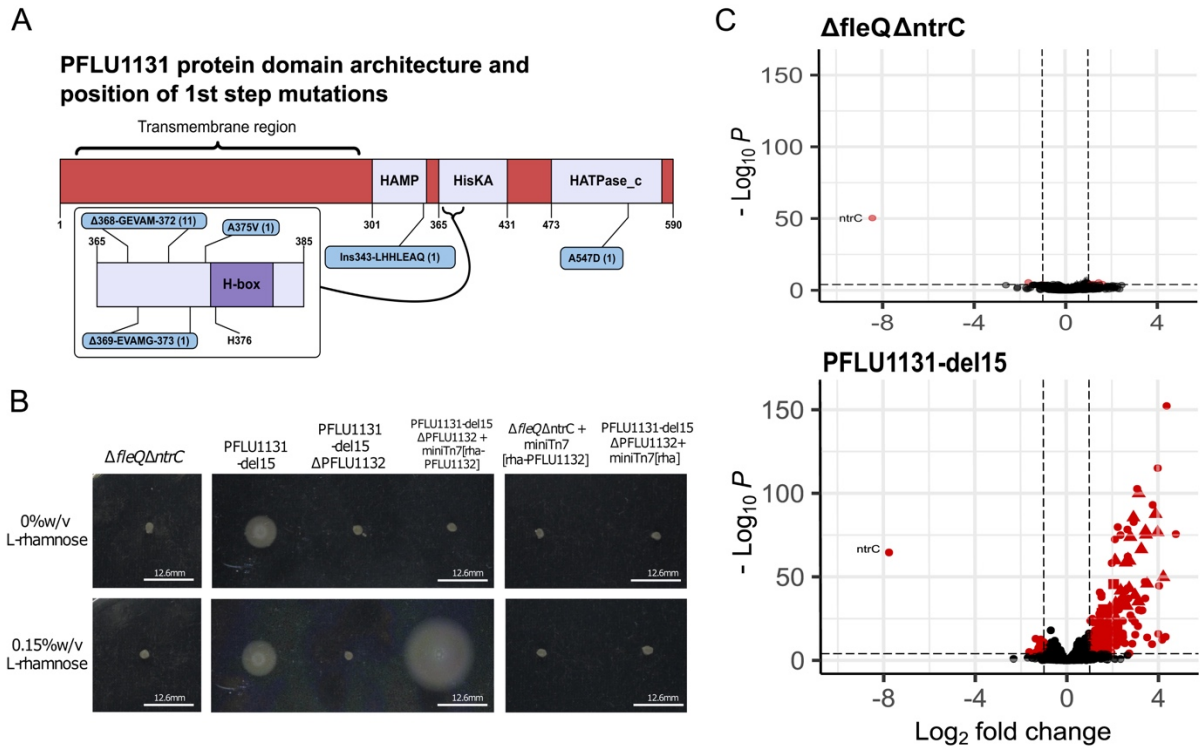
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950 **Figures**
951



952

953 **Figure 1: Rewiring of RpoN-EBP transcription factors to rescue flagellar**
954 **motility through promiscuous control of FleQ-controlled genes.** A) Flagellar
955 motility rescue experiment routmap outlines typical progression of motility evolution.
956 Pathway diagrams display key components of the common (B, Taylor et al., 2015)
957 and unmasked (C, this study) rewiring pathways for rescue of flagellar gene

958 expression. Genes are coloured in white, protein components in green. Mutational
959 targets and their effects are shown in red, and whether a mutation occurs in the first
960 or second evolutionary step indicated.
961
962
963
964
965



966
967

968 **Figure 2: First-step motility rescue mutations in *PFLU1131* A:** Diagram of
969 PFLU1131 protein, indicated positions of first-step motility-rescuing mutations.
970 Predicted protein domains (pfam) are indicated in lilac, amino acid positions are
971 indicated as numbers below each domain. Mutations are indicated in blue boxes,
972 with the number of replicate lines gaining each mutation indicated in brackets. **B)**
973 FleQ-homolog encoding gene *PFLU1132* is essential for rescued flagellar motility in
974 the *PFLU1131*-del15 mutant strain and depends on the presence of the kinase
975 mutation. Scale bars (white) = 12.6mm. Transcription factor gene *PFLU1132* is
976 deleted and reintroduced as a single-copy chromosomal insertion expressed from an
977 L-rhamnose inducible promoter system (*rhaSR-PrhaBAD*). The same
978 complementation lacking the *PFLU1131*-del15 mutation, as well as an empty
979 expression system transposon were included as further controls. Photographs of
980 motility after 1 day incubation in 0.25% agar LB plates supplemented with or without
981 0.15% L-rhamnose for induction of transcription factor expression. **C)** Volcano plots
982 indicating impact of *PFLU1131*-del15 mutation on the transcriptome relative to the

983 Δ*fleQ* ancestor. Red points indicate significantly differentially expressed genes.
984 Triangles indicate flagellar genes, Squares indicate *PFLU1131/2* and adjacent
985 genes, and Circles indicate all other genes.
986

Figure 3: Log2Fold changes in gene expression for RpoN-EBP controlled genes after first-step motility mutations, relative to Δ*fleQ* ancestor. Δ*fleQ*Δ*ntrC* and Δ*fleQ*Δ*ntrC* first-step differentially expressed genes were determined by RNA sequencing (this work). Δ*fleQ* first-step differentially expressed genes were identified by RNA microarray and originally reported by Taylor *et al.* 2015. List of RpoN-EBP controlled genes was obtained from Jones *et al.,* 2007. Groups of genes by function are indicated by the coloured bars above the plot: Red – Flagellum, Blue – Nitrogen, Teal – Other/unknown.

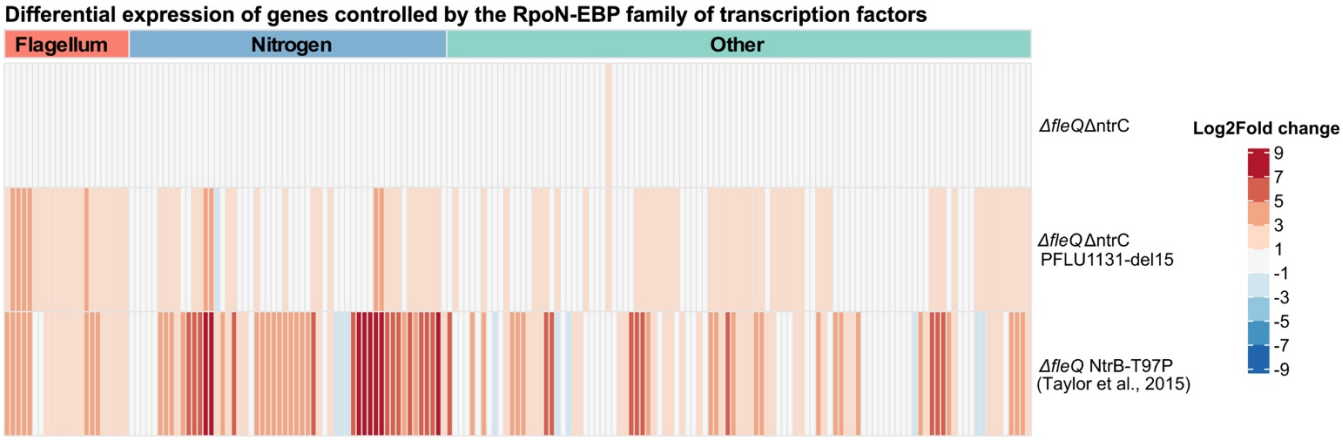
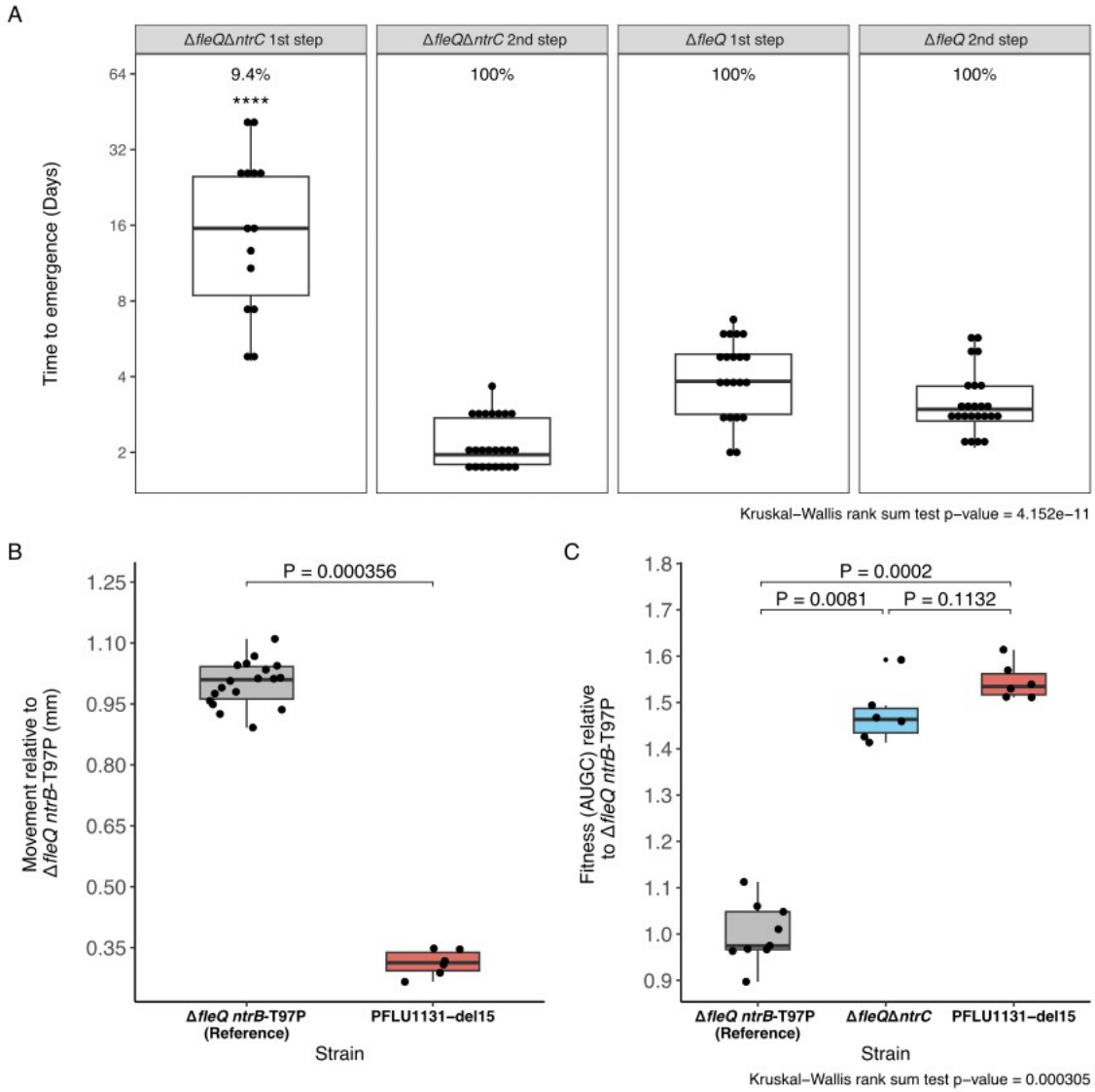
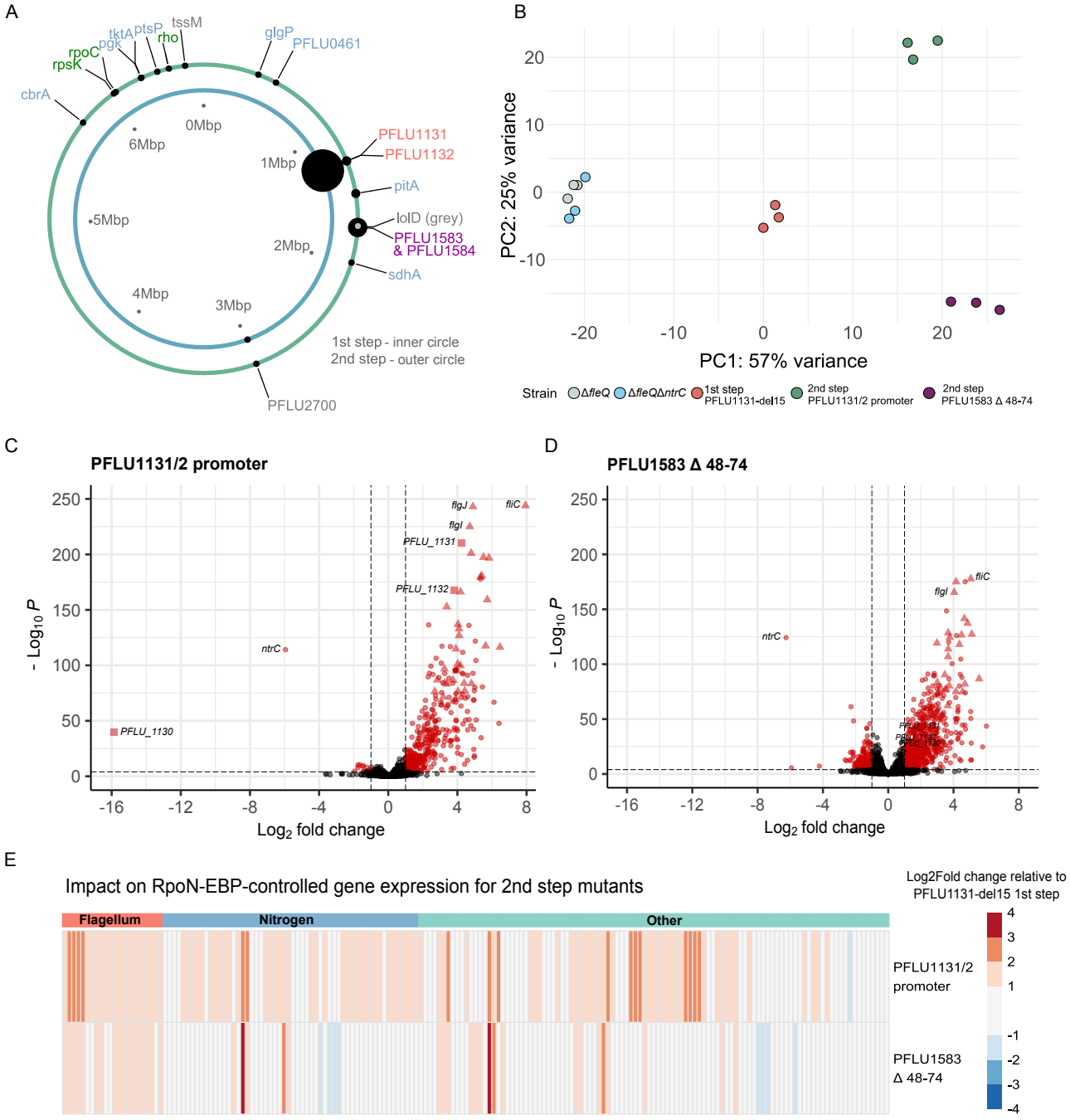


Figure 4: Phenotypes of motility rescue mutants. A) Time to emergence of motile zone (Days) for first-and second-step mutants in the Δ*fleQ*Δ*ntrC* and Δ*fleQ* backgrounds. Percentages above each plot indicate proportion of replicate evolving that step within 6 week assay cutoff. **** indicates significant difference in mean time to emergence from all other groups (Dunn test). **B)** Race assay as measure of motility fitness. Distance moved over 24hrs in 0.25% Agar LB plates measured relative to the Δ*fleQ ntrB*-T97P mutant. P-value indicated is generated from a two-sample Wilcox test. **C)** Fitness in LB broth measured as area under the 24hr growth curve (AUGC) relative to Δ*fleQ ntrB*-T97P. P-values above plots generated from Dunn tests. AUGC values and growth curve plots are provided in supplementary Fig.S3. For all boxplots – box represents first to third quartile range, middle line represents median value, whiskers range from quartiles to maxima and minima.

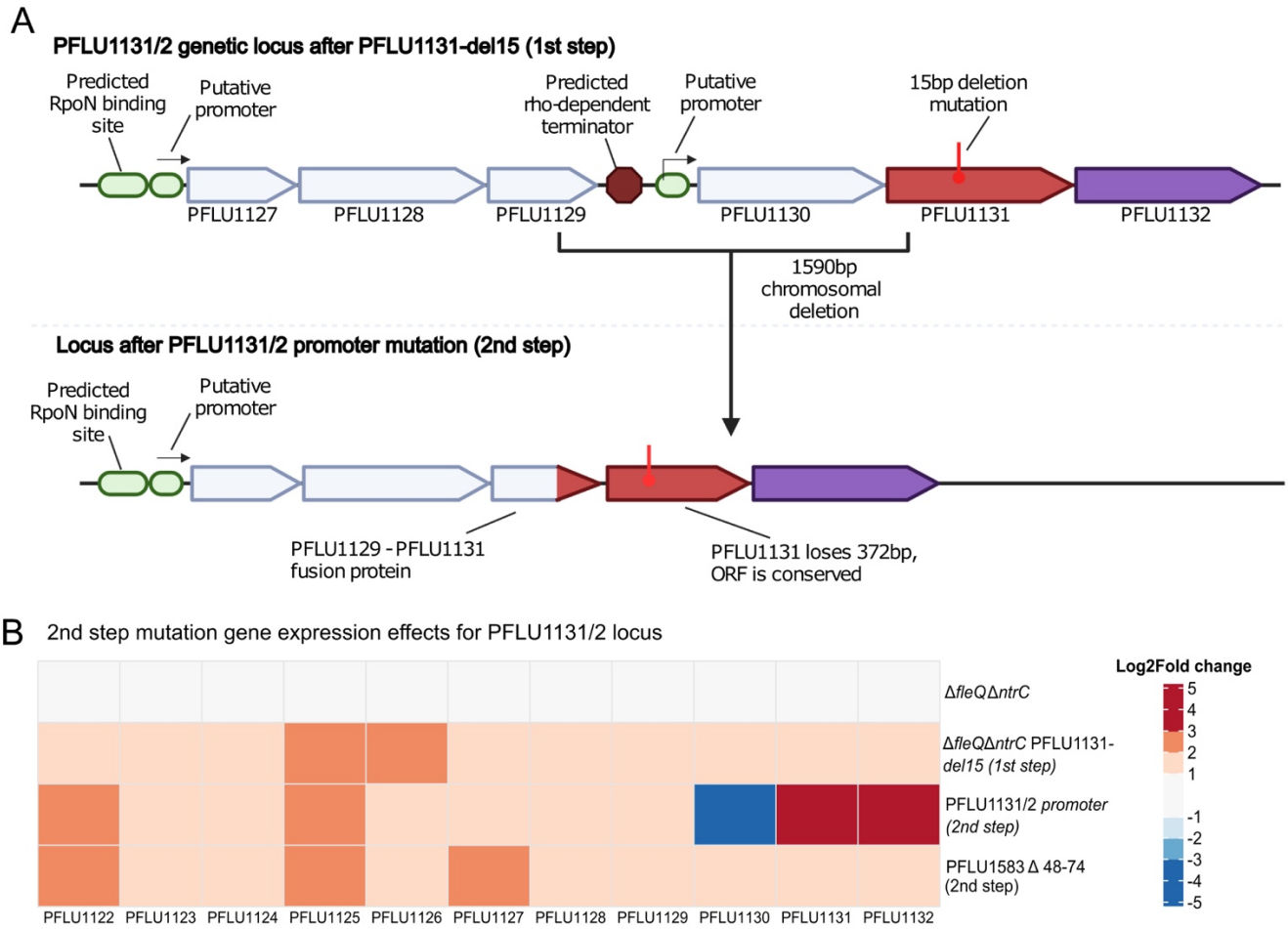


1012 **Figure 5: Diverse second-step mutations indicate multiple global regulatory**
1013 **strategies for enhancing promiscuity through PFLU1132. A)** Diagram of
1014 mutational distribution across the SBW25 chromosome for first and second step
1015 motility mutants. Each ring represents the chromosome, with black dots indicating
1016 non-synonymous mutations. The size of each dot is proportional to the number of
1017 mutations occurring in that locus in independent replicates. Mutational target names
1018 are coloured by functional category: Red – *PFLU1131/2*, Purple – *PFLU1583/4*, Blue
1019 – Metabolic, Green – Global regulatory, Grey – other. **B)** Principle component
1020 analysis of transcriptomic data for ancestral, first-step and second-step motility
1021 mutants. **C and D)** Volcano plots of differentially expressed genes for second-step
1022 mutations relative to Δ*fleQ* ancestor. Red points are significantly differentially
1023 expressed. Triangles indicate flagellar genes, Squares indicate *PFLU1131/2* and
1024 adjacent genes, and Circles indicate all other genes. **E)** Log2Fold change in gene
1025 expression for RpoN-EBP controlled genes after *PFLU1583* Δ48-74 mutation –
1026 relative to the first-step *PFLU1131*-del15 mutant. RpoN-EBP controlled genes
1027 grouped by function are indicated by the coloured bars above the plot: Red –
1028 Flagellum, Blue – Nitrogen, Teal – Other/unknown
1029
1030


A
tssM
tktA
ptsP
rho
pgk
rpoC
rpsK
glgP
PFLU0461
cbrA
0Mbp
6Mbp
1Mbp
PFLU1131
PFLU1132
pitA
5Mbp
lolD (grey)
PFLU1583
& PFLU1584
2Mbp
sdhA
4Mbp
3Mbp
1st step - inner circle
2nd step - outer circle
PFLU2700
B
PC2: 25% variance
PC1: 57% variance
Strain
ΔfleQ
ΔfleQΔntrC
1st step PFLU1131-del15
2nd step PFLU1131/2 promoter
2nd step PFLU1583 Δ 48-74
C
PFLU1131/2 promoter
- Log10 P
Log2 fold change
flgJ
fliC
flgI
PFLU_1131
PFLU_1132
ntrC
PFLU_1130
D
PFLU1583 Δ 48-74
- Log10 P
Log2 fold change
fliC
flgI
ntrC
E
Impact on RpoN-EBP-controlled gene expression for 2nd step mutants
Log2Fold change relative to PFLU1131-del15 1st step
Flagellum
Nitrogen
Other
PFLU1131/2 promoter
PFLU1583 Δ 48-74

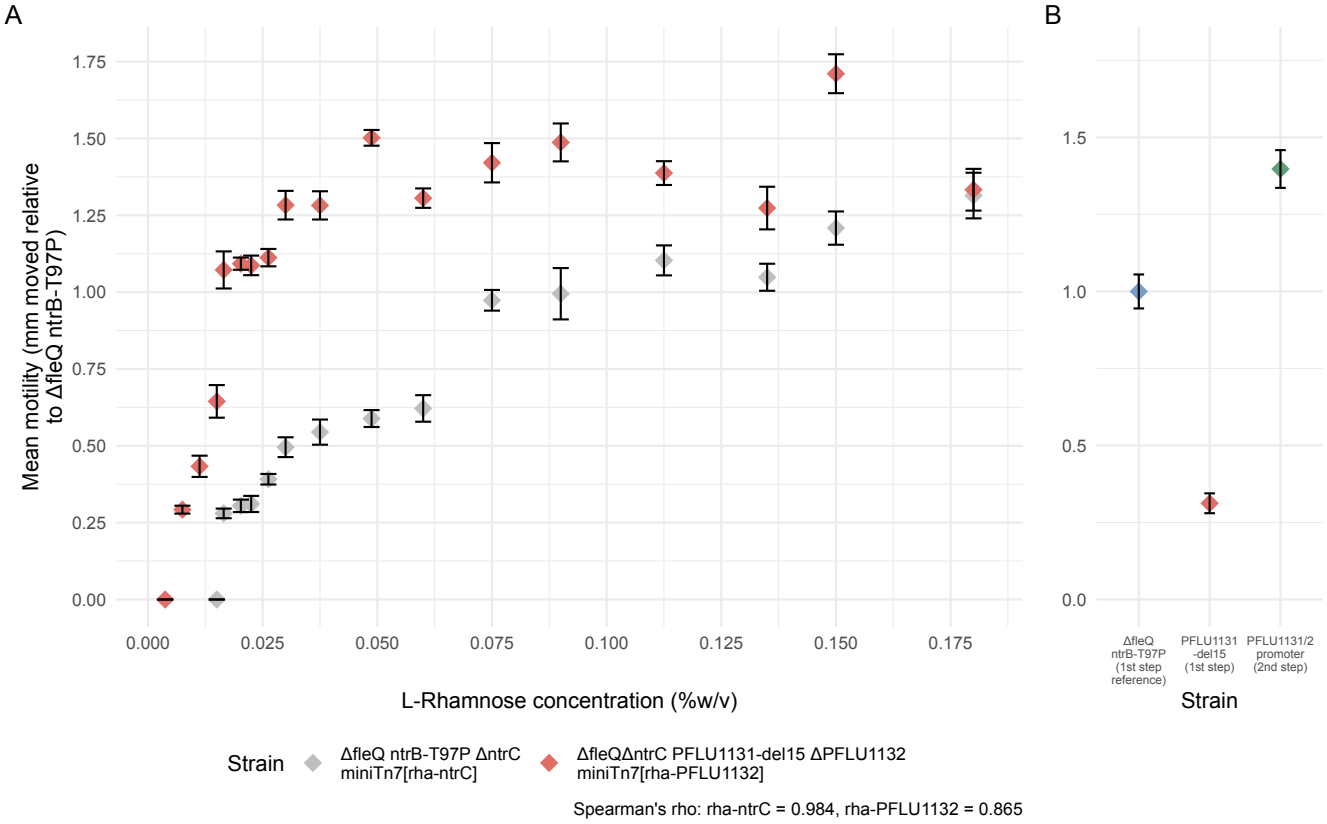
1031
1032

1033 **Figure 6: Second-step *PFLU1131/2* promoter mutation increases *PFLU1131/2***
1034 **expression. A)** *PFLU1130/1/2* genetic locus in the SBW25 chromosome, and
1035 impact of the 1.59kbp deletion identified in one second-step motile mutant. Location
1036 of the RpoN binding site upstream of PFLU1127/8/9 was predicted by Jones et al.,
1037 (2007). The presence of the rho-dependent terminator was determined by use of
1038 RhoTermPredict[71] **B)** Log2Fold change in gene expression (relative to Δ*fleQ*
1039 ancestor) heatmap of *PFLU1130/1/2* and nearby RpoN-EBP controlled genes,
1040 including *PFLU1127/8/9* in the ancestral Δ*fleQ*Δ*ntrC* strain, first and second step
1041 motility mutants.



1042
1043

Figure 7: Impact of RpoN-EBP gene expression on motility speed. A) Mean motility speed (mm relative to Δ*fleQ ntrB*-T97P) plotted against increasing L-Rhamnose concentration (%w/v). Mean of six biological replicates for each point. first-step motility mutants Δ*fleQ ntrB*-T97P and Δ*fleQ*Δ*ntrC PFLU1131*-del15 had their respective RpoN-EBPs knockout out, and reintroduced on a miniTn7 transposon with the RpoN-EBP under control of the PrhaBAD and *rhaSR* rhamnose inducible expression system. In full, these were – Δ*fleQ* ntrB-T97P Δ*ntrC* miniTn7[*rhaSR*-PrhaBAD-stRBS-ntrC] (rha-*ntrC*) and Δ*fleQ*Δ*ntrC PFLU1131*-del15 ΔPFLU1132 miniTn7[*rhaSR*-PrhaBAD-stRBS -PFLU1132] (rha-1132). **B)** Motility speed (mm relative to Δ*fleQ ntrB*-T97P) of common pathway ntrB-T97P first step mutant, *PFLU1131*-del15 first-step and PFLU1131/2 promoter second-step mutants. Whiskers represent standard deviation above and below the mean value.

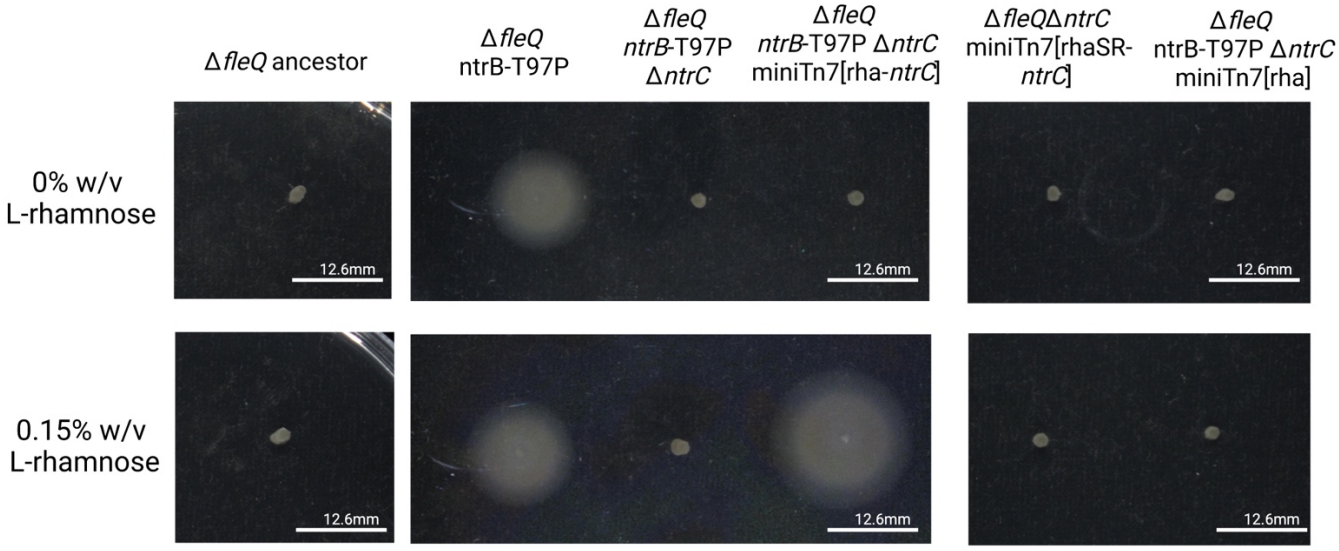


1058

1059

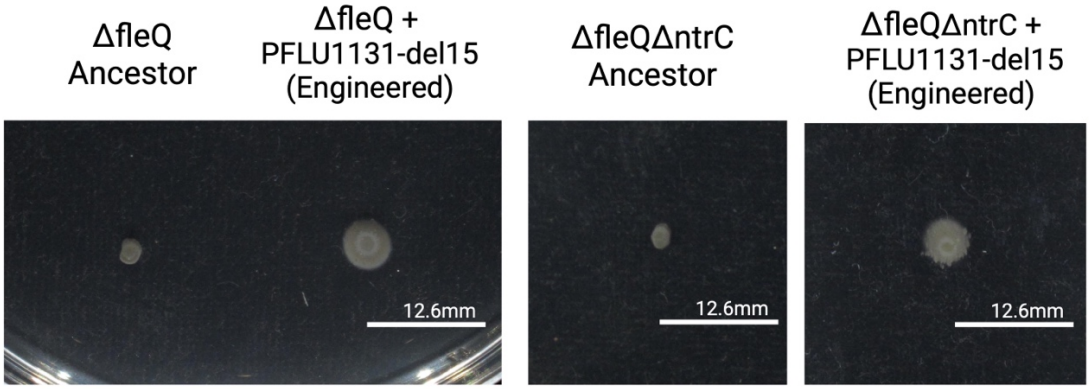
1060 **Supplementary figures**

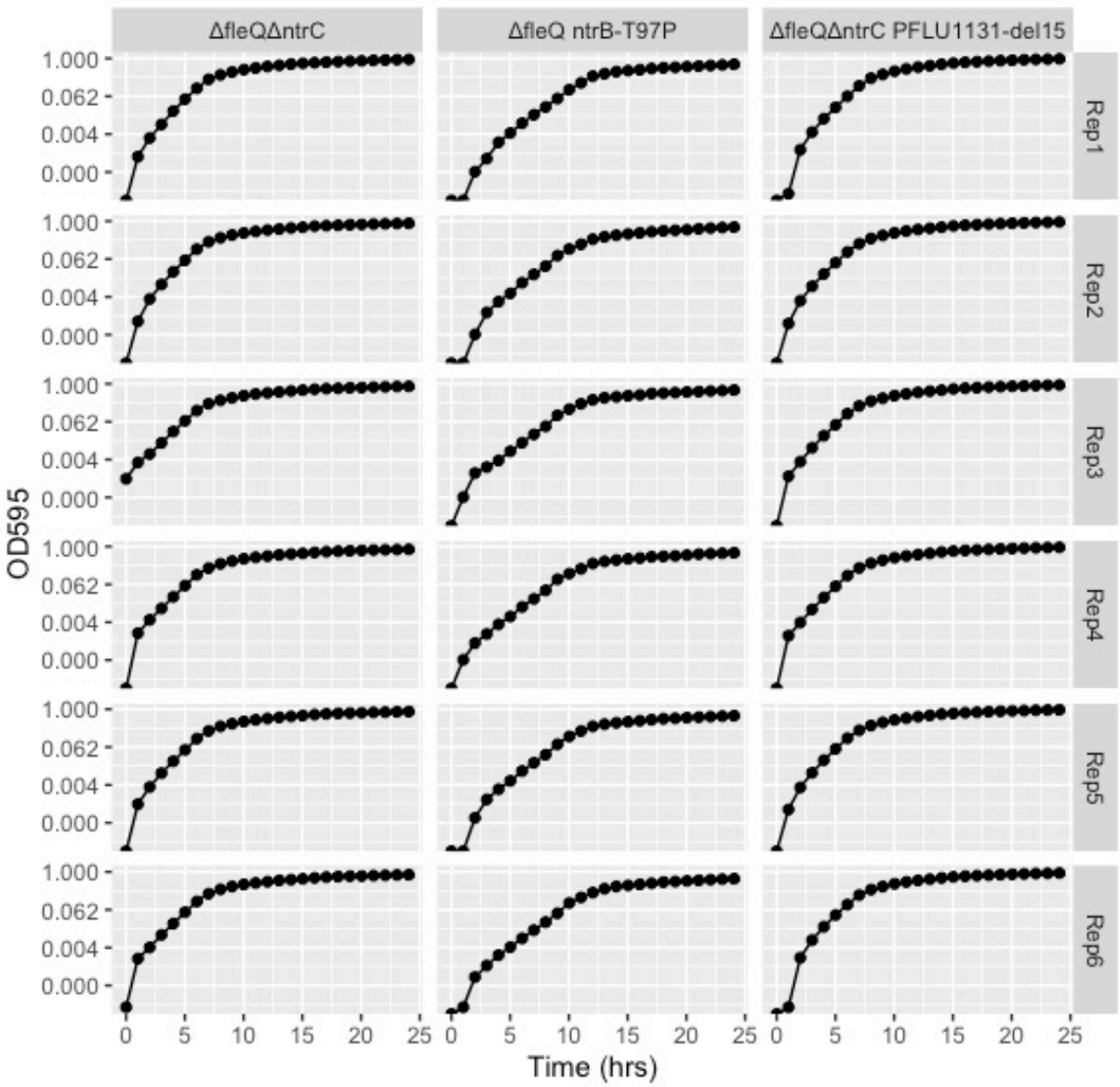
1061 ***Supplementary FigS1****:* FleQ-homolog encoding gene *ntrC* is essential for rescued
1062 flagellar motility in *ntrB* mutant strain and depends on the presence of a kinase
1063 mutation. Transcription factor gene NtrC was deleted, and then reintroduced as a
1064 single-copy chromosomal insertion expressed from an L-rhamnose inducible
1065 promoter system (*rhaSR-PrhaBAD*). The same complementation lacking the ntrB-
1066 T97P mutation, as well as an empty expression system transposon were included as
1067 further controls. Photographs of motility after 1 day incubation in 0.25% agar LB
1068 plates supplemented with or without 0.15% L-rhamnose for induction of transcription
1069 factor expression.



1070
1071

Supplementary Figure S2: PFLU1131-del15 mutation engineered into the Δ*fleQ* and Δ*fleQ*Δ*ntrC* ancestral lines provides motility in both conditions. Indicates that rescue of motility by this mutation does not depend on Δ*ntrC* and is the sole mutation required to do so.

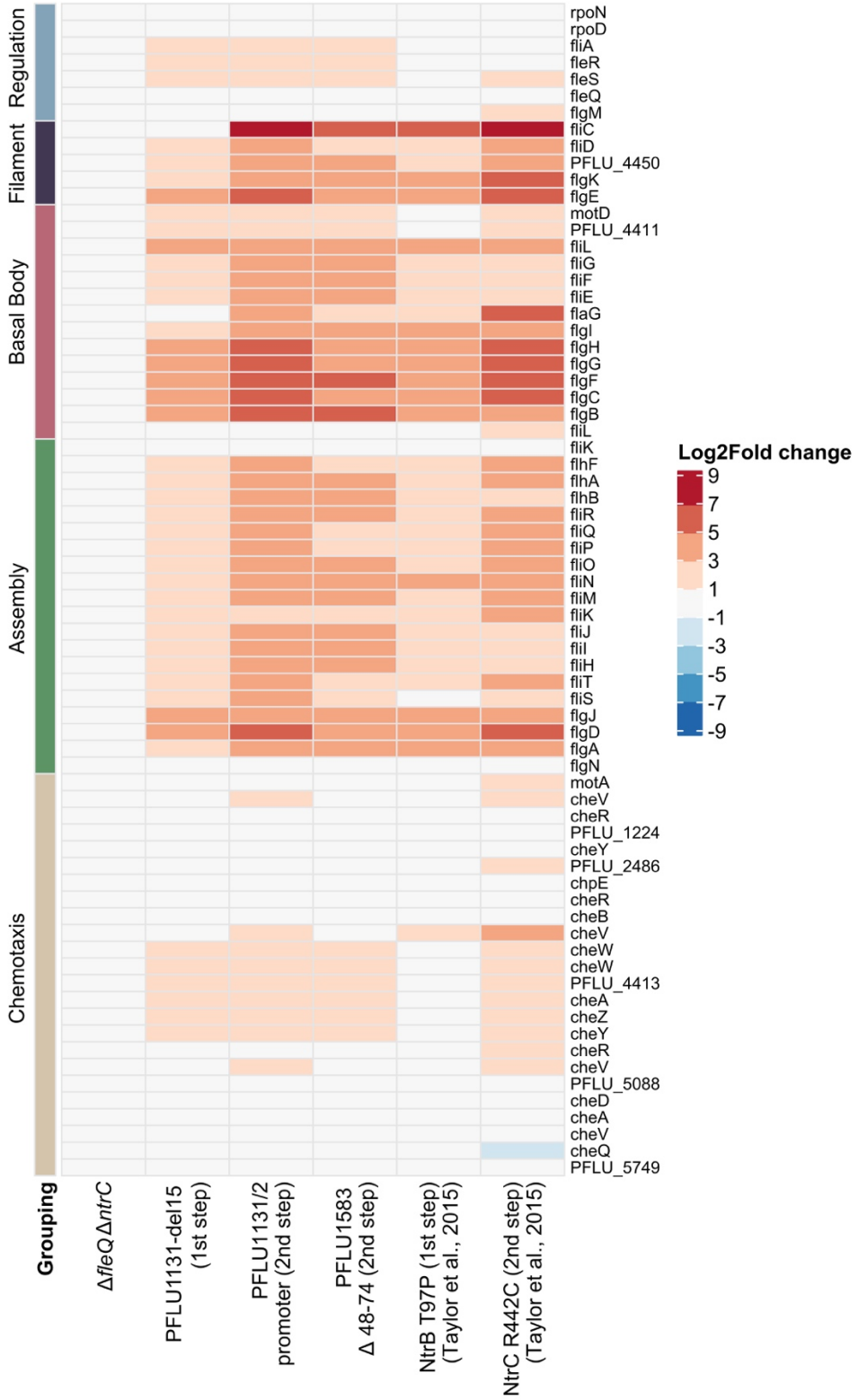




Strains (Area under growth curve - AUGC)			
Δ*fleQ*Δ*ntrC*	*ntrB*-T97P	PFLU1131-del15	Replicate
11.658605	6.393025	11.3905	Rep1
11.513825	6.358134	10.943426	Rep2
11.596943	6.674162	10.755457	Rep3
11.385187	6.434133	10.448776	Rep4
11.830055	6.381795	10.696076	Rep5
12.16704	5.922087	10.351345	Rep6

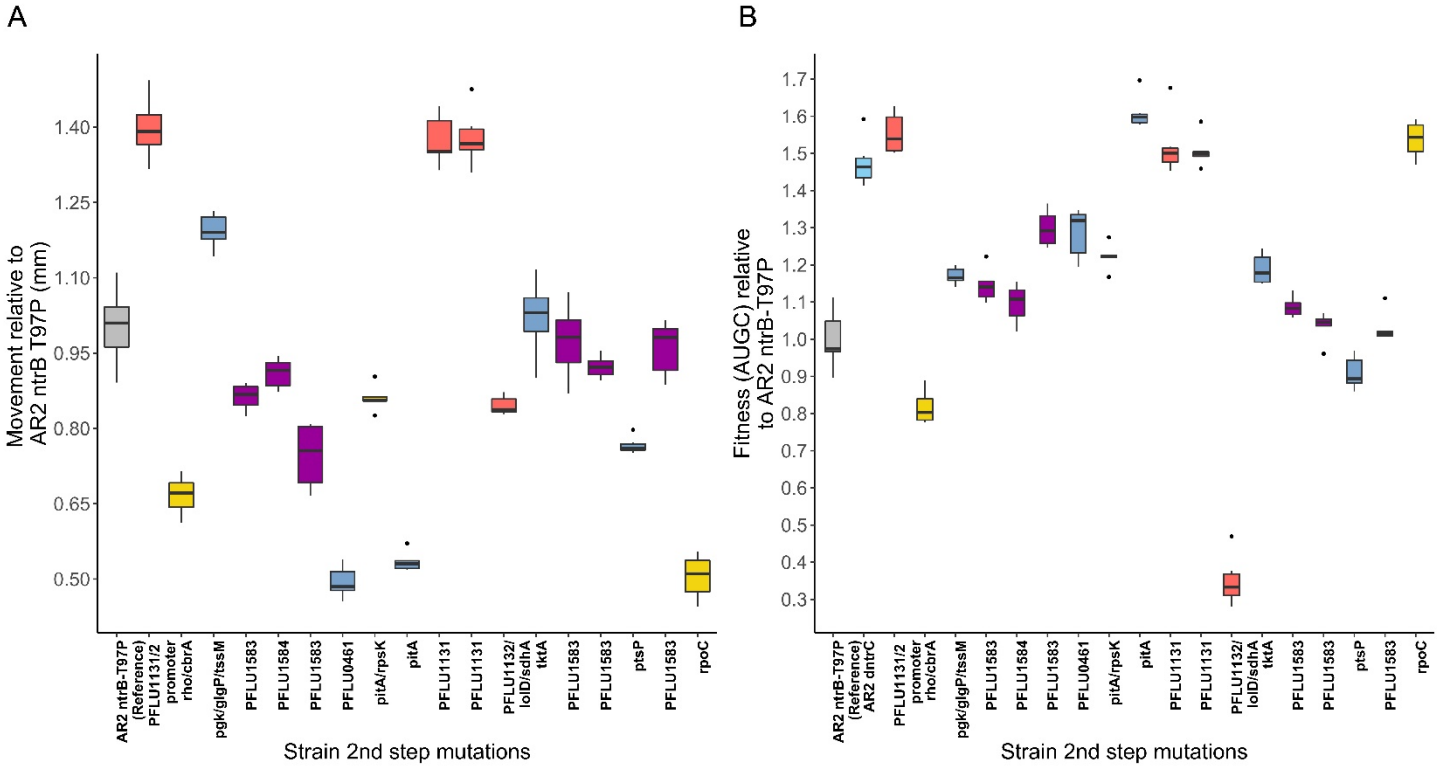
Supplementary Figure S3: Individual microbial growth curves that provide area under the growth curve (AUGC) values that are plotted in Fig.4C. Six biological replicates for each of the three strains were assayed for change in OD_{595} over 24 hours of incubation in shaking LB broth. AUGC values corresponding to each curve are displayed in the table below the graphs.

1120 ***Supplementary Figure S4:*** Impact of motility rescue mutations of flagellar and
1121 chemotaxis gene expression. Log2Fold changes in gene expression relative to the
1122 Δ*fleQ* ancestor are shown for all genes associated with flagellar motility in *P.*
1123 *fluorescens* SBW25. Functional groups of genes are indicated by the coloured bars
1124 and labels to the left of the plot.

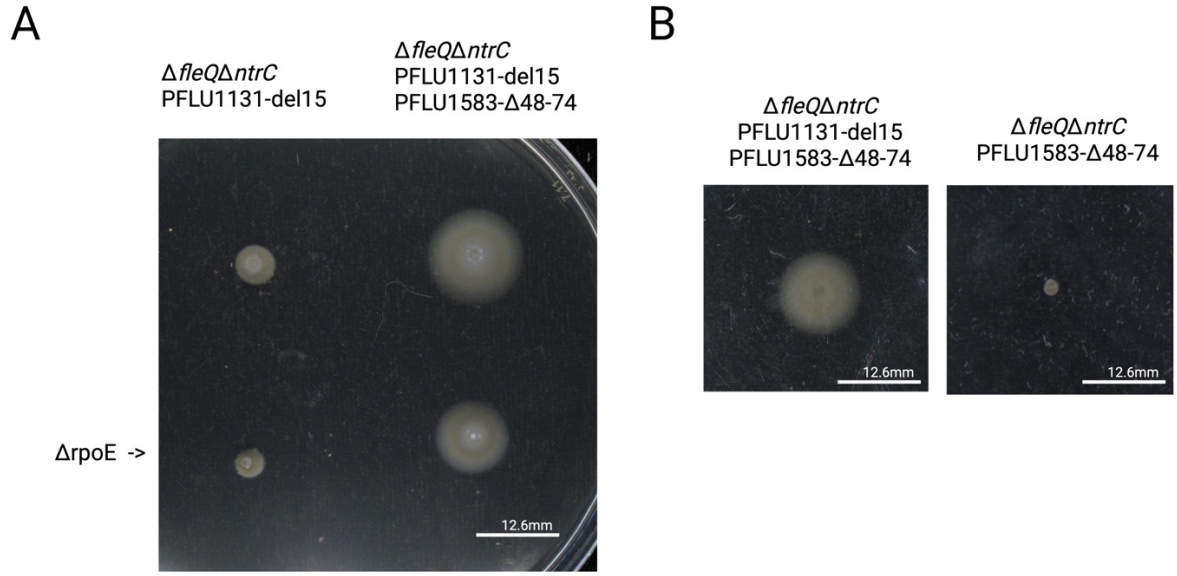
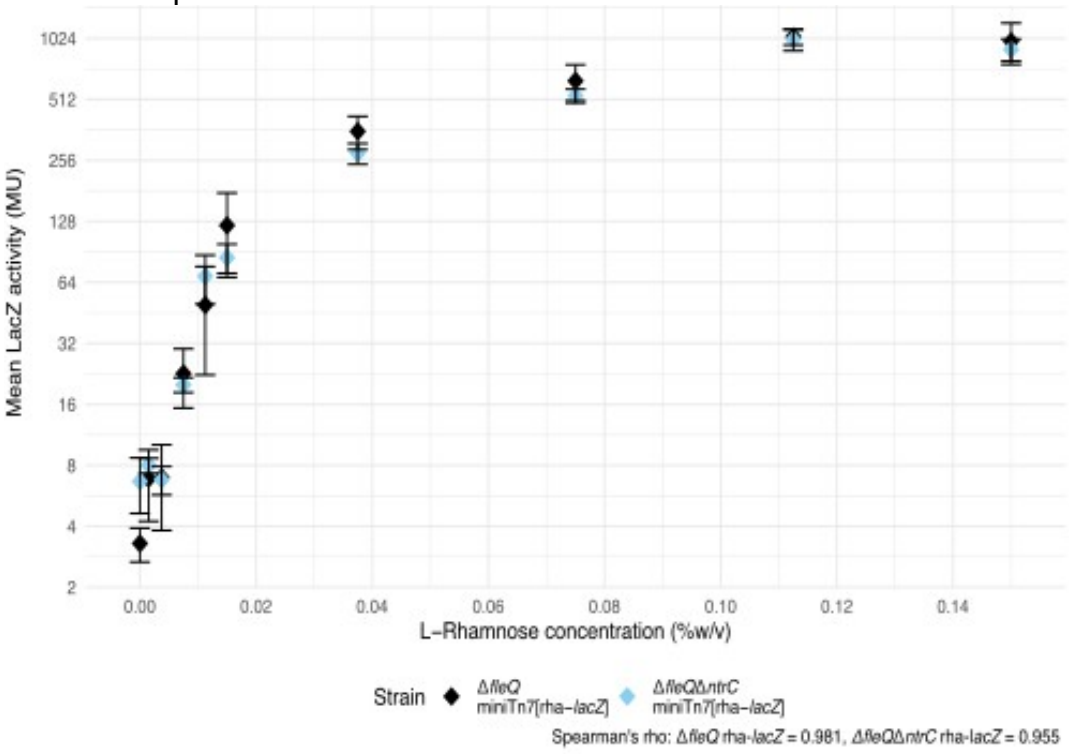


1125
1126

Supplementary Figure S5: Phenotypes of motility rescue second-step mutants. **A)** Race assay as measure of motility fitness. Distance moved over 24hrs in 0.25% Agar LB plates measured relative to the Δ*fleQ* ancestor *ntrB*-T97P mutant. **B)** Fitness in LB broth measured as area under the 24hr growth curve (AUGC) relative to Δ*fleQ* ancestor *ntrB*-T97P. In both plots, mutations are coloured by functional category: Red – PFLU1131/2, Purple – PFLU1583/4, Blue – Metabolic, Yellow – Global regulatory, Grey – other. For all boxplots – box represents first to third quartile range, middle line represents median value, whiskers range from quartiles to maxima and minima.



Supplementary Figure S6: PFLU1583 Δ48-74 mutation depends to presence of PFLU1131 mutation, and does not act through sigma factor RpoE. **A)** Motility of PFLU PFLU1583 Δ48-74 with and without accompanying *PFLU1131*-del15 mutation. **B)** *rpoE* knockout does not revert PFLU1583 mutant to first-step motility phenotype.



Supplementary Figure S7: Activity of LacZ expression reporter (Miller units – MU) with increasing L-rhamnose concentration (%w/v) in both the Δ*fleQ* and Δ*fleQ*Δ*ntrC* ancestral genetic backgrounds with miniTn7[rhaSR-PrhaBAD-stRBS-LacZ]. Whiskers represent standard deviation above and below the mean value.