



Citation for published version:

Shepherd, MJ, Horton, JS & Taylor, TB 2022, 'A Near-Deterministic Mutational Hotspot in *Pseudomonas fluorescens* Is Constructed by Multiple Interacting Genomic Features', *Molecular Biology and Evolution*, vol. 39, no. 6, msac132. <https://doi.org/10.1093/molbev/msac132>

DOI:

[10.1093/molbev/msac132](https://doi.org/10.1093/molbev/msac132)

Publication date:

2022

Document Version

Peer reviewed version

[Link to publication](#)

Publisher Rights

CC BY

University of Bath

Alternative formats

If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

1 **A near-deterministic mutational hotspot in *Pseudomonas fluorescens* is constructed by**
2 **multiple interacting genomic features**

3 *Authors: M. J. Shepherd**, *J. S. Horton*[#]*, and *T. B. Taylor[#]*

4 Milner Centre for Evolution, Department of Biology & Biochemistry, University of Bath,
5 Claverton Down, Bath BA2 7AY, UK

6 *These authors contributed to this work equally.

7 **#Corresponding Authors:** James S. Horton, Milner Centre for Evolution, Department of
8 Biology & Biochemistry, University of Bath, Claverton Down, Bath BA2 7AY, UK, +44
9 (0)1225 385116, j.s.horton@bath.ac.uk.

10 Tiffany B. Taylor, Milner Centre for Evolution, Department of Biology & Biochemistry,
11 University of Bath, Claverton Down, Bath BA2 7AY, UK, +44 (0)1225 384398,
12 t.b.taylor@bath.ac.uk

13 The contributions to this work were as follows:

14 **Experimental design** conceived by JSH with assistance from MJS. **Experimental work**
15 performed by MJS with assistance from JSH. **Manuscript written** by MJS and JSH equally.
16 **Manuscript editing** by all authors. **Project supervised** by TBT. **Funding secured** by TBT.

17 **Acknowledgements:** We thank Prof. Laurence Hurst and Prof. Heath Murray for comments
18 on earlier versions of this manuscript. This work was funded by a Royal Society Research
19 Fellows Grant (RG160491; awarded to TBT) supporting MJS; a University of Bath University
20 Research Studentship Account (awarded to TBT) and a BBSRC NI grant (BB/T012994/1;
21 awarded to TBT) supporting JSH; a Royal Society Dorothy Hodgkin Research Fellowship
22 (DH150169; awarded to and supporting TBT). Figure 1 panels A and B were created using
23 BioRender.com. Bioinformatics analyses for the letter was carried out using the Medical
24 Research Council's (MRC) Cloud Infrastructure for Microbial Bioinformatics (CLIMB), and
25 Illumina Whole-Genome Sequencing conducted by MiGs, Pittsburgh, PA, USA.

26 **Main text word count (inclusive of in-text citations):** 2478.

27 **Abstract**

28 Mutation – whilst stochastic – is frequently biased toward certain loci. When combined with
29 selection this results in highly repeatable and predictable evolutionary outcomes. Immotile
30 variants of the bacterium *Pseudomonas fluorescens* (SBW25) possess a ‘mutational hotspot’
31 that facilitates repeated occurrences of an identical *de novo* single nucleotide polymorphism
32 when re-evolving motility, where $\geq 95\%$ independent lines fix the mutation *ntrB* A289C.
33 Identifying hotspots of similar potency in other genes and genomic backgrounds would prove
34 valuable for predictive evolutionary models, but to do so we must understand the genomic
35 features that enable such a hotspot to form. Here we reveal that genomic location, local
36 nucleotide sequence, gene strandedness and presence of mismatch repair proteins operate in
37 combination to facilitate the formation of this mutational hotspot. Our study therefore provides
38 a framework for utilising genomic features to predict and identify hotspot positions capable of
39 enforcing near-deterministic evolution.

40 **Main text**

41 A growing body of evidence has revealed that mutation bias is a key determinant of
42 evolutionary outcomes (Couce and Tenailon 2019; Cano et al. 2022; Monroe et al. 2022).
43 Contrary to the expectation that mutations simply provide a random supply of genetic diversity
44 for selection to act upon, differences in mutation rates across genomic position (Monroe et al.
45 2022) and mutation types (Cano et al. 2022) can introduce bias into the mutational spectra that
46 has been shown to shape adaptive trajectories. In some circumstances, such biases can be strong
47 enough that a particular position mutates at a rate far higher than expected by chance,
48 generating what is referred to as a mutational hotspot (Zhang et al. 2018; Horton et al. 2021).
49 Mutation bias is of key interest to those wishing to make forecasts of adaptive evolution, as
50 mutational hotspots can facilitate highly repeatable, and by extension reliably predictable,
51 adaptive outcomes. Identifying such hotspots across the bacterial domain and understanding
52 their role in adaptive evolution is therefore an important challenge.

53 In previous work, we found that a mutational hotspot drove repeatable outcomes in immotile
54 variants of *Pseudomonas fluorescens* SBW25 (denoted AR2; see Alsohim et al. 2014; Taylor
55 et al. 2015). During the re-evolution of flagellar motility, $>95\%$ of replicate lines realised an
56 identical *de novo* single nucleotide polymorphism (SNP), A289C in the *ntrB* gene (Horton et
57 al. 2021). Just six synonymous SNPs defined whether strong mutation bias occurred at position

58 289 within the *ntrB* locus, and these changes were theorised to impact the formation of a single
59 stranded DNA secondary structure (Horton et al. 2021). Elucidating biased hotspots of similar
60 potency in other bacterial genomes would inform predictive evolutionary models and bring us
61 closer to forecasting adaptive trajectories. But we must first understand the generalisable
62 genomic principles that allow such hotspots to occur.

63 In this work, we use the *P. fluorescens* mutational hotspot as a model system to elucidate the
64 key genomic features responsible for facilitating heavily biased localised mutation. As this
65 hotspot is sufficiently potent to generate an identical mutation in $\geq 95\%$ of instances, the system
66 provides a unique setting where a hotspot can endure to some degree when partially
67 disassembled. This allows us to augment genomic features known to affect mutation bias in
68 isolation, and in combination, and quantify their individual and combinatorial impact on the
69 mutational hotspot.

70 We began with the hypothesis that the *ntrB* 289 hotspot is reliant on the formation of single-
71 stranded DNA hairpins. If true, then a slew of genomic features can interfere with or facilitate
72 this hotspot-generating mechanism (Trinh and Sinden 1991; Viswanathan et al. 2000; Hoede
73 et al. 2006; Wang and Vasquez 2017). Hairpin formation does not generate mutation on its
74 own, but rather indirectly by way of its interference of DNA polymerase, where it can force
75 the replisome to stall (Voineagu et al. 2008). DNA polymerase fidelity is influenced by
76 replication timing and its correlated genomic position (Dillon et al. 2018), with replisomes
77 being less vulnerable to stalling closer to the origin of replication (OriC). Therefore the
78 hairpin's effect may be weakened when replication fidelity is higher. We investigated this
79 hypothesis by translocating the locus to a genomic position $\sim 348\text{kbp}$ closer to OriC.

80 Translocation was accomplished by knockout of the native *ntrBC* operon, followed by re-
81 introducing *ntrBC* on a miniTn7 transposon insertion system that reliably inserts in a single
82 defined chromosomal site downstream of the *glmS* gene (Choi and Schweizer 2006; Fig.1A).
83 This conserves native operon structure, regulatory elements and strandedness (Fig.1B).
84 Translocating the *ntrBC* operon did not significantly impact expression of either gene
85 according to RT-qPCR expression analysis (Supplementary Table S1). This method measures
86 a population aggregate of expression profiles and therefore we cannot dismiss differences
87 between evolving subpopulations, however the assay does provide confidence that
88 subsequently observed changes in mutation spectra are due to genomic context rather than gene
89 expression. The intact hotspot in its native location ($\sim 376\text{kbp}$ from OriC) leads to 95.2% of

90 motility rescuing mutations being the *ntrB* SNP A289C, and 100% of replicate lines evolving
91 within 6 weeks (Fig. 2). The *ntrBC* translocation strain (AR2 miniTn7[*ntrBC*-Lag], ~28kbp
92 from OriC) displayed a non-significant but increased median time to emergence of motility
93 compared to AR2 from 3.83 days to 5.35 days ($P = 0.035$, Dunn test), accompanied by a
94 reduction in the rate of adaptation, with 95.65% lines evolving in 6 weeks (Fig.2A).

95 Translocating *ntrBC* resulted in decreased potency of the A289C hotspot (Fig. 2B). Alternative
96 *de novo ntrB* mutations occurred at raised frequencies in AR2 miniTn7[*ntrBC*-Lag], including
97 a 12bp deletion $\Delta 410-421$ (22.7%), and the SNP G682A (9.1%). Aside from *ntrB* mutations,
98 9.1% of mutations were observed in *glnK*, a gene encoding an NtrB-repressor (Hervás et al.
99 2009) that has also been observed to permit *ntrBC*-mediated rescue of motility (Taylor et al.
100 2015). These mutations came at the expense of mutation *ntrB* A289C in the translocated lines.
101 The frequency of the SNP dropped from 95.2% in AR2 to 50% in AR2 miniTn7[*ntrBC*-Lag],
102 yielding a significantly different mutation spectrum overall ($P = 0.0015$, Chi-square test).
103 Furthermore, local nucleotide context continued to play a major role in mutation frequency at
104 this site, as the introduction of 6 synonymous changes around position 289 in the novel
105 genomic context (closer to OriC) generated a similar effect to that observed previously (Horton
106 et al. 2021), with A289C frequency falling from 50% to 0% (Fig.2B).

107 These results show that genomic position closer to OriC lowers hotspot potency but does not
108 remove it entirely, until however subsequent synonymous changes are introduced. This
109 supports the broader hypothesis that replisome stalling at a hairpin site facilitates heavily biased
110 mutation at *ntrB* 289. It should, however, be noted that these observations may be explained
111 by alternative interacting mechanisms. DNA polymerase fidelity is dependent on both
112 replication timing (Dillon et al. 2018), as well as local nucleotide context including nucleotide
113 triplets (Long et al. 2014) and possibly flanking dinucleotides (Krawczak et al. 1998).
114 Therefore, the two mechanisms may additively facilitate the hotspot without requiring the
115 formation of a hairpin. Similarly, the rate of stalling may differ between the two genomic
116 locations due to differences in local DNA topology (Postow et al. 2001), rather than by
117 replication timing.

118 Stalled replisomes are vulnerable to generating mutations in part because they are vulnerable
119 to collisions with RNA polymerases (Paul et al. 2013). Head-on collisions occur when genes
120 are encoded on the lagging replicative strand, meaning that the two complexes process DNA
121 in opposing directions, resulting in head-on contact. As *ntrB* is natively encoded on the lagging

122 strand, head-on collisions may increase mutation frequency across the locus. However, if a
123 DNA hairpin is stalling the replisome, then collisions with RNA polymerase may be enriched
124 to occur at the hotspot site (Wang and Vasquez 2017), driving the localised mutation bias.
125 Alternatively, a mechanism independent of collisions may still be strand dependent, as
126 replisome-stalling hairpin structures have more opportunity to form when encoded on the
127 lagging strand (Bikard et al. 2010). Therefore, if hairpin-replisome interactions or collisions
128 are essential for mutagenesis, swapping the strand encoded by the gene should remove the
129 hotspot, even when the local nucleotide sequence that facilitates replisome stalling remains
130 intact. We experimentally examined the effect of gene strandedness by manipulating the
131 encoded position of the *ntrBC* genes from the lagging to the leading DNA strand at the new
132 genomic position (Fig.1B).

133 We first observed that switching the strandedness of the *ntrBC* locus on the miniTn7 transposon
134 (AR2 miniTn7[*ntrBC*-Lead]) impacted the rate of adaptation, increasing median time to
135 emergence relative to AR2 miniTn7[*ntrBC*-Lag] from 5.35 days to 8.75 days ($P = 0.0038$,
136 Dunn test), (Fig.2A). The percentage of replicate populations evolving within 6 weeks also
137 decreased from 95.65% for AR2 miniTn7[*ntrBC*-Lag] to 74.26% for AR2 miniTn7[*ntrBC*-
138 Lead]. The switch from lagging to leading strand additionally eradicated the mutational hotspot
139 effect. The *ntrB* A289C SNP was no longer observed in the mutational spectra, and other *de*
140 *novo ntrB* mutations accounted for only 22.2% of motility rescuing mutations. Mutations were
141 instead observed in other previously identified motility-granting mutational targets *glnK*, *glnA*,
142 and PFLU1131 at frequencies of 66.7%, 7.4% and 3.7% respectively (Fig. 2B), producing a
143 vastly different mutation spectrum to its lagging strand counterpart ($P = 0.00005$, Chi-square
144 test). Loss of the *ntrB* A289C hotspot mutation could not be explained by a drop in viability or
145 motility fitness for this SNP in the altered genomic contexts, as no significant difference in
146 motility speed ($P = 0.2667$, Dunn test) was found when this SNP was engineered on the leading
147 or lagging strand (Supplementary Fig.S1).

148 The mutational data therefore show that gene strandedness is essential to hotspot formation.
149 Additionally, if replisome stalling followed by RNA polymerase collisions are driving hotspot
150 mutagenesis (see Supplementary Table S2), encoding the gene on the leading strand would
151 nullify head-on contact. As such, synonymous mutations that prevent stalling, possibly by
152 abolishing hairpin formation or creating less mutable local sequence, should no longer impact
153 the observed mutational spectrum, as head-on collisions are removed in either case. Neither

154 AR2 miniTn7[*ntfBC*-Lead] or AR2 miniTn7[*ntfBC*-sm-Lead] realised any *ntfB* A289C
155 mutations. Alternative mutations within the locus dropped from 22.2% to 9.1% following
156 synonymous mutation, however *ntfB* mutations seen in the AR2 miniTn7[*ntfBC*-sm-Lead]
157 background were A683G and G682A, which were also seen in other strains in this study. The
158 rest of the mutations for AR2 miniTn7[*ntfBC*-sm-Lead] were 72.7% *glnK*, including multiple
159 observations of Δ 258-272 also observed in the non-sm dataset, 4.5% PFLU1131, and 9.09%
160 *ntfC*. Overall, there was no significant difference in the mutational target on the locus level (P
161 = 0.51, Chi-square test). And despite novel mutational routes being discovered, there was also
162 no significant difference in observed mutation spectra for individual mutations between the
163 synonymous variants (P = 0.061, Chi-square test). This result reinforces that genomic context
164 has a direct impact on the likelihood of mutation at a potential hotspot position. Local
165 nucleotide sequence does not operate in isolation but relies on a prominent interplay with
166 genomic position and gene strandedness to drive the specific occurrence of the A289C SNP.

167 As well as genomic features that are directly involved in the construction of mutational
168 hotspots, it is also important to consider general indirect means by which the mutational spectra
169 can be affected. A prominent example of this is the mismatch repair (MMR) system that is
170 often lacking in mutator strains. MMR systems across numerous bacterial species preferentially
171 correct transition mutations, as mutator strains lacking these genes exhibit transition biases
172 (Schaaper and Dunn 1987; Long et al. 2014). The mutation generated by the *ntfB* 289 hotspot
173 is a transversion mutation, and as such may be more able to dominate the mutational spectrum
174 as the MMR system actively prevents adaptive transitional changes from becoming
175 immortalised in the daughter DNA strands. To test this hypothesis, we constructed and evolved
176 lines of a mismatch defective mutant of AR2 (AR2 Δ *mutS*), which lacks a key part of the
177 mismatch repair protein MutS, the component responsible for binding DNA (Schofield et al.
178 2001).

179 We observed that AR2 Δ *mutS* strains displayed a non-significant reduced mean time to motility
180 (Fig 3), from 4.20 days in AR2 lines to 2.45 days in the mutator lines (P = 0.034, Dunn test).
181 In contrast, the degree of mutational parallelism and spectra across strain backgrounds differed
182 significantly (P = 0.0011, Chi-square test). 35% AR2 Δ *mutS* lines fixed *ntfB* A289C, 40%
183 reported SNPs either elsewhere in *ntfB* or in *glnK*, and 25% harboured unidentified mutations
184 outside of these loci. Mutational repeatability of the *ntfB* A289C mutational hotspot had
185 therefore fallen from 95% to 35% in mutator lines. However, this was not owed to a reduction

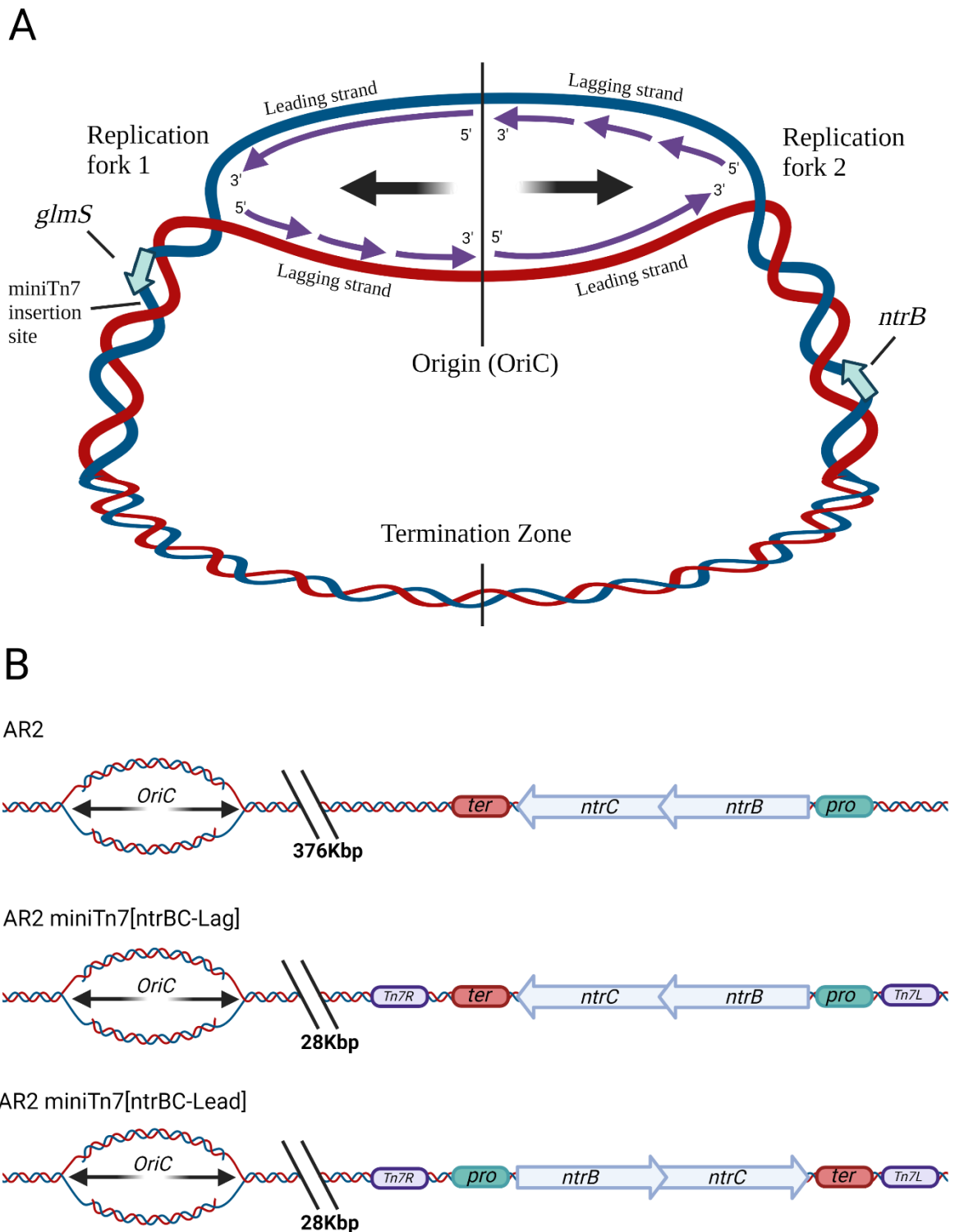
186 in mutation bias operating at the hotspot, but rather an elevation in the realisation of alternative
187 adaptive mutations. *ntrB* A289C mutations were realised sooner in mutator strains ($P = 0.0012$,
188 Wilcoxon test; Fig. 3), but so too were alternative transition mutations (identified mutations
189 are plotted in Fig 3). The 65% non-A289C mutations in AR2 $\Delta mutS$ were realised ≤ 3 days,
190 whereas the 4.8% non-A289C mutations in AR2 were realised ≤ 6 days. Furthermore, all
191 identified non-A289C mutations in AR2 $\Delta mutS$ (40% of total sample) were transition
192 mutations.

193 A single nucleotide can mutate to three alternate nucleotides, two of which are transversion
194 mutations (e.g., $A \rightarrow C$ and $A \rightarrow T$) and the other a transition mutation (e.g., $A \rightarrow G$).
195 Therefore, if we expect transitions to represent 33% of all mutations and assume an equal
196 likelihood of fixation regardless of mutation type, then there is no significant enrichment for
197 either mutation type (transition or transversion) in the mutator lines (Bootstrap test, $n = 1 \times 10^6$,
198 $P > 0.33$). In contrast, there is a significant omission of transitions in an AR2 background where
199 the hotspot transversion remains in effect (Bootstrap test, $n = 1 \times 10^6$, $P < 0.0023$). As such
200 these results show that the mutator strains unlock alternative transition mutations that are
201 suppressed in lines with intact mismatch repair machinery. As the A289C transversion
202 similarly appears more frequently in mutator lines, mismatch repair complexes likely also
203 correct transversion mutations at the hotspot site. Therefore, while the *ntrB* 289 hotspot remains
204 potent in both mutator and non-mutator genomic backgrounds, the mutational spectrum will
205 less heavily favour the hotspot transversion mutation in mutator lines, where alternative
206 mutation types become more common.

207 Together, these results reveal that genomic position, gene strandedness, local genetic sequence
208 and the presence of MMR proteins all operate in concert to generate a near-deterministic
209 mutational hotspot. The interplay of these features may be owed to hairpin formation that stalls
210 the replisome at its position, enriching a collision point for the replisome and RNA polymerase.
211 The bias in mutation spectra is additionally indirectly enforced by MMR proteins, which
212 correct alternate adaptive transition mutations. Mutational hotspots have been argued in some
213 cases to be maintained by natural selection, primarily in evolutionary circumstances where a
214 transient and reversible change in phenotype is beneficial (Moxon et al. 2006). However, the
215 biased mutational event occurring via the mechanisms implicated here is likely not reversible,
216 and therefore the hotspot would degrade under fluctuating selection. Instead, it may well be
217 that the genomic context facilitating the hotspot evolved through neutral evolution, which

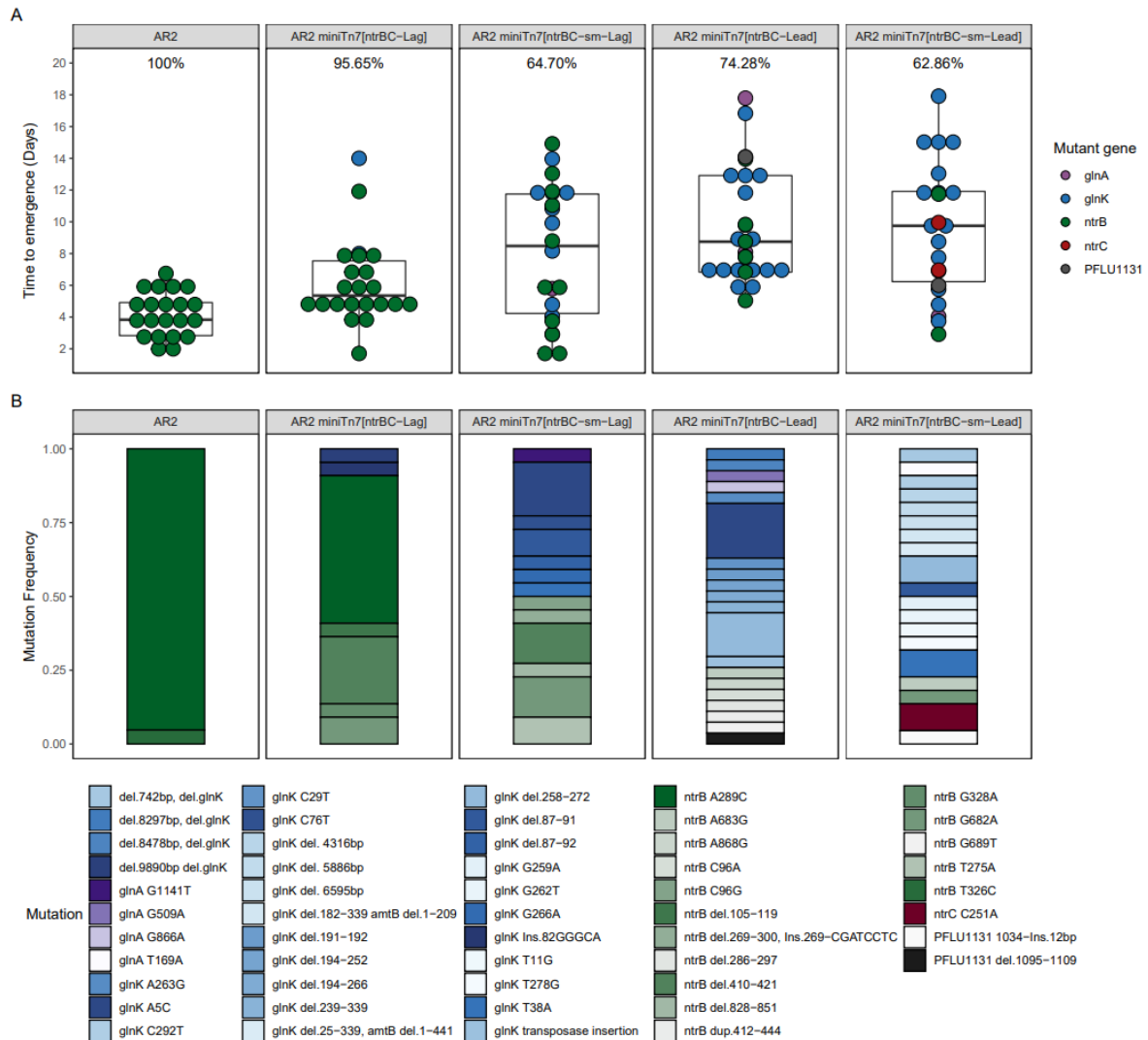
218 generated the potential for a skewed adaptive landscape (Tenailon and Matic 2020). If this is
219 the case, it suggests that hotspots may be found throughout bacterial genomes, and not merely
220 within alleles under transient selection.

221 This work helps expand our knowledge of near-deterministic mutational hotspots away from
222 isolated genomic contexts by highlighting interacting mutable genomic features that are each
223 pervasive throughout the bacterial kingdom. As such, future work that quantifies the impact of
224 these features in model organisms other than *P. fluorescens* will help provide a generalisable
225 genomic framework for searching and identifying similarly potent mutational hotspots
226 throughout bacterial genomes. Therefore, this work will aid in facilitating future accurate
227 forecasts of bacterial evolution and contribute toward our understanding of the role that
228 mutation bias plays in determining adaptive evolutionary outcomes.

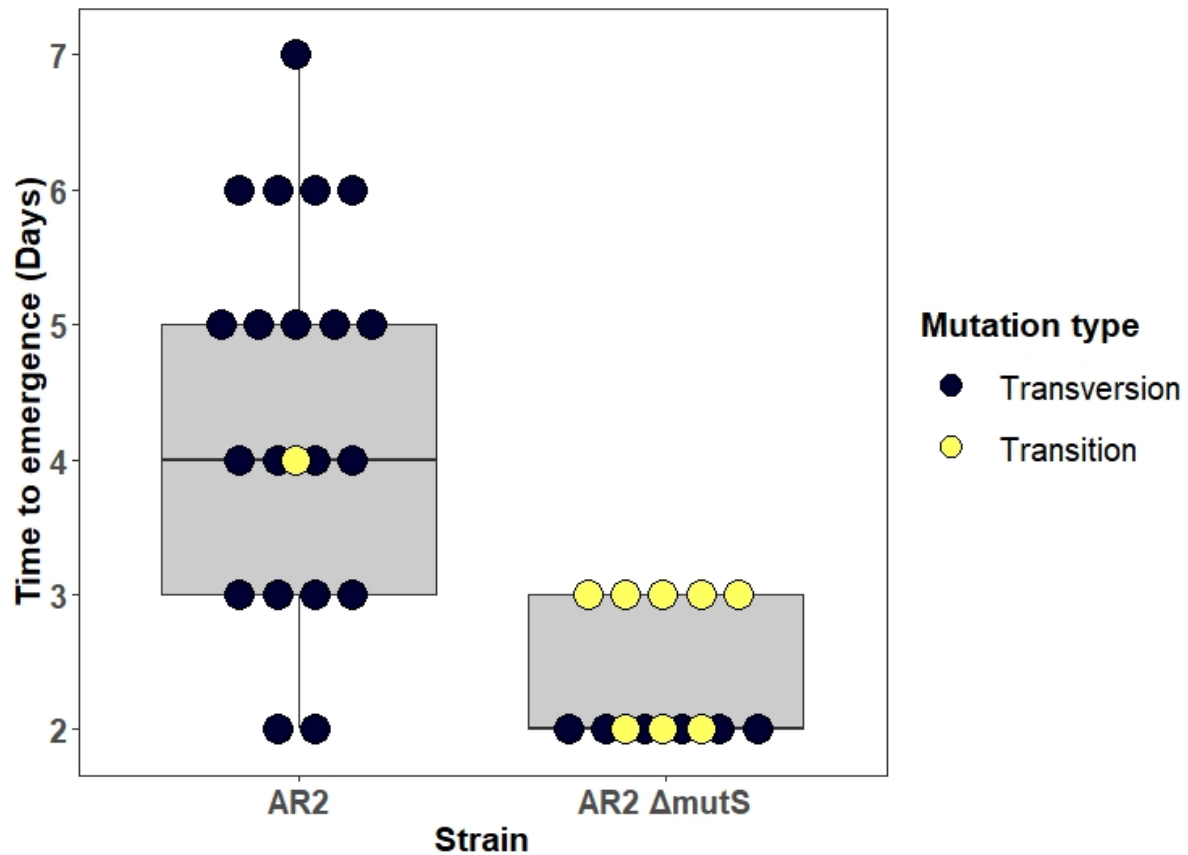


229 **Figure 1:** Manipulation of chromosomal locus and DNA strandedness of *ntrBC*. Black arrows
 230 indicate direction replication fork movement. **A)** Bisymmetrical structure of the circular
 231 bacterial chromosome undergoing theta-replication, with two mirror-image replication forks
 232 moving out from the origin of replication (*OriC*). Synthesis of leading and lagging strands are

233 shown by purple arrows. **B)** Altered genetic contexts of the *ntrBC* locus. Orientation with
234 respect to the replication fork, and distance from OriC show for ancestral (AR2) and the
235 engineered strains AR2 miniTn7[*ntrBC*-Lag] and AR2 miniTn7[*ntrBC*-Lead]. Pro and ter
236 denote *ntrBC* promoter and terminator regions respectively. Tn7R and Tn7L denote miniTn7
237 transposon flanking sites.



238 **Figure 2:** Impact of *ntrBC* translocation, gene strandedness and synonymous variation on
 239 mutation bias for rescuing flagellar motility in AR2-based strains. (*N* for each condition
 240 (evolved/total): AR2 – 21/21, miniTn7[*ntrBC*-Lag] – 22/23, miniTn7[*ntrBC*-sm-Lag] – 22/34,
 241 miniTn7[*ntrBC*-Lead] – 26/35, miniTn7[*ntrBC*-sm-Lead] – 22/35) **A**) Time to emergence of
 242 motility in days for each *ntrBC* strain background. Boxplots display mean and quartile values.
 243 Datapoints for individual replicate lines are shown and coloured by mutant gene identified. The
 244 percentage of replicates evolving motility within 6 weeks is given above each boxplot. **B**)
 245 Frequency of *de novo* mutations identified in motile isolates. Unique mutations have unique
 246 colours. Mutations in the same gene are grouped with shades of the same colour (*ntrB* = greens,
 247 *glnK* = blues, *glnA* = purples, PFLU1131 = greys, *ntrC* = red).



248 **Figure 3:** Removal of a mis-match repair complex uncovers alternative adaptive transition
 249 mutations. Independent replicates of mutator variants (AR2 Δ mutS) realised motility in non-
 250 significantly less time than the AR2 ancestor ($P = 0.034$, Dunn test) but yielded a significantly
 251 different mutational spectra ($P = 0.0011$, Chi-square test). The transversion mutation *ntrB*
 252 A289C was observed in 20/21 cases in the ancestral line, with the remaining observation a
 253 transition mutation *ntrB* T326C. 15/20 mutator lines had identifiable mutations in *ntrB* or *glnK*;
 254 these data points are plotted. 7/15 mutator lines harboured the transversion mutation *ntrB*
 255 A289C. The remaining 8/15 were transition mutations: *ntrB* T323C, T407C, A608G (x2),
 256 A683G, and *glnK* T11C, A131G, A263G.

257 **References**

- 258 Alsohim AS, Taylor TB, Barrett GA, Gallie J, Zhang XX, Altamirano-Junqueira AE, Johnson
259 LJ, Rainey PB, Jackson RW. 2014. The biosurfactant viscosin produced by
260 *Pseudomonas fluorescens* SBW25 aids spreading motility and plant growth promotion.
261 *Environ. Microbiol.* 16:2267–2281.
- 262 Bikard D, Loot C, Baharoglu Z, Mazel D. 2010. Folded DNA in Action: Hairpin Formation
263 and Biological Functions in Prokaryotes. *Microbiol. Mol. Biol. Rev.* 74:570–588.
- 264 Cano A V., Rozhoňová H, Stoltzfus A, McCandlish DM, Payne JL. 2022. Mutation bias
265 shapes the spectrum of adaptive substitutions. *PNAS.* 119:1–11.
- 266 Choi KH, Schweizer HP. 2006. mini-Tn7 insertion in bacteria with single attTn7 sites:
267 Example *Pseudomonas aeruginosa*. *Nat. Protoc.* 1:153–161.
- 268 Couce A, Tenaillon O. 2019. Mutation bias and GC content shape antimutator invasions. *Nat.*
269 *Commun.* 10 (1).
- 270 Dillon MM, Sung W, Lynch M. 2018. Periodic Variation of Mutation Rates in Bacterial
271 Genomes. *MBio* 9:1–15.
- 272 Hervás AB, Canosa I, Little R, Dixon R, Santero E. 2009. NtrC-dependent regulatory
273 network for nitrogen assimilation in *Pseudomonas putida*. *J. Bacteriol.* 191:6123–6135.
- 274 Hoede C, Denamur E, Tenaillon O. 2006. Selection Acts on DNA Secondary Structures to
275 Decrease Transcriptional Mutagenesis. *PLOS Genet.* 2:1697–1701.
- 276 Horton JS, Flanagan LM, Jackson RW, Priest NK, Taylor TB. 2021. A mutational hotspot
277 that determines highly repeatable evolution can be built and broken by silent genetic
278 changes. *Nat. Commun.* 12.
- 279 Krawczak M, Ball E V., Cooper DN. 1998. Neighboring-nucleotide effects on the rates of
280 germ-line single-base- pair substitution in human genes. *Am. J. Hum. Genet.* 63:474–
281 488.
- 282 Long H, Sung W, Miller SF, Ackerman MS, Doak TG, Lynch M. 2014. Mutation rate,
283 spectrum, topology, and context-dependency in the DNA mismatch repair-deficient

- 284 *Pseudomonas fluorescens* ATCC948. *Genome Biol. Evol.* 7:262–271.
- 285 Monroe JG, Srikant T, Carbonell-Bejerano P, Becker C, Lensink M, Exposito-Alonso M,
286 Klein M, Hildebrandt J, Neumann M, Kliebenstein D, et al. 2022. Mutation bias reflects
287 natural selection in *Arabidopsis thaliana*. *Nature* 602:101–105.
- 288 Moxon R, Bayliss C, Hood D. 2006. Bacterial Contingency Loci : The Role of Simple
289 Sequence DNA Repeats in Bacterial Adaptation. *Annu. Rev. Genet.* 40:307–335.
- 290 Paul S, Million-Weaver S, Chattopadhyay S, Sokurenko E, Merrikh H. 2013. Accelerated
291 gene evolution via replication-transcription conflicts. *Nature* 495:1–13.
- 292 Postow L, Crisona NJ, Peter BJ, Hardy CD, Cozzarelli NR. 2001. Topological challenges to
293 DNA replication: Conformations at the fork. *PNAS* 98:8219–8226.
- 294 Schaaper M, Dunn RL. 1987. Spectra of spontaneous mutations in *Escherichia coli* strains
295 defective in mismatch correction : The nature of in vivo DNA replication errors. *PNAS*
296 84:6220–6224.
- 297 Schofield MJ, Brownnewell FE, Nayak S, Du C, Kool ET, Hsieh P. 2001. The Phe-X-Glu
298 DNA Binding Motif of MutS. *J. Biol. Chem.* 276:45505–45508.
- 299 Taylor TB, Mulley G, Dills AH, Alsohim AS, McGuffin LJ, Studholme DJ, Silby MW,
300 Brockhurst MA, Johnson LJ, Jackson RW. 2015. Evolutionary resurrection of flagellar
301 motility via rewiring of the nitrogen regulation system. *Science* 347:1014–1017.
302 Available from: <http://www.sciencemag.org/lookup/doi/10.1126/science.1259145>
- 303 Tenaille O, Matic I. 2020. The Impact of Neutral Mutations on Genome Evolvability. *Curr.*
304 *Biol.* 30:R527–R534.
- 305 Trinh TQ, Sinden RR. 1991. Preferential DNA secondary structure mutagenesis in the
306 lagging strand of replication in *E. coli*. *Nat. Lett.* 352:544–547.
- 307 Viswanathan M, Lacirignola JJ, Hurley RL, Lovett ST. 2000. A Novel Mutational Hotspot in
308 a Natural Quasipalindrome in *Escherichia coli*. *JMB* 302:553–564.
- 309 Voineagu I, Narayanan V, Lobachev KS, Mirkin SM. 2008. Replication stalling at unstable
310 inverted repeats : Interplay between DNA hairpins and fork stabilizing proteins. *PNAS*

311 105:9936–9941.

312 Wang G, Vasquez KM. 2017. Effects of replication and transcription on DNA Structure-
313 Related genetic instability. *Genes*. 8:1.

314 Zhang Xiaolong, Zhang Xuehong, Zhang Xia, Liao Y, Song L, Zhang Q, Li P, Tian J, Shao
315 Y, Li Y, et al. 2018. Spatial Vulnerabilities of the *Escherichia coli* Genome. *Genetics*
316 210:547–558.

317