1 **Chance and pleiotropy dominate genetic diversity**

2 **in complex bacterial environments**

3

4 **Lianet Noda-Garcia¹, Dan Davidi², Elisa Korenblum², Assaf Elazar¹,**

5 **Ekaterina Putintseva³, Asaph Aharoni² & Dan S. Tawfik^{1,*}**

6 ¹Department of Biomolecular Sciences, Weizmann Institute of Science, Rehovot

7 76100, Israel.

8 ²Department of Plant and Environmental Sciences, Weizmann Institute of

9 Science, Rehovot 76100, Israel.

10 ³Institutes of Science and Technology, Klosterneuburg, Austria.

11

12 * Corresponding author; e-mail: dan.tawfik@weizmann.ac.il.

13

14

15

16

17 **Abstract**

18 How does environmental complexity affect the evolution of single genes? Here,
 19 we measured the effects of a set of mutants of *Bacillus subtilis* glutamate
 20 dehydrogenase across 19 different environments – from homogenous single cell
 21 populations in liquid media to heterogeneous biofilms, plant roots and soil
 22 communities. The effects of individual gene mutations on organismal fitness
 23 were highly reproducible in liquid cultures. Strikingly, however, 84% of the
 24 tested alleles showed opposing fitness effects under different **growth conditions**
 25 **(environmental pleiotropy)**. In biofilms and soil samples, different alleles
 26 dominated in parallel replica experiments. Accordingly, we found that in these
 27 **heterogeneous cell communities** the fate of mutations was dictated by a
 28 combination of selection and drift. The latter **relates to** programmed prophage
 29 excisions that occurred along biofilm development. Overall, per individual
 30 condition, by the combined action of selection, pleiotropy and chance, a wide
 31 range of glutamate dehydrogenase mutations persisted and sometimes fixated.
 32 However, across longer periods and multiple environments, nearly all this
 33 diversity would be lost – indeed, considering all environments and conditions we
 34 have tested, wild-type is the fittest allele.

35

36

37

38 The function of most genes may be essential in some conditions, but only
 39 marginally contributing or even **deleterious** in other conditions ¹⁻⁴. The effects of
 40 mutations on organismal fitness are therefore environment-dependent, giving
 41 rise to complex, pleiotropic genotype-by-environment interactions ^{5,6}. (Here, we
 42 refer to different effects of the same mutation as ‘environmental pleiotropy’ ⁷, or
 43 ‘pleiotropy’ for brevity, and opposing effects in different environments as
 44 ‘antagonistic pleiotropy’). Moreover, bacterial populations often do not comprise
 45 single cells, but rather have a structure as in biofilms. Under this complexity:
 46 changing environments and heterogeneous communities (cell or/and species
 47 wise), the fate of mutations could also be dictated by population bottlenecks
 48 (drift) or rapid takeover of beneficial mutations in other genes (selective
 49 sweeps) ⁸⁻¹⁰. Consequently, the frequency of a given gene allele may change
 50 dramatically (from perishing to fixation) with no relation to its **protein** function
 51 ^{11,12}.

52 We aimed at an experimental setup **that** would examine how complex
 53 bacterial growth states and environments might shape protein evolution.
 54 Previous systematic mappings were based on a direct linkage between protein
 55 stability and function and organismal survival, thus enabling measurements of
 56 effects of mutations at the protein level ^{5,13-16}. However, how mutations in a
 57 single gene-protein affect organismal fitness under varying environments and
 58 conditions is largely unexplored ¹⁷. We thus chose as our model *Bacillus subtilis*
 59 NCIB 3610, a non-domesticated strain capable of growing in diverse aquatic and
 60 terrestrial environments ¹⁸. We explored the effects of mutations in different
 61 conditions: in dispersed cells in liquid, but also in biofilms where phenotypic and
 62 genetic variability prevails ¹². We also mapped the effects of mutations during

63 spore formation and germination ¹⁹ and in more complex and close to natural
64 environments including soil, rhizosphere and plant roots.

65 A catabolic glutamate dehydrogenase (GDH) was our model protein. This
66 enzyme is essential when amino acids such as proline serve as sole carbon-
67 nitrogen sources ²⁰. However, in the presence of ammonia and glycolytic sugars,
68 GDH activity is redundant as glutamate must be synthesized rather than
69 catabolized. GDHs therefore respond to changes in carbon-nitrogen sources, and
70 as regulators of glutamate homeostasis, are also associated with biofilm
71 development ^{21,22}. *B. subtilis* has two catabolic GDHs, RocG and GudB. The latter
72 is constitutively expressed, and is regulated via association of its hexameric form
73 ²³. GudB has also regulatory roles ^{24,25} via interactions with the transcriptional
74 activator of glutamate synthase ²⁵ and with a transcription termination factor
75 that modulates the stringent response ²⁶. We explored mutations in the
76 oligomeric interface of GudB, aiming at multilateral effects on GudB's enzymatic
77 and regulatory functions.

78 Altogether, these choices of organism and enzyme allowed us to readily
79 examine and quantify the fate of GudB alleles in a range of different growth
80 conditions and environments, also mimicking natural habitats where strong
81 evolutionary forces act ¹².

82

83 **Experimental setup and data processing**

84 We anticipated that the effects of the explored mutations would be complex and
85 condition-dependent. We thus opted for deep rather than broad coverage and
86 mapped 10 positions within a single ~150 base pairs segment that resides at
87 GudB's oligomeric interface. The mutagenized positions were arbitrarily chosen

88 while aiming to map highly conserved positions (*e.g.* 58, D in all GudB
89 orthologues) as well as diverged ones (*e.g.* 48, or 61; variable amongst GudB
90 orthologues; **Table S1**). The mutagenized codons were diversified to NNS,
91 whereby N represents any of the 4 bases, and S represents G or C. We thus
92 created 10 libraries, each diversifying a single position into 20 different amino
93 acids *plus* a stop-codon. The libraries were incorporated into the chromosome of
94 *B. subtilis* NCIB 3610 under *gudB*'s original promoter and terminator. The
95 combined library contained 320 single mutant alleles, whose genomes differ, in
96 principle, by a single GudB mutation, including 200 different amino acid alleles
97 (including wild-type), 10 stop-codons, and also synonymous alleles whereby the
98 same amino acid could be encoded by 2 or 3 different codons.

99 This starting library (the initial mix, hereafter) was used to inoculate
100 cultures grown in an array of different conditions. We tested 7 different growth
101 states with diverse complexity, from single cells to communities: liquid, pellicles
102 (air-liquid biofilms), spores, germinated spores, biofilms grown on agar
103 including on carbon-nitrogen gradients, and colonized soil. Up to 5 different
104 carbon-nitrogen sources were used that, at least as far as the phenotypes of the
105 GudB knockout indicate, inflict different levels of selection on GudB: Glutamate
106 *plus* ammonia (GA) where Δ GudB has no growth effect; glutamate *plus* glycerol
107 (GG), arginine (A), and arginine *plus* proline (PA), where Δ GudB exhibits a slight
108 growth defect, and proline (P) where Δ GudB exhibits the strongest growth defect
109 (**Fig. S1**). In total, we tested 19 conditions. At each condition, three to five
110 biological replicas were performed by inoculating from the same initial mix. The
111 replicas were grown in parallel, and individually analyzed. Illumina sequencing
112 was applied to determine the frequency of each of the *gudB* alleles in the initial

113 mix and after growth. Following filtering (see Methods), we obtained data for
114 244 up to 269 individual alleles per experiment (**Data S1 & Fig. S2**).

115 The ratio between an allele's frequency at the end of growth and in the
116 initial mix was derived, and this ratio is referred to as the frequency coefficient
117 (FC; **Data S2**). Basically, $FC > 1$ means an enriched, beneficial mutation, and $FC <$
118 1 a purged deleterious one. However, given the experimental error in
119 determining FC values, values between 0.8 and 1.2 were classified as 'neutral', $FC \leq 0.8$
120 assigned a mutation as 'deleterious', and $FC > 1.2$ as 'beneficial' (see
121 Methods). Mutations with $FC \leq 0.1$ were classified as 'highly deleterious', and
122 similarly, $FC \geq 10$ as 'highly beneficial'. FC values reflect the relative frequency of
123 alleles, and therefore relate logarithmically to their relative fitness effects (or
124 selection coefficient, s). Hence, $\log FC$ values were compared throughout. Note,
125 however, that the number of generations differs fundamentally between
126 conditions – e.g. ~50 generations in liquid (following 5 serial transfers into a
127 fresh culture) versus effectively no replication in spores (a dormant non-
128 replicative form of *B. subtilis*). Moreover, in pellicles and biofilms, the number of
129 generations cannot be easily determined, and in biofilms different cell types (i.e.
130 matrix producers, dormant cells, etc.) have different growth rates ^{27,28}. So, while
131 we could not calculate selection coefficients, one should keep in mind that an FC
132 value of 0.8 in spores would effectively mean extinction across 50 generations in
133 liquid ($0.8^{50} = 10^{-5}$).

134

135 **Irreproducibility – selection versus drift**

136 Our first observations indicated two contrasting scenarios. In liquid cultures, for
137 example, we observed highly reproducible FC values between biological replicas

(**Fig. 1a**). Given the small sample numbers (3 replicas **as standard**, 5 in few cases) the observed variance may underestimate the actual variance. However, the repetitively low variance in a range of different liquid conditions, and in other replica measurements in liquid ²⁹, support high reproducibility. In biofilms, however, despite the fact that we did not bottleneck any population upon inoculum, the **poor** correlation between replicas was **evident** (**Fig. 1b**). The reproducibility between biological replicas indicates selection, **suggesting that in reproducible conditions, the fitness of *GudB*'s and of *B. subtilis* are tightly coupled**. In biofilms however, the lack of reproducibility **suggested** the dominance of drift, *i.e.*, random sampling of *GudB* alleles.

To quantify the contribution of selection versus drift in different conditions, we used two criteria. Firstly, we compared the variability in FC values between replicas by calculating the standard deviation (SD) per allele (using, by default, the logarithm of the FC values; see Methods). The average SD value for all alleles in each experiment ($\overline{SD}$) is given for **the different growth states (a growth state, *e.g.* liquid, may include several conditions, *e.g.* different carbon-nitrogen sources; **Fig. 1c; Fig. S3a & Table S2**)**. As can be seen, in liquid, pellicles and spores, the $\overline{SD}$ values between biological replicas were low (< 0.06). In biofilms and bulk soil, however, the $\overline{SD}$ values were > 0.25 indicating low reproducibility. **The Fisher test also indicated that for all tested alleles, the variance between FC values significantly changed between liquid, pellicles and spores when compared to germinated spores, biofilms and bulk soil (p values in the range of 0.048 to 1.18×10^{-27} ; **Data S2**)**.

Secondly, if drift dictates the fate of *GudB* alleles, codons of the same amino acid should exhibit very different FC values. The deviations between

163 synonymous codon alleles of the same amino acid were calculated, averaged for
 164 all alleles in the same experiment, and then for all replicas of the same
 165 experiment ($\overline{SD}_{\text{syn}}$, in log values; **Fig. 1d**; **Fig. S3b** & **Table S3**). The Levene's
 166 test confirmed that the variance between synonymous codons is significantly
 167 different across conditions (p values in the range of 0.01 to 10^{-24} ; **Table S4**).
 168 Note that the $\overline{SD}_{\text{syn}}$ criterion holds within individual replica experiments and is
 169 thus independent of the comparison of $\overline{SD}$ between biological replicas.
 170 Nonetheless, these criteria are clearly correlated (**Fig. 1c & d**). Overall, it appears
 171 that in liquid, pellicles and spores, the FC values report the outcome of selection
 172 acting on GudB alleles at the amino acid level as expected (in few alleles,
 173 selection also acted reproducibly at the codon level, **Fig. S4**). In contrast, in
 174 biofilms and bulk soil we consistently observed higher $\overline{SD}$ as well as higher
 175 $\overline{SD}_{\text{syn}}$ values. In some biofilm experiments, in effect, a single codon had taken
 176 over resulting in $\overline{SD}_{\text{syn}}$ values ≥ 3 (note that logFC values were compared
 177 throughout, and the SD for FC values is therefore $\geq 10^3$).

178 Given that some conditions were selection-dominated and others were
 179 subject to chance, we divided our analysis in two. Firstly, we analyzed selection
 180 dominated conditions (liquid, pellicles and spores) to examine whether and how
 181 GudB mutations exert different fitness effects under different environments.
 182 Secondly, conditions where drift prevailed (germination, biofilms and soil
 183 colonization) were analyzed to reveal the relative contributions of selection
 184 *versus* chance.

185

186 **Pleiotropy - fitness-effects of mutations are condition-dependent**

187 While the FC values, and hence the fitness effects of mutations, were
 188 reproducible under many conditions, their distribution varied widely between
 189 conditions, including between carbon-nitrogen sources (**Fig. S5**). This indicates
 190 pleiotropy – individual GudB alleles have different fitness effects in different
 191 **environments**. To quantify the level of pleiotropy, we compared the FC values of
 192 the same GudB mutation across the 9 individual selection-dominated conditions.
 193 Because the number of generations differs from one condition to another, we
 194 focused on shift from beneficial to deleterious, and vice versa (sign, or
 195 antagonistic pleiotropy) **because** the sign indicates **the overall trend** irrespective
 196 of generation numbers. Representative dot plots comparing the FC values across
 197 3 different liquid conditions are shown (**Fig. 2a**). These indicate that pleiotropy
 198 is common, even when comparing liquid cultures with overlapping carbon-
 199 nitrogen sources. In particular, a significant number of GudB mutations show
 200 antagonistic pleiotropy (dashed squares, **Fig. 2a**). Indeed, the Pearson
 201 correlation values for the 36 possible pair-wise comparisons of **the 9**
 202 **reproducible** conditions were **all** below 0.7, **and many accommodated a negative**
 203 **value indicating an overall anti-correlation (i.e., dominance of antagonistic**
 204 **pleiotropy; Fig. 2b)**. Across all selection-dominated conditions, 84% of alleles
 205 showed antagonistic pleiotropy in at least one of the 36 pair wise comparisons,
 206 and 70% of alleles showed mild or strong antagonistic pleiotropy. These
 207 pleiotropic effects are far beyond experimental noise, as indicated by
 208 comparison to a control sample (**Fig. 2c**).

209 Overall, the dominance of pleiotropy meant that across all conditions
 210 **where selection acts**, 86% of the alleles were beneficial in at least one condition.
 211 However, not a single mutation was beneficial across all conditions. Further, if a

212 mutation were to be considered deleterious if purged under at least one
213 condition, then 98% of the tested GudB mutations were deleterious.

214

215 **Combined action of selection and drift in heterogeneous environments**

216 In biofilms (and also in germination and soil colonization though to a lesser
217 degree), irreproducibility between replicas, variability between codons (**Fig. 1c**
218 & **d**), and the near-fixation of relatively few alleles (**Fig. 3**), all suggested fixation
219 by chance. What is the nature of these few GudB ‘winners’, are they merely
220 lucky?

221 While drift dominated in biofilms and soil colonization, curiously, wild
222 type GudB was enriched in up to 85% of these experiments suggesting that
223 selection may also play a role (**Fig. 3**). To assess the action of selection, we
224 compared the three biofilm areas. There appears a systematic trend, whereby
225 alleles enriched in the edge are more likely to arise from alleles that persisted or
226 even enriched in the center (**Fig. 4a**). Similarly, 75% of the enriched edge alleles
227 were neutral or beneficial under liquid growth with proline, a condition under
228 which GudB experiences the strongest selection (**Fig. 4b**). This suggested that
229 although drift dominated GudB’s fate in biofilms, GudB was under selection at
230 some stage of biofilm development. Accordingly, we found that in biofilm centers
231 the FC values are less skewed and more reproducible than in the edge or
232 wrinkles (**Fig. S5b**) and the center $\overline{SD}_{syn}$ values are half (**Table S3 & Fig. S3b**).
233 The $\overline{SD}_{syn}$ values are obviously much higher in the biofilms’ center compared to
234 liquid cultures, but the trend suggests that at the onset of the biofilm’s
235 development, selection acts on GudB (**Table S3 & Fig. S3b**). Foremost, that wild
236 type GudB was present at high frequency in the vast majority of biofilm

experiments is not a coincidence that relates only to its high frequency in the initial mix. Indeed, a spiking experiment indicated that wild type was enriched by nearly 20-fold even when scarcely present in the initial mix (Fig. 3f).

Similarly, we searched signatures of selection in soil colonization – a process that involves multiple passages, beginning with a change of medium (Hoagland solution rinses; see Methods) and results in colonization of the roots.

As in biofilms, there is a statistically significant trend whereby alleles enriched in the root are more likely to arise from alleles that were enriched in the soil (Fig. 4c). Further, 19 amino acid alleles were found to be enriched in at least 10 out of the 15 sequenced populations, suggesting some degree of reproducibility (Table S5). Selection during soil colonization is also manifested in the variation between biological replicas ($\overline{SD}$ values) of alleles that were enriched in root populations being on average 20% smaller than those that were not (Fig. 4d). Finally, stop codons were purged in all biofilms and soil populations, indicating, as expected, that GudB's activity is required for *B. subtilis*' survival under these conditions^{22,23}.

Altogether, in biofilms in soil colonization both drift and selection determine the fate of GudB alleles. We further examined the biofilms as described in the next section.

Drift in biofilms relates to programmed prophage excisions

Mutagenic rates in biofilms are high and mutations with a selective advantage rapidly take over (genetic sweeps)^{30,31}. Growth in biofilms is also spatially defined, giving rise to segregated lineages whereby an entire segment of the biofilm's edge stems from a single cell in which a beneficial mutation had first

emerged¹². GudB mutations that happen to be in these ‘founder’ cells might therefore fixate along these lineages. In pellicles that are largely considered as biofilms, intercellular matrix is produced³² but spatial segregation is less pronounced than in solid biofilms. Accordingly, we found that in oppose to agar biofilms, in pellicles selection acts reproducibly (**Fig. 1**). To further establish that spatial segregation is a key factor, we divided the edges of the biofilm into small sections, and sequenced them. We found that most sections contained a single GudB allele (**Fig. S6**). Thus, in a way, the GudB allele represents a ‘barcode’ that reports single founder cells giving rise to individual sectors of the biofilm¹².

What might be the mutations driving these genetic sweeps and spatial segregation? We sequenced samples for which enough genomic DNA was available (6 ordinary and 12 gradient biofilms, and for comparison, 2 initial mix, 6 liquid and 4 pellicle samples). A range of single nucleotide polymorphisms (SNPs) in various loci was identified across these samples (**Data S3**). We focused, however, on identifying genomic mutations that were not, or scarcely observed in the initial mix and/or in liquid samples, suggesting that they emerged and enriched in the biofilms.

Foremost, we observed two large genome deletions that occurred in all biofilms with a frequency approaching 100% (**Fig. 5a & 5b**). These deletions correspond to the excision of two mobile genetic elements, or prophages, *skin* and SP-β³³⁻³⁵. Excision of *skin* generates a functional protein: sigK - a sporulation-specific transcription factor essential for cell differentiation in *B. subtilis*³⁶. The excision of SP-β generates another functional protein dubbed SpsM - a protein involved in capsid polysaccharide biosynthesis mediating³⁷ and with relevance to pellicle development³⁸. Nearly all biofilm cells carried one

287 of these variations, and most cells carried both (**Fig. 5a & b; Table S6 & Data**
 288 **S3**). These prophage excisions therefore appear to be under stronger selection
 289 than the GudB mutations. The frequency of these structural variations gradually
 290 increases, from none in the initial mix to 100% in gradient biofilms (**Fig. 5c**), and
 291 so does the signature of GudB's drift (**Fig. 1**). However, at this stage, the
 292 observed link between the prophage excisions and GudB's drift is circumstantial
 293 and further experiments are needed to establish how are these two phenomena
 294 linked. The prophage excisions are also likely to occur in the soil, but the DNA
 295 recovered from these samples was insufficient to allow genome sequencing.

296 Exclusively in biofilms, we also detected 59 enriched SNPs in a conserved
 297 region of 16S rRNA (**Table S6 & Data S3**). However, *B. subtilis* has ten 16S rRNA
 298 gene copies. Since these are essentially identical, we could not determine which
 299 of these 10 paralogues carried mutations. However, per population, 98% of the
 300 16S rRNA mutations occurred in the same Illumina read suggesting that one
 301 paralogue was highly mutated while others remained intact (**Fig. S7**). Large
 302 differences in expression levels of 16S rRNA genes were identified in *P.*
 303 *aeruginosa* biofilms ³⁹, and ribosomal heterogeneity has been linked to biofilm
 304 development in *B. subtilis* ⁴⁰. Yet, to our knowledge, mutations in the 16S rRNA
 305 genes have not been reported in biofilms. At this stage, however, which 16S gene
 306 is inactivated, how do multiple proximal 16S mutations occur, and how
 307 inactivation affects biofilm development, remains unclear. Overall, the 16S rRNA
 308 SNPs, and the structural variations in particular, seem to have a key role in
 309 biofilm development in *B. subtilis*. Accordingly, most of these genetic variations
 310 were reproducible between replica experiments (**Table S6**) suggesting that they

arose during biofilm growth and then enriched by virtue of promoting biofilm formation¹².

Concluding remarks

That the fitness effect of a mutation may vary depending on the environment is generally assumed. However, the magnitude of environmental pleiotropy unraveled here is surprisingly high. Environmental changes, including minute ones like addition of arginine to a proline medium, can completely revert the effect of *GudB* mutations. Overall 84% of the tested *GudB* mutations showed sign reversions. Pleiotropy severely restricts protein sequence space. Extensive pleiotropy has an interesting implication. The so-called wild type sequence of a gene-protein is generally thought to represent just one sequence out of an entire cloud of related sequences that are similarly fit. However, our results indicate that wild-type *GudB*'s sequence is singular in being fit across multiple constrains and environments. Per individual tested conditions, most mutations are either neutral or beneficial (28 - 81%). However, if all tested conditions are considered, only 2% of the tested *GudB* mutations are neutral or beneficial in the 9 reproducible conditions when selection acts on *GudB*.

The pleiotropy of protein mutations across multiple growth and environmental conditions has been rarely measured¹⁷, and, to our knowledge, never for a protein with an intrinsic physiological role. Indeed, the extensive pleiotropy observed here might be the norm in cases of complex relationships between a protein's expression and activity levels ('protein fitness') and organismal fitness, as with *GudB*. The high degree of pleiotropy observed here may also relate to *GudB*'s role as an enzyme and regulator, and also to the

positions explored (oligomer interface). In any case, our results suggest that, as currently performed, laboratory mutational scans broadly underestimate the fraction of mutations that are deleterious in ‘real life’.

Together, pleiotropy and drift dictate **not only** the evolution of short-term polymorphism (micro-evolution) but also the evolution of protein sequences along long evolutionary times and across species (macro-evolution). **Indeed, the** correlation between the effects of mutations in laboratory mappings and their occurrence or absence in natural sequences is limited ¹³. **However, laboratory mappings represent a single condition, and merging of data from multiple conditions could in principle reveal higher correlation. Identifying trends in complex datasets requires an unbiased approach. However, even when several different machine learning approaches were applied, merging of conditions gave no further correlation (Fig. S8).** Thus, along short evolutionary periods, proteins experience variable and opposing selection pressures. Additionally, drift may lead to rapid fixation of alleles that are marginally fit or even deleterious. The effects of drift have been extensively studied initiated by Kimura’s neutral theory ¹¹. Our results quantify its effect in bacterial populations and the potential effect of drift in combination with selection across different environments. For example, nearly 80% of the tested mutations survived or **even enriched during** sporulation, and a single spore **could then initiate** a whole new population. However, once the environment changes, such alleles will be rapidly lost unless compensated **or a priori enabled** by other mutations. **Compensation, or enabling by other mutations, results in epistasis, i.e., in the effect of mutations being dependent on the sequence context in which they occur** ⁴¹. Accordingly, along

360 macro-evolutionary time scales, epistasis dominates gene and genome

361 sequences ⁴².

362

363 **References**

- 364 1. Baba, T. *et al.* Construction of Escherichia coli K-12 in-frame, single-gene
365 knockout mutants: The Keio collection. *Mol. Syst. Biol.* **2**, (2006).
- 366 2. Koo, B. M. *et al.* Construction and Analysis of Two Genome-Scale Deletion
367 Libraries for Bacillus subtilis. *Cell Syst.* **4**, 291–305.e7 (2017).
- 368 3. Pache, R. A., Madan, M. M. & Aloy, P. Exploiting gene deletion fitness effects
369 in yeast to understand the modular architecture of protein complexes
370 under different growth conditions. *BMC Syst. Biol.* **3**, (2009).
- 371 4. Winzeler, E. A. *et al.* Functional characterization of the S. cerevisiae
372 genome by gene deletion and parallel analysis. *Science (80-.)*. **285**, 901–
373 906 (1999).
- 374 5. Civelek, M. & Lusis, A. J. Systems genetics approaches to understand
375 complex traits. *Nature Reviews Genetics* **15**, 34–48 (2014).
- 376 6. Steinberg, B. & Ostermeier, M. Environmental changes bridge evolutionary
377 valleys. *Sci. Adv.* **2**, e1500921 (2016).
- 378 7. Wang, Z., Liao, B.-Y. & Zhang, J. Genomic patterns of pleiotropy and the
379 evolution of complexity. *Proc. Natl. Acad. Sci.* **107**, 18034–18039 (2010).
- 380 8. Fusco, D., Gralka, M., Kayser, J., Anderson, A. & Hallatschek, O. Excess of
381 mutational jackpot events in expanding populations revealed by spatial
382 Luria-Delbrück experiments. *Nat. Commun.* **7**, (2016).
- 383 9. Kimura, M. Genetic variability maintained in a finite population due to
384 mutational production of neutral and nearly neutral isoalleles. *Genet. Res.*
385 (*Camb.*) **89**, 341–363 (2008).
- 386 10. Smith, N. H., Gordon, S. V., de la Rua-Domenech, R., Clifton-Hadley, R. S. &
387 Hewinson, R. G. Bottlenecks and broomsticks: The molecular evolution of
388 Mycobacterium bovis. *Nature Reviews Microbiology* **4**, 670–681 (2006).
- 389 11. Nei, M. Selectionism and neutralism in molecular evolution. *Molecular*
390 *Biology and Evolution* **22**, 2318–2342 (2005).
- 391 12. Steenackers, H. P., Parijs, I., Foster, K. R. & Vanderleyden, J. Experimental
392 evolution in biofilm populations. *FEMS Microbiology Reviews* **40**, 373–397
393 (2016).
- 394 13. Boucher, J. I., Bolon, D. N. A. & Tawfik, D. S. Quantifying and understanding
395 the fitness effects of protein mutations: Laboratory versus nature. *Protein*
396 *Science* 1219–1226 (2016). doi:10.1002/pro.2928
- 397 14. De Visser, J. A. G. M. & Krug, J. Empirical fitness landscapes and the
398 predictability of evolution. *Nature Reviews Genetics* **15**, 480–490 (2014).
- 399 15. Fowler, D. M. & Fields, S. Deep mutational scanning: A new style of protein
400 science. *Nature Methods* **11**, 801–807 (2014).
- 401 16. Steinberg, B. & Ostermeier, M. Shifting Fitness and Epistatic Landscapes
402 Reflect Trade-offs along an Evolutionary Pathway. *J. Mol. Biol.* **428**, 2730–
403 2743 (2016).
- 404 17. Dandage, R. *et al.* Differential strengths of molecular determinants guide
405 environment specific mutational fates. *PLoS Genet.* **14**, (2018).
- 406 18. Earl, A. M., Losick, R. & Kolter, R. Ecology and genomics of Bacillus subtilis.
407 *Trends in Microbiology* **16**, 269–275 (2008).
- 408 19. Branda, S. S., González-Pastor, J. E., Ben-Yehuda, S., Losick, R. & Kolter, R.
409 Fruiting body formation by Bacillus subtilis. *Proc. Natl. Acad. Sci. U. S. A.*
410 **98**, 11621–6 (2001).

- 411 20. Belitsky, B. R. & Sonenshein, A. L. Role and regulation of *Bacillus subtilis*
412 glutamate dehydrogenase genes. *J. Bacteriol.* **180**, 6298–305 (1998).
- 413 21. Gunka, K. & Commichau, F. M. Control of glutamate homeostasis in *Bacillus*
414 *subtilis*: A complex interplay between ammonium assimilation, glutamate
415 biosynthesis and degradation. *Molecular Microbiology* **85**, 213–224
416 (2012).
- 417 22. Liu, J. *et al.* Metabolic co-dependence gives rise to collective oscillations
418 within biofilms. *Nature* **523**, 550–554 (2015).
- 419 23. Noda-Garcia, L., Romero Romero, M. L., Longo, L. M., Kolodkin-Gal, I. &
420 Tawfik, D. S. *Bacilli* glutamate dehydrogenases diverged via coevolution of
421 transcription and enzyme regulation. *EMBO Rep.* **18**, 1139–1149 (2017).
- 422 24. De Jong, L. *et al.* In-Culture Cross-Linking of Bacterial Cells Reveals Large-
423 Scale Dynamic Protein-Protein Interactions at the Peptide Level. *J.*
424 *Proteome Res.* **16**, 2457–2471 (2017).
- 425 25. Stannek, L. *et al.* Evidence for synergistic control of glutamate biosynthesis
426 by glutamate dehydrogenases and glutamate in *Bacillus subtilis*. *Environ.*
427 *Microbiol.* **17**, 3379–3390 (2015).
- 428 26. Mondal, S., Yakhnin, A. V., Sebastian, A., Albert, I. & Babitzke, P. NusA-
429 dependent transcription termination prevents misregulation of global
430 gene expression. *Nat. Microbiol.* **1**, (2016).
- 431 27. Besharova, O., Suchanek, V. M., Hartmann, R., Drescher, K. & Sourjik, V.
432 Diversification of gene expression during formation of static submerged
433 biofilms by *Escherichia coli*. *Front. Microbiol.* **7**, (2016).
- 434 28. Vlamakis, H., Chai, Y., Beauregard, P., Losick, R. & Kolter, R. Sticking
435 together: Building a biofilm the *Bacillus subtilis* way. *Nature Reviews*
436 *Microbiology* **11**, 157–168 (2013).
- 437 29. Mavor, D. *et al.* Determination of ubiquitin fitness landscapes under
438 different chemical stresses in a classroom setting. *Elife* **5**, (2016).
- 439 30. Nguyen, D. & Singh, P. K. Evolving stealth: genetic adaptation of
440 *Pseudomonas aeruginosa* during cystic fibrosis infections. *Proc. Natl. Acad.*
441 *Sci. U. S. A.* **103**, 8305–8306 (2006).
- 442 31. Hallatschek, O., Hersen, P., Ramanathan, S. & Nelson, D. R. Genetic drift at
443 expanding frontiers promotes gene segregation. *Proc. Natl. Acad. Sci.* **104**,
444 19926–19930 (2007).
- 445 32. Lemon, K. P., Earl, A. M., Vlamakis, H. C., Aguilar, C. & Kolter, R. Biofilm
446 development with an emphasis on *Bacillus subtilis*. *Current Topics in*
447 *Microbiology and Immunology* **322**, 1–16 (2008).
- 448 33. Kunkel, B., Losick, R. & Stragier, P. The *Bacillus subtilis* gene for the
449 developmental transcription factor $\sigma(K)$ is generated by excision of a
450 dispensable DNA element containing a sporulation recombinase gene.
451 *Genes Dev.* **4**, 525–535 (1990).
- 452 34. Nicolas, P. *et al.* Condition-dependent transcriptome reveals high-level
453 regulatory architecture in *Bacillus subtilis*. *Science (80-)*. **335**, 1103–1106
454 (2012).
- 455 35. Westers, H. *et al.* Genome Engineering Reveals Large Dispensable Regions
456 in *Bacillus subtilis*. *Mol. Biol. Evol.* **20**, 2076–2090 (2003).
- 457 36. Eichenberger, P. *et al.* The program of gene transcription for a single
458 differentiating cell type during sporulation in *Bacillus subtilis*. *PLoS Biol.* **2**,
459 (2004).

- 460 37. Abe, K. *et al.* Developmentally-Regulated Excision of the SP β Prophage
461 Reconstitutes a Gene Required for Spore Envelope Maturation in *Bacillus*
462 *subtilis*. *PLoS Genet.* **10**, (2014).
- 463 38. Martin, M. *et al.* De novo evolved interference competition promotes the
464 spread of biofilm defectors. *Nat. Commun.* **8**, (2017).
- 465 39. Pérez-Osorio, A. C., Williamson, K. S. & Franklin, M. J. Heterogeneous rpoS
466 and rhlR mRNA levels and 16S rRNA/rDNA (rRNA Gene) ratios within
467 *Pseudomonas aeruginosa* biofilms, sampled by laser capture
468 microdissection. *J. Bacteriol.* **192**, 2991–3000 (2010).
- 469 40. Vesper, O. *et al.* Selective translation of leaderless mRNAs by specialized
470 ribosomes generated by MazF in *Escherichia coli*. *Cell* **147**, 147–157
471 (2011).
- 472 41. Miton, C. M. & Tokuriki, N. How mutational epistasis impairs predictability
473 in protein evolution and design. *Protein Sci.* 1260–1272 (2016).
474 doi:10.1002/pro.2876
- 475 42. Breen, M. S., Kemena, C., Vlasov, P. K., Notredame, C. & Kondrashov, F. A.
476 Epistasis as the primary factor in molecular evolution. *Nature* **490**, 535–
477 538 (2012).
- 478 43. Konkol, M. A., Blair, K. M. & Kearns, D. B. Plasmid-encoded comI inhibits
479 competence in the ancestral 3610 strain of *Bacillus subtilis*. *J. Bacteriol.*
480 **195**, 4085–4093 (2013).
- 481 44. Yi, Y., de Jong, A., Frenzel, E. & Kuipers, O. P. Comparative transcriptomics
482 of *Bacillus mycoides* root exudates reveals different genetic adaptation of
483 endophytic and soil isolates. *Front. Microbiol.* **8**, (2017).
- 484 45. Deatherage, D. E. & Barrick, J. E. Identification of mutations in laboratory-
485 evolved microbes from next-generation sequencing data using breseq.
486 *Methods Mol. Biol.* **1151**, 165–188 (2014).

489 Acknowledgments

490 L.N.G. was supported by the CONACYT grant #203740 and the Martin Kushner
491 Fellowship at the Weizmann Institute of Science. D.S.T. is the Nella and Leon
492 Benozio Professor of Biochemistry. Financial support by the Kahn Center for
493 Systems Biology at the Weizmann Institute of Science is gratefully
494 acknowledged. We are highly grateful to Ron Milo, Sarel Fleishman, Zvi Livneh
495 and Fyodor Kondrashov for support and critical advice, to Einat Segev, Arjan de
496 Visser for critical and insightful comments to the manuscript. We highly
497 appreciate the help of Moshe Hershko in script development for data processing,
498 and of Yinon Bar-On, Shmuel Gleizer in the analysis of genomic sequences. We

499 are grateful to Ron Rotkopf from the Weizmann Life Sciences Core Facilities for
 500 his guidance on the statistical analysis. We are thankful for the services provided
 501 by the Crown Genomics institute of the Nancy and Stephen Grand Israel National
 502 Center for Personalized Medicine, Weizmann Institute of Science.

503

504 **Authors Contribution**

505 L.N.G. and D.S.T. designed experiments and wrote the manuscript. L.N.G., D.D.
 506 and D.S.T. analysed the data. L.N.G. performed all experiments, except selection
 507 in soil colonization that was performed in collaboration with E.K. and A.A. D.D.
 508 and A.E. wrote the scripts used for data analysis and visualization. E.P. applied
 509 machine learning classification.

510 **Fig. 1. Selection *versus* chance-dominated conditions. (A)** Dot-plot indicating
 511 reproducible measurements of frequency coefficients of individual mutations
 512 (FC values) in three parallel replica liquid cultures with proline as carbon-
 513 nitrogen source. $\overline{SD}$ is the average standard deviation between 3 biological
 514 replicas. The S.D. values were calculated per each amino acid allele based on
 515 logFC values and averaged for all alleles in a given condition. **(B)** The same
 516 analysis of three parallel biofilms with arginine as carbon-nitrogen source
 517 indicates low reproducibility. **(C)** The $\overline{SD}$ values categorized by the 7 general
 518 growth states tested here. Each point represents the $\overline{SD}$ value between 3 replicas
 519 of the same experiment (the distributions of SD values per each condition are
 520 shown in **Fig. S3a**). **(D)** $\overline{SD}_{syn}$ represents the standard deviation between the
 521 logFC values of synonymous codons. The standard deviations per allele were
 522 averaged for all synonymous alleles in the same replica experiment, and then
 523 averaged across the 3 replica experiments in a given condition (the distributions
 524 of SD_{syn} values per experiment are shown in **Fig. S3b**).
 525

Fig. 2. The pleiotropic effects of alleles across different conditions. (A) A dot-blot correlation of FC values of individual alleles in three different liquid carbon-nitrogen sources (average values per 3 replicas). The red squares encompass alleles that show sign pleiotropy – i.e., a change from beneficial to deleterious, or vice versa. (B) Pairwise correlation of the FC values in all 9 different reproducible conditions (3 replicas per condition; average FC values of all alleles and replicas). Colors indicate the Pearson correlation values (-1, negative correlation; 0, no correlation; 1, positive correlation). The strongest anti-correlation was found with arginine as carbon-nitrogen source (black square). (C) The distribution of alleles by their level of sign pleiotropy; from pale to dark purple: (i) Weak sign pleiotropy (changes between deleterious and beneficial); (ii) Mild sign pleiotropy (changes from highly deleterious to beneficial, or from highly beneficial to deleterious); and (iii) Strong sign pleiotropy (changes from highly deleterious to highly beneficial, or vice versa). The fraction of alleles showing mild or strong sign pleiotropy is shown above the bars. The control dataset comprises 4 completely independent growth experiments in liquid proline, each inoculated from a different initial mix and grown on separate occasions (Fig. S9). Nonetheless, none of the alleles in this control set exhibited strong pleiotropy. Accordingly, a Chi-squared analysis indicated that the variations between conditions are significantly higher than the variations in the control group (X^2 and p values are shown; degrees of freedom equal 3 in all cases).

549 **Fig. 3. Genetic sweeps in biofilms and soil and the dominance of the wild-**
550 **type allele.** Photographs of 5 days old biofilms: (A) normal biofilms; (B)
551 gradient biofilms; (C) a scheme of soil colonization (shown biofilms with proline
552 as carbon-nitrogen source). (D) The distribution of frequency of individual
553 alleles for different growth states with proline as carbon-nitrogen source. Bar
554 widths represent allele frequency from raw read counts (Rf values; **Data S1**).
555 Magenta corresponds to wild type GudB. (E) Alleles with $R_f \geq 1\%$ were identified
556 and their number and sum of frequencies are shown (averages and standard
557 deviations for all experiments in a given condition). Blue designates selection-
558 dominated conditions and red drift-dominated ones, as in **Fig. 1**. (F) An initial
559 mix of 4 alleles was created including the wild-type allele at a varying frequency
560 from 1 up to 25%. Following growth in normal and gradient biofilm with proline,
561 the frequency of wild-type reached an average of 18% when initiated at 1%, and
562 up to 100% when initiated at 5% (see **Fig. S10** for the entire dataset).
563

564 **Fig. 4. The combined action of selection and chance in biofilms (red) and**
565 **soil colonization (green). (A)** Alleles that enriched in the edge of the biofilms
566 are more likely to arise from alleles that were neutral or enriched in the center.
567 The distribution of categorized FC values from all biofilm centers (grey)
568 compared to the distribution of FC values of center alleles that were enriched in
569 the edge (red). **(B)** The distribution of categorized FC values of all alleles in all
570 liquid conditions (grey) compared to the distribution of FC values of liquid
571 alleles that were enriched in the edge of biofilms (red). **(C)** Alleles enriched in
572 the root are more likely to arise from alleles that were enriched in the bulk soil.
573 The distribution of categorized allele FC values in all soil samples (grey)
574 compared to the distribution of FC values of soil alleles that were enriched in the
575 root (green). **(D)** The distribution of SD values (variability between replica
576 experiments, as in **Fig. 1c**) of alleles enriched in one or more root populations
577 compared to alleles that were never enriched in the roots. **T-tests were**
578 **computed per each FC category. p-values indicating significance ($p < 0.05$) are**
579 **presented above the bars (details of all the T-tests are provided in Table S7).**
580

581 **Fig. 5. Programmed genomic excisions drive GudB's drift in biofilms.**

582 Schematic representation of *B. subtilis* genomic organization before and after the
 583 excision of the prophage mobile elements SP- β (**A**) and *skin* (**B**) and their
 584 position in the genome. (**C**) These excisions were absent in the initial mix yet
 585 dominated biofilms and went to near fixation in the edge of gradient biofilms (for
 586 frequencies in individual experiments see **Table S6**). Excision of the mobile
 587 elements occurred in two different genomic locations within the same
 588 experiment. The values were summed and averaged according to the general
 589 condition shown. The details of the excisions (location and frequency) per
 590 experiment are shown in **Data S3**.

591

592 **Materials and Methods**

593 **Strains**

594 *B. subtilis* NCIB 3610 DS7187 (**kindly** gifted by Dr. Daniel B. Kearns ⁴³) that lacks
595 the ComI peptide and has high competence capacity similar to domesticated *B.*
596 *subtilis* strains was recruited to this study. *Bacillus subtilis* NCIB 3610 *gudB::tet*
597 strain ²³ genomic DNA was transformed into *B. subtilis* NCIB 3610 DS7187. *B.*
598 *subtilis* NCIB 3610 Δ *comI gudB::tet* was thus isolated, and was phenotypically
599 and genetically tested.

600

601 **GudB allele library construction**

602 We performed site directed mutagenesis in 10 codons (amino acids: M46, L48,
603 K52, D58, D59, S61, K63, T66, Y68, S75) of the *gudB* gene cloned in the
604 **pDG_GudB plasmid, which was modified from the pDG1728 backbone vector** ²³.
605 The codons were mutated to NNS (N = all bases & S = C or G) whereby the 20
606 standard amino acids and 1 stop codon is encoded. The codon mutagenesis was
607 done in one step PCR protocol and independently for each position. Thus, we
608 created 10 libraries, each containing 20 different amino acid alleles (non-
609 synonymous, missense mutations), 1 stop-codon (nonsense), and 11
610 synonymous alleles (alternative codons encoding the same amino acid). All
611 mutagenic PCRs were performed with Kapa HiFi HotStart Ready Mix (Kapa
612 Biosystems) following manufacturers conditions (**Table S8** shows the sequence
613 of all primers). The 10 PCR products were purified and used to transform the *E.*
614 *coli* T10 strain (Thermo Fisher Scientific). Clones were pulled together after an
615 overnight growth on LB + Ampicillin (100 µg/ml) agar plates at 37°C. At this
616 stage, 4 to 6 clones per library were isolated and analyzed by sequencing. Total
617 plasmid DNA from these library transformations was extracted and also
618 analyzed by sequencing. Each of the 10 libraries contained, after transformation,
619 at least 10⁵ clones, corresponding to ≥ 1000-fold coverage per allele.
620 Approximately 10 µg of plasmid DNA, from each library, was linearized (XhoI,
621 New England Biolabs, following manufactures conditions), purified, and used to
622 transform the *B. subtilis* NCIB 3610 *gudB::tet* Δ *comI* strain. Transformations were
623 performed as described ²³. After transformation, overnight growth on in +
624 Spectinomycin (100µg/ml) + Glucose (0.5 mg/ml) agar plates was used as

selection. The resulting cells were pulled together and kept at -20°C in 50% glycerol. In total, 10 *B. subtilis* libraries were constructed in parallel and each contained, after transformation, at least 10⁴ clones (≥100-fold coverage per allele). Genomic DNA extraction of each library was performed (GenElute - Sigma). The integrity of the mutagenic process was verified by sanger sequencing the *amyE::gudB* locus indicating that mutations were observed only in the diversified codon.

Selection and growth conditions

10 ml of LB (1% tryptone, 5% yeast extract and 1% NaCl) with Glucose (0.5%), ammonium sulfate (0.5%) and spectinomycin (100 µg/ml) cultures were inoculated with 1 ml of each library stock. The cultures were grown overnight at 37°C with shaking. 500 µl of the overnight culture was used to inoculate 3 ml of LB plus glucose (0.5%) and ammonium sulfate (0.5%). The cultures were incubated at 37°C with shaking and once the O.D.₆₀₀ reached 0.8 they were mixed equally and used as the starting population (initial mix). A fraction of the cells at this stage were harvested by centrifugation and stored for genomic DNA purification. In total, three different initial mixes were used for the experiments described here. Initial Mix #1 was used to inoculate most liquid conditions (4 carbon-nitrogen sources), pellicles and gradient biofilms. Initial Mix #2 was used to inoculate 1 liquid condition, spores, germination and biofilms, and initial mix #3 was used to inoculate bulk soil (**Data S1**). Detailed selection conditions are listed below:

For selection under liquid serial passages 100 µl of the initial mix was used to inoculate 10 ml cultures of MS medium (5 mM potassium phosphate, 100 mM MOPS pH 7.1, 2 mM MgCl₂, 700 µM CaCl₂, 50 µM MnCl₂, 50 µM FeCl₃, 1 µM ZnCl₂, 2 µM thiamine, 50 µg/ml tryptophan, 50 µg/ml phenylalanine and 50 µg/ml threonine)²³ with glucose (0.5%) plus ammonium sulfate (0.5%), glutamate (0.5%) plus glycerol (0.5%), proline (0.5%), arginine (0.5%) or proline (0.25%) plus arginine (0.25%). The cultures were incubated at 30°C with shaking until O.D.₆₀₀ reached 1 – 1.5, after which 100 µl was used to inoculate 10 mL of fresh medium. The serial passages were done every 24 hours when proline (0.5%), arginine (0.5%) or proline (0.25%) plus arginine (0.25%) were used as carbon-

nitrogen sources, and every 12 hours when glucose (0.5%) plus ammonium sulfate (0.5%), or Glutamate (0.5%) plus glycerol (0.5%), were applied. In total, all liquid passages were maintained for approximately 50 generations.

For selection in pellicles, 100 ml of media (b), (c) and (d) were inoculated with 100 µl of the initial mix cells. The culture was incubated at 30°C without shaking, for 5 days.

For selection of spores and germinated spores, three ml of the initial mix was used to inoculate 25 ml of Difco Sporulation Medium (DSM) in 250 ml flasks and incubated at 37°C with 150 rpm shaking until O.D.600 reached 0.4. This culture was used to inoculate 250 ml of fresh DSM in 1L flasks. The cultures were incubated 48h at 37°C with 150 rpm shaking. Cells were subsequently harvested by centrifugation and stored at 4°C over night. After, cells were re-suspended with 200 ml of cold deionized sterile water (dW) and incubated for 30 min at 4°C. Cells were harvested and re-suspended with 200 ml of cold distilled water (dW) and incubated overnight at 4°C with slow orbital agitation, to kill all planktonic of vegetative cells. The culture was harvested, re-suspended in 30 ml of dW and heated to 80°C for 20 min. Finally, spores were harvested, re-suspended in 10 ml of dW, and stored at -20°C. To germinate these spores, they were diluted 1000 times in phosphate-buffered saline solution and 100 µl of this suspension was used to inoculate LB plus glucose (0.5%) agar plates (10 plates). Approximately 10,000 colonies were obtained and pulled together.

For selection in biofilms, MS agar (1.5%) plates supplemented with different carbon-nitrogen sources were prepared. For gradient biofilms, gradient agar plates were prepared. First, square plates (12x12 cm) with MS agar (1.5%) medium were poured. After the agar solidified, an area of 2x14 cm was removed from the top of the plate. In this area, a solution of either Proline 5%, Arginine 5%, monosodium glutamate 5% or Glycerol 5% in 1.5% agar was poured into the removed section. For the glutamate plus glycerol gradient biofilm, two opposite areas of the agar plate were removed. Into one, a solution of monosodium glutamate (5%) in 1.5% agar was poured, and into the other, glycerol (5%) plus 1.5% agar solution (see **Fig. S11a** for a graphic representation of the agar plates preparation). All gradient agar plates were incubated for 24 h at room temperature before use. We also calibrated the place

in the gradient plate where we inoculated the cells such that we observed growth after 1 night incubation at 30°C (**Fig. S11b**). For growth in biofilms and gradient biofilms, 5 µl of the initial mix were used as inoculum. Plates were incubated for 4 days at 30°C and 2 more days at room temperature. The colony was then dissected in 3 areas (center, wrinkle and edge) for normal biofilms, and in 2 areas (center and upper) for gradient biofilms (illustrated in **Fig. S11c-g**). After selection in all the above-mentioned conditions the biomass was harvested and storage at -20°C. All growth experiments were performed in triplicate by inoculating with the same initial mix.

For selection in soil and plant roots, the initial mix was generated as above-mentioned except that the process was scaled up (instead of 3 ml, 10 ml of culture was prepared per library). In total, 200 ml of the initial mix (O.D.₆₀₀ = 0.8) was applied. This LB culture was washed three times (by means of centrifugation and re-suspension) with 100 ml half strength Hoagland solution ⁴⁴. After the final wash, the cells were re-suspended in half strength Hoagland solution to a final O.D.₆₀₀ of 0.1. Since Hoagland's solution is not isotonic, the washes resulted in death of about a third of the *B. subtilis* cells. Thus, handling the samples at this stage was performed as fast as possible. **The Hoagland solution imposes some selection pressure on the initial mix population although the loss in population size is relatively small (≤30%). The soil colonization FC values therefore result from the entire process that begins the rinses with the Hoagland solution and ends with the colonization of the roots.** Natural soil was collected at the Ha-Masrek Reserve, Israel (31.793 N, 35.042 E), sifted through 2 mm sieve and autoclaved three times for 30 min at 121°C. A total of five pots (size 10 x 8 x 5 cm) with autoclaved natural soil were drenched with the initial mix suspended in half strength Hoagland Solution ⁴⁴. These potted soils drenched with bacterial suspensions were used to plant tomato seedlings grown first in sterile conditions. Seeds of tomato (*Solanum lycopersicum* L.; cv. Micro-Tom) were surface-sterilized with 70% ethanol for 5 minutes and, 10 minutes with 3% bleach with 0.01% Tween 20. Surface-sterile seeds were germinated on sterile filter paper (Whatman, catalog # 1001-085) saturated with half strength Hoagland Solution for 7 days (23°C and 16 hours photoperiod). Six tomato seedlings were transferred to each pot and grown for one month (21°C, 16h

light, 8h dark) with drenching with half strength Hoagland twice a week. Plants were subsequently harvested from the five pots. Roots and rhizosphere samples were collected for each replica experiment consisting a pool of six roots. First, the plants were carefully removed from the soil. Roots were then cut out from the plants and vortexed in 20 ml of washing solution (0.85% NaCl) for 30 s. This step was repeated one more time with a fresh washing solution. The combined root washing solutions (40 ml) was centrifuged for 30 min at 3000 rpm and the resulted pelleted samples corresponding to the rhizosphere were frozen in liquid nitrogen and stored at -80°C. The washed roots were blotted in filter paper and stored at -80°C until further use. Finally, bulk soil without roots was also stored at -80°C.

Genomic DNA extraction

All samples, including pellicle, spores, biofilm and gradient biofilm samples, were defrosted and re-suspended in 10 ml of dW. The samples were sonicated at 40% power, VibraCell, Sonics, for 10 min at 60 s intervals. Cells debris was harvested by centrifugation (13,000 g for 20 min). Genomic DNA from all samples was extracted using the GenElute Bacterial Genomic DNA Kit (Sigma-Aldrich) generally following the manufacturer's instructions, with the exception of the soil, rhizosphere soil and plant roots samples. For these samples, the PowerSoil DNA Isolation kit of Mo Bio was used, following its manufacturer's instructions.

Illumina sample preparations

The mutagenized *gudB* fragment (from amino acids 45 to 81) was amplified using the primers GudB_In_For (5'-CTCTTTCCCTACACGACGCTCTTCCGATCTnnnnnnCCCGAAGAGGTATACGAATTGT TAAAAGAG), and GudB_In_Rev (5'-CTGGAGTTCAGACGTGTGCTCTTCCGATCTCGCCTTTCGTTGGACCGAC). To the GudB_In_For primer, 6 N's were added to increase the sequence variability between amplicons. PCRs were performed with the KapaHiFi HotStart Ready Mix (Kapa Biosystems) using approximately 100 ng of genomic DNA as template and following manufacturer's instructions. Using 10 µl of the PCR as template, a second PCR was performed to add the Illumina adaptor sequence, using primers

757 GudB_Out_For (5'-
758 AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGC) and
759 GudB_Out_Rev (5'-
760 CAAGCAGAAGACGGCATACGAGATTGTTATACGTGACTGGAGTTCAGACGTGTGCTCTTCTT).

761 The Illumina index (underlined) was changed in the GudB_Out_Rev primer to
762 different Illumina indexes. Each condition was differently barcoded. All PCRs
763 were purified using the Agencourt AMPure XP (Beckman Coulter). The
764 concentration of PCR products was verified using Qu-bit assay (Life
765 Technologies).

766

767 **Analysis of the Illumina reads**

768 DNA samples were run using the Illumina NextSeq 150-bp paired-end kit. The
769 FASTQ sequence files were obtained for each run and customized using MatLab
770 8.0 and Python 3.6 scripts designed to count the number of each individual allele
771 in each sequenced sample. We filtered the reads to exclude any reads that have
772 mutations outside the mutagenized codons. All codons encoding for the wild-
773 type amino acid were summed in one and assigned as WT. All other codons were
774 counted independently. The unprocessed read counts are shown in **Data S1**.
775 Further filtering excluded alleles with < 100 counts in the initial mix to avoid
776 statistical uncertainty with respect to FC values. In total, we obtained data for up
777 to 269 individual alleles per condition out of the originally introduced 320
778 alleles. Per condition, a minimum of 380,000 reads was obtained. Thus, in
779 average, we obtained 1500 reads per allele.

780

781 **Data Analysis**

782 The frequency of each allele (f_i) was calculated as the ratio between the number
783 of reads for allele i divided by the total number of reads. The allele frequency
784 coefficient (FC_i) was subsequently calculated as the ratio of after selection (f_i)
785 divided by the frequency of the same allele in the initial mix (**Fig. S2 & Data S2**).
786 Normalization by the number of wild-type reads rather than by the total number
787 of reads gave essentially identical FC values for the majority of samples.
788 However, in the few samples where wild-type frequency was significantly
789 reduced after selection, normalization resulted in high noise and large biases

including large changes in sign (higher sign pleiotropy). FC values were therefore derived from the unnormalized frequency (fraction of reads for a given allele out of the total number of reads). FC values relate to fitness logarithmically, and thus logFC values were compared. To this end, all FC's equal to zero had to be changed, and we opted for a tenth of the minimum FC value found amongst all experiments. For the liquid, pellicles, biofilms, spores and germinated spores experiments (**Data S2, sheet 1**) the zeros were changed to 4.2×10^{-6} . For the bulk soil experiments (**Data S2, sheet 2**) zeros were changed to 1.14×10^{-5} . The logarithm of all FC values was calculated and was also used to derive mean FC values. The logFC values were then used to calculate: (i) the standard deviation for all alleles across conditions ($\overline{SD}$; the standard deviation between logFC values observed per each allele in replica experiments were averaged for all alleles measured in a given condition); (ii) the standard deviation between synonymous codons within the same replica experiment (deviations between logFC values of synonymous codons of the same amino acid allele were calculated, averaged for all alleles in the same experiment, and then for all replica experiments per condition). The T, F, X^2 tests and Pearson correlation values were obtained using the PRISM software. The Levene's test was performed using R. In addition, **Table S9** from supplementary material shows the sample size for every test performed in this study.

Defining the limits of neutrality

From all conditions tested here, only in glucose plus ammonia the GudB knockout had no growth effect (**Fig. S1**). Hence, this condition is largely neutral, and the variation observed in FC values would primarily be the outcome of noise. The standard deviation between 3 biological replicas was calculated per allele, and these values spanned over the range of 0.002 to 0.199. We rounded this number to 0.2. Thus, by the strictest measure, FC values between 0.8 and 1.2 were classified as 'neutral'. Accordingly, $FC \leq 0.8$ unambiguously assigned a mutation as 'deleterious', and $FC > 1.2$ as 'beneficial'.

Genome sequencing

We sequenced the genomic DNA of all biofilm populations for which we had ≥ 1 μg of DNA after extraction (6 normal and 12 gradient biofilm). For comparison, we also sequenced initial mix populations 1 and 2, 6 Liquid and 4 pellicle populations. The Illumina HiSeq2500 platform was used, with 2x125 base pairs read length. We obtained a total of 300 million reads. The reads were assembled using as reference the *B. subtilis* NCIB 3610 genome (NCBI Accession number: CP020102). Overall, 95% of all reads were successfully mapped to the reference genome with minimal coverage of x300 for all samples analyzed. The Breseq program was used to identify genomic variants, including single nucleotide polymorphisms (SNPs) and insertion-deletion polymorphisms (INDELs) ⁴⁵ (**Data S3 & Table S6**).

Comparison of FC values and to GudB's natural sequence variability

We examined whether the FC values for individual mutations, in individual conditions, **and in combinations thereof**, might predict whether or not a certain sequence exchange is observed, or not, amongst the sequences of **naturally occurring** GDHs. To this end, we constructed a number of different support vector machines (SVM) classification models with a variety of kernels (such as linear, Gaussian, polynomial etc.). The feature vector of each GudB allele was composed from the normalized FC values from specific condition. The values from replica experiments of the highly reproducible liquid conditions were averaged prior to training. Based on the multiple sequence alignment containing 1013 GDH sequences, we divided the GudB mutations in our dataset into 3 categories, which were then utilized as the prediction labels: (1) mutations seen in less than 5 natural GDH sequences (classified as 'not present', 66% of mutations), (2) mutations observed in 5 - 49 sequences ('rare', 19%) and (3) mutations present in ≥ 50 sequences ('frequent', 15%). Introducing class weights into the loss function compensated the unbalanced nature of the dataset. For each feature combination of a varied length, we built an SVM classification model and assessed its accuracy using 3-fold cross validation. Additionally, in order to reduce noise, assuming that our data belong to linear space, we extracted the first ten principal components of the feature matrix and used them as the new feature vectors for a model construction. To examine if our relatively high (>0.6)

855 model accuracy was distributed uniformly across different classes, for each
856 model and genotype, we recorded the predicted values during 3-fold cross-
857 validation. Moreover, for each condition combination, and for each kernel, we
858 built 100 different models and recorded the number of times each of the
859 genotypes was predicted correctly.
860
861

A

B

C

A **B** **C** **D** **E** **F**

A **B** **C** **D**

A **B** **C**