

Chance and pleiotropy dominate genetic diversity in complex bacterial environments

Lianet Noda-García¹, Dan Davidi², Elisa Korenblum², Assaf Elazar¹, Ekaterina Putintseva³,
Asaph Aharoni² and Dan S. Tawfik^{1*}

How does environmental complexity affect the evolution of single genes? Here, we measured the effects of a set of *Bacillus subtilis* glutamate dehydrogenase mutants across 19 different environments—from phenotypically homogeneous single-cell populations in liquid media to heterogeneous biofilms, plant roots and soil populations. The effects of individual gene mutations on organismal fitness were highly reproducible in liquid cultures. However, 84% of the tested alleles showed opposing fitness effects under different growth conditions (sign environmental pleiotropy). In colony biofilms and soil samples, different alleles dominated in parallel replica experiments. Accordingly, we found that in these heterogeneous cell populations the fate of mutations was dictated by a combination of selection and drift. The latter relates to programmed prophage excisions that occurred during biofilm development. Overall, for each condition, a wide range of glutamate dehydrogenase mutations persisted and sometimes fixated as a result of the combined action of selection, pleiotropy and chance. However, over longer periods and in multiple environments, nearly all of this diversity would be lost—across all the environments and conditions that we tested, the wild type was the fittest allele.

The function of most genes may be essential in some conditions, but make only a marginal contribution or even be deleterious in others^{1–4}. The effects of mutations on organismal fitness are therefore environment dependent, giving rise to complex, pleiotropic genotype-by-environment interactions^{5,6}. (Here, we refer to different effects of the same mutation as environmental pleiotropy⁷, or pleiotropy for brevity, and opposing effects in different environments as sign environmental pleiotropy). Moreover, bacterial populations often do not comprise single cells but instead have a structure, as in biofilms. In the presence of this complexity (changing environments and heterogeneous communities of cells and/or species), the fate of mutations could also be dictated by population bottlenecks (drift) or the rapid takeover of beneficial mutations in other genes (selective sweeps)^{8–10}. Consequently, the frequency of a given gene allele may change dramatically (from perishing to fixation) with no relation to its protein function^{11,12}.

We aimed to develop an experimental set-up that would examine how complex bacterial growth states and environments might shape protein evolution. Previous systematic mappings were based on a direct link between protein stability and function and organismal survival, thus enabling the measurement of the effects of mutations at the protein level^{5,13–16}. However, how mutations in a single gene affect organismal fitness under varying environments and conditions is largely unexplored¹⁷. We thus chose *Bacillus subtilis* NCIB 3610 as our model, a non-domesticated strain capable of growing in diverse aquatic and terrestrial environments¹⁸. We explored the effects of mutations in different conditions: in dispersed cells in liquid cultures, but also in biofilms where phenotypic and genetic variability prevails¹². We also mapped the effects of mutations during spore formation and germination¹⁹ and in more complex and close-to-natural environments including soil, the rhizosphere and plant roots.

A catabolic glutamate dehydrogenase (GDH) was our model protein. This enzyme is essential when amino acids such as proline

serve as the sole carbon/nitrogen source²⁰. However, in the presence of ammonia and glycolytic sugars, GDH activity is redundant as glutamate must be synthesized rather than catabolized. GDHs therefore respond to changes in carbon and nitrogen sources and, as regulators of glutamate homeostasis, are also associated with biofilm development^{21,22}. *B. subtilis* has two catabolic GDHs, RocG and GudB. The latter is constitutively expressed and is regulated via the association of its hexameric form²³. GudB also has regulatory roles^{24,25} via its interactions with the transcriptional activator of glutamate synthase²⁵ and with a transcription termination factor that modulates the stringent response²⁶. We explored mutations in the oligomeric interface of GudB, with a focus on multilateral effects on GudB's enzymatic and regulatory functions.

Together, these choices of organism and enzyme allowed us to readily examine and quantify the fate of *gudB* alleles in a range of different growth conditions and environments, including some that mimic natural habitats where strong evolutionary forces act¹².

Experimental set-up and data processing

We anticipated that the effects of the explored mutations would be complex and condition dependent. We thus opted for deep rather than broad coverage and mapped 10 positions within a single ~150-base-pair segment that resides at the oligomeric interface of GudB. The mutagenized positions were arbitrarily chosen, with the aim of mapping highly conserved positions (for example 58, which is D in all GudB orthologues) as well as divergent ones (for example 48 or 61, which are variable amongst GudB orthologues; Supplementary Table 1). The mutagenized codons were diversified to NNS, where N represents any of the four bases and S represents G or C. We thus created 10 libraries, each diversifying a single position into 20 different amino acids plus a stop codon. The libraries were incorporated into the chromosome of *B. subtilis* NCIB 3610 under the original *gudB* promoter and terminator. The combined library contained 320 single-mutant alleles whose genomes differ, in

¹Department of Biomolecular Sciences, Weizmann Institute of Science, Rehovot, Israel. ²Department of Plant and Environmental Sciences, Weizmann Institute of Science, Rehovot, Israel. ³Institutes of Science and Technology, Klosterneuburg, Austria. *e-mail: dan.tawfik@weizmann.ac.il

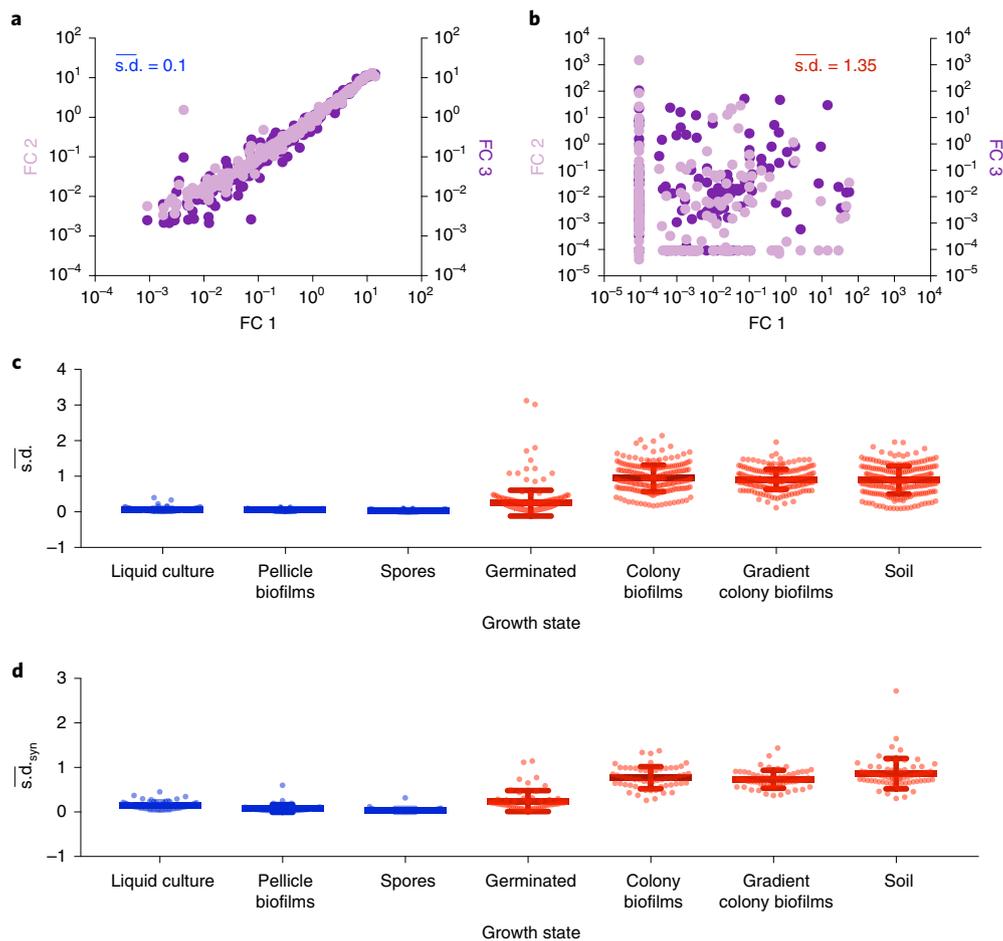


Fig. 1 | Selection versus chance-dominated conditions. **a**, A dot plot showing reproducible measurements of FC values for individual mutations in three parallel replica liquid cultures with proline as the carbon/nitrogen source. $\overline{s.d.}$ is the average $s.d.$ between three biological replicas. The $s.d.$ values were calculated for each amino acid allele from $\log[FC]$ values and averaged for all alleles in a given condition. **b**, A similar analysis of three parallel colony biofilms with arginine as the carbon/nitrogen source indicates low reproducibility. **c**, The $\overline{s.d.}$ values categorized by the seven general growth states tested here are shown. Each point represents the $s.d.$ value between three replicas of the same experiment (the distributions of $s.d.$ values for each condition are shown in Supplementary Fig. 3a). **d**, $\overline{s.d.}_{syn}$ represents the $s.d.$ between the $\log[FC]$ values of synonymous codons. The $s.d.$ values for each allele were averaged for all synonymous alleles in the same replica experiment and then averaged across the three replica experiments in a given condition (the distributions of $s.d._{syn}$ values for each experiment are shown in Supplementary Fig. 3b). In **c** and **d**, error bars indicate $s.d.$ Selection-dominated conditions are shown in blue and drift-dominated conditions in red.

principle, by a single *GudB* mutation, including 200 different amino acid alleles (including wild type), 10 stop codons and synonymous alleles in which the same amino acid could be encoded by 2 or 3 different codons.

This starting library (the initial mix, hereafter) was used to inoculate cultures grown in an array of different conditions. We tested seven different growth states with diverse complexity, from single cells to communities: liquid cultures, biofilms grown at the liquid–air interface (pellicle biofilms), spores, germinated spores, biofilms grown on agar (colony biofilms), including on carbon/nitrogen gradients, and colonized soil. Up to five different carbon/nitrogen sources were used that, at least as far as the phenotypes of the *GudB* knockout indicate, inflict different levels of selection on *GudB*: glutamate plus ammonia (GA), where $\Delta GudB$ has no growth effect; glutamate plus glycerol (GG), arginine (A) and arginine plus proline (PA), where $\Delta GudB$ exhibits a slight growth defect; and proline (P), where $\Delta GudB$ exhibits the strongest growth defect (Supplementary Fig. 1). In total, we tested 19 conditions. For each condition, three to five biological replicas were performed by inoculating from the same initial mix. The replicas were grown in parallel, and individually

analysed. Illumina sequencing was used to determine the frequency of each of the *gudB* alleles in the initial mix and after growth. After filtering (see Methods), we obtained data for 244–269 individual alleles for each experiment (Supplementary Data 1 and Supplementary Fig. 2).

The ratio between an allele's frequency at the end of growth and in the initial mix was derived, and this ratio is referred to as the frequency coefficient (FC; Supplementary Data 2). Basically, $FC > 1$ means an enriched, beneficial mutation and $FC < 1$ a purged, deleterious one. However, given the experimental error in determining FC values, values between 0.8 and 1.2 were classified as neutral, values ≤ 0.8 designated a mutation as deleterious and values > 1.2 as beneficial (see Methods). Mutations with $FC \leq 0.1$ were classified as highly deleterious and, similarly, mutations with $FC \geq 10$ as highly beneficial. FC values reflect the relative frequency of alleles and therefore relate logarithmically to their relative fitness effects (or selection coefficient, s). Hence, $\log[FC]$ values were compared throughout. However, the number of generations differs fundamentally between conditions—for example, 35 generations in liquid cultures (following 5 serial transfers into a fresh culture) versus effectively no

replication in spores (a dormant non-replicative form of *B. subtilis*). Moreover, in pellicle biofilms and colony biofilms, the number of generations cannot be determined easily: biofilms have different cell types (matrix producers, dormant cells, spores and so on). These different cell types also have different growth rates^{27,28}. So, while we could not calculate selection coefficients, it should be noted that an FC value of 0.8 in spores would effectively mean extinction across multiple generations in a liquid culture ($0.8^{50} = 10^{-5}$).

Irreproducibility: selection versus drift

Our first observations indicated two contrasting scenarios. In liquid cultures, for example, we observed highly reproducible FC values between biological replicas (Fig. 1a). Given the small sample numbers (three replicas as standard, five in a few cases), the observed variance may underestimate the actual variance. However, the consistently low variance in a range of different liquid conditions, and in other replica measurements in liquid cultures²⁹, supports high reproducibility. In colony biofilms, on the other hand, despite the fact that we did not bottleneck any of the populations on inoculum, the poor correlation between replicas was evident (Fig. 1b). The reproducibility between biological replicas indicates selection, suggesting that in reproducible conditions, the fitnesses of *GudB* and *B. subtilis* are tightly coupled. In colony biofilms, however, the lack of reproducibility suggested the dominance of drift—that is, random sampling of *gudB* alleles.

To quantify the contribution of selection versus drift in different conditions, we used two criteria. First, we compared the variability in FC values between replicas by calculating the s.d. for each allele (using, by default, the logarithm of the FC values; see Methods). The average s.d. value for all alleles in each experiment (s.d.) is given for the different growth states: a growth state (a liquid culture, for example) may include several conditions, such as different carbon/nitrogen sources (Fig. 1c; Supplementary Fig. 3a and Supplementary Table 2). As can be seen, in liquid cultures, pellicle biofilms and spores, the s.d. values between biological replicas were low (<0.06). In colony biofilms and bulk soil, however, the s.d. values were >0.25 , indicating low reproducibility. Indeed, both the *F*-test (variance ratio) and Levene's test indicated that for all tested alleles, the variance between FC values significantly changed between liquid cultures, pellicle biofilms and spores when compared to germinated spores, colony biofilms and bulk soil (the *P* values of the *F*-test were in the range $0.048-1.18 \times 10^{-27}$; the *P* values of Levene's test were in the range $0.0258-2.58 \times 10^{-16}$; Supplementary Data 2).

Second, if drift dictates the fate of *gudB* alleles, codons of the same amino acid should exhibit very different FC values. Indeed, both the *F*-test and Levene's test showed that, for all tested alleles, the variance between the synonymous FC values significantly changed between reproducible conditions (liquid culture, pellicle biofilms and spores) and non-reproducible ones (germinated spores, colony biofilms and bulk soil; the *P* values of the *F*-test were in the range $2.87 \times 10^{-8}-1.75 \times 10^{-69}$ and the *P* values of Levene's test were in the range $0.01-1 \times 10^{-24}$; Supplementary Table 4). The deviations between synonymous codon alleles of the same amino acid were also calculated, averaged for all alleles in the same experiment and then for all replicas of the same experiment ($\overline{s.d.}_{syn}$, in log values; Fig. 1d; Supplementary Fig. 3b and Supplementary Table 3). Note that the $\overline{s.d.}_{syn}$ criterion holds within individual replica experiments and is thus independent of the comparison of s.d. between biological replicas. Nonetheless, these criteria are clearly correlated (Fig. 1c,d). Overall, it appears that in liquid cultures, pellicle biofilms and spores, the FC values report the outcome of selection acting on *gudB* alleles at the amino acid level as expected (in a few alleles, selection also acted reproducibly at the codon level; Supplementary Fig. 4). In contrast, in colony biofilms and bulk soil we consistently observed higher s.d. as well as higher $\overline{s.d.}_{syn}$ values. In some colony biofilm experiments, a single codon had effectively

taken over, resulting in $\overline{s.d.}_{syn}$ values ≥ 3 (note that $\log[FC]$ values were compared throughout and the s.d. for FC values is therefore $\geq 10^3$). However, the reproducible and non-reproducible conditions differ not only in growth conditions but also in the number of generations. Liquid culture and colony biofilms are two extremes in this sense. Thus, we further measured whether the reproducibility of FC values might be low in the earlier liquid passages (where the number of generations is much lower) compared to the later ones. However, as expected, the s.d. values of the FCs in the first passages were consistently lower than the later ones (Supplementary Fig. 5).

Given that some conditions were selection dominated and others were subject to chance, we divided our analysis in two. First, we analysed selection-dominated conditions (liquid culture, pellicle biofilms and spores) to examine whether and how *GudB* mutations exert different fitness effects in different environments. Second, conditions where drift prevailed (germination, colony biofilms and soil colonization) were analysed to reveal the relative contributions of selection versus chance.

Pleiotropy: condition-dependent fitness effects

While the FC values, and hence the fitness effects of mutations, were reproducible under many conditions, their distribution varied widely between conditions, including between carbon/nitrogen sources (Supplementary Fig. 6). This indicates pleiotropy—individual *gudB* alleles have different fitness effects in different environments. To quantify the level of pleiotropy, we compared the FC values of the same *GudB* mutation across the nine individual selection-dominated conditions. Because the number of generations differs from one condition to another, we focused on the shift from beneficial to deleterious and vice versa (sign environmental pleiotropy) because the sign indicates the overall trend irrespective of generation numbers. Representative dot plots comparing the FC values across three different liquid conditions are shown in Fig. 2a. These indicate that pleiotropy is common, even when comparing liquid cultures with overlapping carbon/nitrogen sources. In particular, a significant number of *GudB* mutations show sign environmental pleiotropy (dashed red squares, Fig. 2a). Indeed, the Pearson correlation values for the 36 possible pairwise comparisons of the 9 reproducible conditions were all below 0.7, and many accommodated a negative value indicating an overall anti-correlation (that is, the dominance of sign environmental pleiotropy; Fig. 2b). Across all selection-dominated conditions, 84% of alleles showed sign environmental pleiotropy in at least one of the 36 pairwise comparisons and 70% of alleles showed mild or strong sign environmental pleiotropy. These pleiotropic effects greatly exceed the experimental noise, as indicated by comparison with a control sample (Fig. 2c).

Overall, the dominance of pleiotropy meant that across all the conditions in which selection acts, 86% of the alleles were beneficial in at least one condition. However, no single mutation was beneficial across all conditions. Furthermore, if a mutation were to be considered deleterious if purged under at least one condition, then 98% of the tested *GudB* mutations would be deleterious.

Combined selection and drift in heterogeneous environments

In colony biofilms (and also in germination and soil colonization, to a lesser degree), irreproducibility between replicas, variability between codons (Fig. 1c,d) and the near fixation of relatively few alleles (Fig. 3) all suggested fixation by chance. What is the nature of these few *GudB* 'winners' in the colony biofilms: are they merely lucky?

Although drift dominated in colony biofilms and soil colonization, wild-type *GudB* was enriched in up to 85% of these experiments, suggesting that selection may also play a role (Fig. 3). To assess the action of selection, we compared the three colony biofilm

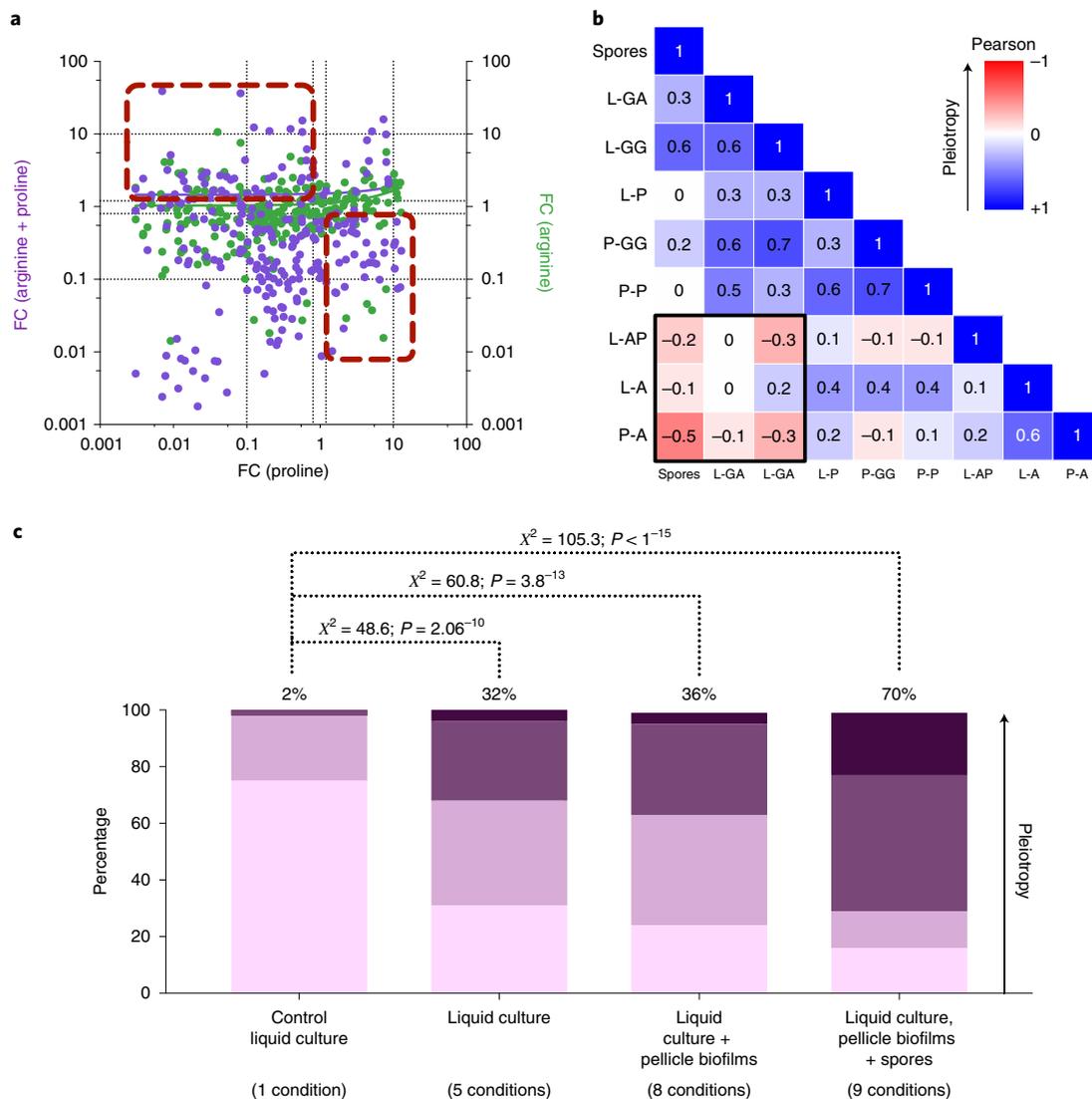


Fig. 2 | The pleiotropic effects of alleles across different conditions. **a**, A dot plot correlation of FC values of individual alleles in three different liquid carbon/nitrogen sources (average values per three replicas). The dashed red squares encompass alleles that show sign environmental pleiotropy. **b**, Pairwise correlation of the FC values in all nine different reproducible conditions (three replicas per condition; average FC values of all alleles and replicas). The colours show the Pearson correlation values (−1, negative correlation; 0, no correlation; +1, positive correlation). The strongest anti-correlation was found with arginine as the carbon/nitrogen source (black square). In the abbreviations, the first part indicates the growth state (L, liquid culture; P, pellicle biofilms) and the second part, the carbon source. **c**, The distribution of alleles by their level of sign environmental pleiotropy, from pale to dark purple: weak sign environmental pleiotropy (changes between deleterious and beneficial); mild sign environmental pleiotropy (changes from highly deleterious to beneficial or from highly beneficial to deleterious); and strong sign environmental pleiotropy (changes from highly deleterious to highly beneficial, or vice versa). The percentage of alleles showing mild or strong sign environmental pleiotropy is shown above the bars. The control dataset comprises four completely independent growth experiments in liquid proline, each inoculated from a different initial mix and grown on separate occasions (Supplementary Fig. 10). Nonetheless, none of the alleles in this control set exhibited strong pleiotropy. Accordingly, a χ^2 analysis performed on the binned FC values indicated that the variations between conditions are significantly higher than the variations in the control group (χ^2 and P values are shown; d.f. = 3 in all cases).

areas. There seems to be a systematic trend, whereby alleles that are enriched in the edges are more likely to arise from alleles that persisted or were even enriched in the centre (Fig. 4a). Similarly, 75% of the enriched edge alleles were neutral or beneficial under liquid growth with proline, a condition under which GudB experiences the strongest selection (Fig. 4b). This suggested that although drift dominated the fate of GudB in colony biofilms, GudB was under selection at some stage of the colony biofilm development. Accordingly, we found that the FC values are less skewed and more reproducible in the colony biofilm centres than in the edges

or wrinkles (Supplementary Fig. 6b) and the centre $\overline{s.d.}_{syn}$ values are half those in the edges or wrinkles (Supplementary Table 3 and Supplementary Fig. 3b). The $\overline{s.d.}_{syn}$ values are obviously much higher in the centre of colony biofilms compared to liquid cultures, but the trend suggests that at the onset of the colony biofilm's development, selection acts on GudB (Supplementary Table 3 and Supplementary Fig. 3b). Foremost, the presence of wild-type GudB at a high frequency in the vast majority of the colony biofilm experiments is not a coincidence that relates only to its high frequency in the initial mix. Indeed, a spiking experiment indicated that the wild

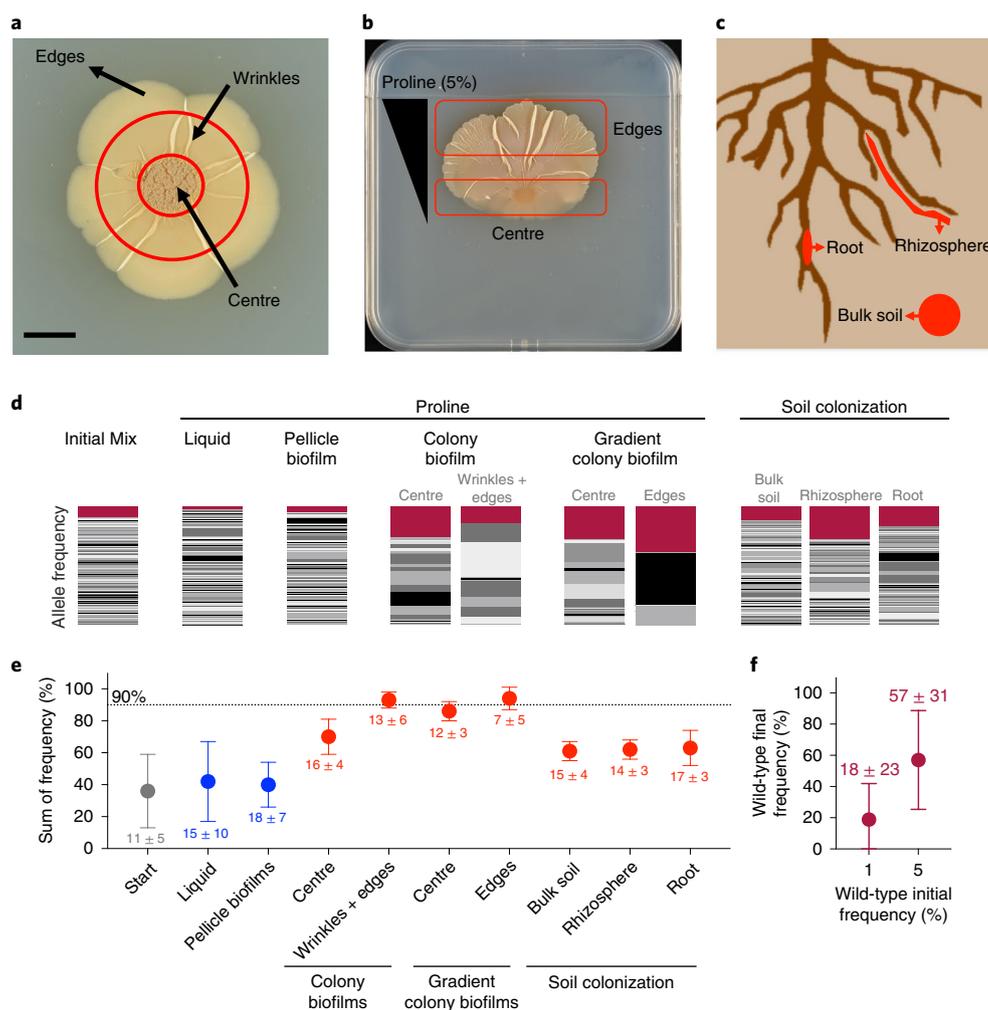


Fig. 3 | Genetic sweeps in colony biofilms and soil and the dominance of the wild-type allele. a, b, Photographs of five-day-old colony biofilms: a normal colony biofilm (**a**) (scale bar, 1 cm) and a gradient colony biofilm (**b**) with proline as the carbon/nitrogen source are shown. **c**, A scheme of soil colonization; red colouring indicates the sample types that were analysed. **d**, The distribution of the frequency of individual alleles for different growth states with proline as the carbon/nitrogen source. Bar thicknesses represent allele frequencies from raw read counts (Rf values; Supplementary Data 1). Magenta corresponds to wild-type *GudB*. **e**, Alleles with Rf $\geq 1\%$ were identified and their number and sum of frequencies are shown (dots indicate the average and the error bars the s.d. for all experiments in a given condition). Selection-dominated conditions are shown in blue and drift-dominated conditions in red; as in Fig. 1. **f**, An initial mix of 4 alleles was created, including the wild-type allele, at a frequency varying from 1–25%. The edges of two normal and two gradient colony biofilms that used proline as the carbon/nitrogen source were dissected in three parts (12 samples). The dots indicate the mean and the error bars the s.d. for all 12 dissected edges. The frequency of the wild type reached a mean of 18% when initiated at 1% and up to 100% when initiated at 5% (see Supplementary Fig. 11 for the entire dataset).

type was enriched by nearly 20-fold even when barely present in the initial mix (Fig. 3f).

Similarly, we searched for 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 colony 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). Furthermore, 19 amino acid alleles were found to be enriched in at least 10 of the 15 sequenced populations, suggesting some degree of reproducibility (Supplementary Table 5). Selection during soil colonization is also manifested in the variation between biological replicas (s.d. values) of alleles that were enriched in root populations is on average 20% smaller than those that were not (Fig. 4d). Finally, stop codons were purged in all colony biofilms and soil populations, indicating, as expected, that the activity of *GudB* is required for *B. subtilis* to survive under these conditions^{22,23}.

Overall, in colony biofilms and in soil colonization both drift and selection determine the fate of *gudB* alleles. We further examined the colony biofilms as described in the next section.

Drift in biofilms relates to programmed prophage excisions

Mutagenic rates in colony biofilms are high and mutations with a selective advantage rapidly take over^{30,31}. Growth in colony biofilms is also spatially defined, giving rise to segregated lineages in which 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 pellicle biofilms, spatial segregation is expected to be less pronounced than in colony biofilms^{32,33}. Accordingly, we found that, in contrast to colony biofilms, selection acts reproducibly in pellicle biofilms (Fig. 1). To further establish that spatial segregation is a key factor, we divided the edges of the colony biofilm into small sections and sequenced them. We found that most

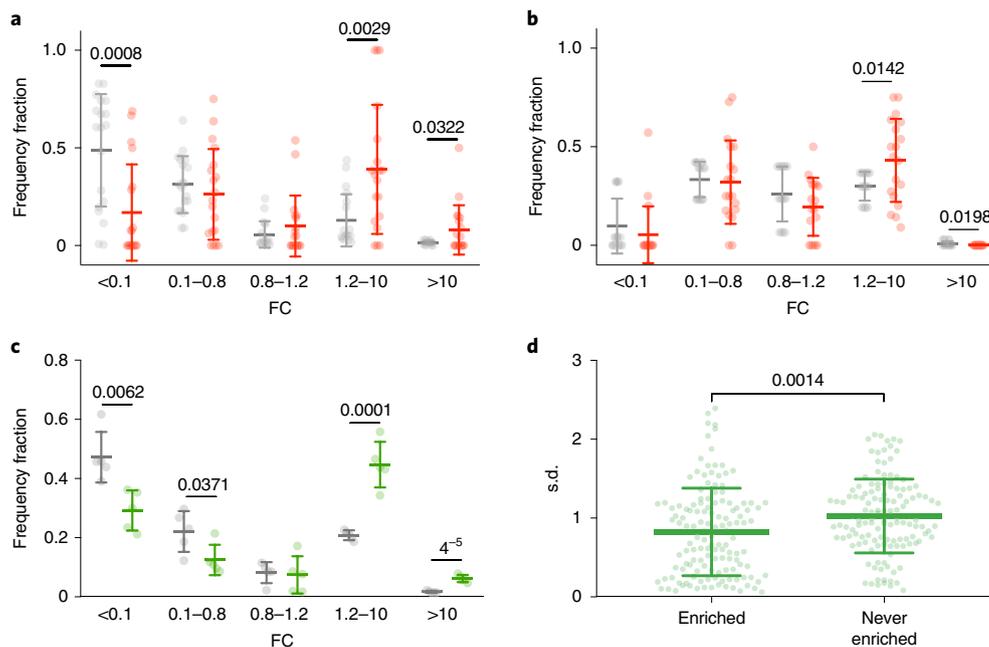


Fig. 4 | The combined action of selection and chance in colony biofilms and soil colonization. **a**, Alleles that were enriched in the edges of the colony biofilms are more likely to arise from alleles that were neutral or enriched in the centre. The distribution of categorized FC values from all colony biofilm centres (grey) compared to the distribution of FC values of centre alleles that were enriched in the edges (red) is shown. **b**, The distribution of categorized FC values of all alleles in all liquid conditions (grey) compared to the distribution of FC values of liquid alleles that were enriched in the edges of colony biofilms (red) is shown. **c**, Alleles enriched in the root are more likely to arise from alleles that were enriched in the bulk soil. The distribution of categorized allele FC values in all soil samples (grey) compared to the distribution of FC values of soil alleles that were enriched in the root (green) is shown. **d**, The distribution of s.d. values (variability between replica experiments, as in Fig. 1c) of alleles enriched in one or more root populations compared to alleles that were never enriched in the roots. The horizontal lines reflect the mean from 19 different normal and gradient colony biofilms samples (**a,b**) and the mean from 5 different root samples (**c,d**). In all panels, the error bars indicate the s.d. Student's *t*-tests were computed to compare the log[FC] values for each category (raw data comparison) for **a–c**, and to compare the s.d. values between replica log[FC] values for **d**. *P* values indicating significance ($P < 0.05$) are presented above the bars (details of all the *t*-tests are provided in Supplementary Table 8).

sections contained a single *gudB* allele (Supplementary Fig. 7 and Supplementary Table 6). Thus, in a way, the *gudB* allele represents a 'barcode' that reports single founder cells giving rise to individual sectors of the colony biofilm¹².

What might be the mutations driving these genetic sweeps and the spatial segregation? We sequenced samples for which enough genomic DNA was available (6 ordinary and 12 gradient colony biofilms and, for comparison, 2 initial mix, 6 liquid culture and 4 pellicle biofilm samples). A range of single nucleotide polymorphisms (SNPs) in various loci was identified across these samples (Supplementary Data 3). We focused, however, on identifying genomic mutations that were not observed, or rarely observed, in the initial mix and/or in liquid samples, suggesting that they emerged and were enriched in the colony biofilms.

We observed two large genome deletions that occurred in all colony biofilms with a frequency approaching 100% (Fig. 5a,b). These deletions correspond to the excision of two mobile genetic elements, or prophages: skin and SP- β ^{34–36}. Excision of skin generates a functional protein: sigK, a sporulation-specific transcription factor essential for cell differentiation in *B. subtilis*³⁷. The excision of SP- β generates another functional protein dubbed SpsM—a protein involved in capsid polysaccharide biosynthesis³⁸ and with relevance to biofilm thickness³⁹. Nearly all colony biofilm cells carried one of these variations and most cells carried both (Fig. 5a,b; Supplementary Table 7 and Supplementary Data 3). These prophage excisions therefore seem to be under a stronger selection pressure than the *GudB* mutations. The frequency of these structural variations gradually increases, from none in the initial mix to 100% in

gradient colony biofilms (Fig. 5c), as does the *GudB* drift signature (Fig. 1). However, at this stage, the observed link between the prophage excisions and the drift of *GudB* is circumstantial and further experiments are needed to establish how these two phenomena are linked. The prophage excisions are also likely to occur in the soil samples, but insufficient DNA was recovered from these samples to allow genome sequencing.

We detected 59 enriched SNPs in a conserved region of 16S ribosomal RNA exclusively in colony biofilms (Supplementary Table 7 and Supplementary Data 3). However, *B. subtilis* has ten 16S rRNA gene copies. Since these are essentially identical, we could not determine which of these ten paralogues carried mutations. For each population, 98% of the 16S rRNA mutations occurred in the same Illumina read, suggesting that one paralogue was highly mutated while others remained intact (Supplementary Fig. 8). Large differences in the expression levels of 16S rRNA genes were identified in *Pseudomonas aeruginosa* biofilms⁴⁰ and ribosomal heterogeneity has been linked to biofilm development in *B. subtilis*⁴¹. Yet mutations in the 16S rRNA genes have not been reported in biofilms. At this stage, however, which 16S gene is inactivated, how multiple proximal 16S mutations occur and how inactivation affects biofilm development remain unclear. Overall, the 16S rRNA SNPs, and the structural variations in particular, seem to have a key role in biofilm development in *B. subtilis*. Accordingly, most of these genetic variations were reproducible between replica experiments (Supplementary Table 7) suggesting that they arose during biofilm growth and were then enriched by virtue of promoting biofilm formation¹².

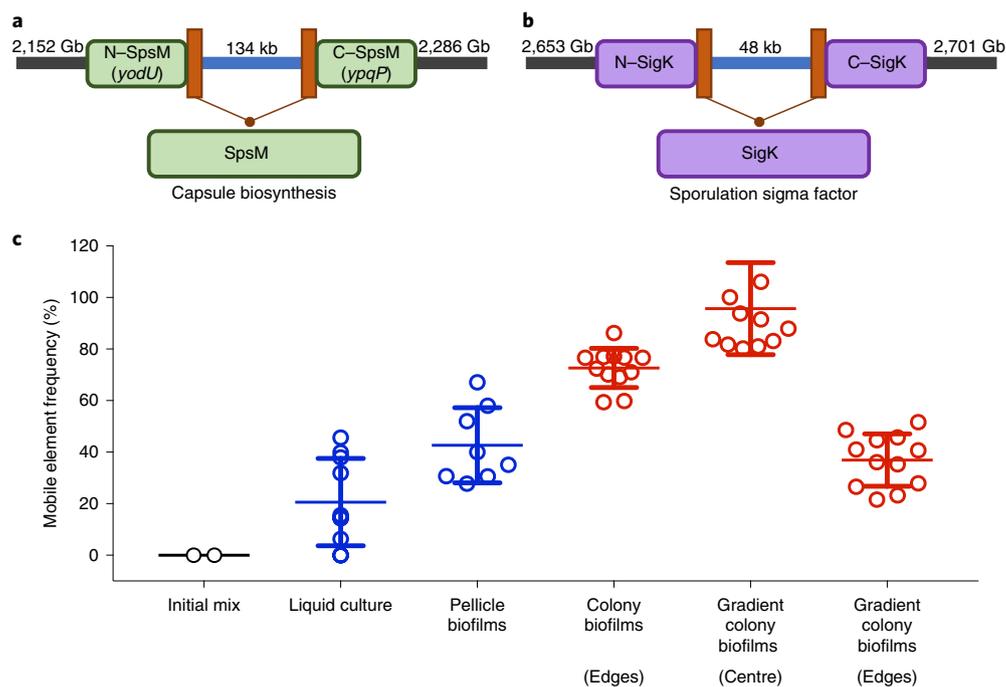


Fig. 5 | Programmed genomic excisions drive *GudB*'s drift in colony biofilms. **a, b**, A schematic representation of *B. subtilis*'s genomic organization before and after the excision of the prophage mobile elements SP- β (**a**) and skin (**b**), and the position of these mobile elements in the genome. **c**, These excisions were absent in the initial mix yet dominated colony biofilms and reached near fixation in the edges of gradient colony biofilms (for frequencies in individual experiments see Supplementary Table 7). Excision of the mobile elements occurred in two different genomic locations within the same experiment. The values were summed and averaged according to the general condition shown. The centre values (long horizontal lines) reflect the mean, and the error bars, the s.d. (sample sizes: initial mix, 2; liquid culture, 6; pellicle biofilms, 4; normal colony biofilms, 6; gradient colony biofilms, 12). The details of the excisions (location and frequency) for each experiment are shown in Supplementary Data 3.

Conclusions

It is generally assumed that the fitness effect of a mutation may vary depending on the environment. However, the magnitude of environmental pleiotropy unravelled here is surprisingly high. Environmental changes, including minute ones like the 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 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 the wild-type *GudB* sequence is singular in being fit across multiple constraints and environments. For each of the individual tested conditions, most mutations are either neutral or beneficial (28–81%). However, if all the tested conditions are considered, only 2% of the tested *GudB* mutations are neutral or beneficial in the nine reproducible conditions when selection acts on *GudB*.

The pleiotropy of protein mutations across multiple growth and environmental conditions has rarely been measured¹⁷, particularly regarding proteins with intrinsic physiological roles. 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 the role of *GudB* as an enzyme and regulator and also to the positions explored (the 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 (microevolution) but also the evolution of protein sequences over long evolutionary periods and across species (macroevolution). The correlation between the effects of mutations in laboratory mappings and their occurrence or absence in natural sequences is limited¹³. However, laboratory mappings represent a single condition and merging data from multiple conditions could in principle reveal higher correlation. Identifying trends in complex datasets requires an unbiased approach, but even when several different machine learning approaches were applied, merging conditions gave no further correlation (Supplementary Fig. 9). Thus, over short evolutionary periods, proteins experience variable and opposing selection pressures. Furthermore, 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 theory¹¹. 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 were 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 for or a priori enabled by other mutations. Compensation (or enabling by other mutations) results in epistasis, that is, the effect of mutations being dependent on the sequence context in which they occur⁴². Accordingly, over macroevolutionary time scales, epistasis dominates gene and genome sequences⁴³.

Methods

Strains. *B. subtilis* NCIB 3610 DS7187 (from D. B. Kearns⁴⁴) that lacks the ComI peptide and has a high competence capacity similar to that of domesticated

B. subtilis strains was used in this study. *B. subtilis* NCIB 3610 *gudB::tet* strain²³ genomic DNA was transformed into *B. subtilis* NCIB 3610 DS7187. *B. subtilis* NCIB 3610 Δ *comI gudB::tet* was thus isolated and phenotypically and genetically tested.

***gudB* allele library construction.** We performed site-directed mutagenesis in ten codons (amino acids: M46, L48, K52, D58, D59, S61, K63, T66, Y68, S75) of the *gudB* gene cloned in the pDG_GudB plasmid, which was modified from the pDG1728 backbone vector²³. The codons were mutated to NNS whereby the 20 standard amino acids and 1 stop codon are encoded. Codon mutagenesis was performed independently for each position using a one-step PCR protocol. Thus, we created 10 libraries, each containing 20 different amino acid alleles (non-synonymous, missense mutations), 1 stop codon (nonsense) and 11 synonymous alleles (alternative codons encoding the same amino acid). All mutagenic PCR assays were performed using KAPA HiFi HotStart Ready Mix (KAPA Biosystems) following the manufacturer's protocol (Supplementary Table 9 shows the sequences of all the primers). The ten PCR products were purified and used to transform the *E. coli* T10 strain (Thermo Fisher Scientific). Clones were pooled together after an overnight growth on LB plus ampicillin (100 μ g ml⁻¹) agar plates at 37 °C. At this stage, four to six clones per library were isolated and analysed by sequencing. Total plasmid DNA from these library transformations was extracted and also analysed by sequencing. Each of the 10 libraries contained, after transformation, at least 1 \times 10⁵ clones, corresponding to \geq 1,000-fold coverage for each allele. Approximately 10 μ g of plasmid DNA, from each library, was linearized (XhoI, New England Biolabs, following the manufacturer's protocol), purified and used to transform the *B. subtilis* NCIB 3610 *gudB::tet* Δ *comI* strain. Transformations were performed as described²³. After transformation, overnight growth on in plus spectinomycin (100 μ g ml⁻¹) plus glucose (0.5 mg ml⁻¹) agar plates was used for the selection. The resulting cells were pooled together and kept at -20 °C in 50% glycerol. In total, ten *B. subtilis* libraries were constructed in parallel and each contained, after transformation, at least 1 \times 10⁴ clones (\geq 100-fold coverage for each allele). Genomic DNA extraction was performed for each library (GenElute, Sigma). The integrity of the mutagenic process was verified by Sanger sequencing the *amyE::gudB* locus, which indicated that mutations were observed only in the diversified codon.

Selection and growth conditions. 10 ml samples of LB (1% tryptone w/v, 5% yeast extract w/v and 1% NaCl w/v) with glucose (0.5% w/v), ammonium sulphate (0.5% w/v) and spectinomycin (100 μ g ml⁻¹) were inoculated with 1 ml of each library stock. The cultures were grown overnight at 37 °C with shaking. The overnight culture (500 μ l) was used to inoculate 3 ml of LB plus glucose (0.5%) and ammonium sulphate (0.5%). The cultures were incubated at 37 °C with shaking and once the OD₆₀₀ reached 0.8 they were mixed equally and used as the starting populations (initial mixes). A fraction of the cells at this stage was 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 (four carbon/nitrogen sources), pellicle biofilms and gradient colony biofilms; initial mix 2 was used to inoculate one liquid condition, spores, germination and colony biofilms; and initial mix 3 was used to inoculate bulk soil (Supplementary Data 1). Detailed selection conditions are listed below.

For the selection under liquid serial passages, 100 μ l of the initial mix was used to inoculate 10 ml cultures of basic MS medium (5 mM potassium phosphate, 100 mM MOPS pH 7.1, 2 mM MgCl₂, 700 μ M CaCl₂, 50 μ M MnCl₂, 50 μ M FeCl₃, 1 μ M ZnCl₂, 2 μ M thiamine, 50 μ g ml⁻¹ tryptophan, 50 μ g ml⁻¹ phenylalanine and 50 μ g ml⁻¹ threonine)²³ supplemented with either: glucose (0.5%) plus ammonium sulphate (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 the OD₆₀₀ reached a maximum of 1.5, after which 100 μ l was used to inoculate 10 ml of fresh medium. When proline (0.5%), arginine (0.5%), or proline (0.25%) plus arginine (0.25%) were used as carbon/nitrogen sources, the serial passages were performed every 24 h. After 18 h, the cultures reached OD₆₀₀ 0.8–1 (corresponding to 7 generations) and were kept for another 6 h in the stationary phase until the cultures reached a maximum of 1.5 at OD₆₀₀. When glucose (0.5%) plus ammonium sulphate (0.5%) or glutamate (0.5%) plus glycerol (0.5%) were used, the serial passages were performed every 12 h. After 10 h, the cultures reached OD₆₀₀ 0.8–1 (corresponding to seven generations) and were kept for another 2 h in the stationary phase until the cultures reached a maximum OD₆₀₀ of 1.5. In total, all the liquid passages were maintained for 35 generations in the exponential phase and an unknown number of generations in the stationary phase.

For selection in pellicle biofilms, 100 ml of media (glutamate + glycerol, proline and arginine) were inoculated with 100 μ l of the initial mix cells. The culture was incubated at 30 °C without shaking, for 5 d.

For selection of spores and germinated spores, 3 ml of the initial mix was used to inoculate 25 ml of Difco Sporulation Medium in 250 ml flasks, which was then incubated at 37 °C with 150 r.p.m. shaking until OD₆₀₀ reached 0.4. This culture was used to inoculate 250 ml of fresh Difco Sporulation Medium in 11 flasks. The cultures were incubated for 48 h at 37 °C with 150 r.p.m. shaking. Cells were

subsequently harvested by centrifugation and stored at 4 °C overnight. Afterwards, cells were resuspended with 200 ml of cold deionized sterile water (dW) and incubated for 30 min at 4 °C. All cells were harvested and resuspended with 200 ml of cold dW and incubated overnight at 4 °C with slow orbital agitation, to kill all planktonic and vegetative cells. The culture was harvested, resuspended in 30 ml of dW and heated to 80 °C for 20 min. Finally, spores were harvested, resuspended in 10 ml of dW and stored at -20 °C. To germinate these spores, they were diluted 1,000 times in phosphate-buffered saline solution and 100 μ l of this suspension was used to inoculate LB plus glucose (0.5%) agar plates (ten plates). Approximately 10,000 colonies were obtained and pooled together.

For selection in colony biofilms, MS agar (1.5%) plates supplemented with different carbon/nitrogen sources were prepared. For gradient colony biofilms, gradient agar plates were prepared. First, square plates (12 cm \times 12 cm) with MS agar (1.5%) medium were poured. After the agar solidified, an area of 2 cm \times 14 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 colony 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 Supplementary Fig. 12a for a graphic representation of the agar plate preparation). For all the gradients, 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 of incubation at 30 °C (Supplementary Fig. 12b). For growth in colony biofilms and gradient colony biofilms, 5 μ l of the initial mix was used as the inoculum. Plates were incubated for 4 d at 30 °C and an additional 2 d at room temperature. The colony was then dissected in three areas (centre, wrinkles and edges) for normal colony biofilms and in two areas (centre and upper) for gradient colony biofilms (illustrated in Supplementary Fig. 12c–g). After selection in all of the previously mentioned conditions, the biomass was harvested and stored 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 previously 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 (OD₆₀₀ = 0.8) was applied. This LB culture was washed three times (by means of centrifugation and resuspension) with 100 ml of half-strength Hoagland solution⁴⁵. After the final wash, the cells were resuspended in half-strength Hoagland solution to a final OD₆₀₀ of 0.1. Since Hoagland solution is not isotonic, the washes resulted in the death of about a third of the *B. subtilis* cells. Thus, the handling of 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 (\leq 30%). The soil colonization FC values therefore result from the entire process that begins with the Hoagland solution rinses 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 a 2 mm sieve and autoclaved three times for 30 min at 121 °C. A total of five pots (10 cm \times 8 cm \times 5 cm) containing autoclaved natural soil were drenched with the initial mix suspended in half-strength Hoagland solution⁴⁵. These potted soils drenched with bacterial suspensions were used to plant tomato seedlings grown first in sterile conditions. Tomato seeds (*Solanum lycopersicum* L.; cultivar Micro-Tom) were surface sterilized with 70% ethanol (5 min) and 3% bleach with 0.01% Tween 20 surfactant (10 min). Surface-sterile seeds were germinated on sterile filter paper (Whatman, catalogue no. 1001-085) saturated with half-strength Hoagland solution for 7 d (23 °C; 16 h photoperiod). Six tomato seedlings were transferred to each pot and grown for one month (21 °C; 16 h light; 8 h dark) and drenched with half-strength Hoagland solution twice a week. Plants were subsequently harvested from the five pots. Roots and rhizosphere samples were collected for each replica experiment, which consisted of a pool of six roots. First, the plants were carefully removed from the soil. Roots were then cut from the plants and vortexed in 20 ml of washing solution (0.85% NaCl) for 30 s. This step was repeated once with a fresh washing solution. The combined root washing solutions (40 ml) were centrifuged for 30 min at 3,000 r.p.m. and the resulting pelleted samples that corresponded to the rhizosphere were frozen in liquid nitrogen and stored at -80 °C. The washed roots were blotted with 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 biofilms, spores, colony biofilms and gradient colony biofilms, were defrosted and resuspended in 10 ml of dW. The samples were sonicated at 40% power (VibraCell, Sonics) for 10 min at 60 s intervals. Cell debris was harvested by centrifugation (13,000g for 20 min). Genomic DNA from all the samples was extracted using the GenElute Bacterial Genomic DNA Kit (Sigma) generally following the manufacturer's instructions, with the exception of the soil, rhizosphere soil and plant root samples. For these samples, the PowerSoil DNA Isolation kit (MO BIO) was used, following the 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'-CTCTTCCCTACACGACGCTCTCCGATCTNNNNNNCCC-GAAGAGGTATACGAATTGTTAAAGAG) and GudB_In_Rev (5'-CTGGAGTTCAGACGTGTGCTCTCCGATCTCGCCTTCGTGGACCGAC). Six Ns were added to the GudB_In_For primer to increase the sequence variability between the amplicons. PCR assays were performed using KAPA HiFi HotStart Ready Mix (KAPA Biosystems) using ~100 ng of genomic DNA as the template and following the manufacturer's instructions. Using 10 µl of the PCR products as a template, a second PCR was performed to add the Illumina adaptor sequence, using primers GudB_Out_For (5'-AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGC) and GudB_Out_Rev (5'-CAAGCAGAAGACGGCATACGAGATGTTATACGTTGACTGGAGTTCAGACGTGTGTC). The Illumina index (underlined) was changed in the GudB_Out_Rev primer to different Illumina indexes. Each condition was barcoded differently. All PCR products were purified using the Agencourt AMPure XP (Beckman Coulter). The concentration of PCR products was verified using a Qubit assay (Life Technologies).

Analysis of the Illumina reads. DNA samples were run using the Illumina NextSeq 150-base-pair paired-end kit. The FASTQ sequence files were obtained for each run and customized using MatLab 8.0 and Python 3.6 custom scripts designed to count the number of each individual allele in each sequenced sample. We filtered the reads to exclude any that had mutations outside the mutagenized codons. All codons encoding the wild-type amino acid were counted together as one group and assigned as wild type. All other codons were counted independently. The unprocessed read counts are shown in Supplementary Data 1. Further filtering excluded alleles with <100 counts in the initial mix to avoid statistical uncertainty with respect to FC values. In total, we obtained data for up to 269 individual alleles for each condition out of the 320 alleles originally introduced. For each condition, a minimum of 380,000 reads were obtained. Thus, on average, we obtained 1,500 reads per allele.

Data analysis. The frequency of each allele (f_i) was calculated as the ratio between the number of reads for allele i divided by the total number of reads. The allele FC (FC_{*i*}) was subsequently calculated as the ratio of the after-selection f_i to f_i of the same allele in the initial mix (Supplementary Fig. 2 and Supplementary Data 2). Normalization by the number of wild-type reads rather than by the total number of reads gave essentially identical FC values for the majority of the samples. However, in the few samples where wild-type frequency was significantly reduced after selection, normalization resulted in high noise and large biases, including large changes in sign (higher sign environmental 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 log[FC] values were compared. To this end, all FC values 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 culture, pellicle biofilm, colony biofilm, spore and germinated spore experiments (Supplementary Data 2, sheet 1) the zeros were changed to 4.2×10^{-6} . For the bulk soil experiments (Supplementary Data 2, sheet 2) zeros were changed to 1.14×10^{-5} . The logarithm of all the FC values was calculated and was also used to derive mean FC values. We then performed the F -test and Levene's test to compare, for each allele, the log[FC] values for the liquid culture, pellicle biofilms and spores (reproducible conditions) versus the log[FC] values for the germinated spores, colony biofilms and bulk soil (irreproducible conditions). Furthermore, the log[FC] values were then used to calculate s.d.; the s.d. values between log[FC] values observed for each allele in replica experiments were averaged for all alleles measured in a given condition; and the s.d. values between synonymous codons in the same replica experiment (deviations between log[FC] 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 in each condition). The t -, F - and χ^2 tests and Pearson correlation values were obtained using the PRISM software. Levene's test was performed using R. Supplementary Table 10 shows the sample size for every test performed in this study.

Defining the limits of neutrality. Of all the conditions tested here, it was only in the glucose plus ammonia condition that the GudB knockout had no growth effect (Supplementary Fig. 1). Hence, this condition is largely neutral and the variation observed in the FC values would primarily be the outcome of noise. The s.d. between three biological replicas was calculated for each allele and these values ranged from 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 designated a mutation as deleterious and $FC > 1.2$ assigned a mutation as beneficial.

Genome sequencing. We sequenced the genomic DNA of all the colony biofilm populations for which we had ≥ 1 µg of DNA after extraction (6 normal and 12 gradient colony biofilms). For comparison, we also sequenced initial mix populations 1 and 2, six liquid culture and four pellicle biofilm populations. The Illumina HiSeq2500 platform was used, with a read length of 2×125 base pairs. We obtained a total of 300 million reads. The reads were assembled using the *B. subtilis* NCIB 3610 genome (NCBI accession number: CP020102; <https://www.ncbi.nlm.nih.gov/ncbi>).

[ncbi.nlm.nih.gov/nuccore/1162900886/](https://www.ncbi.nlm.nih.gov/nuccore/1162900886/)) as reference. Overall, 95% of the reads were successfully mapped to the reference genome with minimal coverage of $\times 300$ for all the samples analysed. The breseq program was used to identify genomic variants, including SNPs and insertion-deletion polymorphisms⁴⁶ (Supplementary Data 3 and Supplementary Table 7).

Comparison of FC values and 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 amongst the sequences of naturally occurring GDHs. To this end, we constructed a number of different support vector machine classification models with a variety of kernels (such as linear, Gaussian, polynomial and so on). The feature vector of each *gudB* allele was composed of the normalized FC values for a specific condition. The values from the replica experiments of the highly reproducible liquid conditions were averaged prior to training. Based on the multiple sequence alignment containing 1,013 GDH sequences, we divided the GudB mutations in our dataset into three categories, which were then utilized as the prediction labels: (1) mutations seen in <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 for the unbalanced nature of the dataset. For each feature combination of a varied length, we built a support vector machine classification model and assessed its accuracy using threefold cross-validation. To reduce noise, and assuming that our data belong to linear space, we also extracted the first ten principal components of the feature matrix and used them as the new feature vectors for model construction. To examine whether our relatively high (>0.6) model accuracy was distributed uniformly across different classes for each model and genotype, we recorded the predicted values during threefold cross-validation. Moreover, for each condition combination, and for each kernel, we built 100 different models and recorded the number of times each of the genotypes was predicted correctly.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data that support the findings of this study are available from the corresponding author on request.

Received: 29 May 2018; Accepted: 14 February 2019;
Published online: 1 April 2019

References

- Baba, T. et al. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**, 2006.0008 (2006).
- Koo, B. M. et al. Construction and analysis of two genome-scale deletion libraries for *Bacillus subtilis*. *Cell Syst.* **4**, 291–305 (2017).
- Pache, R. A., Madan, M. M. & Aloy, P. Exploiting gene deletion fitness effects in yeast to understand the modular architecture of protein complexes under different growth conditions. *BMC Syst. Biol.* **3**, 74 (2009).
- Winzeler, E. A. et al. Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* **285**, 901–906 (1999).
- Civelek, M. & Lusi, A. J. Systems genetics approaches to understand complex traits. *Nat. Rev. Genet.* **15**, 34–48 (2014).
- Steinberg, B. & Ostermeier, M. Environmental changes bridge evolutionary valleys. *Sci. Adv.* **2**, e1500921 (2016).
- Wang, Z., Liao, B.-Y. & Zhang, J. Genomic patterns of pleiotropy and the evolution of complexity. *Proc. Natl Acad. Sci. USA* **107**, 18034–18039 (2010).
- Fusco, D., Gralka, M., Kayser, J., Anderson, A. & Hallatschek, O. Excess of mutational jackpot events in expanding populations revealed by spatial Luria-Delbrück experiments. *Nat. Commun.* **7**, 12760 (2016).
- Kimura, M. Genetic variability maintained in a finite population due to mutational production of neutral and nearly neutral isoalleles. *Genet. Res.* **89**, 341–363 (2008).
- Smith, N. H., Gordon, S. V., de la Rua-Domenech, R., Clifton-Hadley, R. S. & Hewinson, R. G. Bottlenecks and broomsticks: the molecular evolution of *Mycobacterium bovis*. *Nat. Rev. Microbiol.* **4**, 670–681 (2006).
- Nei, M. Selectionism and neutralism in molecular evolution. *Mol. Biol. Evol.* **22**, 2318–2342 (2005).
- Steenackers, H. P., Parijs, I., Foster, K. R. & Vanderleyden, J. Experimental evolution in biofilm populations. *FEMS Microbiol. Rev.* **40**, 373–397 (2016).
- Boucher, J. I., Bolon, D. N. A. & Tawfik, D. S. Quantifying and understanding the fitness effects of protein mutations: laboratory versus nature. *Protein Sci.* **25**, 1219–1226 (2016).
- De Visser, J. A. G. M. & Krug, J. Empirical fitness landscapes and the predictability of evolution. *Nat. Rev. Genet.* **15**, 480–490 (2014).
- Fowler, D. M. & Fields, S. Deep mutational scanning: a new style of protein science. *Nat. Methods* **11**, 801–807 (2014).

16. Steinberg, B. & Ostermeier, M. Shifting fitness and epistatic landscapes reflect trade-offs along an evolutionary pathway. *J. Mol. Biol.* **428**, 2730–2743 (2016).
17. Dandage, R. et al. Differential strengths of molecular determinants guide environment specific mutational fates. *PLoS Genet.* **14**, e1007419 (2018).
18. Earl, A. M., Losick, R. & Kolter, R. Ecology and genomics of *Bacillus subtilis*. *Trends Microbiol.* **16**, 269–275 (2008).
19. Branda, S. S., González-Pastor, J. E., Ben-Yehuda, S., Losick, R. & Kolter, R. Fruiting body formation by *Bacillus subtilis*. *Proc. Natl Acad. Sci. USA* **98**, 11621–11626 (2001).
20. Belitsky, B. R. & Sonenshein, A. L. Role and regulation of *Bacillus subtilis* glutamate dehydrogenase genes. *J. Bacteriol.* **180**, 6298–6305 (1998).
21. Gunka, K. & Commichau, F. M. Control of glutamate homeostasis in *Bacillus subtilis*: a complex interplay between ammonium assimilation, glutamate biosynthesis and degradation. *Mol. Microbiol.* **85**, 213–224 (2012).
22. Liu, J. et al. Metabolic co-dependence gives rise to collective oscillations within biofilms. *Nature* **523**, 550–554 (2015).
23. Noda-Garcia, L., Romero Romero, M. L., Longo, L. M., Kolodkin-Gal, I. & Tawfik, D. S. *Bacilli* glutamate dehydrogenases diverged via coevolution of transcription and enzyme regulation. *EMBO Rep.* **18**, 1139–1149 (2017).
24. De Jong, L. et al. In-culture cross-linking of bacterial cells reveals large-scale dynamic protein-protein interactions at the peptide level. *J. Proteome Res.* **16**, 2457–2471 (2017).
25. Stannek, L. et al. Evidence for synergistic control of glutamate biosynthesis by glutamate dehydrogenases and glutamate in *Bacillus subtilis*. *Environ. Microbiol.* **17**, 3379–3390 (2015).
26. Mondal, S., Yakhnin, A. V., Sebastian, A., Albert, I. & Babitzke, P. NusA-dependent transcription termination prevents misregulation of global gene expression. *Nat. Microbiol.* **1**, 15007 (2016).
27. Besharova, O., Suchanek, V. M., Hartmann, R., Drescher, K. & Sourjik, V. Diversification of gene expression during formation of static submerged biofilms by *Escherichia coli*. *Front. Microbiol.* **7**, 1568 (2016).
28. Vlamakis, H., Chai, Y., Beauregard, P., Losick, R. & Kolter, R. Sticking together: building a biofilm the *Bacillus subtilis* way. *Nat. Rev. Microbiol.* **11**, 157–168 (2013).
29. Mavor, D. et al. Determination of ubiquitin fitness landscapes under different chemical stresses in a classroom setting. *eLife* **5**, e15802 (2016).
30. Driffield, K., Miller, K., Bostock, J. M., O'Neill, A. J. & Chopra, I. Increased mutability of *Pseudomonas aeruginosa* in biofilms. *J. Antimicrob. Chemother.* **61**, 1053–1056 (2008).
31. Hallatschek, O., Hersen, P., Ramanathan, S. & Nelson, D. R. Genetic drift at expanding frontiers promotes gene segregation. *Proc. Natl Acad. Sci. USA* **104**, 19926–19930 (2007).
32. Hölscher, T. et al. Monitoring spatial segregation in surface colonizing microbial populations. *J. Vis. Exp.* **116**, e54752 (2016).
33. Van Gestel, J., Weissing, F. J., Kuipers, O. P. & Kovács, Á. T. Density of founder cells affects spatial pattern formation and cooperation in *Bacillus subtilis* biofilms. *ISME J.* **8**, 2069–2079 (2014).
34. Kunkel, B., Losick, R. & Stragier, P. The *Bacillus subtilis* gene for the developmental transcription factor sigma K is generated by excision of a dispensable DNA element containing a sporulation recombinase gene. *Genes Dev.* **4**, 525–535 (1990).
35. Nicolas, P. et al. Condition-dependent transcriptome reveals high-level regulatory architecture in *Bacillus subtilis*. *Science* **335**, 1103–1106 (2012).
36. Westers, H. et al. Genome engineering reveals large dispensable regions in *Bacillus subtilis*. *Mol. Biol. Evol.* **20**, 2076–2090 (2003).
37. Eichenberger, P. The program of gene transcription for a single differentiating cell type during sporulation in *Bacillus subtilis*. *PLoS Biol.* **2**, e328 (2004).
38. Abe, K. Developmentally-regulated excision of the SPβ prophage reconstitutes a gene required for spore envelope maturation in *Bacillus subtilis*. *PLoS Genet.* **10**, e1004636 (2014).
39. Sanchez-Vizuete, P. et al. Identification of ypqP as a new *Bacillus subtilis* biofilm determinant that mediates the protection of *Staphylococcus aureus* against antimicrobial agents in mixed-species communities. *Appl. Environ. Microbiol.* **81**, 109–118 (2015).
40. Pérez-Osorio, A. C., Williamson, K. S. & Franklin, M. J. Heterogeneous rpoS and rhlR mRNA levels and 16S rRNA/rDNA (rRNA gene) ratios within *Pseudomonas aeruginosa* biofilms, sampled by laser capture microdissection. *J. Bacteriol.* **192**, 2991–3000 (2010).
41. Vesper, O. et al. Selective translation of leaderless mRNAs by specialized ribosomes generated by MazF in *Escherichia coli*. *Cell* **147**, 147–157 (2011).
42. Miton, C. M. & Tokuriki, N. How mutational epistasis impairs predictability in protein evolution and design. *Protein Sci.* **25**, 1260–1272 (2016).
43. Breen, M. S., Kemena, C., Vlasov, P. K., Notredame, C. & Kondrashov, F. A. Epistasis as the primary factor in molecular evolution. *Nature* **490**, 535–538 (2012).
44. Konkol, M. A., Blair, K. M. & Kearns, D. B. Plasmid-encoded ComI inhibits competence in the ancestral 3610 strain of *Bacillus subtilis*. *J. Bacteriol.* **195**, 4085–4093 (2013).
45. Yi, Y., de Jong, A., Frenzel, E. & Kuipers, O. P. Comparative transcriptomics of *Bacillus mycoides* root exudates reveals different genetic adaptation of endophytic and soil isolates. *Front. Microbiol.* **8**, 1487 (2017).
46. Deatherage, D. E. & Barrick, J. E. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods Mol. Biol.* **1151**, 165–188 (2014).

Acknowledgements

L.N.-G. was supported by CONACYT grant no. 203740 and the Martin Kushner Fellowship at the Weizmann Institute of Science. D.S.T. is the Nella and Leon Benozio Professor of Biochemistry. Financial support from the Kahn Centre for Systems Biology at the Weizmann Institute of Science is gratefully acknowledged. We thank R. Milo, S. Fleishman, Z. Livneh and F. Kondrashov for their support and critical advice and E. Segev and A. de Visser for their critical and insightful comments on the manuscript. We appreciate the help of M. Hershko with script development for data processing and of Y. Bar-On and S. Gleizer with the analysis of genomic sequences. We are grateful to R. Rotkopf from Weizmann Life Sciences Core Facilities for his guidance on the statistical analysis. We are thankful for the services provided by the Crown Genomics Institute of the Nancy and Stephen Grand Israel National Centre for Personalized Medicine, Weizmann Institute of Science.

Author contributions

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

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41564-019-0412-y>.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to D.S.T.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated
- Clearly defined error bars
State explicitly what error bars represent (e.g. SD, SE, CI)

Our web collection on [statistics for biologists](#) may be useful.

Software and code

Policy information about [availability of computer code](#)

Data collection A computer code, which will be provided, was written to process the illumina sequence data.

Data analysis Excel, R and Prism softwares were used to analyze the data.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Figure 1, 2 and 4 derive from Data S2, Figure 3 derives from Data S1 and Figure 5 derives from Data S3

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/authors/policies/ReportingSummary-flat.pdf](https://www.nature.com/authors/policies/ReportingSummary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Typically for these experiments (quantification of the fitness effect of point mutations) only one replica is performed, in this paper we used 3 to 5 replicas to measure the reproducibility of the fitness effects of each mutation analyzed. The sample size was defined arbitrarily.
Data exclusions	The initial bacterial population (referred to as the Initial Mix in the manuscript), not selected yet in any condition were sequenced. The alleles that contained less than 100 reads in these initial populations were excluded from further analysis.
Replication	The reproducibility levels for each experiment performed is a finding described to detail in the manuscript. Some conditions are not reproducible, and this is discussed and further analyzed. We provide a molecular explanation for the lack of reproducibility.
Randomization	n/a
Blinding	n/a

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Unique biological materials
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants

Methods

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging