Genetics of adaptation of the ascomycetous fungus *Podospora anserina* to submerged cultivation

Olga A. Kudryavtseva[1¶], Ksenia R. Safina[2,3¶*], Olga A. Vakhrusheva[2,3], Maria D. Logacheva[2,4], Aleksey A. Penin[3,4,5,6], Tatiana V. Neretina[4], Viktoria N. Moskalenko[7], Elena S. Glagoleva[8], Georgii A. Bazykin[2,3], Alexey S. Kondrashov[9]

[1] Department of Mycology and Phycology, Faculty of Biology, Lomonosov Moscow State University, Moscow, Russia

[2] Department of Life Sciences, Skolkovo Institute of Science and Technology, Skolkovo, Russia

[3] Sector of Molecular Evolution, Institute for Information Transmission Problems of the Russian Academy of Sciences, Moscow, Russia

[4] Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University,

Moscow, Russia

[5] Department of Genetics, Faculty of Biology, Lomonosov Moscow State University, Moscow, Russia

[6] Extreme Biology Laboratory, Institute of Fundamental Medicine and Biology, Kazan Federal University, Kazan, Russia

[7] Zoological Museum, Lomonosov Moscow State University, Moscow, Russia

[8] Department of Plant Physiology, Faculty of Biology, Lomonosov Moscow State University, Moscow, Russia

[9] Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan, USA

* Corresponding author

E-mail: k.r.safina@gmail.com

¶ These authors contributed equally to this work.

Abstract

Podospora anserina is a model ascomycetous fungus which shows pronounced phenotypic senescence when grown on solid medium but possesses unlimited lifespan under submerged cultivation. In order to study the genetic aspects of adaptation of *P. anserina* to submerged cultivation, we initiated a long-term evolution experiment. In the course of the first four years of the experiment, 125 single-nucleotide substitutions and 23 short indels were fixed in eight independently evolving populations. Six proteins that affect fungal growth and development evolved in more than one population; in particular, the G-protein alpha subunit FadA evolved in seven out of eight experimental populations. Parallel evolution at the level of genes and pathways, an excess of nonsense and missense substitutions, and an elevated conservation of proteins and their sites where the changes occurred suggest that many of the observed allele replacements were adaptive and driven by positive selection.

Author summary

Living beings adapt to novel conditions that are far from their original environments in different ways. Studying mechanisms of adaptation is crucial for our understanding of evolution. The object of our interest is a multicellular fungus *Podospora anserina*. This fungus is known for its pronounced senescence and a definite lifespan, but it demonstrates an unlimited lifespan and no signs of senescence when grown under submerged conditions. Soon after transition to submerged cultivation, the rate of growth of *P. anserina* increases and its pigmentation changes. We wanted to find out whether there are any genetic changes that contribute to adaptation of *P. anserina* to these novel conditions and initiated a long-term evolutionary experiment on eight

independent populations. Over the first four years of the experiment, 148 mutations were fixed in these populations. Many of these mutations lead to inactivation of the part of the developmental pathway in *P. anserina*, probably reallocating resources to vegetative proliferation in liquid medium. Our observations imply that strong positive selection drives changes in at least some of the affected protein-coding genes.

Introduction

Over the last several decades, experimental evolution became an essential part of evolutionary biology, providing an opportunity to observe and study changes that occur in controlled populations in the course of many generations. In particular, the *E. coli* long-term evolution experiment led by Richard Lenski since 1988 yielded a number of important results, including that on morphological evolution of cells (1,2), long-term fitness gain under a constant environment (1,3), and ecological specialization (4,5). The existence of twelve replicas made it possible to study changes in genotypes and phenotypes that occurred in parallel in independently evolving bacterial populations (6–8).

Being easily cultured and fast-growing, *E. coli* serves as a convenient model organism for experimental evolution. Still, the range of questions that can be studied on it and on other prokaryotes is limited by their relative simplicity. Eukaryotes such as protozoans or fungi possess more complex gene regulation, morphology and physiology and have also been used in the studies of experimental evolution (9–12).

Podospora anserina is a model ascomycete fungus that possesses a definite lifespan with pronounced phenotypic senescence when grown on solid medium (13). In contrast, under conditions of submerged cultivation, *P. anserina* displays unlimited lifespan with no signs of

81 senescence (14,15). Isolates taken from such cultures revert to senescence, indicating that the
82 unlimited lifespan is not associated with genetic changes, instead being a product of epigenetic
83 changes or environmental influences. Still, such isolates demonstrate increased lifespan when
84 plated on solid medium and become sterile, only rarely forming mostly nonfunctional
85 microconidia and protoperithecia (14,15). After some period of submerged cultivation,
86 *P. anserina* cultures acquire higher rates of biomass accumulation and change their
87 pigmentation. The mechanisms that alter *P. anserina* senescence program and affect its
88 phenotype under such conditions are unknown.

89 In order to study the genetic aspects of adaptation that takes place during the submerged
90 cultivation of *P. anserina*, we started in 2012 a long-term evolutionary experiment. We used two
91 wild strains of *P. anserina* to establish eight independent experimental populations. Here, we
92 analyse the genetic changes that took place in the first four years of this experiment.

93

94 ## Results

95 ### Allele replacements

96 We used two founder genotypes, denoted A and B, to establish eight replicate
97 experimental populations (five from A and three from B, referred to as A populations and B
98 populations). Populations were cultivated via serial passages in the standard synthetic medium
99 M2 as described in Materials and Methods. After the short period of initial adaptation to the
100 conditions of the submerged cultivation, all the populations displayed identical phenotypic
101 changes, including changes in colony morphology and pigmentation (15). In order to identify
102 genetic changes acquired by the populations during the course of the experiment, we sequenced
103 both the founder genotypes and each of the populations at four time points (Fig 1).

104 The summary of the observed allele replacements is provided in Table 1 and their full list
105 is presented in Table S1. Our set of allele replacements comprises 85 (73% transitions) and 40
106 (53% transitions) nucleotide substitutions and 13 and 10 indels in A and B populations,
107 respectively. Among the 148 identified allele replacements, 52 result in an amino acid
108 substitution, 11 are single-nucleotide substitutions leading to a premature inframe stop codon,
109 and 13 cause a frameshift, comprising 35.1%, 7.4% and 8.8% of all allele replacements,
110 respectively.

111 **Table 1. Summary of allele replacements that occurred in the experimental populations.**

Type of a replacement	Experimental populations A	Experimental populations B
Total	98	50
By mutation type (single nucleotide mutations):		
Transition	62	21
Transversion	23	19
By mutation effect:		
Missense nucleotide substitutions	31	21
Synonymous nucleotide substitutions	6	7
Nonsense nucleotide substitutions	8	3
Nucleotide substitutions within noncoding regions	26	5
Nucleotide substitutions within genome regions that cannot be aligned with the reference genome	9	4
Nucleotide substitutions within introns	3	0

Nucleotide substitutions in incomplete proteins	2	0
Short insertions/deletions (including frameshifts)	13 (9)	10 (4)

The dynamics of accumulation of allele replacements is summarized in Fig 2 and Table S2. There are no obvious deviations from their uniform rate in the course of the experiments (linear regression's time variable p-value is 0.0613). Furthermore, there were no mutators: the numbers of the mutations in each population were all drawn from the same barrel and there were no difference between A and B populations (two-way ANOVA test, see Table S2).

Distribution of allele replacements within the genome

We then focused on the analysis of single nucleotide substitutions and first computed dN/dS ratios for populations A and B. dN/dS ratio equals 1.77 and 1.11 for populations A and B, respectively (Table S3). While dN/dS > 1 is commonly thought to indicate positive selection, our estimates may be unreliable because of the scarcity of the data. We next investigated how allele replacements are distributed in the *P. anserina* genome. We produced null-distributions for the number of missense and nonsense substitutions, using permutation procedure (see Materials and Methods). Null-distributions were obtained separately for A and B populations, assuming that mutations strike the genome and become fixed at random, with the uniform per site rates. There are significantly more nonsense substitutions than expected by chance in both A and B populations, and significantly more missense substitutions in B populations (Fig 3).

Moreover, despite the fact that there were very few mutations overall, a substantial proportion of protein-altering allele replacements affected the same gene and protein in more

132 than one population. Totally, six proteins were subject to such parallel evolution (Table 2). The
133 probability of observing by chance the same or a higher number of genes carrying just two
134 single-nucleotide protein-altering substitutions by chance is very low (Fig 4), and four out of six
135 proteins that underwent parallel evolution accepted more than two allele replacements. Clearly,
136 positive natural selection was responsible for this phenomenon.
137
138
139 **Table 2. Proteins affected independently in more than one population.**

Protein identifier and its function	The number of independent allele replacements
Pa_1_10140, Putative transcriptional regulatory protein pro-1	2 nonsense substitutions
Pa_1_23950, Guanine nucleotide-binding protein alpha subunit encoded by the mod-D gene	1 nonsense substitutions, 2 missense mutations
Pa_3_400, putative protein with unknown function	6 frameshifts
Pa_3_6550, Putative protein homologous to VeA of *Aspergillus nidulans*	2 missense substitutions
Pa_6_3340, Putative developmental regulator flbA	2 nonsense substitutions, 1 missense substitutions
Pa_7_7970, Putative guanine nucleotide-binding protein alpha-1 subunit	7 missense substitutions

140

141 This parallelism cannot be an artefact of ancestral polymorphism segregating within our
142 lineages, since we conservatively excluded a number of positions with identical parallel
143 mutations in multiple populations founded from the same ancestral lineage. For the same reason,
144 we could lose some actual functional parallelism, making our estimates for the frequency of
145 parallel substitutions conservative. We did observe two identical nucleotide mutations between
146 two lineages formed from two different ancestral lineages (Table S1); those are unlikely to be
147 due to ancestral polymorphism.

148

149 **Proteins carrying missense and nonsense substitutions are more conservative**

150 To study the rate of long-term evolution of proteins that evolved in our experiment, we
151 searched for pairs of orthologs pairs encoded by *P. anserina* and *Neurospora crassa* genomes.
152 Out of 10230 proteins in the reference *P. anserina* genome, 4713 have an ortholog in *N. crassa*
153 genome, such that their alignments contain no more than 20% of gaps. Both nonsense and, to a
154 lesser extent, missense allele replacements preferentially occurred in those *P. anserina* proteins
155 that have reliable orthologs in *N. crassa* (Table 3).

156 Moreover, missense allele replacements within such proteins preferentially occurred at
157 amino acid sites conserved between the two species. While the overall proportion of such sites
158 for 4713 alignments of reliable orthologs is 62%, 87% (27 out of 31) of replaced amino acids
159 occupied conservative sites (p = 0.003).

160 **Table 3. Statistics of orthologization of *P. anserina* и *N. crassa* genomes.**

	Total # of proteins	# of proteins carrying	# of proteins carrying missense

		nonsense substitutions	substitutions
# of proteins in *P. anserina* genome	10230	9	44
# of such proteins that have a reliable ortholog in *N. crassa*	4713	9 (0.000940)	27 (0.048609)

Values specified in parentheses correspond to p-values of Fisher's exact test

Possible phenotypic effects of parallel allele replacements

Proteins that underwent allele replacements in more than one experimental population (Table 2) are of particular interest. Indeed, their evolution must have been adaptive (Fig 4) and, therefore, must have led to changes in the phenotype. Moreover, they probably affected the protein structures radically. Indeed, the differences in physicochemical properties (computed as Miyata's distance) for pairs of amino acids were higher for missense substitutions that occurred during our experiment than for those that occurred in the course of divergence between *P. anserina* and *N. crassa* (mean values 1.979811 vs. 1.472333, two-tail t-test p-value 0.002).

Let us discuss what these changes could be.

The most striking example of parallel evolution is protein Pa_7_7970 which experienced nucleotide substitutions in seven independent populations out of eight. Pa_7_7970 is a putative alpha-1 subunit of a guanine nucleotide-binding protein (Table 2). There are three classes of G protein alpha subunits in ascomycetes fungi (16), each affecting fungal growth and development differently. Pa_7_7970 belongs to class 1 and is 93% identical to its well-studied orthologous protein FadA of *Aspergillus nidulans*. FadA stimulates vegetative growth (17) and probably suppresses sexual reproduction (18), which agrees well with intensive vegetative growth during the experiment suggesting that constitutive activation of Pa_7_7970 may be adaptive under the

180 conditions of submerged cultivation.

181 Only four sites in Pa_7_7970 evolved in the course of the experiment: arg-178 was
182 substituted in three experimental populations, gly-183 was substituted twice, and ser-206 and
183 lys-210 were each substituted once. All these sites have already been studied. Arg-178 is a
184 catalytic residue crucial for GTP hydrolysis and alpha subunit inactivation (19) and its mutations
185 result in constitutive activation of alpha subunit (20). Gly-183 takes part in the interaction
186 between alpha subunit and regulator of G protein signalling (RGS) that activates GTP hydrolysis,
187 and Gly->Ser substitution (observed in one case) hinders this process drastically (21). Ser-206
188 and Lys-210 are also involved in the interaction with RGS (22,23) although no specific effects of
189 their mutations have yet been described.

190 Remarkably, we also observed one missense and two nonsense substitutions in the
191 ortholog of RGS (Pa_6_3340) which belongs to the same pathway as FadA and likely
192 contributes to *P. anserina* adaptation in similar fashion.

193 Another protein that is worth mentioning is a transcription factor pro-1 (Pa_1_10140)
194 which acquired two loss-of-function allele replacements. pro-1 interacts with promoters of many
195 developmental genes including *esdC* gene that stimulates sexual development and is suppressed
196 by FadA (22,23), and deletion of pro-1 results in sexual sterility. Since *P. anserina* proliferates
197 only vegetatively during the experiment, defects in sexual development pathway may be
198 beneficial.

199 We observed two missense substitutions in VeA homolog (Pa_3_6550). Although we
200 cannot suggest the effect of two missense substitutions on protein function, VeA in *A. nidulans*
201 positively regulates expression of the previously mentioned *esdC* gene (18) and might affect *P.*
202 *anserina* metabolism in some way.

203 We also observed one protein with a surprising amount of parallel evolution via

206 frameshifts. Pa_3_400 protein experienced frameshift in six independent populations. Its
207 homologs in *N. crassa* (24) and *A. nidulans* are involved in asexual reproduction (25). As
208 *P. anserina* lacks asexual reproduction, we cannot suggest the role of Pa_3_400 inactivation in
209 *P. anserina* adaptation, although in *P. anserina* it can be involved in other developmental
210 pathways.

211 As for Pa_1_23950, it belongs to class 3 alpha subunits of fungal G proteins. Mutations
212 in this protein cause defects in vegetative proliferation in *P. anserina* (26) making it difficult to
213 interpret the observed substitutions in this protein in the context of adaptation.

214 All proteins discussed above are involved in growth and development processes in
215 various fungi. Taken with the excess of coding substitutions and the number of protein changed
216 in parallel, it favors the possible adaptiveness of some of the observed allele replacements
217 toward the conditions of submerged cultivation. Yet mutant construction is needed to claim it
218 with certainty.

219

220 **Discussion**

221 In populations of *P. anserina*, transition to prolonged submerged cultivation inevitably
222 results in the same set of morphophysiological changes (14). These changes always occur early
223 in the course of an experiment (15) and therefore are most likely caused by modifications of
224 epigenetic patterns that alter gene expression. Yet at a longer timescale, one can also expect
225 *P. anserina* to adapt to the conditions that are far from its native environment through fixations
226 of newly acquired mutations.

227 We analysed small-scale changes that occurred in the *P. anserina* genome in the course
228 of 268 experimental passages. We observed and confirmed the appearance of 125 single-

229 nucleotide substitutions and 23 short indels. It is hard to tell which of these mutations are
230 advantageous for submerged cultivation. Experimental populations of *P. anserina* cannot
231 reproduce sexually under conditions of the experiment, and *P. anserina* genus itself lacks the
232 ability to form asexual reproductive structures, macroconidia. Thus, the propagation of our
233 experimental populations occurred exclusively through vegetative proliferation, so that they must
234 possess clonal population structure. Thus, hitch-hiking of neutral and even slightly deleterious
235 mutations that arose in high-fitness genotypes destined to fixation was likely common (27).
236 Nevertheless, two observations clearly indicate that positive selection drove many of the
237 observed allele replacements which, therefore, contributed to adaptation to the experimental
238 environment.

239 First, evolution of our eight experimental populations often occurred in parallel, with the
240 same gene and protein being affected by a mutation up to seven times (Table 2). Parallel
241 evolution of phenotypes is common in nature, although it does not necessarily involve parallel
242 genetic changes (Steiner et al. 2009). In contrast, in evolutionary experiments parallel genetic
243 changes have being reported several times (28–31). Parallelisms at nucleotide level are rare, and
244 we observed only two such events in Pa_7_7970 gene, which might be mutational hotspots but
245 considering the importance of the mutated protein sites clearly indicate their adaptive role.

246 Some of the six proteins that underwent parallel changes in our experiment are involved
247 in the same pathways, and all of them seem to be associated with growth and development
248 processes in fungi. A prominent role of nonsense and frameshift alleles, as well as the nature of
249 the fixed missense alleles, indicate that this parallel evolution mostly, if not exclusively, led to
250 inactivation of the affected proteins. At least five missense substitutions in Pa_7_7970, as well as
251 nonsense substitutions in Pa_6_3340, alter developmental pathways probably stimulating
252 vegetative growth and inhibiting sexual reproduction at a very early stage. A similar pattern was

253 observed in experimental yeast populations, where parallel evolution affected genes of the
254 mating pathway and small GTPases of the Ras protein family (29). Loss-of-function mutations of
255 large effects can provide a selective advantage inactivating parts of regulatory/metabolic network
256 that became useless under the current environmental conditions (32). For example, in yeasts
257 mutations resulting in sexual sterility provide advantage in growth rate (33).

258 Still, the extent of parallel evolution in our experiment is remarkable. Fungi tend to not
259 differentiate in liquid medium (34), so that sexual reproduction was already suppressed before
260 experimental evolution ever begun. Thus, here the situation is very different from parallel loss of
261 eyes in many cave populations of *Astyanax mexicana*, where the cost of maintaining even a
262 useless capacity to see was estimated as up to 15% of resting metabolism (35), which makes its
263 elimination obviously beneficial. In contrast, we observed that positive selection strongly
264 favored mutations that damage a top-level metabolic pathway which already lacks a salient
265 phenotypic expression. Perhaps, before loss-of-function mutations in regulatory proteins occur,
266 some cryptic activity of this pathway remains, and fixation of such mutations allows cells to
267 reallocate their resources to vegetative proliferation. Thus, the role of positive selection in
268 degenerative evolution may be even more prominent than thought previously.

269 The second reason to believe that many of the observed allele replacements were
270 adaptive is overrepresentation among them of nonsense and missense mutations which must have
271 disproportionally strong impact on function (Fig 3). Moreover, among missense mutations, those
272 that involve radical amino acid replacements and those that occurred at conservative sites and in
273 more conservative proteins are also more common than expected. Because such mutations have a
274 reduced chance of being effectively neutral, this pattern implies that positive selection drove
275 most of their fixations.

276 Surprisingly, there is still no clarity regarding the role of functionally radical mutations in

277 the adaptive evolution in nature. The radical-conservative ratio of missense substitutions has
278 been studied in the context of both positive and negative selection, and while there are some
279 hints that it correlates with the strength of negative selection (36–38), its relation to positive
280 selection remains ambiguous. Radical-conservative ratio has been proposed as a measure to
281 detect positive selection almost three decades ago (39). However, this measure is vulnerable to
282 various factors not related to selection (40) and studies of different species did not find enough
283 evidences for the excess of radical compared to conservative substitutions at positively selected
284 sites (36,41,42).

285 A prominent role of mutations at conservative sites and proteins in adaptation to the
286 experimental environment was reported for *E. coli* (43). An obvious possibility is that these, as
287 well as ours, data reflect how adaptation occurs in nature. Indeed, substitutions at conservative
288 sites that occur in the course of natural evolution are enriched with those driven by positive
289 selection (44). Alternatively, a disproportional contribution of radical mutations to experimental
290 evolution may also be due to simplified laboratory environments, under which adaptation mostly
291 involves shedding functions that are no longer needed. Clearly, this issue deserves to be studied
292 further.

293 At the moment, there is no obvious trend in mutation accumulation rate. Experimental
294 populations acquired two times less mutations between time points 3 and 4 than they did in the
295 previous time interval. However, the reason for that is yet unclear, and further observations
296 might shed some light on this tendency. The experiment lasts for six years already, and to our
297 knowledge, there are not much long-lasting evolution experiments conducted on multicellular
298 organisms. We are planning to continue the experiment and analyse changes in experimental
299 populations of *P. anserina*.

300

301 # Material and Methods

302 **Evolutionary experiment**

303 Two homokaryotic vegetatively incompatible wild strains *s* and *S* (denoted here as A and
304 B) were kindly provided by Annie Sainsard-Chanet and Carole H. Sellem (Département Biologie
305 Cellulaire et Intégrative, Centre de Génétique Moléculaire, CNRS, Gif-sur-Yvette Cedex,
306 France) and gave rise to five and three independent experimental populations, respectively (Fig
307 1). Experimental populations were maintained via serial passages and were grown in the dark on
308 the standard synthetic medium M2 (15). Submerged cultivation was carried out in 750ml
309 Erlenmeyer flasks with 100 ml of the medium at 27 ± 1°C on a 200rpm rotor shaker. Cultures
310 were transferred to fresh medium every 4-7 days. During the experiment, mycelium samples
311 were plated on agarized (20 g/l) M2 medium in Petri dishes. Initial founder strains are stored on
312 ararized medium.

313 We denoted 64, 130, 200, and 268 passages as the first, the second, the third, and the
314 fourth time points of the experiment, respectively. Whole-genome sequencing was performed for
315 both the founder populations, for populations B at the points 1, 2, and 4, and for populations A
316 only at the point 4. The allele replacements that were observed from whole genome sequencing
317 data were subsequently verified by Sanger sequencing (see below). Several samples of passage
318 64 were lost due to contamination; for that reason, we substituted passage 64 with the closest
319 available passage 75 and refer to passage 75 as the first time point in the Results section. The
320 scheme of the experiment is shown in Fig 1.

321

322 **Whole-genome sequencing**

323 DNA was extracted using CTAB-based method (45). For library preparation for the first

324 series of samples (founder genotype B and experimental populations B at the first time point), we
325 used Truseq DNA sample preparation kit (Illumina), for the second series of samples
326 (experimental populations B at the second time point) we used NEBNext Ultra DNA kit (New
327 England Biolabs) according to manufacturer instructions. For the third series (founder genotype
328 A and all eight experimental populations at the fourth time point), we constructed PCRfree
329 libraries using Accel-NGS 2S PCR-Free DNA Library Kit (Swift Biosciences). The first and
330 second series were sequenced on HiSeq2000 (Illumina) using Truseq v.3 reagents. The third
331 series was sequenced using Nextseq500 and v.2 reagents.
332
333 **De novo genome assembly**
334 Pair-end reads from HiSeq (100bp) and NextSeq (150bp) were trimmed using
335 Trimmomatic (46) with options (ILLUMINACLIP:adapters:2:30:10 LEADING:7 TRAILING:7
336 SLIDINGWINDOW:5:10 MINLEN:50). Duplicate reads were excluded using fastq-mcf (with -
337 D 50 option).
338 We performed *de novo* genome assembly of both founder genotypes with SPAdes (47)
339 (with -only-assembler option). The initial assemblies were then filtered using the published
340 reference genome (48), v. 6.31, last accessed ~ Dec 2014). We used blastn (49) to find the best
341 hits for each contig in published genome and kept only those contigs that had a local alignment
342 of at least 5kb with published genome with identity of at least 95%. Finally, we performed
343 correction of assemblies using ICORN2 (50). Statistics on assemblies is provided in Table S4.
344
345 **Variant calling**
346 For each of the sequenced samples, we mapped its pair-end reads onto the corresponding
347 founder genome using bwa (51) and removed reads that had non-unique alignments. Filtered

348 mappings were used for variant calling. We used samtools (52) to call SNPs and InDels. The
349 resulting calls were filtered by coverage (>=10) and variant frequency (>=80%) to produce high-
350 quality variant set for each of the experimental populations. We then excluded non-unique
351 variants that were supported by at least two reads in more than one experimental population due
352 to extremely low probability of mutations affecting the same base pair in several independent
353 populations. We further masked low complexity regions of founder genomes by repeatmasker
354 (53) and excluded indels that fell into the masked sites and adjacent regions ten base pairs around
355 them. In populations A4 and B2, we observed a couple of mutations in adjacent genomic sites;
356 we excluded these variants as probable dinucleotide mutations.
357
358 **Sanger sequencing**
359 Sanger sequencing was used to check the presence of the called variants (138 SNPs and
360 24 indels) at each of the four time points. Several samples of passage 64 were lost due to
361 *Penicillium* contamination, so we used the closest available passage 75 as the first time point
362 passage. As each of the mutations observed at earlier time points in B populations were always
363 present in later populations, we considered a mutation as verified if it was confirmed by Sanger
364 sequencing in any of four time points (assuming that the mutation that appeared once became
365 fixed at later time points). For B populations, we believed in NGS data for passages 64 and 130
366 if the mutation was identified by Sanger sequencing at any time point. For A populations, we had
367 to check each of the four time points.
368 Primers for Sanger sequencing were designed using primer3 (54,55) and tntblast (56)
369 software and synthesized in Evrogen. PCR was performed using Encyclo kit (Evrogen).
370 Amplification products were sequenced in both directions on ABI Prism 3500 sequencer
371 (Applied Biosystems). Chromatograms were validated manually in CodonCode Aligner (57).

372 The detailed results of Sanger sequencing are included in Table S1. The number of
373 confirmed mutations is provided in Table S2. We were unable to confirm neither the presence
374 nor the absence of 15 mutations and excluded them from any further analyses. The majority of
375 mutations that we failed to validate comprise mutations in noncoding/not annotated regions.
376 Also, there were three mutations in A populations for which we couldn't track the time of their
377 appearance.
378
379 **Variant annotation**
380 We mapped reference CDS sequences (48) (v. 6.31, last accessed ~ Dec 2014) onto the
381 reference genome and obtained coding regions coordinates using Splign (58). Out of 10,635
382 reference protein-coding genes, we were able to identify 10230 genes in the reference genome
383 that coded complete proteins with proper start and stop codons. We then performed pairwise
384 genome alignment of A and B founder genotype assemblies with published genome using Multiz
385 (59) and identified coding regions coordinates in the assembled genomes using Splign annotation
386 results. 9732 and 9732 (numbers are indeed the same but protein sets differ) complete genes
387 were identified in A and B founder genotypes, respectively, with a tiny amount of incomplete
388 genes (627 and 530 genes for A and B populations, respectively). Based on the produced
389 annotations, the identified genomic variants were classified as exonic, intronic or intergenic.
390 Exonic variants in complete genes (i.e., genes with all of their exons completely identified in the
391 founder genome) were further classified as missense, nonsense or silent mutations in case of
392 SNPs and frameshift or amino acid insertion/deletion mutations in case of indels.
393
394 **Permutations**
395 In order to estimate the probability to obtain the observed number of protein-altering

396 variants by chance, we generated null distributions for the number of missense and nonsense
397 mutations in A and B founder genomes using permutations. A permutation procedure was
398 performed as follows. 85 and 40 observed SNP mutations were thrown at random onto A and B
399 genomes, respectively, and annotated as described above. To control for possible confounding
400 effects of mutation context, we added three-nucleotide sequence contexts to permutations. The
401 procedure was repeated 10,000 times and produced null distributions for the expected number of
402 nonsense and missense mutations in each of the founder genomes (Fig 3). To obtain the expected
403 number of proteins evolving in parallel, we combined permutation results from A and B
404 populations and counted the number of proteins that acquired at least two single-nucleotide
405 protein-changing mutations in eight populations during the permutation procedure.
406
407 **Protein conservation analysis**
408 We used OrthoMCL (60) to identify orthologous groups between protein sets from
409 *P. anserina* and *N. crassa* (61) reference genomes. Orthologous protein pairs were then aligned
410 using MAFFT (62) (with -localpair option), and for each protein carrying ortholog in *N. crassa*
411 genome, we calculated its conservation as a fraction of conserved positions in the alignment. We
412 also measured the conservation of mutated positions as a fraction of conserved positions among
413 the whole set of mutated positions.
414
415 **References**

416 1. Lenski RE, Travisano M. Dynamics of adaptation and diversification: a 10,000-generation
417 experiment with bacterial populations. Proc Natl Acad Sci U S A. 1994 Jul 19;91(15):6808–14.
418 2. Philippe N, Pelosi L, Lenski RE, Schneider D. Evolution of penicillin-binding protein 2

419 concentration and cell shape during a long-term experiment with Escherichia coli. J Bacteriol. 2009
420 Feb;191(3):909–21.
421 3. Lenski RE, Wiser MJ, Ribeck N, Blount ZD, Nahum JR, Morris JJ, et al. Sustained fitness gains and
422 variability in fitness trajectories in the long-term evolution experiment with Escherichia coli
423 [Internet]. 2015. Available from: http://dx.doi.org/10.1101/027391
424 4. Blount ZD, Barrick JE, Davidson CJ, Lenski RE. Genomic analysis of a key innovation in an
425 experimental Escherichia coli population. Nature. 2012;489(7417):513–8.
426 5. Leiby N, Marx CJ. Metabolic Erosion Primarily Through Mutation Accumulation, and Not
427 Tradeoffs, Drives Limited Evolution of Substrate Specificity in Escherichia coli. PLoS Biol.
428 2014;12(2):e1001789.
429 6. Lenski RE, Travisano M. Dynamics of adaptation and diversification: a 10,000-generation
430 experiment with bacterial populations. Proc Natl Acad Sci U S A. 1994 Jul 19;91(15):6808–14.
431 7. Woods R, Schneider D, Winkworth CL, Riley MA, Lenski RE. Tests of parallel molecular evolution
432 in a long-term experiment with Escherichia coli. Proceedings of the National Academy of Sciences.
433 2006;103(24):9107–12.
434 8. Cooper TF, Remold SK, Lenski RE, Schneider D. Expression profiles reveal parallel evolution of
435 epistatic interactions involving the CRP regulon in Escherichia coli. PLoS Genet. 2008 Feb;4(2):e35.
436 9. Gray JC, Cutter AD. Mainstreaming Caenorhabditis elegans in experimental evolution. Proc Biol
437 Sci. 2014 Mar 7;281(1778):20133055.
438 10. McDonald MJ, Rice DP, Desai MM. Sex speeds adaptation by altering the dynamics of molecular
439 evolution. Nature. 2016 Mar 10;531(7593):233–6.
440 11. Rundle HD, Chenoweth SF, Doughty P, Blows MW. Divergent selection and the evolution of signal
441 traits and mating preferences. PLoS Biol. 2005 Nov;3(11):e368.
442 12. Dettman JR, Anderson JB, Kohn LM. Divergent adaptation promotes reproductive isolation among
443 experimental populations of the filamentous fungus Neurospora. BMC Evol Biol. 2008 Jan 31;8:35.
444 13. Lorin S, Dufour E, Sainsard-Chanet A. Mitochondrial metabolism and aging in the filamentous

445 fungus Podospora anserina. Biochimica et Biophysica Acta (BBA) - Bioenergetics. 2006;1757(5-
446 6):604–10.
447 14. Turker MS, Cummings DJ. Podospora anserina does not senesce when serially passaged in liquid
448 culture. J Bacteriol. 1987 Feb;169(2):454–60.
449 15. Kudryavtseva OA, Mazheika IS, Solovchenko AE, Kamzolkina OV. Genetic instability of the short-
450 living ascomycetous fungus Podospora anserina induced by prolonged submerged cultivation.
451 Microbiology. 2011;80(6):784–96.
452 16. Bölker M. Sex and crime: heterotrimeric G proteins in fungal mating and pathogenesis. Fungal
453 Genet Biol. 1998 Dec;25(3):143–56.
454 17. Krijgsheld P, Bleichrodt R, van Veluw GJ, Wang F, Müller WH, Dijksterhuis J, et al. Development
455 in Aspergillus. Stud Mycol. 2013 Mar 15;74(1):1–29.
456 18. Han K-H, Kim JH, Moon H, Kim S, Lee S-S, Han D-M, et al. The Aspergillus nidulans esdC (early
457 sexual development) gene is necessary for sexual development and is controlled by veA and a
458 heterotrimeric G protein. Fungal Genet Biol. 2008;45(3):310–8.
459 19. Coleman DE, Berghuis AM, Lee E, Linder ME, Gilman AG, Sprang SR. Structures of active
460 conformations of Gi alpha 1 and the mechanism of GTP hydrolysis. Science. 1994 Sep
461 2;265(5177):1405–12.
462 20. Landis CA, Masters SB, Spada A, Pace AM, Bourne HR, Vallar L. GTPase inhibiting mutations
463 activate the α chain of Gs and stimulate adenylyl cyclase in human pituitary tumours. Nature.
464 1989;340(6236):692–6.
465 21. Lan KL, Sarvazyan NA, Taussig R, Mackenzie RG, DiBello PR, Dohlman HG, et al. A point
466 mutation in Galphao and Galphai1 blocks interaction with regulator of G protein signaling proteins. J
467 Biol Chem. 1998 May 22;273(21):12794–7.
468 22. Khafizov K, Lattanzi G, Carloni P. G protein inactive and active forms investigated by simulation
469 methods. Proteins. 2009 Jun;75(4):919–30.
470 23. Clark MA, Lambert NA. Endogenous regulator of G-protein signaling proteins regulate the kinetics

471 of Galphaq/11-mediated modulation of ion channels in central nervous system neurons. Mol
472 Pharmacol. 2006 Apr;69(4):1280–7.
473 24. Springer ML, Yanofsky C. A morphological and genetic analysis of conidiophore development in
474 Neurospora crassa. Genes Dev. 1989 Apr;3(4):559–71.
475 25. Chung D-W, Greenwald C, Upadhyay S, Ding S, Wilkinson HH, Ebbole DJ, et al. acon-3, the
476 Neurospora crassa ortholog of the developmental modifier, medA, complements the conidiation
477 defect of the Aspergillus nidulans mutant. Fungal Genet Biol. 2011 Apr;48(4):370–6.
478 26. Loubradou G, Bégueret J, Turcq B. MOD-D, a Galpha subunit of the fungus Podospora anserina, is
479 involved in both regulation of development and vegetative incompatibility. Genetics. 1999
480 Jun;152(2):519–28.
481 27. Barton NH. Genetic hitchhiking. Philos Trans R Soc Lond B Biol Sci. 2000 Nov
482 29;355(1403):1553–62.
483 28. Woods R, Schneider D, Winkworth CL, Riley MA, Lenski RE. Tests of parallel molecular evolution
484 in a long-term experiment with Escherichia coli. Proceedings of the National Academy of Sciences.
485 2006;103(24):9107–12.
486 29. Lang GI, Rice DP, Hickman MJ, Sodergren E, Weinstock GM, Botstein D, et al. Pervasive genetic
487 hitchhiking and clonal interference in forty evolving yeast populations. Nature. 2013 Aug
488 29;500(7464):571–4.
489 30. Wichman HA, Badgett MR, Scott LA, Boulianne CM, Bull JJ. Different trajectories of parallel
490 evolution during viral adaptation. Science. 1999 Jul 16;285(5426):422–4.
491 31. Lang GI, Desai MM. The spectrum of adaptive mutations in experimental evolution. Genomics.
492 2014 Dec;104(6 Pt A):412–6.
493 32. Hottes AK, Freddolino PL, Khare A, Donnell ZN, Liu JC, Tavazoie S. Bacterial Adaptation through
494 Loss of Function. PLoS Genet. 2013;9(7):e1003617.
495 33. Lang GI, Murray AW, Botstein D. The cost of gene expression underlies a fitness trade-off in yeast.
496 Proc Natl Acad Sci U S A. 2009 Apr 7;106(14):5755–60.

497 34. Osiewacz HD. Molecular Biology of Fungal Development. CRC Press; 2002. 608 p.
498 35. Moran D, Softley R, Warrant EJ. The energetic cost of vision and the evolution of eyeless Mexican
499 cavefish. Sci Adv. 2015 Sep;1(8):e1500363.
500 36. Smith NGC. Are radical and conservative substitution rates useful statistics in molecular evolution?
501 J Mol Evol. 2003 Oct;57(4):467–78.
502 37. Wernegreen JJ. Reduced selective constraint in endosymbionts: elevation in radical amino acid
503 replacements occurs genome-wide. PLoS One. 2011 Dec 14;6(12):e28905.
504 38. Popadin K, Polishchuk LV, Mamirova L, Knorre D, Gunbin K. Accumulation of slightly deleterious
505 mutations in mitochondrial protein-coding genes of large versus small mammals. Proc Natl Acad Sci
506 U S A. 2007 Aug 14;104(33):13390–5.
507 39. Hughes AL, Ota T, Nei M. Positive Darwinian selection promotes charge profile diversity in the
508 antigen-binding cleft of class I major-histocompatibility-complex molecules. Mol Biol Evol. 1990
509 Nov;7(6):515–24.
510 40. Dagan T, Talmor Y, Graur D. Ratios of Radical to Conservative Amino Acid Replacement are
511 Affected by Mutational and Compositional Factors and May Not Be Indicative of Positive
512 Darwinian Selection. Mol Biol Evol. 2002;19(7):1022–5.
513 41. Hanada K, Gojobori T, Li W-H. Radical amino acid change versus positive selection in the evolution
514 of viral envelope proteins. Gene. 2006 Dec 30;385:83–8.
515 42. Suzuki Y. Inferring natural selection operating on conservative and radical substitution at single
516 amino acid sites. Genes Genet Syst. 2007 Aug;82(4):341–60.
517 43. Maddamsetti R, Hatcher PJ, Green AG, Williams BL, Marks DS, Lenski RE. Core Genes Evolve
518 Rapidly in the Long-term Evolution Experiment with Escherichia coli. Genome Biol Evol [Internet].
519 2017 Apr 4; Available from: http://dx.doi.org/10.1093/gbe/evx064
520 44. Bazykin GA, Kondrashov AS. Major role of positive selection in the evolution of conservative
521 segments of Drosophila proteins. Proc Biol Sci. 2012 Sep 7;279(1742):3409–17.
522 45. Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue.

523 Phytochemical Bulletin. 1987;19:11–5.
524 46. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data.
525 Bioinformatics. 2014 Aug 1;30(15):2114–20.
526 47. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new
527 genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012
528 May;19(5):455–77.
529 48. Podospora anserina Genome Project [Internet]. [cited 2018 Jun 16]. Available from:
530 http://podospora.i2bc.paris-saclay.fr/download.php
531 49. Altschul S. Basic Local Alignment Search Tool. J Mol Biol. 1990;215(3):403–10.
532 50. Otto TD, Sanders M, Berriman M, Newbold C. Iterative Correction of Reference Nucleotides
533 (iCORN) using second generation sequencing technology. Bioinformatics. 2010 Jul 15;26(14):1704–
534 7.
535 51. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform.
536 Bioinformatics. 2009 Jul 15;25(14):1754–60.
537 52. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map
538 format and SAMtools. Bioinformatics. 2009 Aug 15;25(16):2078–9.
539 53. RepeatMasker Home Page [Internet]. [cited 2018 Jun 16]. Available from:
540 http://www.repeatmasker.org
541 54. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3—new
542 capabilities and interfaces. Nucleic Acids Res. 2012;40(15):e115–e115.
543 55. Koressaar T, Remm M. Enhancements and modifications of primer design program Primer3.
544 Bioinformatics. 2007 May 15;23(10):1289–91.
545 56. Gans JD, Wolinsky M. Improved assay-dependent searching of nucleic acid sequence databases.
546 Nucleic Acids Res. 2008 Jul;36(12):e74.
547 57. Website [Internet]. [cited 2018 Jun 16]. Available from: www.codoncode.com
548 58. Kapustin Y, Souvorov A, Tatusova T, Lipman D. Splign: algorithms for computing spliced

549 alignments with identification of paralogs. Biol Direct. 2008;3(1):20.
550 59. Blanchette M, Kent WJ, Riemer C, Elnitski L, Smit AFA, Roskin KM, et al. Aligning multiple
551 genomic sequences with the threaded blockset aligner. Genome Res. 2004 Apr;14(4):708–15.
552 60. Li L. OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. Genome Res.
553 2003;13(9):2178–89.
554 61. Krumlau R, Marzluf GA. Characterization of the sequence complexity and organization of the
555 Neurospora crassa genome. Biochemistry. 1979;18(17):3705–13.
556 62. Katoh K. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier
557 transform. Nucleic Acids Res. 2002;30(14):3059–66.
558

559 ## Figures

560 **Fig 1. Scheme of the experiment.**

561 Eight experimental populations were established from two founder lineages A and B. After 75,
562 130, 200, and 268 passages, we determined genotypes of the experimental populations (dashed
563 lines; see Methods).

564 **Fig 2. Accumulation of allele replacements during the experiment.**

565 Each dot reflects a number of replacements observed in a population at a particular time point.
566 Populations are shown as thin lines. Thick blue line corresponds to the linear regression line with
567 confidence interval depicted as shaded area.

568 **Fig 3. The observed number of missense and nonsense substitutions is greater than**
569 **expected by chance.**

570 Distributions of the expected number of nonsense (A) and missense (B) substitutions in A (left
571 panel) and B (right panel) experimental populations in 10,000 permutations. Red line
572 corresponds to the observed number of nonsense or missense substitutions on (A) and (B),

573 respectively. *p* (p-value) is the area to the right of the red line.

574 **Fig 4. The excess of the number of proteins with parallel changes.**

575 Distribution of the number of genes with two protein-altering single-nucleotide substitutions in
576 10,000 random permutations. Red line shows the observed number of genes with two or more
577 such substitutions occurred.

578

579 Supplementary Information

580 **Table S1. The summary of the observed allele replacements.** Sheets A1-A5 and B1-B3
581 contain the results of the Sanger sequencing validation procedure. The two last sheets comprise
582 annotation of the allele replacements in A and B populations.

583 **Table S2. Accumulation of allele replacements.** Each entry provides the number of newly
584 observed replacements, relatively to the previous time point.

585 **Table S3. dN/dS calculation.**

586 **Table S4. Statistics on founder genotypes assemblies.**

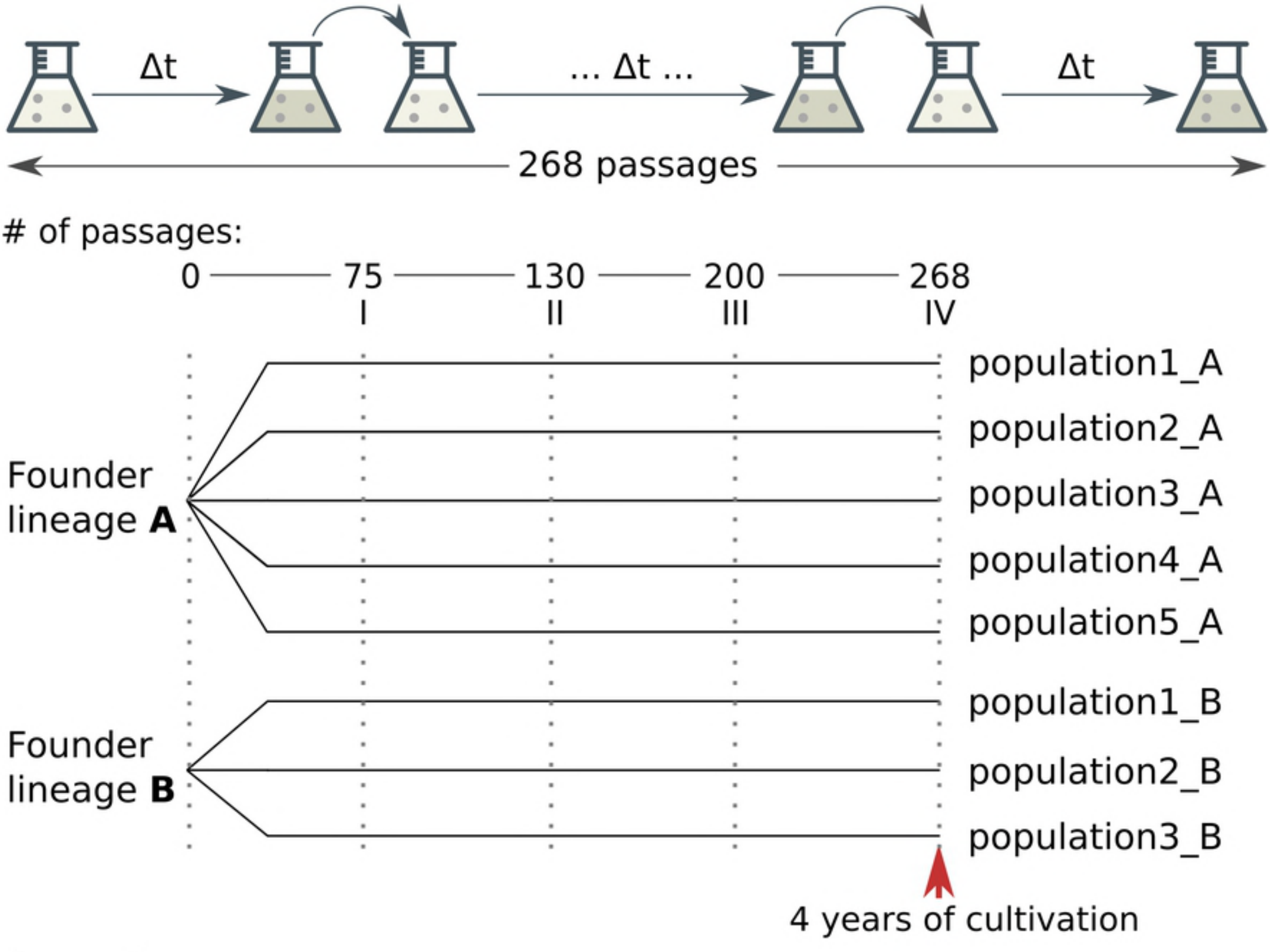
Δt
... Δt ...
Δt
268 passages
# of passages:
0
75
130
200
268
I
II
III
IV
Founder lineage A
population1_A
population2_A
population3_A
population4_A
population5_A
Founder lineage B
population1_B
population2_B
population3_B
4 years of cultivation

Figure 1

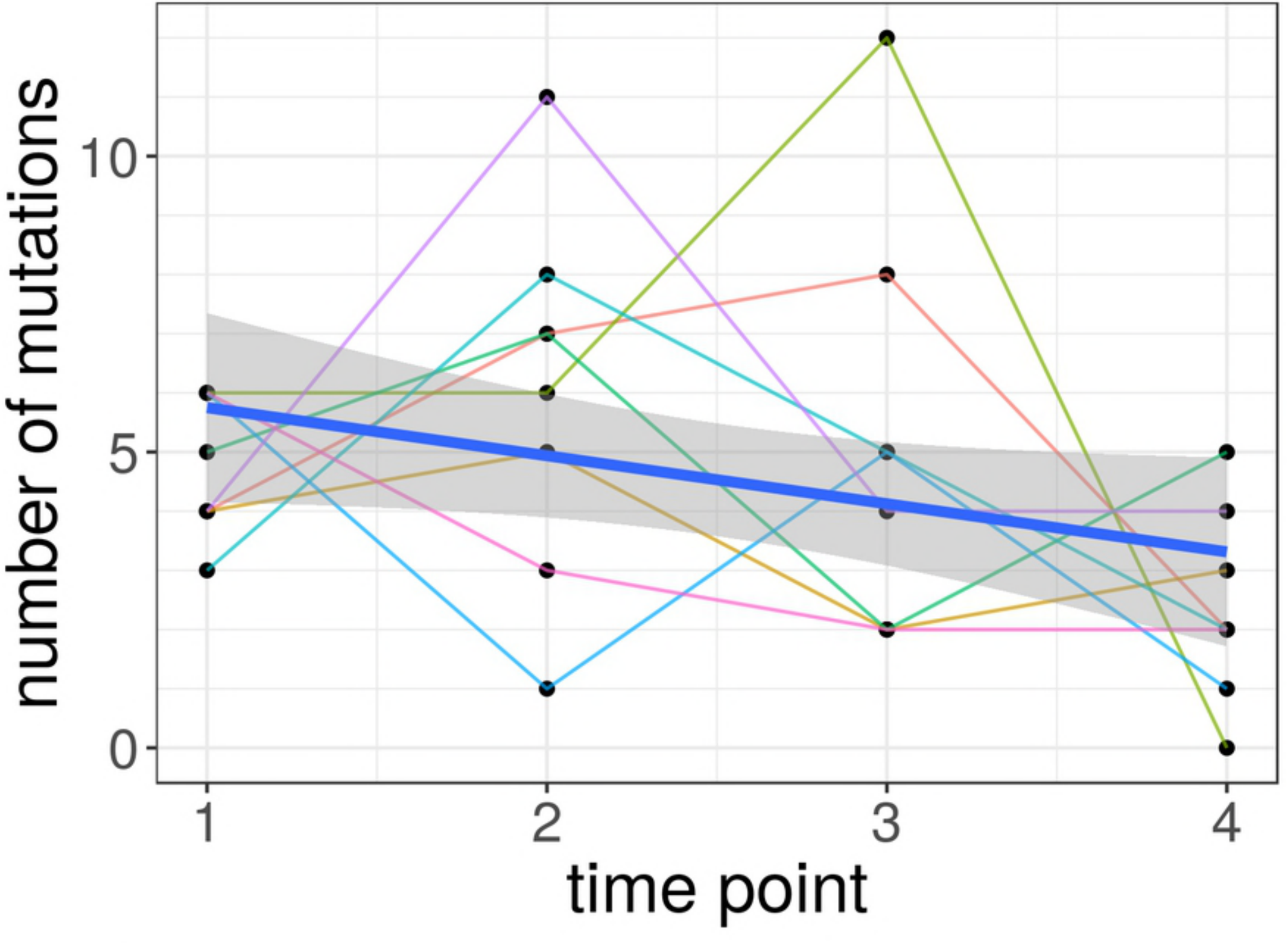

Figure 2

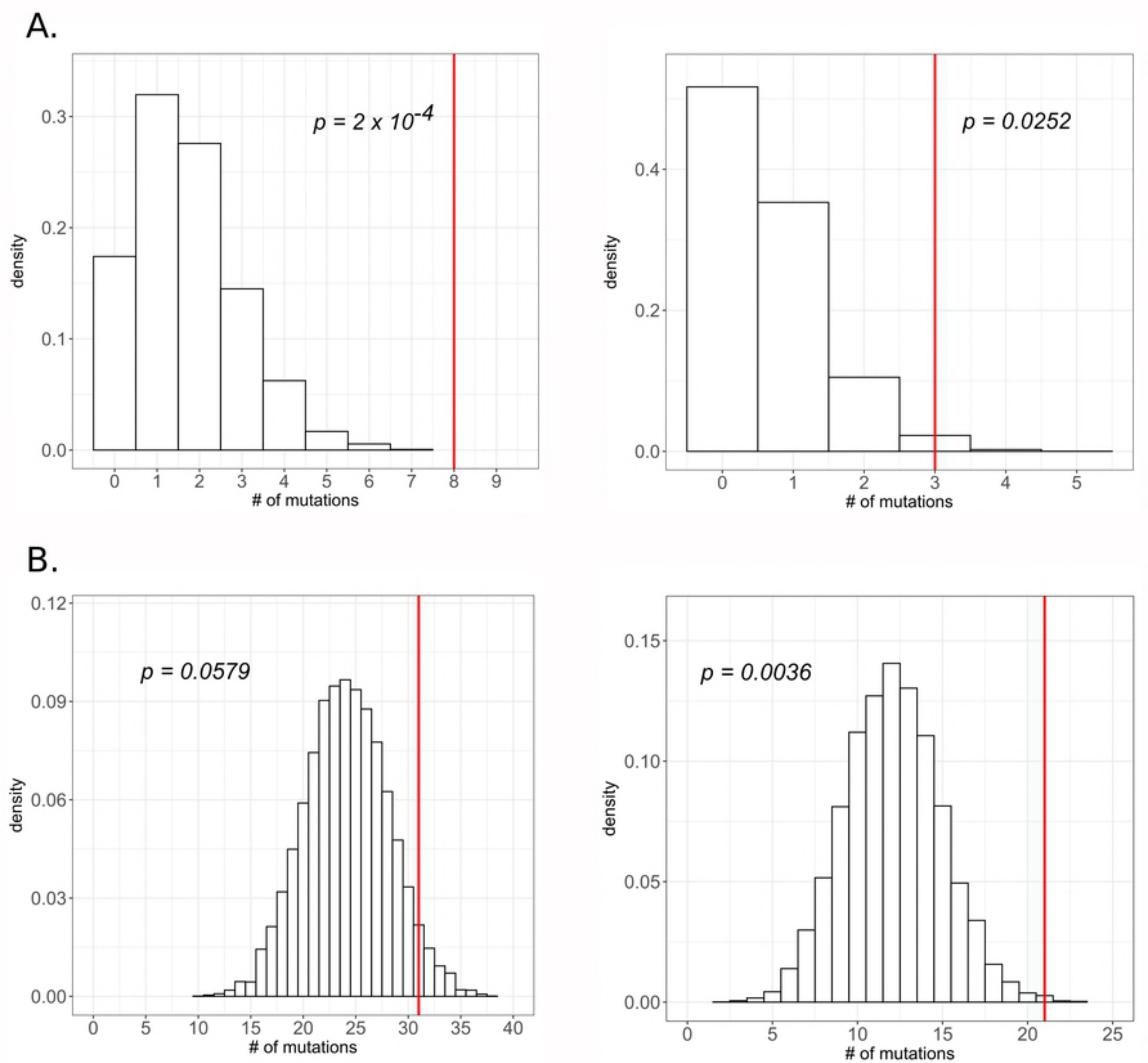
A.
p = 2 x 10^-4
density
# of mutations
p = 0.0252
density
# of mutations
B.
p = 0.0579
density
# of mutations
p = 0.0036
density
# of mutations

Figure 3

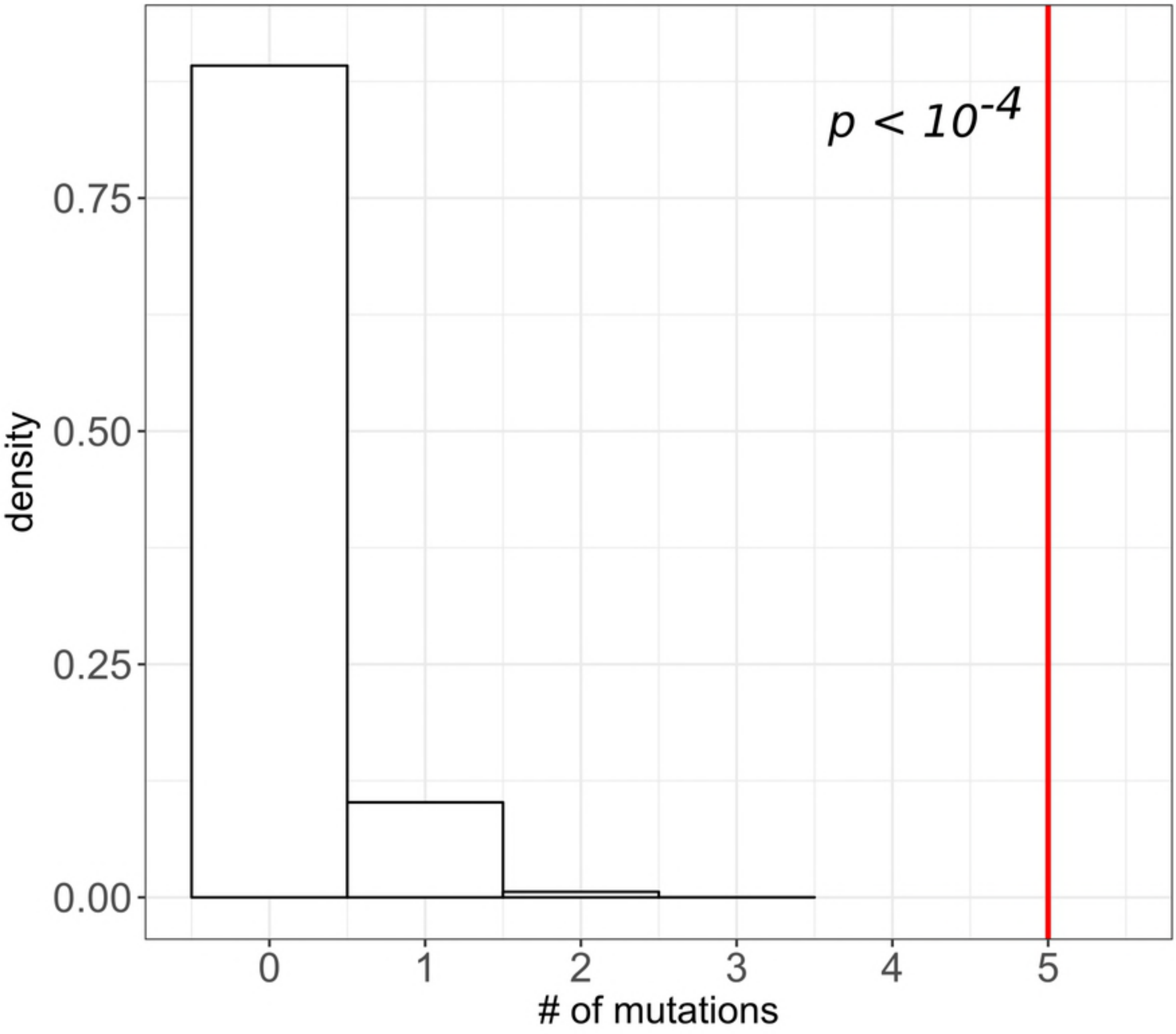

Figure 4