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.

Data Availability

Genome sequence data have been deposited at DDBJ/ENA/GenBank under accessions QHKV00000000 (founder genotype A; version QHKV01000000) and QHKU00000000 (founder genotype B; version QHKU01000000), with the respective BioSample accessions SAMN09270751 and SAMN09270757, under BioProject PRJNA473312. Sequencing data have been deposited at the SRA with accession numbers SRR7233712-SRR7233727, under the same BioProject.

Funding

Experimental work and sequencing were supported by the Russian Foundation for Basic Research (grants no. 16-04-01845a and 18-04-01349a). Bioinformatic analysis was supported by the Russian Science Foundation (grant no. 16-14-10173). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Introduction

Over the last several decades, experimental evolution became an essential part of

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

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

93 *Podospora anserina* is a model ascomycete fungus that possesses a definite lifespan with
94 pronounced phenotypic senescence when grown on solid medium (13). In contrast, under
95 conditions of submerged cultivation, *P. anserina* displays unlimited lifespan with no signs of
96 senescence (14,15). Isolates taken from such cultures revert to senescence, indicating that the
97 unlimited lifespan is not associated with genetic changes, instead being a product of epigenetic
98 changes or environmental influences. Still, such isolates demonstrate increased lifespan when
99 plated on solid medium and become sterile, only rarely forming mostly nonfunctional
100 microconidia and protoperithecia (14,15). After some period of submerged cultivation,
101 *P. anserina* cultures acquire higher rates of biomass accumulation and change their
102 pigmentation. The mechanisms that alter *P. anserina* senescence program and affect its
103 phenotype under such conditions are unknown.

104 In order to study the genetic aspects of adaptation that takes place during the submerged

105 cultivation of *P. anserina*, we started in 2012 a long-term evolutionary experiment. We used two
106 wild strains of *P. anserina* to establish eight independent experimental populations. Here, we
107 analyse the genetic changes that took place in the first four years of this experiment.
108

109 ## Results

110 **Allele replacements**

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

119 The summary of the observed allele replacements is provided in Table 1 and their full list
120 is presented in Table S1. Our set of allele replacements comprises 85 (73% transitions) and 40
121 (53% transitions) nucleotide substitutions and 13 and 10 indels in A and B populations,
122 respectively. Among the 148 identified allele replacements, 52 result in an amino acid
123 substitution, 11 are single-nucleotide substitutions leading to a premature inframe stop codon,
124 and 13 cause a frameshift, comprising 35.1%, 7.4% and 8.8% of all allele replacements,
125 respectively.

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

Type of a replacement	Experimental	Experimental

	populations A	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).

133

134 **Distribution of allele replacements within the genome**

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

145 Moreover, despite the fact that there were very few mutations overall, a substantial
146 proportion of protein-altering allele replacements affected the same gene and protein in more
147 than one population. Totally, six proteins were subject to such parallel evolution (Table 2). The
148 probability of observing by chance the same or a higher number of genes carrying just two
149 single-nucleotide protein-altering substitutions by chance is very low (Fig 4), and four out of six
150 proteins that underwent parallel evolution accepted more than two allele replacements. Clearly,
151 positive natural selection was responsible for this phenomenon.

152

153

154 **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

155

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

163

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

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

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

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

	Total # of proteins	# of proteins carrying nonsense substitutions	# of proteins carrying missense 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)

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

177

178 **Possible phenotypic effects of parallel allele replacements**

179 Proteins that underwent allele replacements in more than one experimental population
180 (Table 2) are of particular interest. Indeed, their evolution must have been adaptive (Fig 4) and,

181 therefore, must have led to changes in the phenotype. Moreover, they probably affected the
182 protein structures radically. Indeed, the differences in physicochemical properties (computed as
183 Miyata's distance) for pairs of amino acids were higher for missense substitutions that occurred
184 during our experiment than for those that occurred in the course of divergence between
185 *P. anserina* and *N. crassa* (mean values 1.979811 vs. 1.472333, two-tail t-test p-value 0.002).

186 Let us discuss what these changes could be.

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

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

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

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

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

218 We also observed one protein with a surprising amount of parallel evolution via
219 frameshifts. Pa_3_400 protein experienced frameshift in six independent populations. Its
220 homologs in *N. crassa* (24) and *A. nidulans* are involved in asexual reproduction (25). As
221 *P. anserina* lacks asexual reproduction, we cannot suggest the role of Pa_3_400 inactivation in
222 *P. anserina* adaptation, although in *P. anserina* it can be involved in other developmental
223 pathways.

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

227 All proteins discussed above are involved in growth and development processes in
228 various fungi. Taken with the excess of coding substitutions and the number of protein changed

229 in parallel, it favors the possible adaptiveness of some of the observed allele replacements
230 toward the conditions of submerged cultivation. Yet mutant construction is needed to claim it
231 with certainty.
232

233 **Discussion**

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

240 We analysed small-scale changes that occurred in the *P. anserina* genome in the course
241 of 268 experimental passages. We observed and confirmed the appearance of 125 single-
242 nucleotide substitutions and 23 short indels. It is hard to tell which of these mutations are
243 advantageous for submerged cultivation. Experimental populations of *P. anserina* cannot
244 reproduce sexually under conditions of the experiment, and *P. anserina* genus itself lacks the
245 ability to form asexual reproductive structures, macroconidia. Thus, the propagation of our
246 experimental populations occurred exclusively through vegetative proliferation, so that they must
247 possess clonal population structure. Thus, hitch-hiking of neutral and even slightly deleterious
248 mutations that arose in high-fitness genotypes destined to fixation was likely common (27).
249 Nevertheless, two observations clearly indicate that positive selection drove many of the
250 observed allele replacements which, therefore, contributed to adaptation to the experimental
251 environment.

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

259 Some of the six proteins that underwent parallel changes in our experiment are involved
260 in the same pathways, and all of them seem to be associated with growth and development
261 processes in fungi. A prominent role of nonsense and frameshift alleles, as well as the nature of
262 the fixed missense alleles, indicate that this parallel evolution mostly, if not exclusively, led to
263 inactivation of the affected proteins. At least five missense substitutions in Pa_7_7970, as well as
264 nonsense substitutions in Pa_6_3340, alter developmental pathways probably stimulating
265 vegetative growth and inhibiting sexual reproduction at a very early stage. A similar pattern was
266 observed in experimental yeast populations, where parallel evolution affected genes of the
267 mating pathway and small GTPases of the Ras protein family (29). Loss-of-function mutations of
268 large effects can provide a selective advantage inactivating parts of regulatory/metabolic network
269 that became useless under the current environmental conditions (32). For example, in yeasts
270 mutations resulting in sexual sterility provide advantage in growth rate (33).

271 Still, the extent of parallel evolution in our experiment is remarkable. Fungi tend to not
272 differentiate in liquid medium (34), so that sexual reproduction was already suppressed before
273 experimental evolution ever begun. Thus, here the situation is very different from parallel loss of
274 eyes in many cave populations of *Astyanax mexicana*, where the cost of maintaining even a
275 useless capacity to see was estimated as up to 15% of resting metabolism (35), which makes its

272 elimination obviously beneficial. In contrast, we observed that positive selection strongly
273 favored mutations that damage a top-level metabolic pathway which already lacks a salient
274 phenotypic expression. Perhaps, before loss-of-function mutations in regulatory proteins occur,
275 some cryptic activity of this pathway remains, and fixation of such mutations allows cells to
276 reallocate their resources to vegetative proliferation. Thus, the role of positive selection in
277 degenerative evolution may be even more prominent than thought previously.

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

285 Surprisingly, there is still no clarity regarding the role of functionally radical mutations in
286 the adaptive evolution in nature. The radical-conservative ratio of missense substitutions has
287 been studied in the context of both positive and negative selection, and while there are some
288 hints that it correlates with the strength of negative selection (36–38), its relation to positive
289 selection remains ambiguous. Radical-conservative ratio has been proposed as a measure to
290 detect positive selection almost three decades ago (39). However, this measure is vulnerable to
291 various factors not related to selection (40) and studies of different species did not find enough
292 evidences for the excess of radical compared to conservative substitutions at positively selected
293 sites (36,41,42).

294 A prominent role of mutations at conservative sites and proteins in adaptation to the
295 experimental environment was reported for *E. coli* (43). An obvious possibility is that these, as

296 well as ours, data reflect how adaptation occurs in nature. Indeed, substitutions at conservative
297 sites that occur in the course of natural evolution are enriched with those driven by positive
298 selection (44). Alternatively, a disproportional contribution of radical mutations to experimental
299 evolution may also be due to simplified laboratory environments, under which adaptation mostly
300 involves shedding functions that are no longer needed. Clearly, this issue deserves to be studied
301 further.

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

309

310 Material and Methods

311 Evolutionary experiment

312 Two homokaryotic vegetatively incompatible wild strains *s* and *S* (denoted here as A and
313 B) were kindly provided by Annie Sainsard-Chanet and Carole H. Sellem (Département Biologie
314 Cellulaire et Intégrative, Centre de Génétique Moléculaire, CNRS, Gif-sur-Yvette Cedex,
315 France) and gave rise to five and three independent experimental populations, respectively (Fig
316 1). Experimental populations were maintained via serial passages and were grown in the dark on
317 the standard synthetic medium M2 (15). Submerged cultivation was carried out in 750ml
318 Erlenmeyer flasks with 100 ml of the medium at 27 ± 1°C on a 200rpm rotor shaker. Cultures

319 were transferred to fresh medium every 4-7 days. During the experiment, mycelium samples
320 were plated on agarized (20 g/l) M2 medium in Petri dishes. Initial founder strains are stored on
321 ararized medium.

322 We denoted 64, 130, 200, and 268 passages as the first, the second, the third, and the
323 fourth time points of the experiment, respectively. Whole-genome sequencing was performed for
324 both the founder populations, for populations B at the points 1, 2, and 4, and for populations A
325 only at the point 4. The allele replacements that were observed from whole genome sequencing
326 data were subsequently verified by Sanger sequencing (see below). Several samples of passage
327 64 were lost due to contamination; for that reason, we substituted passage 64 with the closest
328 available passage 75 and refer to passage 75 as the first time point in the Results section. The
329 scheme of the experiment is shown in Fig 1.

330

331 **Whole-genome sequencing**

332 DNA was extracted using CTAB-based method (45). For library preparation for the first
333 series of samples (founder genotype B and experimental populations B at the first time point), we
334 used Truseq DNA sample preparation kit (Illumina), for the second series of samples
335 (experimental populations B at the second time point) we used NEBNext Ultra DNA kit (New
336 England Biolabs) according to manufacturer instructions. For the third series (founder genotype
337 A and all eight experimental populations at the fourth time point), we constructed PCRfree
338 libraries using Accel-NGS 2S PCR-Free DNA Library Kit (Swift Biosciences). The first and
339 second series were sequenced on HiSeq2000 (Illumina) using Truseq v.3 reagents. The third
340 series was sequenced using Nextseq500 and v.2 reagents.

341

342 **De novo genome assembly**

343 Pair-end reads from HiSeq (100bp) and NextSeq (150bp) were trimmed using
344 Trimmomatic (46) with options (ILLUMINACLIP:adapters:2:30:10 LEADING:7 TRAILING:7
345 SLIDINGWINDOW:5:10 MINLEN:50). Duplicate reads were excluded using fastq-mcf (with -
346 D 50 option).

347 We performed *de novo* genome assembly of both founder genotypes with SPAdes (47)
348 (with -only-assembler option). The initial assemblies were then filtered using the published
349 reference genome (48), v. 6.31, last accessed ~ Dec 2014). We used blastn (49) to find the best
350 hits for each contig in published genome and kept only those contigs that had a local alignment
351 of at least 5kb with published genome with identity of at least 95%. Finally, we performed
352 correction of assemblies using ICORN2 (50). Statistics on assemblies is provided in Table S4.

353

354 **Variant calling**

355 For each of the sequenced samples, we mapped its pair-end reads onto the corresponding
356 founder genome using bwa (51) and removed reads that had non-unique alignments. Filtered
357 mappings were used for variant calling. We used samtools (52) to call SNPs and InDels. The
358 resulting calls were filtered by coverage (>=10) and variant frequency (>=80%) to produce high-
359 quality variant set for each of the experimental populations. We then excluded non-unique
360 variants that were supported by at least two reads in more than one experimental population due
361 to extremely low probability of mutations affecting the same base pair in several independent
362 populations. We further masked low complexity regions of founder genomes by repeatmasker
363 (53) and excluded indels that fell into the masked sites and adjacent regions ten base pairs around
364 them. In populations A4 and B2, we observed a couple of mutations in adjacent genomic sites;

365 we excluded these variants as probable dinucleotide mutations.

366

367 **Sanger sequencing**

368 Sanger sequencing was used to check the presence of the called variants (138 SNPs and
369 24 indels) at each of the four time points. Several samples of passage 64 were lost due to
370 *Penicillium* contamination, so we used the closest available passage 75 as the first time point
371 passage. As each of the mutations observed at earlier time points in B populations were always
372 present in later populations, we considered a mutation as verified if it was confirmed by Sanger
373 sequencing in any of four time points (assuming that the mutation that appeared once became
374 fixed at later time points). For B populations, we believed in NGS data for passages 64 and 130
375 if the mutation was identified by Sanger sequencing at any time point. For A populations, we had
376 to check each of the four time points.

377 Primers for Sanger sequencing were designed using primer3 (54,55) and tntblast (56)
378 software and synthesized in Evrogen. PCR was performed using Encyclo kit (Evrogen).
379 Amplification products were sequenced in both directions on ABI Prism 3500 sequencer
380 (Applied Biosystems). Chromatograms were validated manually in CodonCode Aligner (57).

381 The detailed results of Sanger sequencing are included in Table S1. The number of
382 confirmed mutations is provided in Table S2. We were unable to confirm neither the presence
383 nor the absence of 15 mutations and excluded them from any further analyses. The majority of
384 mutations that we failed to validate comprise mutations in noncoding/not annotated regions.
385 Also, there were three mutations in A populations for which we couldn't track the time of their
386 appearance.

387

388 **Variant annotation**

389 We mapped reference CDS sequences (48) (v. 6.31, last accessed ~ Dec 2014) onto the
390 reference genome and obtained coding regions coordinates using Splign (58). Out of 10,635
391 reference protein-coding genes, we were able to identify 10230 genes in the reference genome
392 that coded complete proteins with proper start and stop codons. We then performed pairwise
393 genome alignment of A and B founder genotype assemblies with published genome using Multiz
394 (59) and identified coding regions coordinates in the assembled genomes using Splign annotation
395 results. 9732 and 9732 (numbers are indeed the same but protein sets differ) complete genes
396 were identified in A and B founder genotypes, respectively, with a tiny amount of incomplete
397 genes (627 and 530 genes for A and B populations, respectively). Based on the produced
398 annotations, the identified genomic variants were classified as exonic, intronic or intergenic.
399 Exonic variants in complete genes (i.e., genes with all of their exons completely identified in the
400 founder genome) were further classified as missense, nonsense or silent mutations in case of
401 SNPs and frameshift or amino acid insertion/deletion mutations in case of indels.

402

403 **Permutations**

404 In order to estimate the probability to obtain the observed number of protein-altering
405 variants by chance, we generated null distributions for the number of missense and nonsense
406 mutations in A and B founder genomes using permutations. A permutation procedure was
407 performed as follows. 85 and 40 observed SNP mutations were thrown at random onto A and B
408 genomes, respectively, and annotated as described above. To control for possible confounding
409 effects of mutation context, we added three-nucleotide sequence contexts to permutations. The
410 procedure was repeated 10,000 times and produced null distributions for the expected number of

411 nonsense and missense mutations in each of the founder genomes (Fig 3). To obtain the expected
412 number of proteins evolving in parallel, we combined permutation results from A and B
413 populations and counted the number of proteins that acquired at least two single-nucleotide
414 protein-changing mutations in eight populations during the permutation procedure.
415
416 **Protein conservation analysis**
417 We used OrthoMCL (60) to identify orthologous groups between protein sets from
418 *P. anserina* and *N. crassa* (61) reference genomes. Orthologous protein pairs were then aligned
419 using MAFFT (62) (with -localpair option), and for each protein carrying ortholog in *N. crassa*
420 genome, we calculated its conservation as a fraction of conserved positions in the alignment. We
421 also measured the conservation of mutated positions as a fraction of conserved positions among
422 the whole set of mutated positions.
423

424 ## References

425 1. Lenski RE, Travisano M. Dynamics of adaptation and diversification: a 10,000-generation
426 experiment with bacterial populations. Proc Natl Acad Sci U S A. 1994 Jul 19;91(15):6808–14.
427 2. Philippe N, Pelosi L, Lenski RE, Schneider D. Evolution of penicillin-binding protein 2
428 concentration and cell shape during a long-term experiment with Escherichia coli. J Bacteriol. 2009
429 Feb;191(3):909–21.
430 3. Lenski RE, Wiser MJ, Ribeck N, Blount ZD, Nahum JR, Morris JJ, et al. Sustained fitness gains and
431 variability in fitness trajectories in the long-term evolution experiment with Escherichia coli
432 [Internet]. 2015. Available from: http://dx.doi.org/10.1101/027391
433 4. Blount ZD, Barrick JE, Davidson CJ, Lenski RE. Genomic analysis of a key innovation in an
434 experimental Escherichia coli population. Nature. 2012;489(7417):513–8.

435 5. Leiby N, Marx CJ. Metabolic Erosion Primarily Through Mutation Accumulation, and Not
436 Tradeoffs, Drives Limited Evolution of Substrate Specificity in Escherichia coli. PLoS Biol.
437 2014;12(2):e1001789.
438 6. Lenski RE, Travisano M. Dynamics of adaptation and diversification: a 10,000-generation
439 experiment with bacterial populations. Proc Natl Acad Sci U S A. 1994 Jul 19;91(15):6808–14.
440 7. Woods R, Schneider D, Winkworth CL, Riley MA, Lenski RE. Tests of parallel molecular evolution
441 in a long-term experiment with Escherichia coli. Proceedings of the National Academy of Sciences.
442 2006;103(24):9107–12.
443 8. Cooper TF, Remold SK, Lenski RE, Schneider D. Expression profiles reveal parallel evolution of
444 epistatic interactions involving the CRP regulon in Escherichia coli. PLoS Genet. 2008 Feb;4(2):e35.
445 9. Gray JC, Cutter AD. Mainstreaming Caenorhabditis elegans in experimental evolution. Proc Biol
446 Sci. 2014 Mar 7;281(1778):20133055.
447 10. McDonald MJ, Rice DP, Desai MM. Sex speeds adaptation by altering the dynamics of molecular
448 evolution. Nature. 2016 Mar 10;531(7593):233–6.
449 11. Rundle HD, Chenoweth SF, Doughty P, Blows MW. Divergent selection and the evolution of signal
450 traits and mating preferences. PLoS Biol. 2005 Nov;3(11):e368.
451 12. Dettman JR, Anderson JB, Kohn LM. Divergent adaptation promotes reproductive isolation among
452 experimental populations of the filamentous fungus Neurospora. BMC Evol Biol. 2008 Jan 31;8:35.
453 13. Lorin S, Dufour E, Sainsard-Chanet A. Mitochondrial metabolism and aging in the filamentous
454 fungus Podospora anserina. Biochimica et Biophysica Acta (BBA) - Bioenergetics. 2006;1757(5-
455 6):604–10.
456 14. Turker MS, Cummings DJ. Podospora anserina does not senesce when serially passaged in liquid
457 culture. J Bacteriol. 1987 Feb;169(2):454–60.
458 15. Kudryavtseva OA, Mazheika IS, Solovchenko AE, Kamzolkina OV. Genetic instability of the short-
459 living ascomycetous fungus Podospora anserina induced by prolonged submerged cultivation.
460 Microbiology. 2011;80(6):784–96.

461 16. Bölker M. Sex and crime: heterotrimeric G proteins in fungal mating and pathogenesis. Fungal
462 Genet Biol. 1998 Dec;25(3):143–56.
463 17. Krijgsheld P, Bleichrodt R, van Veluw GJ, Wang F, Müller WH, Dijksterhuis J, et al. Development
464 in Aspergillus. Stud Mycol. 2013 Mar 15;74(1):1–29.
465 18. Han K-H, Kim JH, Moon H, Kim S, Lee S-S, Han D-M, et al. The Aspergillus nidulans esdC (early
466 sexual development) gene is necessary for sexual development and is controlled by veA and a
467 heterotrimeric G protein. Fungal Genet Biol. 2008;45(3):310–8.
468 19. Coleman DE, Berghuis AM, Lee E, Linder ME, Gilman AG, Sprang SR. Structures of active
469 conformations of Gi alpha 1 and the mechanism of GTP hydrolysis. Science. 1994 Sep
470 2;265(5177):1405–12.
471 20. Landis CA, Masters SB, Spada A, Pace AM, Bourne HR, Vallar L. GTPase inhibiting mutations
472 activate the α chain of Gs and stimulate adenylyl cyclase in human pituitary tumours. Nature.
473 1989;340(6236):692–6.
474 21. Lan KL, Sarvazyan NA, Taussig R, Mackenzie RG, DiBello PR, Dohlman HG, et al. A point
475 mutation in Galphao and Galphai1 blocks interaction with regulator of G protein signaling proteins. J
476 Biol Chem. 1998 May 22;273(21):12794–7.
477 22. Khafizov K, Lattanzi G, Carloni P. G protein inactive and active forms investigated by simulation
478 methods. Proteins. 2009 Jun;75(4):919–30.
479 23. Clark MA, Lambert NA. Endogenous regulator of G-protein signaling proteins regulate the kinetics
480 of Galphaq/11-mediated modulation of ion channels in central nervous system neurons. Mol
481 Pharmacol. 2006 Apr;69(4):1280–7.
482 24. Springer ML, Yanofsky C. A morphological and genetic analysis of conidiophore development in
483 Neurospora crassa. Genes Dev. 1989 Apr;3(4):559–71.
484 25. Chung D-W, Greenwald C, Upadhyay S, Ding S, Wilkinson HH, Ebbole DJ, et al. acon-3, the
485 Neurospora crassa ortholog of the developmental modifier, medA, complements the conidiation
486 defect of the Aspergillus nidulans mutant. Fungal Genet Biol. 2011 Apr;48(4):370–6.

487 26. Loubradou G, Bégueret J, Turcq B. MOD-D, a Galpha subunit of the fungus Podospora anserina, is
488 involved in both regulation of development and vegetative incompatibility. Genetics. 1999
489 Jun;152(2):519–28.
490 27. Barton NH. Genetic hitchhiking. Philos Trans R Soc Lond B Biol Sci. 2000 Nov
491 29;355(1403):1553–62.
492 28. Woods R, Schneider D, Winkworth CL, Riley MA, Lenski RE. Tests of parallel molecular evolution
493 in a long-term experiment with Escherichia coli. Proceedings of the National Academy of Sciences.
494 2006;103(24):9107–12.
495 29. Lang GI, Rice DP, Hickman MJ, Sodergren E, Weinstock GM, Botstein D, et al. Pervasive genetic
496 hitchhiking and clonal interference in forty evolving yeast populations. Nature. 2013 Aug
497 29;500(7464):571–4.
498 30. Wichman HA, Badgett MR, Scott LA, Boulianne CM, Bull JJ. Different trajectories of parallel
499 evolution during viral adaptation. Science. 1999 Jul 16;285(5426):422–4.
500 31. Lang GI, Desai MM. The spectrum of adaptive mutations in experimental evolution. Genomics.
501 2014 Dec;104(6 Pt A):412–6.
502 32. Hottes AK, Freddolino PL, Khare A, Donnell ZN, Liu JC, Tavazoie S. Bacterial Adaptation through
503 Loss of Function. PLoS Genet. 2013;9(7):e1003617.
504 33. Lang GI, Murray AW, Botstein D. The cost of gene expression underlies a fitness trade-off in yeast.
505 Proc Natl Acad Sci U S A. 2009 Apr 7;106(14):5755–60.
506 34. Osiewacz HD. Molecular Biology of Fungal Development. CRC Press; 2002. 608 p.
507 35. Moran D, Softley R, Warrant EJ. The energetic cost of vision and the evolution of eyeless Mexican
508 cavefish. Sci Adv. 2015 Sep;1(8):e1500363.
509 36. Smith NGC. Are radical and conservative substitution rates useful statistics in molecular evolution?
510 J Mol Evol. 2003 Oct;57(4):467–78.
511 37. Wernegreen JJ. Reduced selective constraint in endosymbionts: elevation in radical amino acid
512 replacements occurs genome-wide. PLoS One. 2011 Dec 14;6(12):e28905.

513 38. Popadin K, Polishchuk LV, Mamirova L, Knorre D, Gunbin K. Accumulation of slightly deleterious
514 mutations in mitochondrial protein-coding genes of large versus small mammals. Proc Natl Acad Sci
515 U S A. 2007 Aug 14;104(33):13390–5.
516 39. Hughes AL, Ota T, Nei M. Positive Darwinian selection promotes charge profile diversity in the
517 antigen-binding cleft of class I major-histocompatibility-complex molecules. Mol Biol Evol. 1990
518 Nov;7(6):515–24.
519 40. Dagan T, Talmor Y, Graur D. Ratios of Radical to Conservative Amino Acid Replacement are
520 Affected by Mutational and Compositional Factors and May Not Be Indicative of Positive
521 Darwinian Selection. Mol Biol Evol. 2002;19(7):1022–5.
522 41. Hanada K, Gojobori T, Li W-H. Radical amino acid change versus positive selection in the evolution
523 of viral envelope proteins. Gene. 2006 Dec 30;385:83–8.
524 42. Suzuki Y. Inferring natural selection operating on conservative and radical substitution at single
525 amino acid sites. Genes Genet Syst. 2007 Aug;82(4):341–60.
526 43. Maddamsetti R, Hatcher PJ, Green AG, Williams BL, Marks DS, Lenski RE. Core Genes Evolve
527 Rapidly in the Long-term Evolution Experiment with Escherichia coli. Genome Biol Evol [Internet].
528 2017 Apr 4; Available from: http://dx.doi.org/10.1093/gbe/evx064
529 44. Bazykin GA, Kondrashov AS. Major role of positive selection in the evolution of conservative
530 segments of Drosophila proteins. Proc Biol Sci. 2012 Sep 7;279(1742):3409–17.
531 45. Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue.
532 Phytochemical Bulletin. 1987;19:11–5.
533 46. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data.
534 Bioinformatics. 2014 Aug 1;30(15):2114–20.
535 47. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new
536 genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012
537 May;19(5):455–77.
538 48. Podospora anserina Genome Project [Internet]. [cited 2018 Jun 16]. Available from:

539 http://podospora.i2bc.paris-saclay.fr/download.php
540 49. Altschul S. Basic Local Alignment Search Tool. J Mol Biol. 1990;215(3):403–10.
541 50. Otto TD, Sanders M, Berriman M, Newbold C. Iterative Correction of Reference Nucleotides
542 (iCORN) using second generation sequencing technology. Bioinformatics. 2010 Jul 15;26(14):1704–
543 7.
544 51. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform.
545 Bioinformatics. 2009 Jul 15;25(14):1754–60.
546 52. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map
547 format and SAMtools. Bioinformatics. 2009 Aug 15;25(16):2078–9.
548 53. RepeatMasker Home Page [Internet]. [cited 2018 Jun 16]. Available from:
549 http://www.repeatmasker.org
550 54. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3—new
551 capabilities and interfaces. Nucleic Acids Res. 2012;40(15):e115–e115.
552 55. Koressaar T, Remm M. Enhancements and modifications of primer design program Primer3.
553 Bioinformatics. 2007 May 15;23(10):1289–91.
554 56. Gans JD, Wolinsky M. Improved assay-dependent searching of nucleic acid sequence databases.
555 Nucleic Acids Res. 2008 Jul;36(12):e74.
556 57. Website [Internet]. [cited 2018 Jun 16]. Available from: www.codoncode.com
557 58. Kapustin Y, Souvorov A, Tatusova T, Lipman D. Splign: algorithms for computing spliced
558 alignments with identification of paralogs. Biol Direct. 2008;3(1):20.
559 59. Blanchette M, Kent WJ, Riemer C, Elnitski L, Smit AFA, Roskin KM, et al. Aligning multiple
560 genomic sequences with the threaded blockset aligner. Genome Res. 2004 Apr;14(4):708–15.
561 60. Li L. OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. Genome Res.
562 2003;13(9):2178–89.
563 61. Krumlau R, Marzluf GA. Characterization of the sequence complexity and organization of the
564 Neurospora crassa genome. Biochemistry. 1979;18(17):3705–13.

565 62. Katoh K. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier
566 transform. Nucleic Acids Res. 2002;30(14):3059–66.
567

568 ## Figures

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

570 Eight experimental populations were established from two founder lineages A and B. After 75,
571 130, 200, and 268 passages, we determined genotypes of the experimental populations (dashed
572 lines; see Methods).

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

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

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

579 Distributions of the expected number of nonsense (A) and missense (B) substitutions in A (left
580 panel) and B (right panel) experimental populations in 10,000 permutations. Red line
581 corresponds to the observed number of nonsense or missense substitutions on (A) and (B),
582 respectively. *p* (p-value) is the area to the right of the red line.

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

584 Distribution of the number of genes with two protein-altering single-nucleotide substitutions in
585 10,000 random permutations. Red line shows the observed number of genes with two or more
586 such substitutions occurred.
587

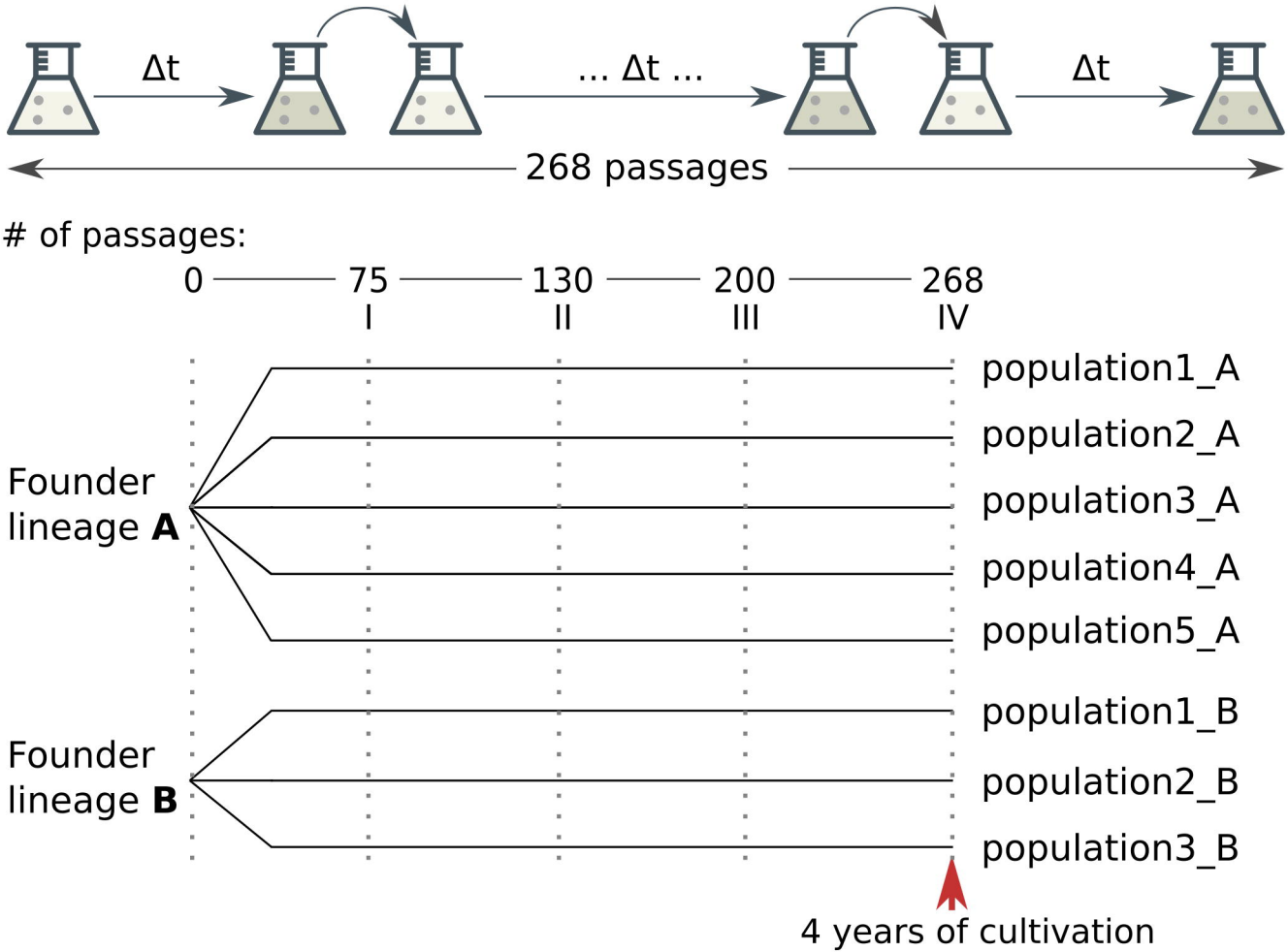
588 ## Supplementary Information

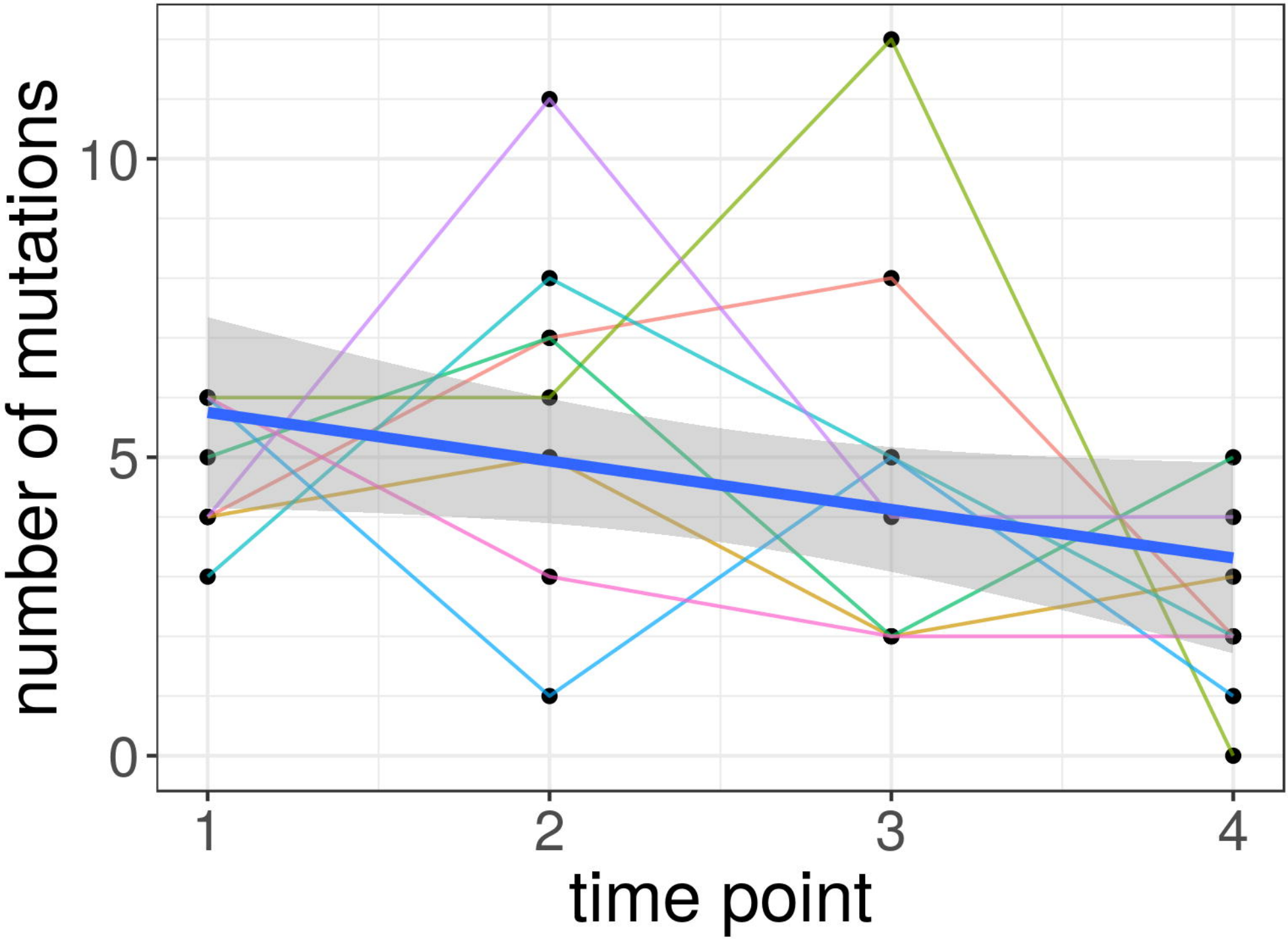
589 **Table S1. The summary of the observed allele replacements.** Sheets A1-A5 and B1-B3
590 contain the results of the Sanger sequencing validation procedure. The two last sheets comprise
591 annotation of the allele replacements in A and B populations.

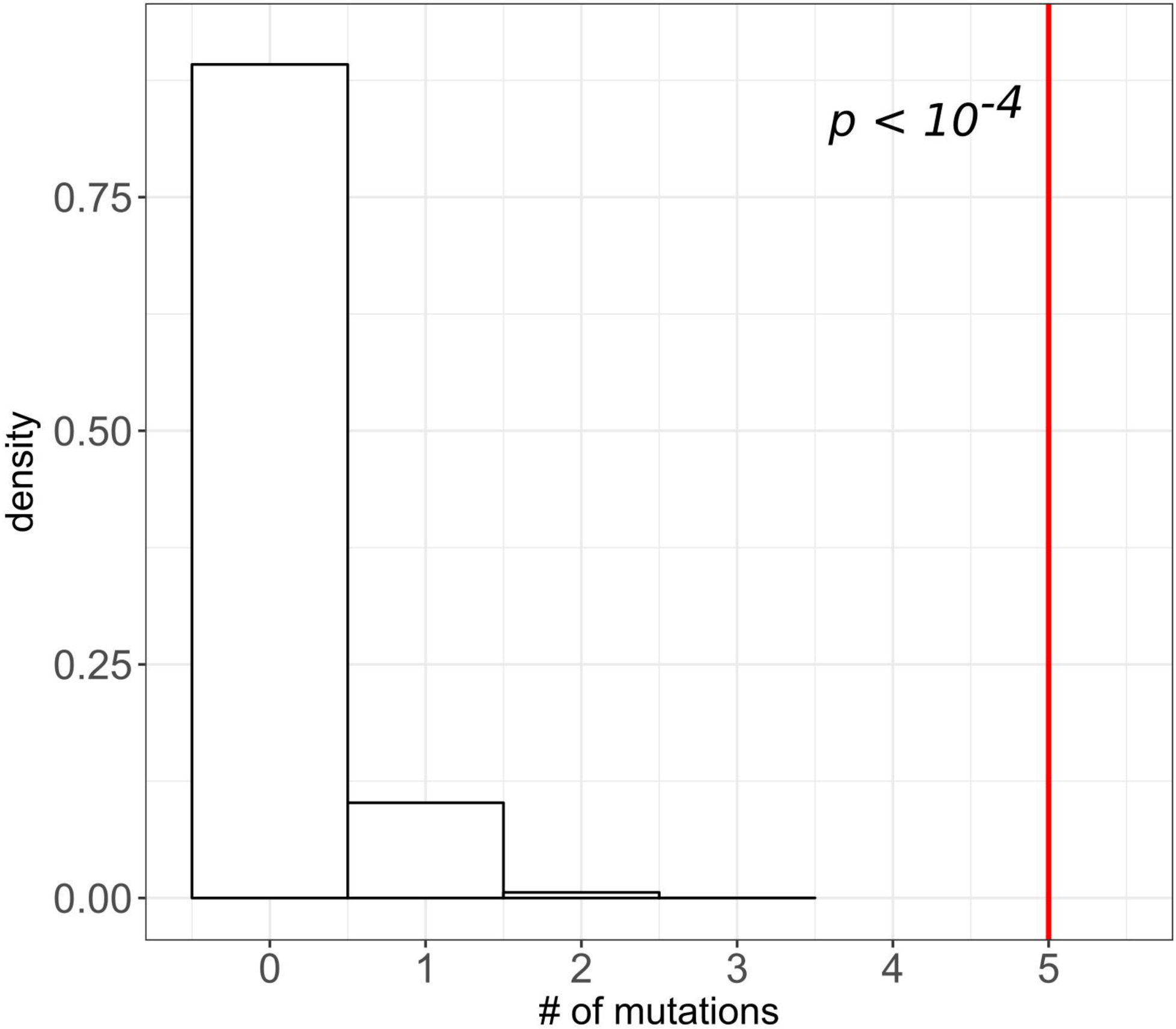
592 **Table S2. Accumulation of allele replacements.** Each entry provides the number of newly
593 observed replacements, relatively to the previous time point.

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

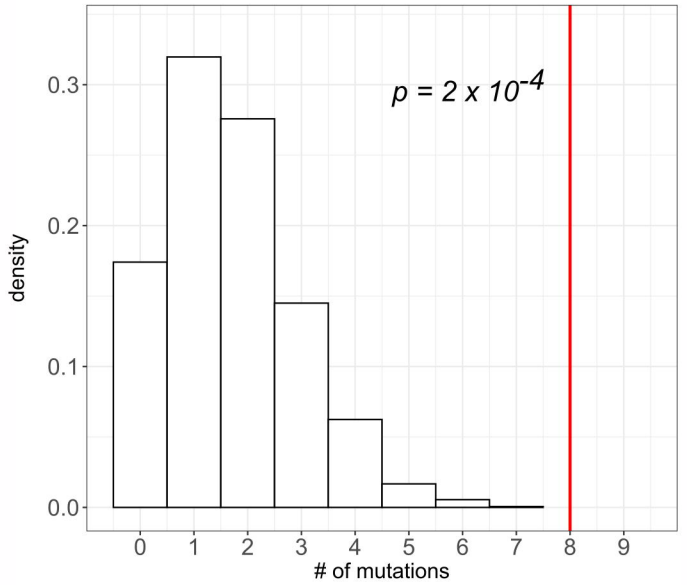
595 **Table S4. Statistics on founder genotypes assemblies.**

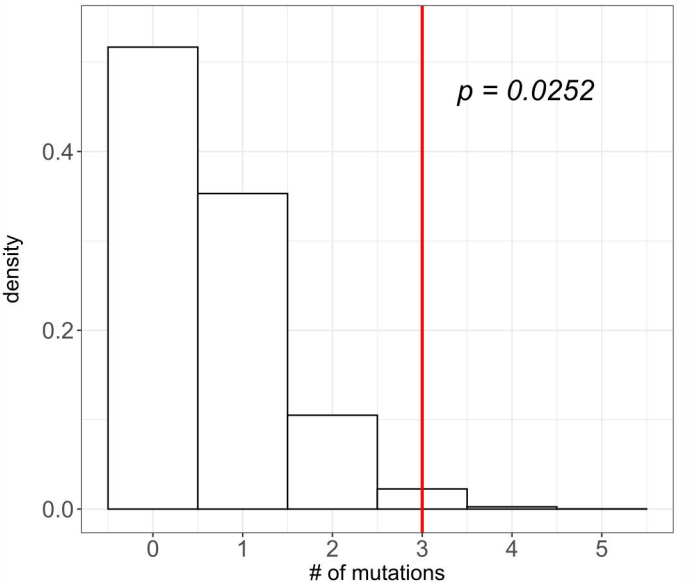

Δt
... Δt ...
Δt
268 passages
# of passages:
0
75
I
130
II
200
III
268
IV
Founder lineage A
Founder lineage B
population1_A
population2_A
population3_A
population4_A
population5_A
population1_B
population2_B
population3_B
4 years of cultivation


number of mutations
10
5
0
1
2
3
4
time point

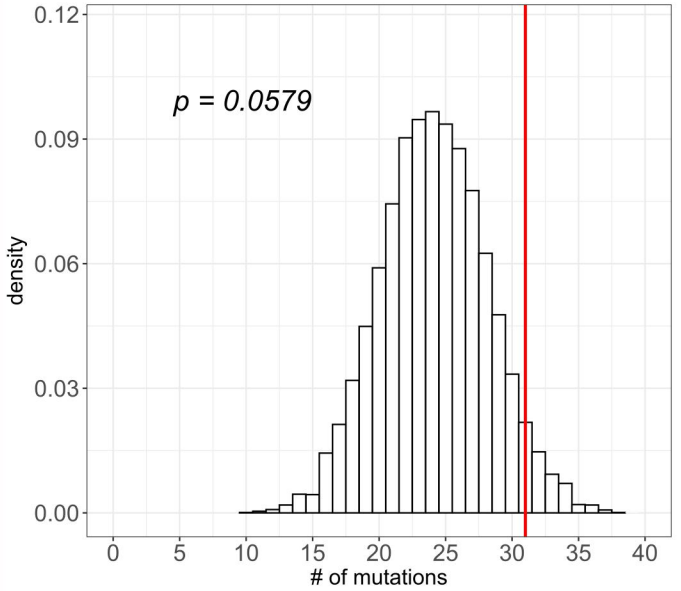

$p < 10^{-4}$
0.75
0.50
0.25
0.00
density
0
1
2
3
4
5
# of mutations

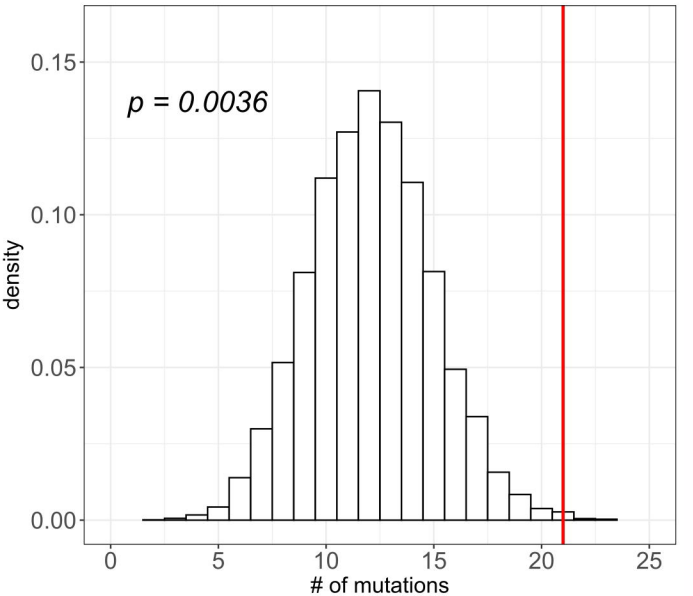
A.


p = 2 x 10^-4
density
0.0
0.1
0.2
0.3
0 1 2 3 4 5 6 7 8 9
# of mutations


p = 0.0252
density
0.0
0.2
0.4
0 1 2 3 4 5
# of mutations

B.


p = 0.0579
density
0.00
0.03
0.06
0.09
0.12
0 5 10 15 20 25 30 35 40
# of mutations


p = 0.0036
density
0.00
0.05
0.10
0.15
0 5 10 15 20 25
# of mutations