

1 **Title:** Polymorphism and the Predictability of Evolution in Fisher's Geometric Model

2 Sandeep Venkataram, Diamantis Sellis, Dmitri A. Petrov

3 *Department of Biology, Stanford University, Stanford, CA 94305*

4 **Running Title:** Predictability of Evolution

5 **Key Words:** heterozygote advantage, fitness landscape, selection, population genetics

6 **Corresponding Author:**

7 Dmitri A. Petrov

8 Stanford University Department of Biology

9 371 Serra Mall Stanford CA 94305

10 (650) 736 1169

11 dpetrov@stanford.edu

12

ABSTRACT

13 Predicting the future evolutionary state of a population is a primary goal of evolutionary
14 biology. One can differentiate between forward and backward predictability, where forward
15 predictability is the probability of the same adaptive outcome occurring in independent
16 evolutionary trials, and backward predictability is the likelihood of a particular adaptive
17 path given the knowledge of the starting and final states. Most studies of evolutionary
18 predictability assume that alleles along an adaptive walk fix in succession with individual
19 adaptive mutations occurring in monomorphic populations. However, in nature, adaptation
20 generally occurs within polymorphic populations, and there are a number of mechanisms
21 by which polymorphisms can be stably maintained by natural selection. Here we
22 investigate the predictability of evolution in monomorphic and polymorphic situations by
23 studying adaptive walks in diploid populations using Fisher's geometric model, which has
24 been previously found to generate balanced polymorphisms through overdominant
25 mutations. We show that overdominant mutations cause a decrease in forward
26 predictability and an increase in backward predictability relative to diploid walks lacking
27 balanced states. We also show that in the presence of balanced polymorphisms, backward
28 predictability analysis can lead to counterintuitive outcomes such as reaching different final
29 adapted population states depending on the order in which mutations are introduced and
30 cases where the true adaptive trajectory appears inviable. As stable polymorphisms can be
31 generated in both haploid and diploid natural populations through a number of
32 mechanisms, we argue that natural populations may contain complex evolutionary histories
33 that may not be easily inferred without historical sampling.

INTRODUCTION

34

35 Predicting evolution is one of the fundamental challenges of evolutionary biology (reviewed
36 in DE VISSER and KRUG (2014)). This question became particularly prominent with
37 Gould’s famous thought-experiment on “replaying the tape of life” (GOULD, 1990). Gould
38 wondered whether we would regenerate the observed evolutionary history of the world if we
39 reset our evolutionary history to any point in the past and let evolution retake its course
40 from there. More generally, we can ask whether it is possible to predict the path or the
41 final destination of the evolutionary process from a given starting point. It is also possible,
42 however, to ask whether we can reconstruct the true evolutionary trajectory given the final
43 adapted state (WEINREICH *et al.*, 2006). This distinction between types of predictability is
44 rarely made (however see NOURMOHAMMAD *et al.* (2013) and SZENDRO *et al.* (2013)), so
45 we formalize the methods for studying predictability and utilize these distinctions to study
46 the impact of polymorphism on the predictability of evolution.

47 **Forward predictability of evolution:** We define forward predictability as the
48 probability of observing a particular future evolutionary outcome from a known starting
49 state. Previous experimental evolution studies have generally (but not always) focused on
50 the forward predictability of evolution. This type of analysis can be done at a number of
51 levels, including the predictability of overall fitness changes, phenotypic shifts and different
52 levels of genotypic changes (pathways, genes, and individual mutations).

53 For example, FEREÁ *et al.* (1999), COOPER *et al.* (2003) and FONG *et al.* (2005) evolved
54 independent replicates of microbes and observed similar changes in gene expression and
55 growth rate in the evolved clones. A large study of 145 parallel long-term experimental
56 evolutions with *Escherichia coli* grown at elevated temperature showed that the same
57 genes and pathways were repeatedly targeted for mutations in independent populations
58 (TENAILLON *et al.*, 2012) as did a study of 40 replicate *Saccharomyces cerevisiae* batch

59 culture evolutions (LANG *et al.*, 2013) and a study that sequenced clones from 10 replicate
60 evolutions for each of 13 different genetic backgrounds (KRYAZHIMSKIY *et al.*, 2014).
61 TENAILLON *et al.* (2012) also observed a high degree of parallel evolution at the level of
62 individual nucleotides, but nucleotide level parallelism was rarely observed by LANG *et al.*
63 (2013). HERRON and DOEBELI (2013) evolved *E. coli* under multiple carbon sources and
64 repeatedly observed the evolution of two distinct ecotypes with differential ability to grow
65 on each carbon source. By sequencing independent replicate clones of both ecotypes, they
66 found the same genes, and sometimes the same exact mutations invading these replicate
67 populations and differentiating the ecotypes. These studies suggest that evolution is indeed
68 forward predictable to a surprising degree.

69 Repeated evolution has been observed at both the genetic and morphological levels in
70 natural systems as well (reviewed in STERN (2013)). KVITEK *et al.* (2008) showed that
71 highly divergent yeast strains isolated from oak trees had similar growth rates across a
72 panel of diverse growth conditions. Studies of *Anolis* lizards in the Caribbean show
73 repeated independent adaptive radiations into similar niches across the islands (LOSOS,
74 1998). In addition, a study of the adaptive radiation of cichlid fish in Lake Tanganyika
75 showed convergent morphological evolution when the skeletal morphology of the various
76 species was compared to their phylogeny (MUSCHICK *et al.*, 2012).

77 **Backward predictability of evolution:** In addition to Gould's thought-experiment, one
78 can study predictability in a historical manner. Given the current state, we can try to
79 predict the ancestral state or the evolutionary path that resulted in the current state of the
80 study system. We call this backward predictability, as it requires us to look backward in
81 time. For example, one can try to predict exactly how corn or rice became domesticated
82 from one or more wild ancestors (MATSUOKA *et al.*, 2002; MOLINA *et al.*, 2011), identify
83 the ancestral species that gave rise to Darwin's Finches (DARWIN, 1872; SATO *et al.*, 2001),

84 or reconstruct the ancestral state of a particular protein (ORTLUND *et al.*, 2007).

85 Alternatively, if we already know the ancestral state, we can try to predict the particular
86 order of mutations or phenotypic states that led to the evolution of the current state.

87 WEINREICH *et al.* (2006) conducted a seminal study of backward predictability in this
88 sense, using a combinatorially complete reverse genetic study design pioneered by

89 MALCOLM *et al.* (1990). WEINREICH *et al.* (2006) reconstructed every possible

90 combination of five mutations in the beta-lactamase gene in *E. coli* which are known to

91 lead to high levels of resistance to the drug rifampicin. They then assayed each genotype's

92 resistance to the drug, which they used as a proxy for fitness. Using this data, they

93 determined the fitness changes involved in every step of each of the $5! = 120$ possible

94 mutational paths that converts the wild-type genotype to the resistant five-mutant

95 genotype. A mutational path was deemed viable if fitness monotonically increased with

96 every step, that is, there were no mutations along the path that decreased resistance to the

97 drug.

98 WEINREICH *et al.* (2006) found that only 18 of the 120 possible paths were viable,

99 suggesting high backward predictability of evolution. In contrast, KHAN *et al.* (2011)

100 performed an analysis of five adaptive mutations from experimentally evolved bacterial

101 lineages using identical methodology and found that a majority of the orders were viable.

102 Finally, FRANKE *et al.* (2011) studied backward predictability in all subsets of two to six

103 mutations in an empirical eight-locus system and found that the number of viable paths

104 varied widely for a given subset size. For example, they observed both zero and nine viable

105 paths (out of 24 possible) in different four-locus subsets. The varying degrees of backward

106 predictability found in these different systems does not yet allow us to draw general

107 conclusions, and the laborious nature of the experiments makes it challenging to study

108 more than a few mutations at a time. In addition, without knowing the true order in which

109 the mutations arose in the population, it is unclear how accurate backward predictability
110 analysis actually is.

111 **Predictability in Fisher's Geometric Model:** Overall, there seems to be no consensus
112 on whether evolution is backward predictable using the method of WEINREICH *et al.*
113 (2006). It is also unclear how forward and backward predictability are correlated with each
114 other. In principle, one would want to conduct forward evolution and then conduct
115 backward predictability analysis on the same system to understand their relationship.
116 However such studies would be extremely laborious, and given the disparate answers
117 coming out of different experimental systems, a large number of independent experiments
118 in many systems would need to be conducted to give a convincing answer.

119 Another difficulty in experimental evolution studies of predictability are practical
120 limitations in sampling adaptive mutations. As most studies can only afford to sample a
121 few adapted individuals from a given experiment, mutations must be at high frequency to
122 be observed and a common assumption is that each of these mutations fixed in the
123 population in succession (GILLESPIE, 1983, 1984; ORR, 2002; WEINREICH *et al.*, 2006;
124 KHAN *et al.*, 2011; FRANKE *et al.*, 2011). However, we know that mutations can be
125 maintained in a polymorphic state by a number of mechanisms. These include negative
126 frequency-dependent selection (LEVIN *et al.*, 1988; ISERBYT *et al.*, 2013), spatial and
127 temporal fluctuations in selection (RAINEY and TRAVISANO, 1998; KASUMOVIC *et al.*,
128 2008; SALTZ and NUZHIDIN, 2014) and heterozygote advantage (also called overdominance,
129 TAKAHATA and NEI (1990)). Polymorphisms can also be present in an unstable form
130 through clonal interference (DESAI and FISHER, 2007; HERRON and DOEBELI, 2013;
131 KVITEK and SHERLOCK, 2013; LANG *et al.*, 2013). The presence of functionally
132 consequential polymorphisms in a population can in principle significantly alter
133 predictability analysis as the selective effect of a new mutation may be dependent on other

134 alleles segregating in the population (fitness epistasis). Many of these polymorphisms are
135 either lost by the end of the experiment or are not observed in the sampled adapted
136 individuals, leading to incorrect inferences of predictability. Additional complications can
137 arise when estimating predictability as mutations can occur in multiple backgrounds in a
138 given population, so the likelihood of each mutation occurring in a particular background
139 also has to be taken into account, as well as any epistatic interactions the mutation has
140 with the rest of that background.

141 Due to the challenges of isolating sufficient numbers of independent adaptive mutations
142 from experimental populations to study predictability, we utilize a simulation-based
143 approach to study the impact of polymorphisms on forward and backward predictability.
144 We employ Fisher's geometric model (FGM, FISHER (1930)), which is a well-studied (ORR,
145 1999, 2005) phenotypic model that treats individuals and alleles as a phenotype that is a
146 vector in coordinate space with a fitness that is determined by the distance of the
147 individual's phenotype from a predefined optimal phenotype using a gaussian function
148 (Figure 1a). SELLIS *et al.* (2011) showed that adaptive mutations in diploid FGM
149 simulations are frequently overdominant if the mutations are sufficiently large in phenotypic
150 space, resulting in balanced polymorphisms. Such overdominant mutations are stable but
151 can be driven out of the population by subsequent adaptive mutations. As we are
152 interested in the interaction between balanced polymorphic states and the predictability of
153 evolution, we select the distribution of mutational effects such that some evolutionary
154 trajectories contain overdominant mutations, generating stable polymorphisms, and others
155 do not. We then compare both types of trajectories to understand how polymorphisms
156 influence predictability. We conclude that the presence of polymorphic states has a
157 substantial qualitative effect on the predictability of evolution, such that at least in this
158 model, forward and backward predictability are inversely correlated.

METHODS

159

160 **Simulations:** We model adaptive walks in diploid populations with Wright-Fisher
161 simulations using Fisher's geometric model (FGM) as in SELLIS *et al.* (2011). In FGM,
162 alleles are represented as a vector in n-dimensional phenotype space (Figure 1a). The
163 simulations use code modified from Sellis et al. to allow for more than 2 dimensions. We
164 perform 10,000 replicate simulations with population size $N = 5,000$ for 10,000
165 generations. We explore two models, one with two dimensions and one with 25 dimensions.
166 We partition our adaptive walks into those that do and those that do not contain
167 overdominant mutations to study the impact of balanced states on predictability. For the
168 remainder of our analysis, we identify the most frequent allele in each simulated population
169 at the end of 10,000 generations of evolution and study the mutations present on that
170 allele. We limit our analysis to studying the first five mutations of each adaptive walk and
171 ignore simulations with fewer than 5 mutations in order to control for the length of the
172 adaptive walk when studying predictability.

173 **Forward Predictability Analysis:** We calculate the forward predictability of the
174 adaptive trajectory using two metrics. In both of these metrics, we only consider
175 homozygous phenotypes. Our first metric, maximum pairwise distance, considers pairs of
176 adaptive walks. We compute the maximum of the phenotypic distances between the
177 observed single mutant phenotypes of the two adaptive walks, the double mutant
178 phenotypes, the triple mutant phenotypes etc. Our second metric measures the maximal
179 deviation from the optimal trajectory. For each adaptive walk, we compute the maximal
180 phenotypic distance of any encountered (homozygous) phenotype from the line segment
181 connecting the ancestral phenotype and the optimum.

182 **Backward Predictability Analysis:** We compute backward predictability on adaptive
183 walks of exactly five mutations. We calculate the probability of all possible mutational

184 orders for the given set of mutations in a manner similar to WEINREICH *et al.* (2006), but
185 generalized to allow balanced states as the experimental protocol of WEINREICH *et al.*
186 (2006) assumes that every mutation along each mutational order fixes in succession. We
187 summarize the set of possible mutational orders for a given set of mutations through the
188 effective number of trajectories statistic, which we define as

189
$$\frac{1}{\sum p^2}$$

190 where p is the probability of each mutational order possible for a given set of mutations. If
191 no mutational order is viable (has nonzero probability), the effective number of trajectories
192 is defined to be 0. Please see the Supplementary Methods for full methodological details.

RESULTS

193

194 We explore the predictability of evolution in the framework of Fisher’s geometric model
195 (FGM) of adaptation. In FGM, alleles are represented as vectors in coordinate space, with
196 individuals having a phenotype that is the average of the phenotypes of their constituent
197 alleles. Mutations are vectors that modify the phenotype of an allele, and fitness is a
198 gaussian function of the distance of the individual’s phenotype from the optimal phenotype
199 (which we define as the origin).

200 In order to focus on the effect of polymorphic states on the predictability of evolution, we
201 choose a parameter regime that generates simulations both with and without overdominant
202 mutations after a number of trial simulations with various parameter values. We perform
203 10,000 replicate simulations of adaptation under FGM in diploids with $N = 5000$
204 individuals. Mutational magnitudes are drawn from an exponential distribution with mean
205 $= \frac{1}{2}$ and the population is initiated at two units from the optimum. The mutation rate is
206 $5 * 10^{-6}$, which results in a mutation-limited regime (significantly less than one mutation
207 per generation as $2 * N * \mu = 0.05$), in order to minimize the generation of polymorphic
208 states by clonal interference so that we can focus on only those polymorphic states
209 generated by overdominant mutations.

210 We conduct our simulations using an FGM of two dimensions, and show that our
211 qualitative results also hold at 25 dimensions. In the 25 dimension regime, we need to
212 rescale our mutational magnitude mean to 5 in order to obtain a sufficient number of walks
213 with five mutations over our 10,000 generation simulations for statistical analysis. For all
214 of our statistical analyses, we consider only those mutations that are present on the most
215 frequent allele at generation 10,000. Such mutations are typically the only ones available
216 for analysis in a natural system. We additionally limit our analysis to studying the first
217 five mutations of each adaptive walk, and ignore simulations with fewer than five mutations

218 in order to compare adaptive walks of equal lengths. We partition the resulting
219 five-mutation adaptive walks into those that do (n=4975, 1548 in simulations with two and
220 25 dimensions, respectively) and do not (n=1251, 10) contain overdominant mutations to
221 study the impact of balanced polymorphisms on the predictability of evolution. The
222 presence of overdominant mutations in an observed five-mutation adaptive trajectory is
223 detected by the observation of a set of alleles during the FGM simulation that are capable
224 of being maintained as a balanced polymorphism (KIMURA, 1956). For details, please see
225 the Supplementary Methods.

226 **Predictability of Adaptive Walks:** We first consider the forward predictability of
227 phenotypic paths, which we define as the tendency of independent adaptive walks to
228 explore similar portions of phenotypic space. The ability of adaptive walks with
229 overdominant mutations to explore a larger phenotypic space compared to walks without
230 overdominance (α -dip vs γ , Figure 1a) should lead to lower predictability of the phenotypic
231 intermediates along the adaptive walk, which is confirmed by visual inspection of our
232 simulations (Figure 1b,c) and is consistent with the findings of SELLIS *et al.* (2011).

233 We quantify forward predictability by measuring the distribution of maximal phenotypic
234 distances between pairs of independent adaptive trajectories. Pairs of walks with
235 overdominant states are, on average, 40% further apart than walks without overdominant
236 mutations and are therefore less forward predictable (Figure 2, Kolmogorov-Smirnov test
237 $p \ll 10^{-10}$). We also measure forward predictability as the maximal phenotypic distance of
238 each observed trajectory from the optimal trajectory - the vector from the ancestral
239 phenotype to the optimal phenotype. We observe that the presence of overdominant
240 mutations in a walk increases the average distance from the optimal trajectory by 5%
241 (Figure 3, Kolmogorov-Smirnov test $p \ll 10^{-10}$), again suggesting that overdominant
242 mutations decrease forward predictability.

243 We then study backward predictability in a manner similar to WEINREICH *et al.* (2006).
244 As before, we limit our analysis to adaptive walks of exactly five mutations, which is
245 comparable to many recent experimental studies of backward predictability (WEINREICH
246 *et al.*, 2006; KHAN *et al.*, 2011; FRANKE *et al.*, 2011). Backward predictability analysis
247 requires knowledge of the five mutations that occurred during the FGM simulations and
248 computes the likelihood of every possible order of those five mutations in generating the
249 observed adapted five-mutation allele (e.g. see WEINREICH *et al.* (2006) Figure 2). In
250 order to conduct this analysis, we compute the probability of every possible path to the
251 five-mutant state by successively introducing each of the five mutations into the population
252 and assessing the probability of each of these mutations to successfully invade the
253 population (see Supplementary Methods). Although we artificially constrain the available
254 phenotypes to only those generated by combinations of the five mutations under
255 consideration, this analysis is a model for studying predictability in situations where there
256 are only a few possible adaptive mutations, such as the drug resistance mutations used by
257 Weinreich *et al.* We compute the effective number of adaptive trajectories for each
258 adaptive walk, with a higher number suggestive of a lower backward predictability.

259 The results of our backward predictability analysis are shown in Figure 4. We find that in
260 contrast to forward predictability, overdominant states decrease the effective number of
261 paths (and thus increase backward predictability) in a walk by 30%, on average
262 (Kolmogorov-Smirnov test $p \ll 10^{-10}$). In other words, conditional on reaching a particular
263 five-mutant state, it is more probable that independent trials of a walk that experienced at
264 least one overdominant state will use the same mutational order in repeated trials relative
265 to a walk without overdominant states. We also utilize the mean path divergence of
266 LOBKOVSKY *et al.* (2011) to study backward predictability and find that overdominant
267 states resulted in walks that were 10% less divergent (and thus more backward
268 predictable), on average (Kolmogorov-Smirnov test $p \ll 10^{-10}$).

269 **Multiple End States:** In addition to studying the probability of a given mutational order
270 in our backward predictability analysis, we also study the adapted population state that
271 results from each viable mutation order. In particular, we observe that when mutations are
272 introduced in different orders, the population encounters different intermediate alleles,
273 resulting in instances where the final adapted five-mutant allele can balance against
274 different intermediate alleles depending on the order in which the mutations were
275 introduced into the population. We also observe instances where walks that did not
276 experience balanced states in the FGM simulations generate balanced states when
277 introduced in a different order.

278 We find that 53% of all walks have at least two different end population states containing
279 the final adapted allele, with a maximum of 19 different population states for a single set of
280 five mutations. We also find that the presence of overdominant mutations in the FGM
281 simulation has a significant effect on whether there are multiple end states observed. The
282 presence of an overdominant mutation in the observed walk increases the frequency of
283 multiple end states from 30% to 60%. Our results suggest that adaptation occurring in the
284 same genetic background, in response to the same selection pressure and using the same
285 mutations, can result in significantly different final population states depending on the
286 historical order in which the adaptive mutations occurred.

287 **Qualitative categorization with regard to backward predictability:** We analyze
288 our backward predictability results to discern qualitative categorizations of our simulations.
289 We find four broad categorizations of simulations: 1) simulations whose backward
290 predictability reconstructions of the five-mutant allele by introducing the mutations in the
291 order observed in the FGM simulation generate no balanced states, 2) those
292 reconstructions that do generate balanced states, 3) reconstructions where the order of
293 mutations that was observed in the simulation was impossible to reconstruct due to

294 deleterious intermediate states during the reconstructions and 4) reconstructions where
295 every possible order of mutations was impossible due to deleterious intermediate states
296 (which is a subset of category 3).

297 We observe 2326, 3898, 89 and 5 simulations in each of these four categories, respectively.
298 We can further separate these categories by conditioning on our original definitions of
299 whether or not a simulation contained an overdominant intermediate state (i.e. whether
300 there was a set of alleles that could be maintained in a stable balanced state at any point
301 during the FGM simulation before the 5-mutant state reached 5% frequency). We find
302 1187, 62, 2 and 0 simulations in each of these four categories, respectively, among the
303 simulations that we had previously identified as not containing overdominant intermediate
304 states while we observe 1139, 3836, 87 and 5 simulations in each of these four categories,
305 respectively, among simulations that we had previously identified as containing
306 overdominant intermediate states.

307 The presence of backward predictability reconstructions where the observed order (and in a
308 few cases, every order) of mutations is impossible is surprising. We hypothesize that this is
309 due to the presence of adaptive alleles that are generated and stably maintained during a
310 walk that are transient and do not survive until the end of the simulation. We call these
311 “hidden alleles”, as they are hidden from almost all modern experimental studies of
312 adaptation. Lack of knowledge of hidden alleles appear to decrease the computed
313 probability of the true adaptive path observed in the FGM simulations, and in extreme
314 cases, can make the true path impossible to reconstruct. Visual inspection of adaptive
315 trajectories that are unable to be successfully reconstructed confirms this intuition (Figure
316 5). Backward predictability reconstructions that incorporate all mutations present at $\geq 1\%$
317 frequency at any point in the simulation, regardless of whether the mutation was present
318 on the allele sampled at the end of the simulation, can successfully reconstruct the

319 observed adaptive trajectory of this previously impossible evolutionary outcome,
320 confirming the necessity of hidden alleles for the viability of the observed adaptive
321 trajectory in these instances.

322 We then compare the forward and backward predictability metrics described above on the
323 different categories of simulations. In particular, we compare the simulations that were
324 initially defined as not containing overdominant states at any point to those that did not
325 have balanced states in the backward predictability analysis but did have balanced states
326 during the FGM simulation. We find no significant difference between these sets of
327 simulations by any of our predictability metrics (maximum pairwise distance, maximum
328 distance from optimal trajectory and effective number of paths Kolmogorov-Smirnov test
329 $p > 0.05$). This result suggests that the signal in our predictability metrics is being driven
330 by the presence of balanced states between intermediate alleles along the adaptive
331 trajectory to the five-mutant allele rather than a general feature of observing balanced
332 states in our simulations as a whole.

333 **High Dimensionality:** In our implementation of Fisher's Model, balanced states arise
334 when mutations are overdominant. The presence of additional phenotypic dimensions,
335 which seems realistically plausible from observed rates of pleiotropy (DUDLEY *et al.*, 2005;
336 ALBERT *et al.*, 2008), increases the frequency of overdominant mutations (SELLIS *et al.*,
337 2011). However, this concordantly decreases the fitness advantage of the average new
338 beneficial mutation, decreasing the number of adaptive mutations that successfully invade
339 the population over our 10,000 generation FGM simulations. To study the impact of high
340 dimensional landscapes on predictability, we conducted simulations using 25 dimensions
341 with a mean mutation size of 5. The increase in mean mutation size relative to our original
342 two dimensional simulations is necessary to generate a sufficient number of walks containing
343 at least 5 mutations within 10,000 generations. We again partitioned the simulations into

344 those with ($n = 1548$) and without ($n = 10$) overdominant mutations at any point of the
345 FGM simulation before the time when the five-mutant allele reached 5% frequency.

346 We observe the same qualitative results in 25 dimensions as in 2 dimensions (see
347 Supplementary Figures 1-4). In general, it appears that our conclusions about
348 predictability of adaptive walks do not depend on the dimensionality of the system, and
349 only on the presence of overdominant mutations in the adaptive walk.

DISCUSSION

350
351 In this study, we explored the predictability of evolution using Fisher's geometric model.
352 We distinguished between forward and backward predictability, where forward
353 predictability measures the likelihood of the same or a similar adaptive trajectory
354 occurring in independent evolutions, while backward predictability measures the likelihood
355 of a particular order of adaptive mutations given the ultimate adapted state. We knew
356 from prior work that diploids frequently generate overdominant mutations under Fisher's
357 geometric model (SELLIS *et al.*, 2011), so we studied predictability using walks with and
358 without overdominant mutations to understand the impact of balanced polymorphisms on
359 predictability.

360 We found that simulations without overdominant mutations are more forward predictable
361 than simulations with overdominance, while the reverse is true for backward predictability.
362 The anti-correlation between forward and backward predictability can be intuitively
363 understood by considering the the nature of adaptation in Fisher's geometric model. In
364 walks without overdominant mutations, mutations are confined to within γ (Figure 1a),
365 leading to high forward predictability. There is minimal opportunity for deviation from the
366 optimal trajectory, and most of the adaptive mutations that occur during these walks have
367 similar direction vectors to the optimal trajectory. Therefore, regardless of the order of
368 mutations, each step will move the population closer to the optimum, making most of the
369 trajectories viable, and resulting in low backward predictability. The reverse is true in
370 walks with overdominant mutations, which explore a much larger portion of phenotypic
371 space (α_{dip}). Overdominant mutations tend to overshoot the optimum and are frequently
372 followed by compensatory mutations. The larger amount of phenotypic space explored
373 generates lower forward predictability, while the high frequency of compensatory mutations,
374 and thus the importance of the order in which the mutations are introduced, results in high
375 backward predictability. While Fisher's geometric model is a useful tool to consider

376 adaptation under phenotypic stabilizing selection, further work is required to determine the
377 extent to which this anti-correlation is generalizable to biological systems. Nevertheless,
378 the anti-correlation we observe between forward and backward predictability highlights the
379 importance of distinguishing between types of predictability in future studies.

380 In natural populations, stable polymorphisms can be due to overdominance or other types
381 of balancing selection, such as negative frequency dependent selection (LEVIN *et al.*, 1988;
382 ISERBYT *et al.*, 2013), and spatially or temporally variable selection (RAINEY and
383 TRAVISANO, 1998; KASUMOVIC *et al.*, 2008; SALTZ and NUZHDIN, 2014). Transient
384 functional polymorphisms at intermediate frequencies can also be generated via clonal
385 interference (DESAI and FISHER, 2007; HERRON and DOEBELI, 2013; KVITEK and
386 SHERLOCK, 2013; LANG *et al.*, 2013). Both frequency dependent selection and clonal
387 interference can occur in both haploid and diploid populations. Our work shows that the
388 presence of polymorphisms in the population, regardless of source, significantly complicates
389 analysis of adaptive trajectories, and these complications must be considered in all natural
390 systems.

391 One such complication is the existence of simultaneous mutational lineages, which can
392 result in hidden alleles (i.e. alleles that are not present at the end of the evolution) and
393 transient population states that nevertheless significantly impact the future course of
394 evolution. Ignoring hidden alleles can significantly modify the inferred backward
395 predictability, and in extreme cases, can incorrectly suggest that the true order of
396 mutations is impossible. Different orders of mutations can also generate different sets of
397 heterozygous genotypes and different end population states, requiring the consideration of
398 the state of the entire adapted population rather than the presence of a particular adapted
399 allele.

400 Polymorphic states also drastically increase the number of possible adaptive paths. In
401 systems where adaptation proceeds through sequential fixation, one only needs to consider
402 the fitness of the 2^n possible genotypes relative to the ancestral background for an
403 n-mutation system. This is the methodology used in the experimental backward
404 predictability studies of WEINREICH *et al.* (2006), KHAN *et al.* (2011) and FRANKE *et al.*
405 (2011). However, in regimes where polymorphic states are frequently generated, the fitness
406 of an invading mutation can vary depending on the alleles already present in the
407 population. Within each adaptive trajectory, every mutation along the trajectory needs to
408 be introduced into the prior population at low frequency on every available allele and
409 tracked until the frequency of the new mutation has been stabilized in order to establish
410 that the mutation is truly beneficial. Such a study would be extremely laborious, and to
411 our knowledge, has never been conducted in any system.

412 **Experimental Implications:** In an experimental setting, high forward predictability
413 means it is likely that the same set of mutations will be generated in independent adaptive
414 walks, which make the probabilities generated through backward predictability analysis
415 meaningful for predicting future events. This can occur by either a small mutational target
416 size such as mutations that cause resistance to drugs, or a large mutational input into the
417 population which makes rare but extremely beneficial mutations dominate the adaptive
418 process (e.g. DESAI and FISHER (2007); KVITEK and SHERLOCK (2011); GERSTEIN *et al.*
419 (2012); PENNINGS (2012)). A study in FGM also suggests that a multi-locus FGM where
420 each locus only influences a subset of the independent phenotypic dimensions (restricted
421 pleiotropy) also promotes forward predictability, which the authors call parallel evolution
422 (CHEVIN *et al.*, 2010). Despite the large number of replicates required to achieve statistical
423 significance, experimentally determining forward predictability has been shown to be
424 feasible.

425 On the other hand, the possibility of hidden alleles makes accurate estimates of backward
426 predictability impossible in both natural and artificial experimental systems. Since we do
427 not have access to hidden alleles from natural populations, it is impossible to accurately
428 compute the backward predictability of the adaptive walk leading to the current
429 population state. Studying backward predictability using forward evolutions and constant
430 sampling is equally infeasible. Even if we could sample every mutation that rises to
431 reasonable frequency in a population, almost all of these mutations will be lost, and there
432 may be far too many to determine the subset which are non-neutral. As mentioned above,
433 there is also the problem of combinatorially many adaptive walks possible for even a few
434 mutations, making complete experimental analysis of even a five mutation system
435 extremely challenging. As others have mentioned, sampling a few high-fitness mutations
436 and conducting backward predictability experiments may not generate a correct
437 representation of the probability of any particular adaptive walk, as there may be
438 alternative adaptive peaks (WEINREICH *et al.*, 2006). Additionally, there is the possibility
439 of adaptation and potential epistatic interactions at sites not under consideration, and
440 spatial or temporal fluctuations in selection pressures can further complicate accurate
441 assessments of backward predictability in natural systems, and calls into question the
442 accuracy of reconstructed ancestral states.

443 Finally, the impact of hidden alleles on evolutionary trajectories depends on the rate at
444 which stable polymorphic states are generated. RAINEY and TRAVISANO (1998), for
445 example, observed adaptive radiation by niche construction in every replicate evolution
446 experiment they conducted. Under these conditions, we may expect hidden alleles to be
447 frequent in a large evolving population. The adapted state of natural populations may thus
448 experience a strong historical dependence on transient mutations that are eventually lost
449 and impossible to sample, decreasing the forward predictability of evolution and making
450 the inference of backward predictability impossible. The rate at which polymorphic states

451 are generated in natural systems and potential differences between types of polymorphic
452 states and their impact on forward and backward predictability should be further explored
453 to improve our understanding of the predictability of evolution.

Literature Cited

- ALBERT, A. Y. K., S. SAWAYA, T. H. VINES, A. K. KNECHT, C. T. MILLER, *et al.*,
2008 The genetics of adaptive shape shift in stickleback: pleiotropy and effect size.
Evolution; international journal of organic evolution **62**: 76–85.
- CHEVIN, L.-M., G. MARTIN, and T. LENORMAND, 2010 Fisher’s model and the genomics
of adaptation: restricted pleiotropy, heterogenous mutation, and parallel evolution.
Evolution **64**: 3213–31.
- COOPER, T. F., D. E. ROZEN, and R. E. LENSKI, 2003 Parallel changes in gene
expression after 20,000 generations of evolution in *Escherichiacoli*. Proceedings of the
National Academy of Sciences of the United States of America **100**: 1072–7.
- DARWIN, C., 1872 *The Origin of Species*. John Murray, London, 6th edition.
- DE VISSER, J. A. G. M., and J. KRUG, 2014 Empirical fitness landscapes and the
predictability of evolution. Nature reviews. Genetics **15**: 480–90.
- DESAI, M. M., and D. S. FISHER, 2007 Beneficial mutation selection balance and the
effect of linkage on positive selection. Genetics **176**: 1759–98.
- DUDLEY, A. M., D. M. JANSE, A. TANAY, R. SHAMIR, and G. M. CHURCH, 2005 A
global view of pleiotropy and phenotypically derived gene function in yeast. Molecular
systems biology **1**: 2005.0001.
- FEREA, T., D. BOTSTEIN, P. O. BROWN, and R. F. ROSENZWEIG, 1999 Systematic
changes in gene expression patterns following. Proceedings of the National Academy of
Sciences of the United States of America **96**: 9721–9726.
- FISHER, R., 1930 *The genetical theory of natural selection*. Oxford at the Clarendon Press,
Oxford, 1st edition.

- 477 FONG, S. S., A. R. JOYCE, and B. O. PALSSON, 2005 Parallel adaptive evolution
478 cultures of *Escherichia coli* lead to convergent growth phenotypes with different gene
479 expression states. *Genome Research* **15**: 1365–72.
- 480 FRANKE, J., A. KLOZER, J. A. G. M. DE VISSER, and J. KRUG, 2011 Evolutionary
481 Accessibility of Mutational Pathways. *PLoS Computational Biology* **7**.
- 482 GERSTEIN, A. C., D. S. LO, and S. P. OTTO, 2012 Parallel Genetic Changes and
483 Non-parallel Gene-environment Interactions Characterize the Evolution of Drug
484 resistance in Yeast. *Genetics* **192**: 241–252.
- 485 GILLESPIE, J., 1983 A Simple Stochastic gene substitution model. *Theoretical Population*
486 *Biology* **23**: 202–215.
- 487 GILLESPIE, J., 1984 Molecular evolution over the mutational landscape. *Evolution* **38**:
488 1116–1129.
- 489 GOULD, S. J., 1990 *Wonderful Life: The Burgess Shale and the Nature of History*. W. W.
490 Norton & Company.
- 491 HERRON, M. D., and M. DOEBELI, 2013 Parallel Evolutionary Dynamics of Adaptive
492 Diversification in *Escherichia coli*. *PLoS Biology* **11**: e1001490.
- 493 ISERBYT, A., J. BOTS, H. VAN GOSSUM, and T. N. SHERRATT, 2013 Negative
494 frequency-dependent selection or alternative reproductive tactics: maintenance of female
495 polymorphism in natural populations. *BMC evolutionary biology* **13**: 139.
- 496 KASUMOVIC, M. M., M. J. BRUCE, M. C. B. ANDRADE, and M. E. HERBERSTEIN,
497 2008 Spatial and temporal demographic variation drives within-season fluctuations in
498 sexual selection. *Evolution* **62**: 2316–25.

- 499 KHAN, A. I., D. M. DINH, D. SCHNEIDER, R. E. LENSKI, and T. F. COOPER, 2011
500 Negative epistasis between beneficial mutations in an evolving bacterial population.
501 *Science* **332**: 1193–6.
- 502 KIMURA, M., 1956 Rules for testing stability of a selective polymorphism. *Proceedings of*
503 *the National Academy of Sciences of . . .* **1966**: 336–340.
- 504 KRYAZHIMSKIY, S., D. P. RICE, E. R. JERISON, and M. M. DESAI, 2014 Global
505 Epistasis Makes Adaptation Predictable Despite Sequence-Level Stochasticity. *Science*
506 **344**: 1519–1522.
- 507 KVITEK, D. J., and G. SHERLOCK, 2011 Reciprocal Sign Epistasis between Frequently
508 Experimentally Evolved Adaptive Mutations Causes a Rugged Fitness Landscape. *PLoS*
509 *Genetics* **7**: e1002056.
- 510 KVITEK, D. J., and G. SHERLOCK, 2013 Whole Genome, Whole Population Sequencing
511 Reveals That Loss of Signaling Networks Is the Major Adaptive Strategy in a Constant
512 Environment. *PLoS Genetics* **9**: e1003972.
- 513 KVITEK, D. J., J. L. WILL, and A. P. GASCH, 2008 Variations in stress sensitivity and
514 genomic expression in diverse *S. cerevisiae* isolates. *PLoS Genetics* **4**: e1000223.
- 515 LANG, G. I., D. P. RICE, M. J. HICKMAN, E. SODERGREN, G. M. WEINSTOCK, *et al.*,
516 2013 Pervasive genetic hitchhiking and clonal interference in forty evolving yeast
517 populations. *Nature* .
- 518 LEVIN, B. R., J. ANTONOVICS, and H. SHARMA, 1988 Frequency-Dependent Selection in
519 Bacterial Populations [and Discussion]. *Philosophical Transactions of the Royal Society*
520 *B: Biological Sciences* **319**: 459–472.
- 521 LOBKOVSKY, A. E., Y. I. WOLF, and E. V. KOONIN, 2011 Predictability of evolutionary
522 trajectories in fitness landscapes. *PLoS Computational Biology* **7**: e1002302.

- 523 LOSOS, J. B., 1998 Contingency and Determinism in Replicated Adaptive Radiations of
524 Island Lizards. *Science* **279**: 2115–2118.
- 525 MALCOLM, B., K. WILSON, and B. MATTHEWS, 1990 Ancestral lysozymes reconstructed,
526 neutrality tested, and thermostability linked to hydrocarbon packing. *Nature* **345**:
527 86–89.
- 528 MATSUOKA, Y., Y. VIGOUROUX, M. M. GOODMAN, J. SANCHEZ G, E. BUCKLER,
529 *et al.*, 2002 A single domestication for maize shown by multilocus microsatellite
530 genotyping. *Proceedings of the National Academy of Sciences of the United States of*
531 *America* **99**: 6080–4.
- 532 MOLINA, J., M. SIKORA, N. GARUD, J. M. FLOWERS, S. RUBINSTEIN, *et al.*, 2011
533 Molecular evidence for a single evolutionary origin of domesticated rice. *Proceedings of*
534 *the National Academy of Sciences of the United States of America* **108**: 8351–6.
- 535 MUSCHICK, M., A. INDERMAUR, and W. SALZBURGER, 2012 Convergent evolution
536 within an adaptive radiation of cichlid fishes. *Current Biology* **22**: 2362–8.
- 537 NOURMOHAMMAD, A., T. HELD, and M. LÄSSIG, 2013 Universality and predictability in
538 molecular quantitative genetics. *Current opinion in genetics & development* **23**: 684–93.
- 539 ORR, H. A., 1999 The evolutionary genetics of adaptation: a simulation study. *Genetical*
540 *Research* **74**: 207–14.
- 541 ORR, H. A., 2002 The population genetics of adaptation: the adaptation of DNA
542 sequences. *Evolution; international journal of organic evolution* **56**: 1317–30.
- 543 ORR, H. A., 2005 The genetic theory of adaptation: a brief history. *Nature reviews.*
544 *Genetics* **6**: 119–27.

- 545 ORTLUND, E. A., J. T. BRIDGHAM, M. R. REDINBO, and J. W. THORNTON, 2007
546 Crystal structure of an ancient protein: evolution by conformational epistasis. *Science*
547 (New York, N.Y.) **317**: 1544–8.
- 548 PENNING, P. S., 2012 Standing Genetic Variation and the Evolution of Drug Resistance
549 in HIV. *PLoS Computational Biology* **8**: e1002527.
- 550 RAINEY, P., and M. TRAVISANO, 1998 Adaptive radiation in a heterogeneous
551 environment. *Nature* **394**: 69–72.
- 552 SALTZ, J. B., and S. V. NUZHIDIN, 2014 Genetic variation in niche construction:
553 implications for development and evolutionary genetics. *Trends in ecology & evolution*
554 **29**: 8–14.
- 555 SATO, A., H. TICHY, C. O’HUGIN, P. GRANT, B. R. GRANT, *et al.*, 2001 On the origin
556 of Darwin’s finches. *Molecular Biology and ...*: 299–311.
- 557 SELLIS, D., B. CALLAHAN, D. A. PETROV, and P. W. MESSER, 2011 Heterozygote
558 advantage as a natural consequence of adaptation in diploids. *Proceedings of the*
559 *National Academy of Sciences of the United States of America* **2011**: 1–6.
- 560 STERN, D. L., 2013 The genetic causes of convergent evolution. *Nature reviews. Genetics*
561 **14**: 751–64.
- 562 SZENDRO, I. G., J. FRANKE, J. A. G. M. DE VISSER, and J. KRUG, 2013 Predictability
563 of evolution depends nonmonotonically on population size. *Proceedings of the National*
564 *Academy of Sciences of the United States of America* **110**: 571–6.
- 565 TAKAHATA, N., and M. NEI, 1990 Allelic genealogy under overdominant and
566 frequency-dependent selection and polymorphism of major histocompatibility complex
567 loci. *Genetics* **124**: 967–78.

568 TENAILLON, O., A. RODRIGUEZ-VERDUGO, R. L. GAUT, P. McDONALD, A. F.

569 BENNETT, *et al.*, 2012 The Molecular Diversity of Adaptive Convergence. *Science* **335**:

570 457–461.

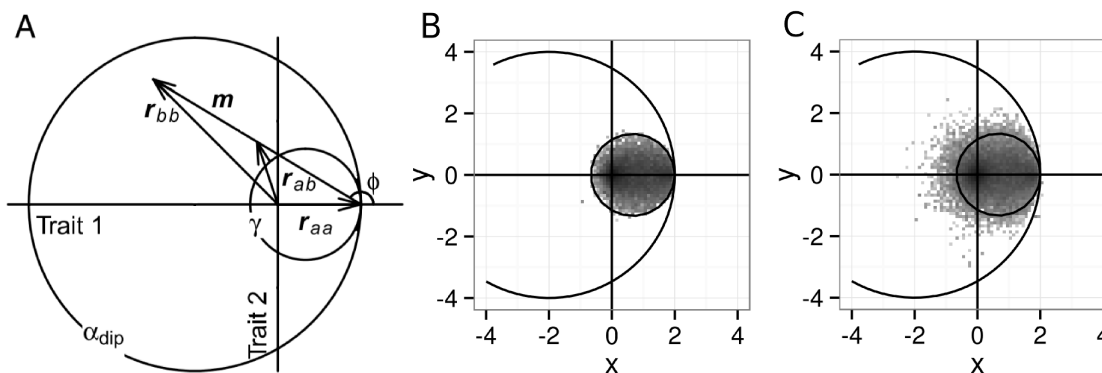
571 WEINREICH, D. M., N. F. DELANEY, M. A. DEPRISTO, and D. L. HARTL, 2006

572 Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science*

573 **312**: 111–4.

574

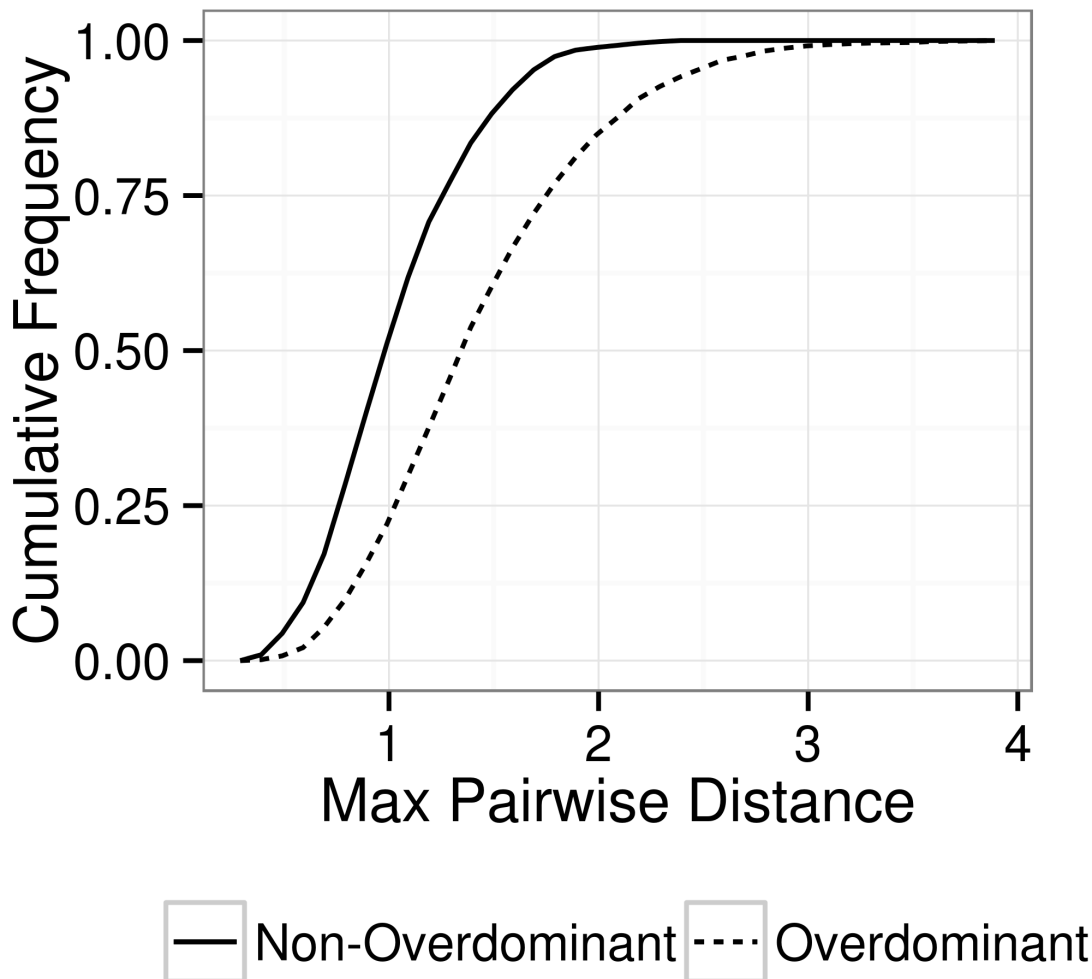
FIGURES



575

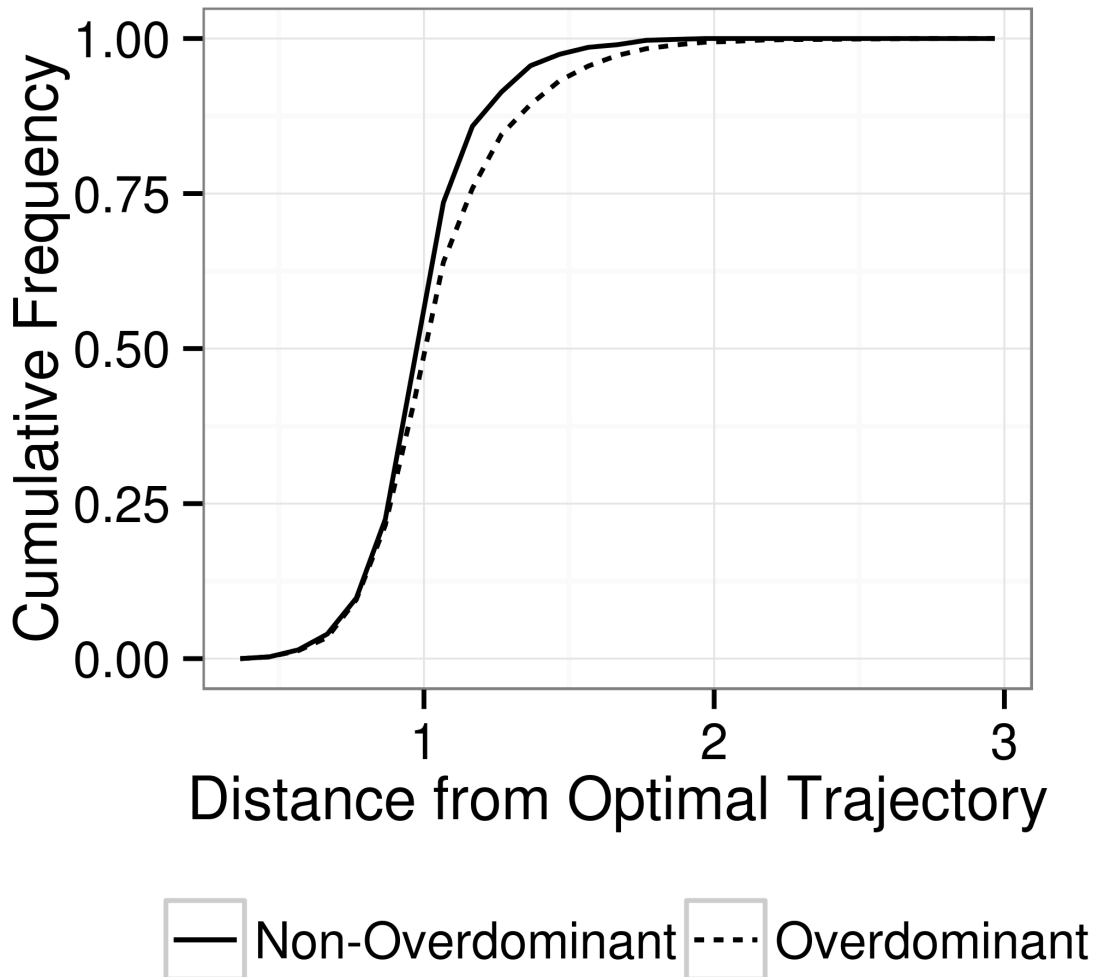
576 **Figure 1. Fisher's geometric model description and confirmation of accurate**
 577 **separation of simulations into those with and without overdominant mutations.**

578 **(A)** Modified from Figure 2A Sellis et al 2011. Two orthogonal axes represent independent
 579 character traits. Fitness is determined by a symmetrical Gaussian function centered at the
 580 origin. Consider a population initially monomorphic for the wild-type allele $r_{aa}^{\vec{}}$. A
 581 mutation m gives rise to a mutant phenotype vector $r_{bb}^{\vec{}} = r_{aa}^{\vec{}} + \vec{m}$. The phenotype of the
 582 mutant heterozygote assuming phenotypic codominance ($h = 1/2$) is $r_{ab}^{\vec{}} = r_{aa}^{\vec{}} + \vec{m}/2$. The
 583 different circles specify the areas in which mutations are adaptive (i.e. successfully invade
 584 the population, α_{dip}) and replacing (i.e. fix in the population, γ) in diploids. **(B)** Density
 585 plot of all phenotypes of homozygous individuals observed in the adaptive walks of FGM
 586 simulations that do not contain overdominant mutations. Note that all observed
 587 phenotypes lie within γ , as all mutations must be replacing and not balancing in this group
 588 of simulations. Circles denote α_{dip} and γ as described in (A). **(C)** Homozygous phenotypes
 589 for simulations that do contain overdominant mutations. Note that a large number of
 590 phenotypes lie outside of γ , as expected for overdominant mutations, confirming that we
 591 are correctly separating walks with and without overdominant mutations. When comparing
 592 B and C, we observe that simulations with overdominant mutations are less forward
 593 predictable than those without such mutations.



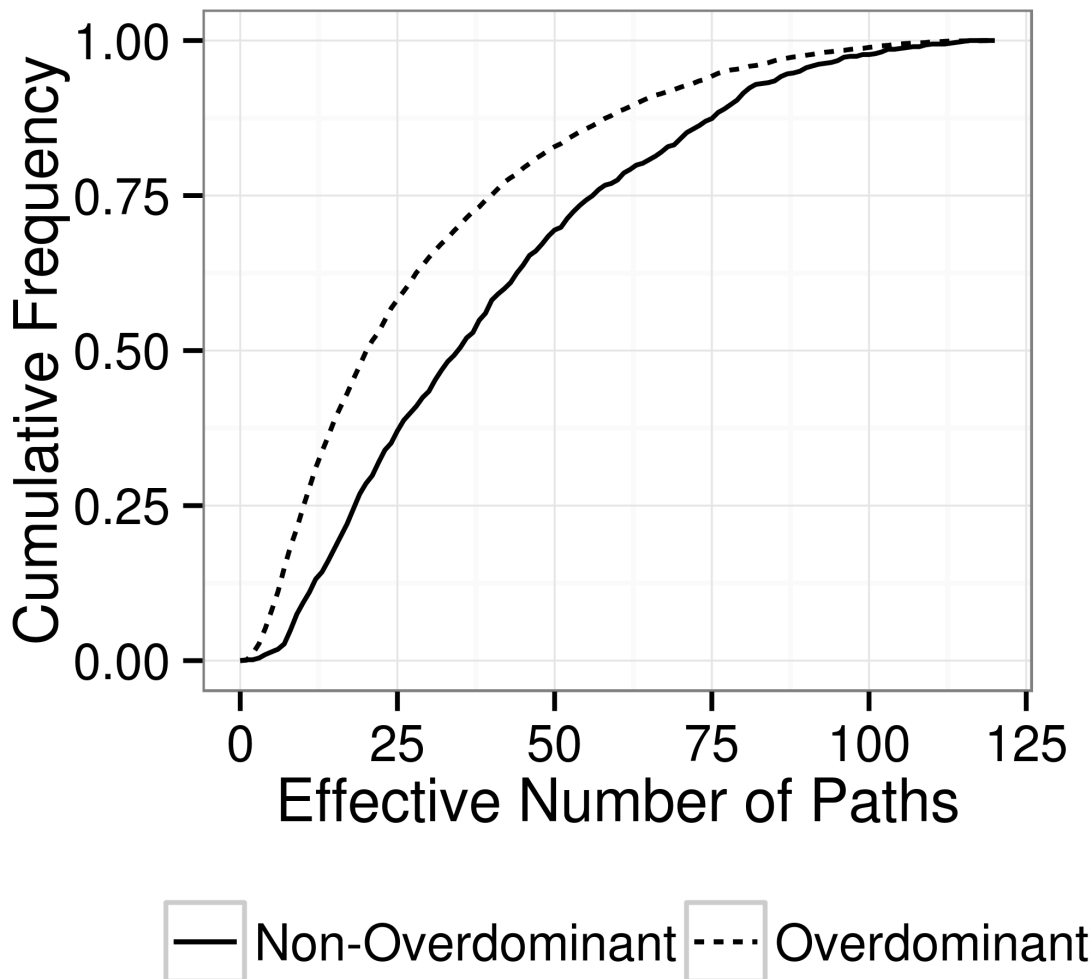
594

595 **Figure 2. Overdominant mutations decrease forward predictability by 40%**
596 **using the maximum pairwise distance metric.** Shown are the cumulative
597 distributions of the maximum phenotypic distance between independent pairs of adaptive
598 walks, excluding the ancestral state. This is a measure of the phenotypic forward
599 repeatability of independent walks on the same evolutionary landscape. The maximum
600 phenotypic distance in simulations without overdominant states is significantly less than in
601 simulations with such states (Kolmogorov-Smirnov test $p \ll 10^{-10}$).



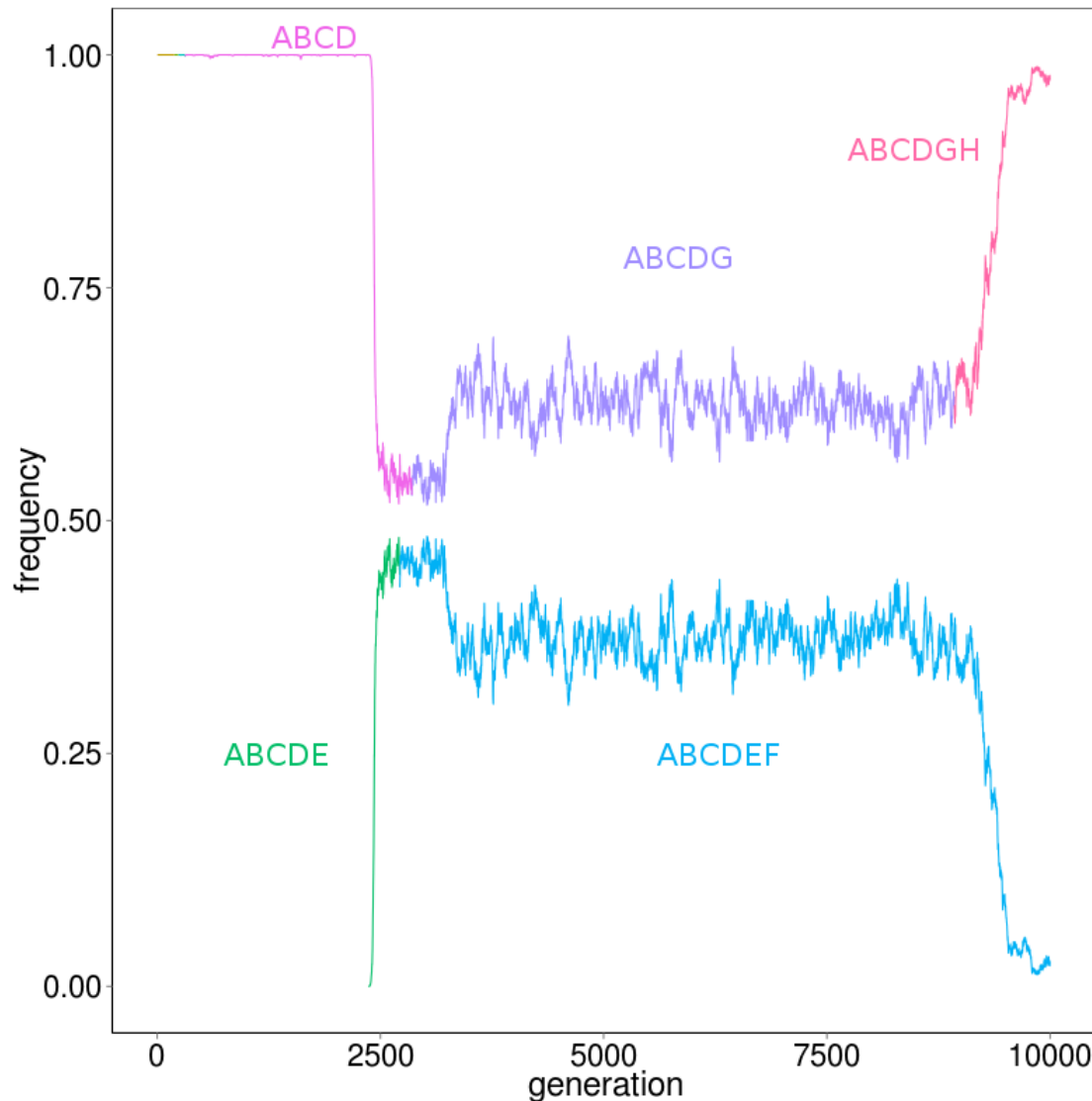
602

603 **Figure 3. Overdominant mutations decrease forward predictability by 5% using**
604 **the maximum distance from the optimal trajectory metric.** Shown are the
605 cumulative distributions of the maximum distance from the optimal trajectory of adaptive
606 walks. This is a measure of the phenotypic forward predictability of walks. The maximum
607 distance from the optimal trajectory in simulations without overdominant mutations is
608 significantly less than those with such mutations (Kolmogorov-Smirnov test $p \ll 10^{-10}$).



609

610 **Figure 4. Overdominant mutations increase backward predictability by 30%**
611 **using the effective number of paths metric.** Shown are the cumulative distributions
612 of the effective number of paths for adaptive walks with five mutations. This is a metric of
613 backward predictability of evolution. Each mutation is introduced into the ancestral
614 background in every possible order, and the number of viable mutational orders, weighted
615 by their probabilities, determines the effective number of paths. The effective number of
616 paths in simulations without overdominant mutations is significantly greater than in
617 simulations with such mutations (Kolmogorov-Smirnov test $p \ll 10^{-10}$).



618

619 **Figure 5. Example simulation with a hidden allele where the observed most**
620 **frequent allele was impossible to reconstruct by our method to compute**
621 **backward predictability.** The frequency of the two mutational lineages that reached at
622 least 1% frequency in the population are shown throughout the 10,000 generations of the
623 simulation. The main lineage, ending with allele ABCDGH, is at high frequency at the end
624 of the simulation, while the minor lineage, ending with allele ABCDEF (a “hidden allele”)
625 is at low frequency at the end of the simulation.

626 In the simulation, four mutations initially fix in quick succession, resulting in allele ABCD

627 fixed in the population. At this point, mutations causing balanced polymorphisms result in
628 branched mutational lineages. Mutation E is the first mutation to occur on allele ABCD,
629 generating a balanced polymorphism between alleles ABCD and ABCDE and allowing
630 both alleles to be stably maintained in the population at intermediate frequency. Mutation
631 F then quickly occurs on the background of allele ABCDE, generating allele ABCDEF
632 which also balances with allele ABCD. Mutation G then occurs on the background of allele
633 ABCD generating allele ABCDG soon afterwards, which balances with allele ABCDEF.
634 Finally, mutation H occurs on allele ABCDG generating allele ABCDGH, which
635 outcompetes all other alleles and is nearly fixed by the end of the simulation.

636 In our backward predictability reconstructions, we consider only the first five mutations of
637 the most frequent allele at the end of the simulation, that is, we consider only mutations A,
638 B, C, D and G as these were the first five mutations on allele ABCDGH. In attempting to
639 reconstruct this observed order of mutations, we find that we can successfully introduce
640 mutations A, B, C and D in order, but mutation G, which results in allele ABCDG, is not
641 beneficial if allele ABCD is the only other allele in the population (data not shown).
642 Therefore, the true order of mutations is impossible to reconstruct in this case when only
643 sampling allele ABCDGH at the end of the simulation. However, if we also consider
644 mutations E and F, we are able to successfully reconstruct the intermediate steps of the
645 observed adaptive trajectory, suggesting that the presence of allele ABCDEF is necessary
646 for allele ABCDG to be beneficial (data not shown).