SUPPLEMENTARY TEXT

Examples of calculating genotypic historical predictability

Haploid example

As a concrete example of how we measure historical predictability, we will use a system of two mutations, m_1 and m_2. Let A^{wt} be the ancestral allele, A^1 be the allele containing mutation m_1, A^2 be the allele with m_2 and $A^{1,2} = A^{der}$ be the the derived allele containing both available mutations m_1 and m_2.

There are $2! = 2$ different orders of mutations that can generate allele A^{der}. In the mutation order under consideration, m_1 occurs first, then m_2:

$M_1 = A^{wt} \xrightarrow{m_1} A^1 \xrightarrow{m_2} A^{der}$

the remaining mutation order is:

$M_2 = A^{wt} \xrightarrow{m_2} A^2 \xrightarrow{m_1} A^{der}$

Let us first consider the method of Weinreich *et al.* (2006), using example data in Table S1, the results of which are shown in Figure S1. Using our recursive procedure, we start with $S_{existing} = A^{wt}$ at fixation. There are two possible mutations in this population, where m_1 occurs on A^{wt} to get A^1 and where m_2 occurs on A^{wt} to get A^2. We compute the unconditioned probability of allele A^1 successfully being generated and invading the population, ρ_{A^1} on $S_{existing}$ and similarly for A^2. These are 0.3 and 0.5, respectively. We then compute the conditioned probabilities of success for these alleles. For A^1, this is $\frac{0.3}{0.3+0.5} = 0.375$, while it is $\frac{0.5}{0.3+0.5} = 0.625$ for A^2 (Table S1). In essence, a single successful mutation in $S_{existing}$ (A^{wt} fixed in the population) will generate an S_{new} of A^1 fixed in the population 0.375 of the time, and an S_{new} of A^2 fixed in the population 0.625 of the time. From here, we can then do the recursive call for the next step of the inference procedure for each of these new population states. These two recursive calls will be: 1) $S_{new} = A^1$ fixed in the population, with $A_{new} = A^{wt}, A^1$ and $P_{new} = 1 * 0.375 = 0.375$ and 2) $S_{new} = A^2$ fixed in the population, with $A_{new} = A^{wt}, A^2$ and $P_{new} = 1 * 0.625 = 0.625$

Let us now consider the first of these recursive calls, when A^1 is the first successful allele to invade the population and $P_{existing}$ for this call is 0.375. In this case, there is only one available mutation, m_2, which

will generate the fully adapted allele A^{der} with an unconditioned probability of 0.6 but a conditioned probability of 1. We now have $S_{new} = A^{der}$ fixed in the population, with a P_{new} of $0.375 * 1 = 0.375$. We then call the recursive condition again with this new S_{new}, where we find that the termination condition of having A^{der} in $S_{existing}$ has been reached. Therefore, we are done, and the unconditioned probability of the mutation order used to get A^{der} this time, namely $A^{wt} \xrightarrow{m_1} A^1 \xrightarrow{m_2} A^{der}$ is 0.375.

A similar procedure with the other initial recursive call, where m_2 was the first mutation, finds that the unconditioned probability of the mutation order $A^{wt} \xrightarrow{m_2} A^2 \xrightarrow{m_1} A^{der}$ is 0.625. Therefore, we find two viable orders of mutations, one with conditioned probability $\frac{0.375}{0.375+0.625} = 0.375$ and one with probability $\frac{0.625}{0.375+0.625} = 0.625$. Note that with the WEINREICH *et al.* (2006) method, the conditioned probability for a mutation order always equals its unconditioned probability since the number of mutations introduced into the population is always equal to the number of mutations in the mutation order. This is not the case in our diploid model, as we will see below.

Table S1

$S_{existing}$	Mutation	New Allele A^n	Invasion Prob $P^i_{A^n}$	ρ_n	$P_{existing}$	P_{new}	S_{new}	Mutation Order for Adapted Allele A^{der}
A^{wt}	m_1 on A^{wt}	A^1	0.3	0.3	1	0.375	A^1, freq = 1	
A^{wt}	m_2 on A^{wt}	A^2	0.5	0.5	1	0.625	A^2, freq = 1	
A^1	m_2 on A^1	A^{der}	0.6	0.6	0.375	0.375	A^{der}, freq = 1	$A^{wt} \xrightarrow{m_1} A^1 \xrightarrow{m_2} A^{der}$
A^2	m_1 on A^2	A^{der}	0.2	0.2	0.625	0.625	A^{der}, freq = 1	$A^{wt} \xrightarrow{m_2} A^2 \xrightarrow{m_1} A^{der}$

Diploid example

We then turn to measuring historical predictability in diploids, using a different pair of mutations with example data in Table S2 and the results in Figure S2. In this case, let us suppose that the A^1 can successfully invade the ancestral population consisting of A^{wt} to result in a balanced state consisting of A^1 and A^{wt} at intermediate frequencies. Meanwhile, A^2 can also successfully invade the ancestral population, but it fixes, resulting in S_{new} consisting of A^2 at frequency 1. Given their relative invasion probabilities and the fact that A^{wt} was initially fixed, we find that the conditioned probability of A^1 invading A^{wt} and resulting in a balanced state = 1 (freq of A^{wt}) * 0.2 (invasion probability of A^1) / $(1 * 0.2 + 1 * 0.35)$ = 0.36, while the probability of A^2 being the next mutation in A^{wt} is 0.64.

For the next recursion step, let us consider the $S_{existing}$ of A^1 and A^{wt} at intermediate frequencies. There are two possible mutations in this scenario, in which mutation m_2 can occur on either A^1 or A^{wt} to generate alleles A^{der} and A^2, respectively. The successful invasion of A^2 results in a S_{new} containing a balanced state consisting of both A^1 and A^2. The new allele A^{der} can also successfully invade the population and results in a stable polymorphism as well. Mutation m_1 is not allowed to occur on A^{wt}, since that would regenerate allele A^1 which has already been observed in this trajectory so far. The conditioned probability of A^2 succeeding in this population is $\frac{0.7*0.14}{0.7*0.14+0.3*0.6} = 0.35$, while the conditioned probability of A^{der} succeeding is 0.65. The running probability of these two mutation orders after two mutations have been introduced in the population are $0.36 * 0.35 = 0.126$ and $0.36 * 0.65 = 0.234$, respectively.

Now let us consider the mutations on the $S_{existing}$ where both A^1 and A^2 exist as a balanced polymorphism. In this situation, there are two possible mutations, where m_1 can arise on A^2 to give the adapted allele A^{der}, and m_2 can arise on A^1 to also give the adapted allele A^{der}. Even though this is the same allele being generated by the two mutations, the initial frequency of A^1 and A^2 are different, giving rise to different unconditioned probabilities of their occurrence. The m_1 mutation has a conditioned probability of $\frac{0.8*0.4}{0.8*0.4+0.2*0.4} = 0.8$, while the m_2 mutation has a conditioned probability of 0.2. The running probability after each of these mutations are $0.36 * 0.35 * 0.8 = 0.1008$ and $0.36 * 0.35 * 0.2 = 0.0252$, respectively.

The final possible trajectory, where m_2 occurred first on $S_{existing} = A^{wt}$ and resulted in the fixation of A^2 has only one possible mutation. This is mutation m_1 on A^2 resulting in the allele A^{der}. Supposing that A^{der} is deleterious in this situation, it cannot invade and therefore has 0 probability of occurring. We then

terminate this recursion as there are no valid beneficial mutations available to this population.

Finally, we now need to compute the conditioned likelihoods of each mutation order. We managed to successfully get A^{der} in 3 different ways when considering the mutations introduced into the population, but only 2 different ways when considering the mutations introduced onto the allele that generated A^{der}. The unconditioned probabilities of these two different mutation orders are: $0.234 + 0.0252 = 0.2592$ for mutation order $M_1 = A^{wt} \xrightarrow{m_1} A^1 \xrightarrow{m_2} A^{der}$ and 0.1008 for mutation order $M_2 = A^{wt} \xrightarrow{m_2} A^2 \xrightarrow{m_1} A^{der}$. The conditioned probabilities for these two mutation orders are thus $\frac{0.2592}{0.2592+0.1008} = 0.72$ and $\frac{0.1008}{0.2592+0.1008} = 0.28$, respectively.

Table S2

$S_{existing}$	Mutation	A^n	$P^i_{A^n}$	ρ_n	$P_{existing}$	P_{new}	S_{new}	Mutation Order of $A^{der} \in S_{new}$
A^{wt}	m_1	A^1	0.2	0.2	1	0.36	A^1 freq = 0.3, A^{wt} freq = 0.7	
A^{wt}	m_2	A^2	0.35	0.35	1	0.64	A^2 freq = 1	
A^1 freq = 0.3, A^{wt} freq = 0.7	m_2 on A^{wt}	A^2	0.14	0.098	0.36	0.126	A^1 freq = 0.2, A^2 freq = 0.8	
A^1 freq = 0.3, A^{wt} freq = 0.7	m_2 on A^1	A^{der}	0.6	0.18	0.36	0.234	A^{der} freq = 0.8, A^{wt} freq = 0.2	$A^{wt} \xrightarrow{m_1} A^1 \xrightarrow{m_2} A^{der}$
A^1 freq = 0.2, A^2 freq = 0.8	m_1 on A^2	A^{der}	0.4	0.32	0.126	0.1008	A^{der} freq = 1	$A^{wt} \xrightarrow{m_2} A^2 \xrightarrow{m_1} A^{der}$
A^1 freq = 0.2, A^2 freq = 0.8	m_2 on A^1	A^{der}	0.4	0.08	0.126	0.0252	A^{der} freq = 1	$A^{wt} \xrightarrow{m_1} A^1 \xrightarrow{m_2} A^{der}$
A^2 freq = 1	m_1 on A^2	A^{der}	0	0	0.64	0	A^2 freq = 1	

SUPPLEMENTARY FIGURES

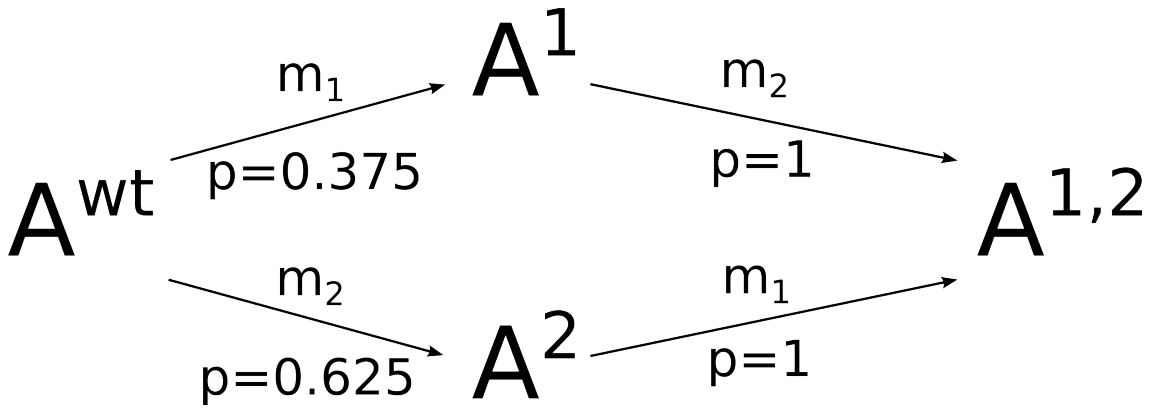

Figure S1. Along with table S1, a simple example of measuring genotypic historical predictability in haploids. Arrows represent transitions after the introduction of an available mutation into the population, with the mutation above the arrow and the conditioned probability of the mutation successfully being generated and invading the population below the arrow.

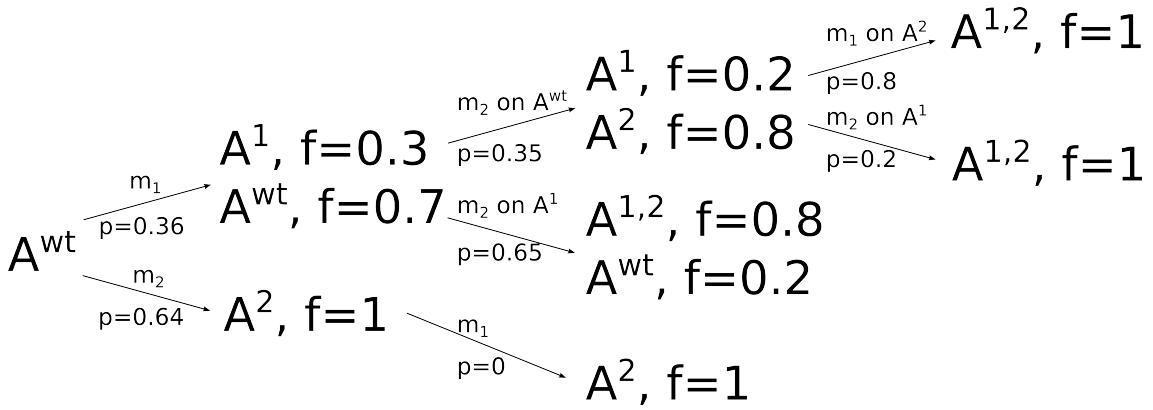

Figure S2. Along with table S2, a simple example of measuring genotypic historical predictability in diploids. Note that this example uses a different set of mutations than the example for the haploid method. Arrows represent transitions after the introduction of an available mutation into the population, with the mutation above the arrow and the conditioned probability of the mutation successfully being generated and invading the population below the arrow. Successful mutations that result in a balanced polymorphism are represented by the presence of multiple alleles each at some frequency (f).

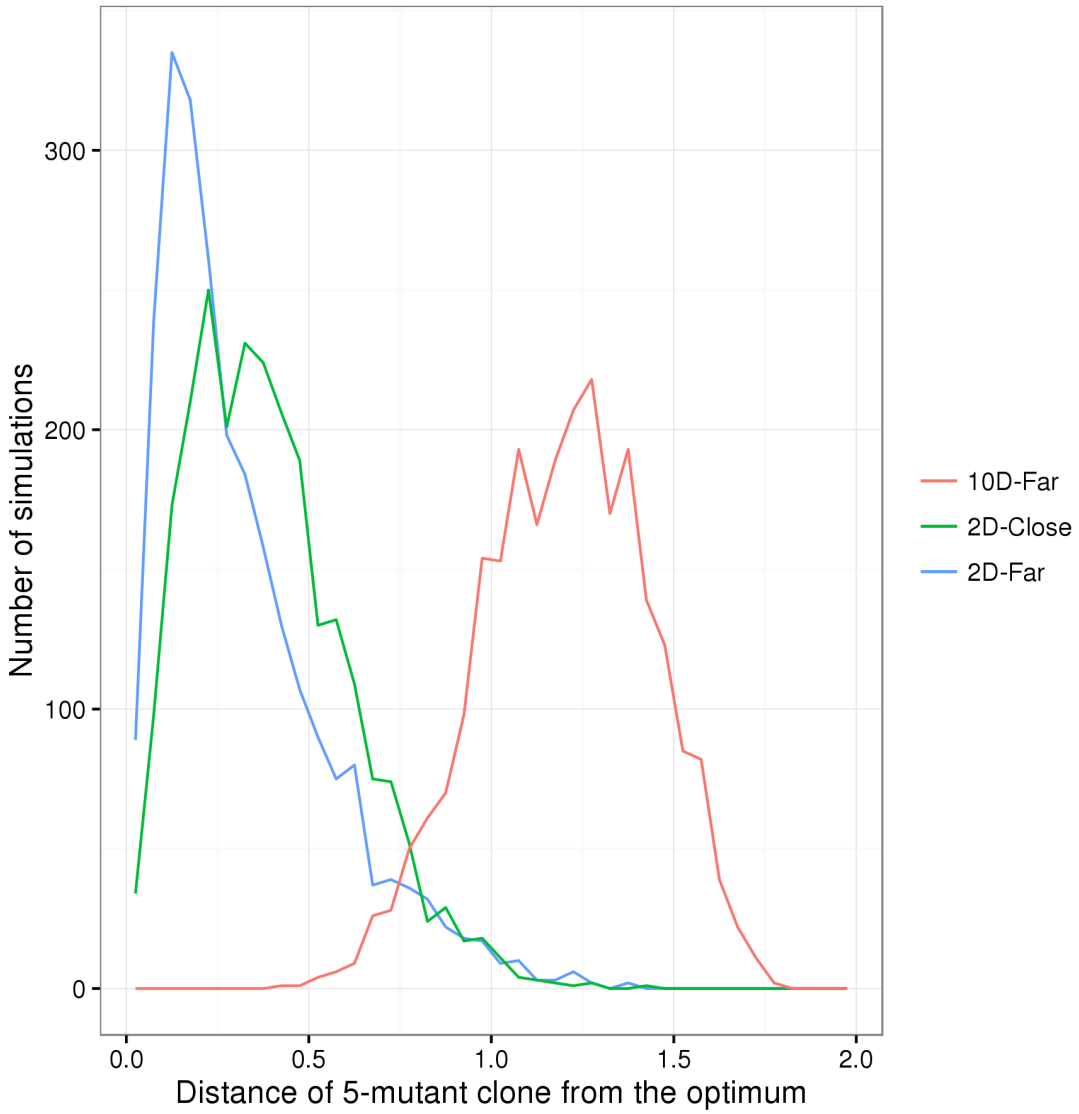

Figure S3. Distribution of the distance to the optimum of the 5-mutant genotype from the haploid simulations across all parameter regimes. A lower dimensionality allows the population to be phenotypically closer to the optimum (10D-Far vs 2D-Far), as does evolving on a more peaked landscape (2D-Far vs 2D-Close), consistent with prior studies (Blanquart *et al.* 2014).

98 Blanquart, F., G. Achaz, T. Bataillon, and O. Tenaillon, 2014 Properties of selected mutations
99 and genotypic landscapes under Fisher's geometric model. Evolution (N. Y.). **68**: 3537–3554.

100 Weinreich, D. M., N. F. Delaney, M. A. Depristo, and D. L. Hartl, 2006 Darwinian evolution can
101 follow only very few mutational paths to fitter proteins. Science **312**: 111–4.