To evaluate whether linkage patterns between sites could generate *the patterns observed in the data*, we ran Wright-Fisher simulations with SLiM (ref). First, we simulated the behavior of a single SNP evolving under fluctuating selection. In a population of constant size $N = 1000$ diploid individuals, we have introduced a target SNP A at an initial frequency p_0. Then, we traced allele frequency of A over 13 generations. The first 9 generations corresponded to the expansion phase, and the fitness of A during expansion was labeled $s_{\text{expansion}}$. The last 4 generations corresponded to the truncation phase, with the fitness of A labeled $s_{\text{truncation}}$. In our simulations, we considered every possible combination of fitnesses $(s_{\text{expansion}}, s_{\text{truncation}})$, where $s_{\text{expansion}}$ was varied between 1.0 (neutral) and 1.2 (highly advantageous), and $s_{\text{truncation}}$ was varied between 1.0 (neutral) and 0.8 (highly deleterious). In all simulations here and below, we assumed no dominance with $h = 0.5$. In total, we ran 250 replicate simulations for each combination of two selection parameters $(s_{\text{expansion}}, s_{\text{truncation}})$, and in each replicate simulation, the initial frequency p_0 was chosen randomly in the interval $(0.34, 0.66)$, ensuring that we simulate SNPs at "intermediate frequencies".

We then also ran two-locus simulation in a setup similar to that described above. In a population of constant size $N = 1000$, we have considered two loci, labeled A and B, located d nucleotides apart. We seeded the genotypes of inividuals in the initial generation by fixing initial frequencies of alleles, $p_0^{(A)}$ and $p_0^{(B)}$, as well as their linkage state. The linkage was determined by the magnitude of the LD measure r^2 between A and B, with $r^2 < 0.01$ meaning that A and B are "unlinked", and $r^2 > 0.05$ meaning that A and B are "highly linked". However, we further distinguished between the two different ways in which A and B could be highly linked by using the sign of r – the "raw" correlation between A and B that when squared yields r^2. When $r > 0$ and $r^2 > 0.05$, we say that A and B are highly linked in "attraction", meaning that derived alleles — the ones that are assigned fitness effects in simulations — at both loci are assorted together. Conversely, when $r < 0$ and $r^2 > 0.05$, A and B are highly linked in "repulsion", with derived allele of A assorted together with the reference allele of B, and vice versa. To summarise, we have distinguished three linkage states: "unlinked", "highly linked in attraction", and "highly linked in repulsion".

As the range of possible values of r and r^2 depends on the initial frequencies of A and B (VanLiere and Rosenberg, 2008), in each simulation we have sampled the value of r uniformly from the its range of possible values. For example, when trying to simulate loci in attraction, we chose r uniformly in the interval $\left(\sqrt{0.05}, r_{\max}(p_0^{(A)}, p_0^{(B)})\right)$. When trying to simulate loci in repulsion, we chose r uniformly in the interval $\left(r_{\min}(p_0^{(A)}, p_0^{(B)}), -\sqrt{0.05}\right)$. Finally, for unlinked loci, we chose r randomly in the interval $(-\sqrt{0.05}, \sqrt{0.05})$. The functions $r_{\min}$ and $r_{\max}$ were derived following the approach of VanLiere and Rosenberg (2008). After r was chosen, we computed the numbers of individuals possessing haplotypes AB as

$$p_0^{(AB)} = p_0^{(A)} p_0^{(B)} + r\sqrt{p_0^{(A)} p_0^{(B)} (1 - p_0^{(A)})(1 - p_0^{(B)})}$$

and used this value in setting the initial generation of the simulation.

After seeding the simulation by fixing initial frequencies of A and B and their linkage state in the first generation, we have run the simulation for 13 generations, with first 9 generations of "expansion" and then 4 generations of "truncation". In all simulations, we set $s_{\text{expansion}}^{(B)} = s_{\text{truncation}}^{(A)} = 0$. The recombination rate ρ was equal to $\rho = 2.39 \times 10^{-8}$ per generation per basepair during expansion, and to $\rho = 0$ during truncation. The initial frequencies $p_0^{(A)}$ and $p_0^{(B)}$ were chosen randomly in each simulation from the interval $(0.34, 0.66)$. For rest of the parameters, we formed a parameter grid and considered every possible combination of parameters (linkage state, $d, s_{\text{expansion}}^{(A)}, s_{\text{truncation}}^{(B)}$), where $d \in \{0.5, 1, 2, 4, 20\}$ measured in Megabases, $s_{\text{expansion}}^{(A)} \in \{1.0, 1.05, 1.1, 1.15, 1.2\}$, $s_{\text{truncation}}^{(B)} \in \{1.0, 0.95, 0.9, 0.85, 0.8\}$, and linkage state was one of "unlinked", "highly linked in attraction", and "highly linked in repulsion". We ran 1000 replicate simulations for each combination of parameter values. The seeding procedure described above ensures that we are quanfitying average behavior of alleles in each linkage category with varying values of r^2 and initial frequency.

To summarize the results of the simulation, we have estimated "empirically observed" fitness value $\widetilde{s}$ for A and B directly from the allele frequency change in the first generation of truncation by using the formula

$$\widetilde{s} = 1 + \frac{\Delta p}{1 - p - \Delta p},$$

where p is the frequency at the end of expansion phase, and Δp is the allele frequency change in the first generation of truncation phase.

Finally, to asses the net effect of linkage on the behavior of allele frequency of SNP A, we have combined replicate simulation runs for all three linkage categories, and computed fitness $\widetilde{s}$ of SNP A during the first generation of truncation from all combined simulations. The same number of runs was made for attraction and repulsion scenarios, which ensures that in the combined set of simulations attraction and repulsion is present in equal proportions, matching what we expect intuitively to see in real data because r^2 is broken down quickly with time, and spurious (transient) linkage has no reason to be biased towards repulsion or attraction specifically, as it is just an issue of naming alleles.