RESEARCH ARTICLE

Detection of hard and soft selective sweeps from *Drosophila melanogaster* population genomic data

Nandita R. Garud [1,2] *, **Philipp W. Messer** [3], **Dmitri A. Petrov** [4]

1 Department of Ecology and Evolutionary Biology, University of California, Los Angeles, California, United States of America, **2** Department of Human Genetics, University of California, Los Angeles, California, United States of America, **3** Department of Computational Biology, Cornell University, Ithaca, New York, United States of America, **4** Department of Biology, Stanford University, Stanford, California, United States of America

* ngarud@ucla.edu

Abstract

Whether hard sweeps or soft sweeps dominate adaptation has been a matter of much debate. Recently, we developed haplotype homozygosity statistics that (i) can detect both hard and soft sweeps with similar power and (ii) can classify the detected sweeps as hard or soft. The application of our method to population genomic data from a natural population of *Drosophila melanogaster* (DGRP) allowed us to rediscover three known cases of adaptation at the loci *Ace*, *Cyp6g1*, and *CHKov1* known to be driven by soft sweeps, and detected additional candidate loci for recent and strong sweeps. Surprisingly, all of the top 50 candidates showed patterns much more consistent with soft rather than hard sweeps. Recently, Harris *et al.* 2018 criticized this work, suggesting that all the candidate loci detected by our haplotype statistics, including the positive controls, are unlikely to be sweeps at all and that instead these haplotype patterns can be more easily explained by complex neutral demographic models. They also claim that these neutral non-sweeps are likely to be hard instead of soft sweeps. Here, we reanalyze the DGRP data using a range of complex admixture demographic models and reconfirm our original published results suggesting that the majority of recent and strong sweeps in *D. melanogaster* are first likely to be true sweeps, and second, that they do appear to be soft. Furthermore, we discuss ways to take this work forward given that most demographic models employed in such analyses are necessarily too simple to capture the full demographic complexity, while more realistic models are unlikely to be inferred correctly because they require a large number of free parameters.

Author summary

Whether hard versus soft sweeps dominate adaptation has long been a matter of great debate. Recently, we proposed novel statistics that can identify and differentiate hard and soft sweeps and found that soft sweeps are surprisingly common in North American Drosophila melanogaster. Among our top ranking candidates are three well-known soft

sweeps at the loci *Cyp6g1*, *Ace*, and *CHKov1*. Recently, Harris *et al.* 2018 have claimed that the selective sweeps we identified are false positives and are more likely to be explained by an admixture demographic history. Moreover, they claim that these neutral non-sweeps are more likely to be hard sweeps rather soft sweeps. Here we re-assess our and Harris *et al*'s work and find that given a reasonably well-fitting demographic model, soft sweeps are a dominant mode of adaptation in North American Drosophila melanogaster.

Introduction

Pervasive adaptation has been extensively documented in *Drosophila melanogaster*. Recent studies suggest that (i) ~50% of amino acid changing and non-coding substitutions in *D. melanogaster* evolution were adaptive, and (ii) there are abundant signatures of adaptation in the population genomic data detectable as reductions of neutral diversity in the regions of higher functional divergence and as elevation in the frequencies of derived alleles above neutral expectations [1–11].

In three cases—at the loci *CYP6g1*, *CHKov1*, and *Ace*–we specifically know the causal mutations and have functional hypotheses for the causes of adaptation [12–18]. Intriguingly, in all three cases, there is strong evidence that adaptation was not driven by a single *de novo* adaptive mutation that rose to high frequency, but rather, multiple adaptive mutations. In the case of *Cyp6g1*, adaptive changes leading to resistance to DDT evolved via multiple insertions of *Accord* transposon in the 5' regulatory region of the locus on different genomic backgrounds, as well as a duplication of the entire locus [12,13]. At the *CHKov1* locus, the adaptive change led to a higher resistance to organophosphates and viral infections and evolved by a transposon element insertion in the protein coding region of *CHKov1*, which then segregated in the ancestral populations before rising to high frequency only recently [14,15]. Finally, resistance to pesticides such as carbamates and organophosphates evolved via multiple independent point mutations at four highly conserved sites in the gene *Ace* on different genomic backgrounds on multiple continents [16–18]. Thus, all three well-understood examples of adaptation are, by definition, known soft sweeps (**Fig 1**) in which multiple adaptive alleles have risen to high frequency simultaneously at the same locus [19–21]. This suggests that recent and strong adaptation is not mutation-limited in *D. melanogaster* [16].

These three empirical examples of soft sweeps at *Ace*, *CYP6g1*, and *CHKov1* were all defined experimentally and suggest that soft sweeps might be common or at least not vanishingly rare in *Drosophila*. However, until recently, it was difficult to systematically assess the frequency of soft sweeps in the Drosophila genome given that most scans for detecting selective sweeps were specifically designed to detect signatures of hard sweeps [2,22–29], making it challenging to assess the frequency of soft sweeps [30–33]. Specifically, most methods are focused on the detection of regions of low diversity or the presence of a single common haplotype expected only for hard and not soft sweeps.

To answer this question, we recently introduced novel haplotype homozygosity statistics for the detection and differentiation of hard and soft sweeps that are capable of (i) detecting both hard and soft sweeps with similar power using the haplotype homozygosity statistic H12 and then (ii) to determining whether the detected sweeps are likely to be either hard or soft using the statistic H2/H1 that is conditioned on first detecting high H12 values [34]. Application of these statistics to the Drosophila Genetic Reference Panel (DGRP) [35], composed of 145 whole-genome sequences from a North Carolina *D. melanogaster* population, revealed several putative sweeps with unusually high haplotype homozygosity relative to expectations
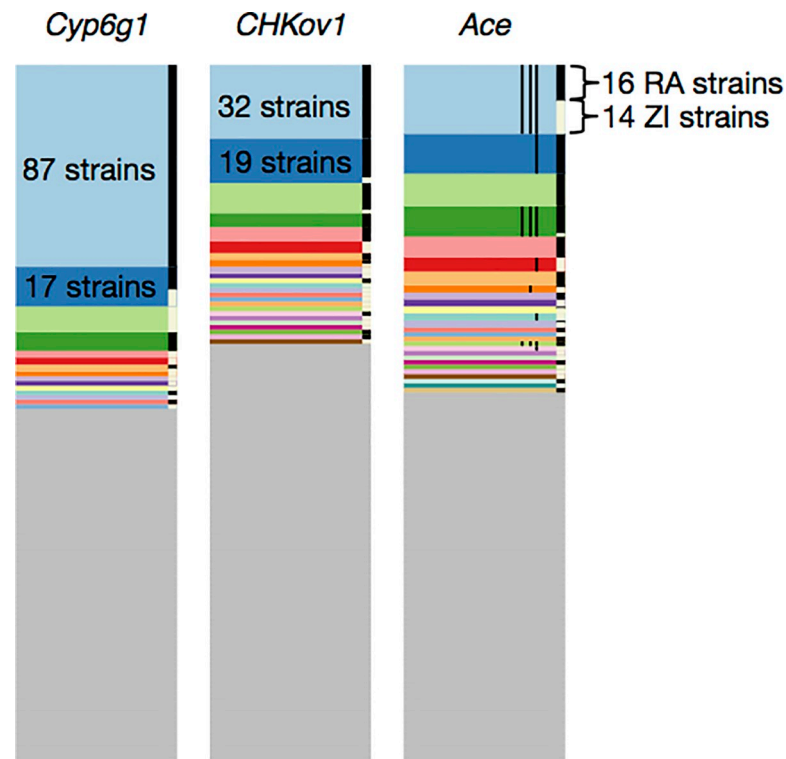
Fig 1. Haplotype frequency spectra at the *Cyp6g1*, *CHKov1*, and *Ace* loci. Recreated from Garud and Petrov 2016 [17]. Haplotype frequency spectra at the three positive controls in a joint dataset, comprised of 300 Raleigh (RA) and Zambian (ZI) strains in 801 SNP windows, centered around the sites of the selective sweeps. 801 SNP windows in the joint data set correspond to slightly smaller analysis window sizes (<10 kb) in terms of base pairs on average than in the Raleigh or Zambian data alone. Each color bar represents a different, unique haplotype, and the height of the bar represents the number of chromosomes sharing the haplotype. The grey bars represent unique, singleton haplotypes in the sample. On the right side of each of the frequency spectra are black and white bars, indicating which strains are from RA and ZI, respectively. At all three positive controls, common haplotypes are shared across the two populations. The thin vertical black lines shown in the haplotype spectrum for *Ace* correspond to the presence of three adaptive mutations that confer pesticide resistance.

under several neutral demographic scenarios (**Fig 2**). The top 50 empirical outliers, which included the rediscovered positive controls at *CYP6g1*, *CHKov1*, and *Ace*, had multiple unusually long haplotypes present at high frequency, consistent with soft sweeps (**Fig 1**). By contrast, site-frequency spectrum statistics like Pi/bp and S/bp did not reveal sharp dips in diversity at the positive controls (**S1** and **S2** **Figs**). We found that simulations of hard sweeps were unable to produce signatures observed in the data, whereas simulations of soft sweeps from *de novo* mutations and standing variation did so naturally. Subsequent studies found that soft sweeps appear to be common not only in this North American population, but also in Sub Saharan populations of *D. melanogaster* [17,18,36]. Finally, these haplotype homozygosity statistics have been applied to several other organisms including pigs [37,38], dogs [37], cattle [39,40], soy beans [41], and humans [42] to identify hard and soft sweeps, and have become standard summary statistics in machine learning methods for detecting selection [43,44].

Harris et al. 2018 [45] recently re-evaluated our analysis of the DGRP data using our statistics and argued that that there was in fact scant evidence for abundant recent strong selective sweeps in the North American *D. melanogaster* population. They claimed that appropriate neutral admixture models naturally generated detected haplotype signatures in the absence of positive selection and thus most of the detected signatures did not in fact correspond to
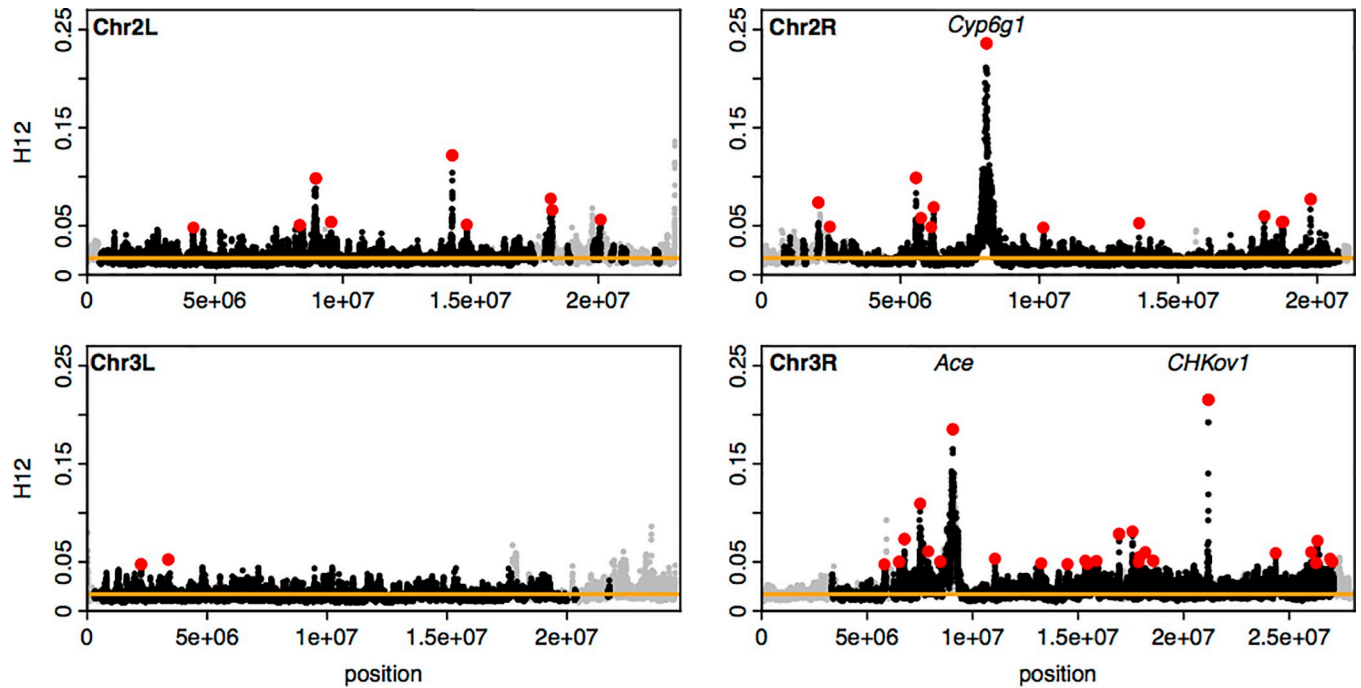
Fig 2. H12 scan of DGRP data. Recreated from Garud *et al*. 2015 [34]. Scan of the four autosomes using the H12 statistic. Each point indicates an H12 value computed in a 401 SNP window. Grey points indicate regions excluded from the analysis with recombination rates lower than 5x10^-7 cM/bp. The orange line represents the 1-per-genome FDR line calculated from simulations of a neutral model with constant population size of 10^6. Red points indicate the top 50 extreme outlier peaks relative to the 1-per-genome FDR line. Three positive controls are indicated at *Ace*, *Cyp6g1*, and *CHKov1*.

selective sweeps at all. They also argued that if these sweeps did exist, then they would be hard rather than soft sweeps.

Here we re-evaluate our own analysis using a range of demographic models and show that our previous findings stand. We then discuss the reasons for the different conclusions from Harris et al. 2018 [45], the best practices for the use of of our haplotype statistics, implications of these additional (re)analyses, and directions for future work.

Results

In this paper we assess the fit of several demographic models to the DGRP data and the ability to identify and distinguish hard and soft selective sweeps. We first summarize our previous findings in Garud *et al*. 2015, and then re-visit the analyses performed in Harris et al. 2018. We also discuss the fit of demographic scenarios proposed by Duchen et al. 2013 and Arguello et al. 2018. Finally, we propose additional demographic scenarios that fit the DGRP data better but not perfectly.

Summary of our previous results in Garud et al. 2015

The increase in frequency of an adaptive allele is expected to also lead to the increase in frequency of the linked haplotype [46–48]. Such an increase of haplotype frequency is expected to elevate levels of haplotype homozygosity (H1) in the vicinity of the selected locus [25,26,28,29,49], where H1 is defined as

$$H1 = \Sigma(p_i)^2,$$

with p_i being the frequency of the ith most common haplotype. While H1 is expected to be

elevated for both hard and soft selective sweeps, the hard sweeps should have higher H1 values than soft sweeps, given that soft sweeps bring multiple haplotypes to high frequency.

In Garud *et al.* 2015 [34] we define a similar haplotype homozygosity statistic, which we denoted H12, which combines the frequencies of the first and second most common haplotypes into a single frequency and is defined as follows:

$$H12 = (p_1 + p_2)^2 + \sum_3^n (p_i)^2$$

Using extensive simulations, we showed that H12 has more equal power to detect hard and soft sweeps than H1 with a slight remaining bias in favor of hard sweeps [34].

The application of these statistics requires one to define a window size. Longer windows should have lower false positive rates for distinguishing selection from neutrality, but simultaneously have reduced ability to detect weaker sweeps (See Box 1 for further discussion on

Box 1: Best practices for applying H12 and H2/H1 to genomic data

The application of H12 and H2/H1 to genomic data comes with many implementation choices that depend on the population genetics of the organism and population being studied. Here we will make our recommendations based on applying our statistic to Drosophila [17,34], humans [54], rats [55], and dogs [37]:

1. **Data**. H12 and H2/H1 are intended for the analysis of phased whole genome data. In the absence of phased data, we recommend using G12 and G2/G1, which are unphased analogs of H12 and H2/H1 that have only slightly reduced power and have been successfully applied to humans [54] and rats [55]. All recommendations made below are applicable to both the phased and unphased versions of the statistics.

2. **Sample size.** The ability to distinguish selective sweeps from neutrality persists until sample sizes are as small as 25 individuals [54]. Smaller samples sizes not only make it difficult to distinguish neutrality from selection, but also to assess the softness of a sweep [20].

3. **Window size.** Longer windows are less likely to generate peaks of homozygosity under neutrality. Simultaneously, however, windows should not be so long such that reasonably strong sweeps cannot be detected, including known cases of recent adaptation. Several diagnostics can help with defining a window size: First, the level of linkage disequilibrium should decay and plateau within the length of the window size to ensure that recombination has had sufficient opportunity to break down linkage in neutral regions. Second, longer windows will bias H12 towards detecting stronger sweeps. Specifically, the footprint of a hard selective sweep extends over approximately $s/[\log(Ne^*s)^*rho]$ base pairs, where *Ne* is the population size and *rho* is the recombination rate. As an example, sweeps with $s = 0.05\%$ are likely to generate sweeps spanning 10kb windows when $rho = 5^*10^{\wedge}-7$. As *rho* increases, only those selective sweeps with $s > 0.05\%$ should be observed in 10kb windows.

To circumvent the issue of defining a window size, another approach is to use H-scan [37] or nSL [56] to identify hard and soft selective sweeps, both of which do not require a window and are capable of detecting hard and soft sweeps. However, there is currently no method associated with these statistics for distinguishing hard and soft sweeps.

Finally, scans such as iHS [28] which are expressly designed to detect partial hard sweeps, may still be powerful for detecting soft sweeps. In Garud et al. 2015 [34], several of our top peaks identified with H12 overlap with peaks identified with iHS.

4. **Windows defined in terms of SNPs versus base pairs.** Windows defined in terms of SNPs are all guaranteed to have the same number of SNPs, which then can be used to define the number and frequencies of haplotypes in a window. These SNP-based windows are fully capable of detecting complete hard sweeps [34,37]. SNP-based windows generally are longer in terms of base pairs in regions of low diversity. These longer windows should then have proportionately elevated recombination rates, thereby reducing H12 and making the selection scan more conservative.

Alternatively, windows can be defined in terms of base pairs. However, with this approach, some windows may lack genetic variation due to bottlenecks and other demographic forces, drift, background selection, or low recombination rates. These low diversity windows will generally have fewer haplotypes and might be misinterpreted as sweeps. To ensure that this does not contribute to the false positive rate, windows with extremely low diversity should be excluded from the scan (as was done in [54]), and a permutation analysis in different recombination rate categories can ensure that low recombination is not a driver of low diversity and thus high haplotype homozygosity.

Finally, windows could be defined in terms of centimorgans. The advantage of this is that recombination rates per window are constant across the genome. However, like base-pair defined windows, there may be varying power to detect selection of different strengths due to differing numbers of SNPs in each window.

5. **Peak calling.** To assess whether the observed H12 values calculated in the data are unusually high compared to neutral expectations, a distribution of H12 values can be simulated under realistic demographic models with a program such as MS [57] or SLiM [58]. We advocate for simulating 10x the number of analysis windows observed in the data, and then identifying the tenth highest H12 value in the resulting distribution as a cutoff value ($H12_o$). With this cutoff, 1 false positive per genome can be expected. To call individual sweeps, first all windows with $H12 > H12_o$ are identified. Then consecutive windows with $H12 > H12_o$ are grouped together into a 'peak' as they may belong to the same selective event'. Indeed, if a window truly has high homozygosity due to a sweep, then neighboring windows should provide supporting evidence. The window with the highest H12 value among all windows in a peak is used to represent the H12 values of the entire peak.

6. **Recombination.** Regions of low recombination should be excluded *after* peak calling to avoid including spurious selective sweeps in the candidate set because low recombination can result in high homozygosity. Permutations may be needed to ensure that recombination is not depressed near the edges of chromosomes [59]. Organisms with high rates of recombination are ideal candidates for performing a scan with H12 because selective sweeps will be more easily distinguishable from the rest of the genome.

7. **Choice of demographic model.** How do you know if your demographic model fits the data well? In this paper, we advocate for a fit to multiple summary statistics

including site frequency spectrum (e.g. Pi/bp and S/bp) and LD statistics. However, many organisms and populations lack a well-fitting demographic model. One option is to fit a simple model using software such as DaDi [60] or PSMC [61].

However, all demographic models are probably incorrect to some extent (including the ones presented in this paper). Despite this, the analysis does not need to hinge on a demographic model. First, irrespectively of the model availability, we advocate focusing on the extreme outliers of the scan, as these candidates are the least likely to be sensitive to the choice of demographic model. Second, if known positive controls are among the outliers of the scan, then, some confidence can be placed on there being true positives among the candidate list. Finally, functional classes that *a priori* are hypothesized to be under selection (e.g. Human viral interacting proteins [62,63]) can be tested for elevated H12 values compared to a matched control set.

8. **Application of H2/H1.** Visual inspection of the haplotype frequency spectrum of the top peaks can lend intuition as to whether a sweep resembles a hard or soft sweep. Application of H2/H1 provides a more quantitative approach for assessing the softness of a sweep. When applying H2/H1, it is necessary to condition on high H12 values because otherwise there is no evidence for a sweep to begin with. Additionally, H12 imposes an upper bound on H2/H1, making the two statistics dependent on one another [53]. Thus, *both* H12 and H2/H1 are needed to distinguish hard and soft sweeps since the magnitude of H2/H1 alone provides insufficient information.

As an example, both neutrality and soft sweeps can generate elevated H2/H1 values, but only soft sweeps can also generate elevated H12 values (**Fig 6**). In fact, in an extreme scenario, if every haplotype is unique, then H2/H1 approaches 1 because, homozygosity does not change substantially when excluding the most common haplotype. Therefore, high H2/H1 does not indicate anything interesting unless H12 is high as well.

A sweep can be categorized as hard versus soft by computing the number of simulations that generate matching H12 and H2/H1 values under the two scenarios. Since the exact evolutionary history of each sweep is unknown, the selection strength, partial frequencies, and ages of the simulated sweeps can be drawn from uniform priors. Given the integration over a large range of evolutionary scenarios, a Bayesian inference is most appropriate for quantifying the fit of a hard versus soft sweep model, since a frequentist inference necessitates a point hypothesis test of one evolutionary scenario at a time.

window size). In Garud *et al.* 2015 [34] we used windows of 401 SNPs in length (~10Kb in Drosophila), and we show that this biases our analysis towards detecting sweeps with selection coefficients (s) > = 0.1%. Note that most of the detected sweeps span multiple analysis windows with the peaks ranging from ~11kb to ~870 kb, and with half of the peaks over 100kb (**Fig 2**), suggesting that the identified sweeps were in fact driven by selection substantially stronger on average than $s = 0.1\%$.

We tested a range of neutral models fitting overall polymorphism levels in the data and found that these rarely generate elevated values of H12 on such long length scales [34]. We specifically considered six models of increasing complexity (**Fig 3**). We included four simple models that were fit to site frequency spectrum-based summary statistics measured from short
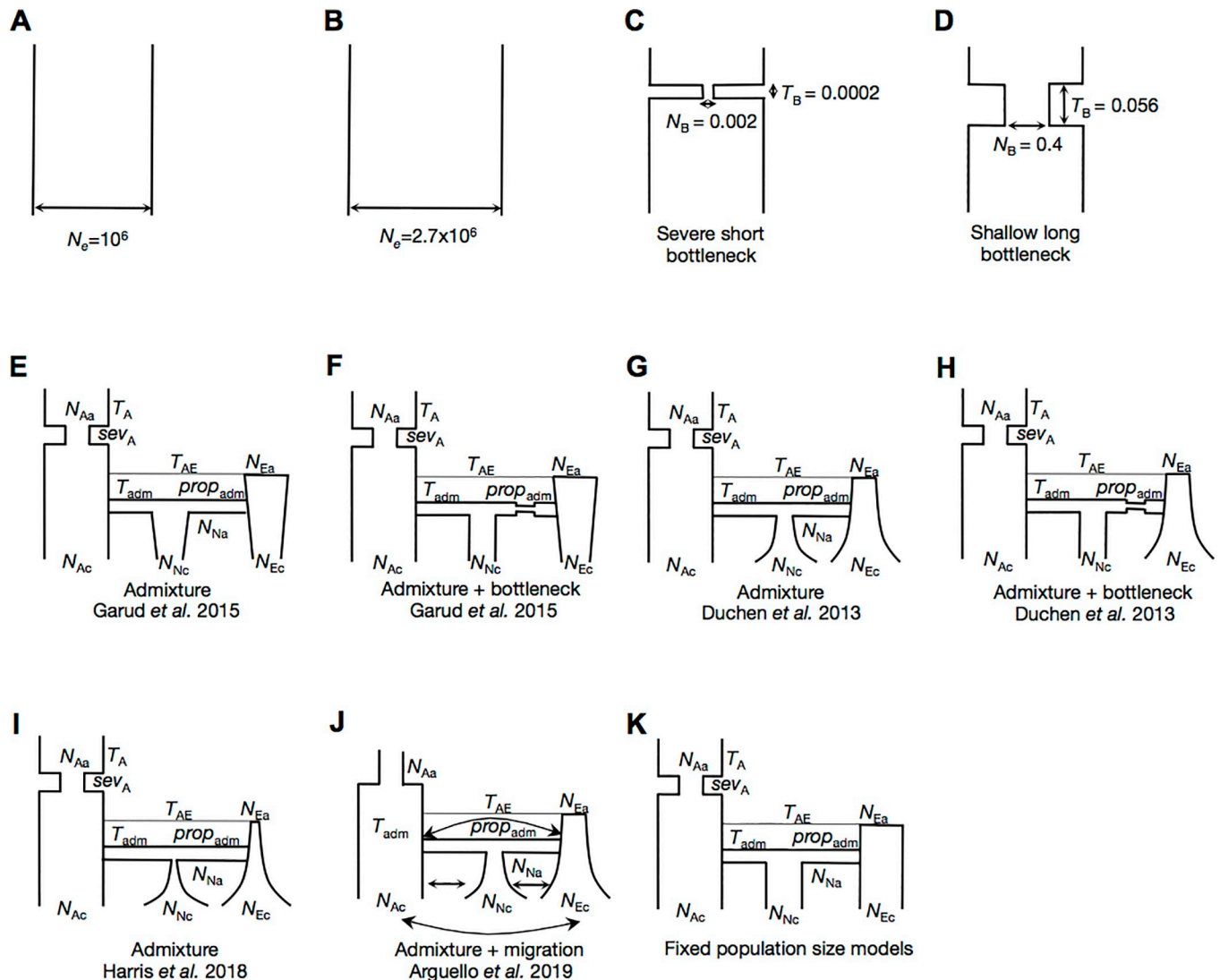
Fig 3. Neutral demographic models. Diversity statistics were measured in simulations of 11 neutral demographic models: (A) A constant $N_e = 10^6$ model (B) A constant $N_e = 2.7 \times 10^6$ model (fit to Watterson's θ_W measured in autosomal short introns in DGRP data) C) A severe short bottleneck model fit to Pi and S in autosomal short introns in DGRP data (D) A shallow long bottleneck model fit to Pi and S in autosomal short introns in DGRP data (E) The implemented admixture model in Garud *et al.* 2015 (F) The implemented admixture + bottleneck model in Garud *et al.* 2015 (G) The admixture model proposed by Duchen *et al.* 2013 (H) The admixture + bottleneck model proposed by Duchen *et al.* 2013 (I) The implemented admixture model in Harris *et al.* 2018 (J) The admixture model proposed by Arguello *et al.* 2019 (K) A variant of the Duchen *et al.* 2013 admixture model where North America, Europe, and Africa have fixed population sizes.

introns in the DGRP data: two constant population size models ($Ne = 10^6$ and $Ne = 2.7 * 10^6$) and two bottleneck models with varying bottleneck durations and sizes. Finally, we included two complex admixture models inferred by Duchen et al. 2013 [50] using an approximate Bayesian computation (ABC) approach and data from 242 short intronic and intergenic fragments from the X-chromosome. These models were fit to both site frequency spectrum-based summary statistics (number of segregating sites, S/bp, and average nucleotide diversity, Pi/bp) and LD measured on short length-scales (~500bp) using Kelly's ZnS. Using their ABC method, Duchen et al. 2013 [50] inferred a posterior distribution for each of the 11 parameters for the admixture models.

Reanalysis of Garud et al. 2015 and Harris et al. 2018

Upon revisiting the Duchen et al. 2013 [50] models for the present paper, we found that the model implemented in Garud et al. 2015 [34] was a variant of the model published in Duchen et al 2013 [50]. Instead of an expansion in North America and Europe, we had implemented roughly constant population sizes in North America and Europe (**Fig 3E and 3F**). Despite this difference from the published model, all six implemented models fit the autosomal DGRP data in terms of S/bp, Pi/bp and decay in short-range LD (**Figs 4 and S3, S6, S9, and S12,** which show the full distirbutions of the statistics in data versus simulations, as well as the associated quantile-quantile plots of the fit of the distributions). Both long-range LD (~10Kb) **Fig 4D)** and H12 (**Figs 5** and **S15**), in simulated models were depressed compared to the observed data. Specifically, the data showed much slower decay of LD on the scale of 10kb and substantially larger median haplotype homozygosity (H12).

The distribution of H12 in the data also had a much longer tail compared to admixture simulations in Garud et al. 2015 (**Figs 5 and S15, S18, and S21**). Indeed, the observed median values of H12 in the data are similar to levels expected only once in the genome under these neutral models. We interpreted the elevation of long-range LD and haplotype homozygosity in long windows in the data as evidence of positive selection. Supporting this point, the H12 values in data do not fit a Gaussian distribution well due to the elevated tails (**S21C Fig**). However, the bulk of the distribution within 1SD around the median can more reasonably fit a Gaussian. As a point of contrast, in **S21D Fig**, we find that the full distribution of a simulated constant Ne model (not just the bulk) fits a Gaussian well.

The lack of the model fit with the bulk of the H12 values in the data presents a problem for identifying selective sweeps with elevated homozygosity. Thus, we elected to be conservative. First, we defined a 1-per-genome false discovery rate for each demographic model by performing 1.3×10^5 simulations (>10 times the number of analysis windows observed in the data) for each model. The FDRs, corresponding to the 10th highest H12 value in the distributions, were approximately equal to the median H12 value observed in the data. Several genomic regions had 'peaks' of elevated H12 values that were especially unlikely to be generated by neutrality. These regions corresponded to our candidate selective sweeps (**Fig 2**). We then focused on the 50 empirical outliers to further characterize as hard or soft. These peaks were defined by identifying the window with the highest H12 value and finding all consecutive windows in both directions with H12 values exceeding the 1-per-genome FDR. These candidates all had maximum H12 at least eleven standard deviations away from the median H12 value in the data after fitting a Gaussian distribution to the bulk of the data (**S21 Fig**) (Methods). The top 3 outliers were the positive controls, *Ace*, *Cyp6g1*, *CHKov1*, confirming that H12 has the ability to detect known soft selective sweeps that arose from multiple *de novo* mutations or standing genetic variation.

Can admixture generate elevated haplotype homozygosity?

Harris et al. [45] claim that the admixture model proposed by Duchen et al. 2013 [50] can easily generate all the elevated H12 values in the data, suggesting that the selective sweeps identified by H12 are false positives.

Given that our original implementation of the admixture model in Garud et al. 2015 [34] was a variant of the Duchen et al. 2013 [50] model, we tested Harris et al.'s [45] claim by implementing the model specified in their supplement, which also differs from the Duchen et al. 2013 [50] model (methods). Despite our best efforts, our implementation of the Harris et al. model does not generate the summary statistics of S/bp, Pi/bp, and H12 presented in Harris et al. The supplemental document released by Harris et al. provides a template but not the

Legend:

- DGRP data
- Constant $Ne=10^6$
- Constant $Ne=2.7*10^6$
- Severe, short bottleneck
- Shallow, long bottleneck
- Admixture (mode), Garud *et al.* 2015 implementation
- Admixture + bottleneck (mode), Garud *et al.* 2015 implementation
- Admixture (mode), Duchen *et al.* 2013
- Admixture + bottleneck (mode), Duchen *et al.* 2013
- Admixture (posterior distribution), Duchen *et al.* 2013
- Admixture (posterior distribution), Harris *et al.* 2018 implementation
- Arguello *et al.* 2019
- Admixture, Eur=$7*10^5$, Nac=$1.11*10^6$
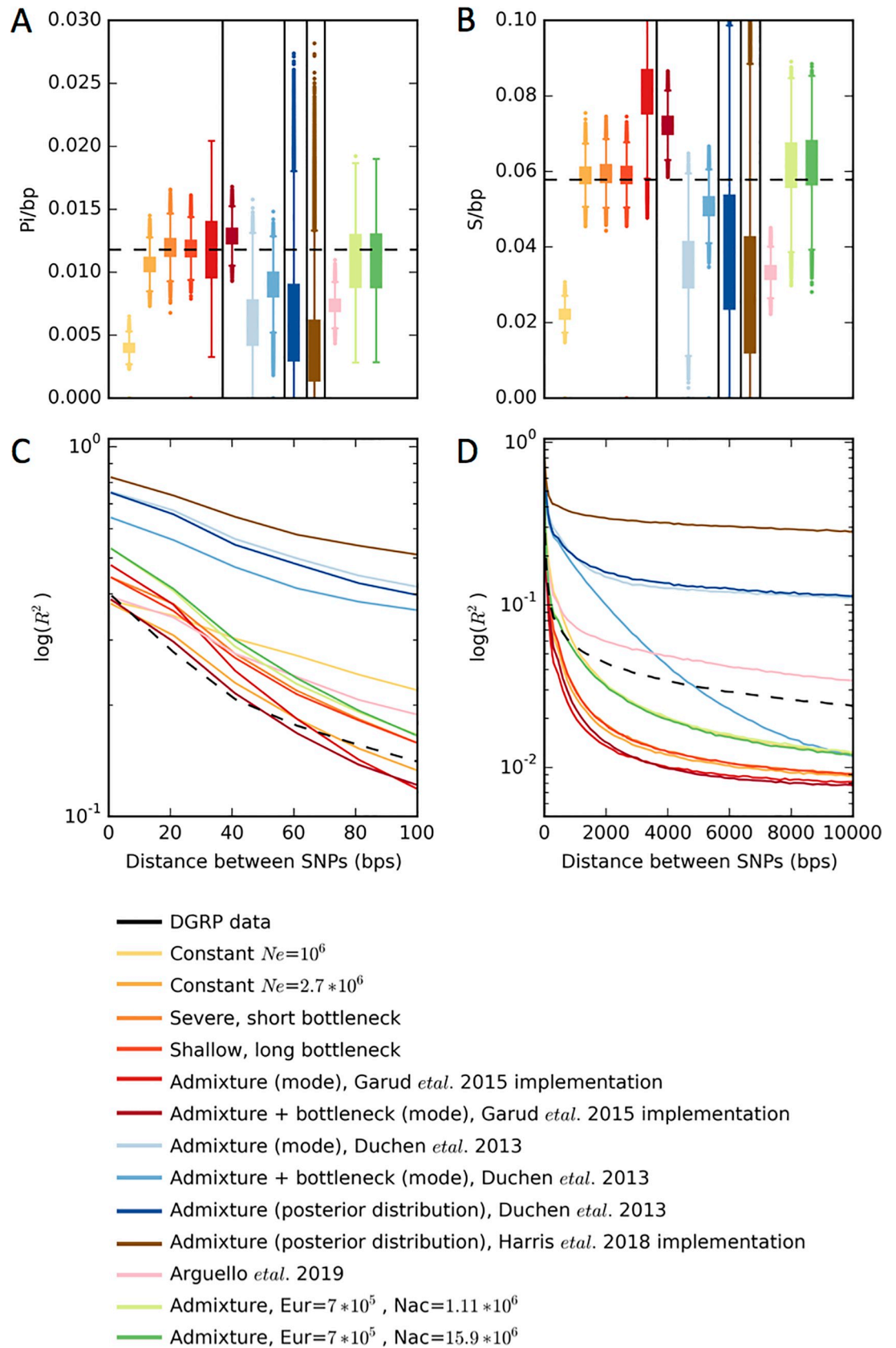- Admixture, Eur=$7*10^5$, Nac=$15.9*10^6$

Fig 4. Distributions of Pi, S, and linkage disequilibrium in data and simulations. Distributions of (A) Pi/bp, (B) S/bp, (C) short range LD (R^2), and (D) long range LD measured in DGRP data and simulated neutral demographic models. In Figures A and B, models belonging to the following categories are delineated with a vertical line: models implemented in Garud et al., models implemented in Duchen et al, model specified by Harris et al., the model proposed by Arguello et al, and finally, models proposed in this paper. Simulations were generated with a recombination rate $\rho = 5 \times 10^{-7}$ cM/bp. Diversity statistics were calculated in DGRP data in genomic regions with $\rho \geq 5 \times 10^{-7}$ cM/bp. The horizontal dashed lines in (A) and (B) depict the median Pi/bp, S/bp, and H12 values measured in DGRP data. For each model, statistics from 1.3×10^5 simulations are plotted in (A) and (B). The dashed black lines in (C) and (D) correspond to mean LD values computed in DGRP data. LD in simulations was estimated from 1×10^7 pairs of SNPs. Histograms and quantile-quantile plots of the full distributions of Pi/bp and S/bp are shown in **S3**–**S14 Figs**.

https://doi.org/10.1371/journal.pgen.1009373.g004

actual code used. Given the lack of code, it is impossible to track down the exact source of the discrepancy.

To more broadly consider the Harris et al.'s [45] claim that appropriate demographic models can easily generate the distribution of H12 values observed in the data, we also tested the Duchen et al. 2013 [50] model both with the mode of the posterior distributions of the 11 parameters for the admixture model and by drawing parameters from the 95 CIs of the posterior distributions, as was done in Harris et al. 2018 [45]. Note that Harris et al. 2018 [45] did not use a joint posterior distribution, and thus nor did we in our implementation of the Harris et al. 2018 model, leading to the possibility that many of the parameter combinations may not correspond to realistic scenarios. We also tested a variant of the admixture model proposed by Duchen et al. 2013 [50] that included a bottleneck in the founding European population (**Fig 3H**). Finally, we also tested a new admixture model with migration between Africa, Europe, and North America (**Fig 3J**) recently proposed by Arguello et al 2019 [51].

The Duchen et al. 2013 [50] model and Harris et al. 2018 [45] implementation of the model generate S/bp and Pi/bp values that are 2-fold lower than the median values measured from short introns in the DGRP data (**Figs 4 and S4, S7, S10, and S13**). More strikingly, however, H12 is extremely elevated compared to values observed in the DGRP data (**Figs 5 and S16 and S18**), e.g the bulk of the distribution of values generated in simulations is non-overlapping with the bulk of the distribution of values from genome-wide data. This elevation is likely due to the sharp bottlenecks specified in the models, especially in the Harris et al. 2018 implementation where the bottleneck size is 4 times smaller than reported by Duchen et al. 2013 [50]. The elevation is even more pronounced when drawing parameters from the 95 CIs. In some cases, H12 almost approaches 1, implying that these models predict essentially no variation in the DGRP in the large ~10kb window sizes used for our analysis. Consistent with elevated haplotype homozygosity, the Duchen et al. 2013 [50] models produced elevated pairwise LD compared to observations in the data (**Fig 4C and 4D**). The mismatch between the expected values of all the statistics in the Arguello et al. 2019 [51] model and the data is less pronounced compared to the Duchen et al. 2013 [50] model, presumably because migration replenishes some of the diversity lost during the extreme bottlenecks, but nevertheless there is a significant mismatch.

At first glance, the elevated haplotype homozygosity produced by the Duchen et al. 2013 [50] model might suggest that the peaks observed in the DGRP data (**Fig 2**) could be explained by the admixture model. However, the S/bp, Pi/bp, H12, and LD values produced by this admixture model deviate significantly from genome-wide summary statistics in the data. In particular, the distribution of H12 values in the data has a very specific distribution that the simulations of neutrality cannot match (**Figs 5 and S15**–**S21**). Almost 80% of the analysis windows in the DGRP data have H12 values within 2 standard deviations from the median, after fitting a Gaussian to the bulk of the distribution (**Methods, S21 Fig**). This is followed by a long tail of H12 values that includes the values for the top 50 peaks, which are > = 11 standard

DGRP data
Constant $Ne=10^6$
Constant $Ne=2.7*10^6$
Severe, short bottleneck
Shallow, long bottleneck
Admixture (mode), Garud *et al.* 2015 implementation
Admixture + bottleneck (mode), Garud *et al.* 2015 implementation
Admixture (mode), Duchen *et al.* 2013
Admixture + bottleneck (mode), Duchen *et al.* 2013
Admixture (posterior distribution), Duchen *et al.* 2013
Admixture (posterior distribution), Harris *et al.* 2018 implementation
Arguello *et al.* 2019
Admixture, Eur=$7*10^5$, Nac=$1.11*10^6$
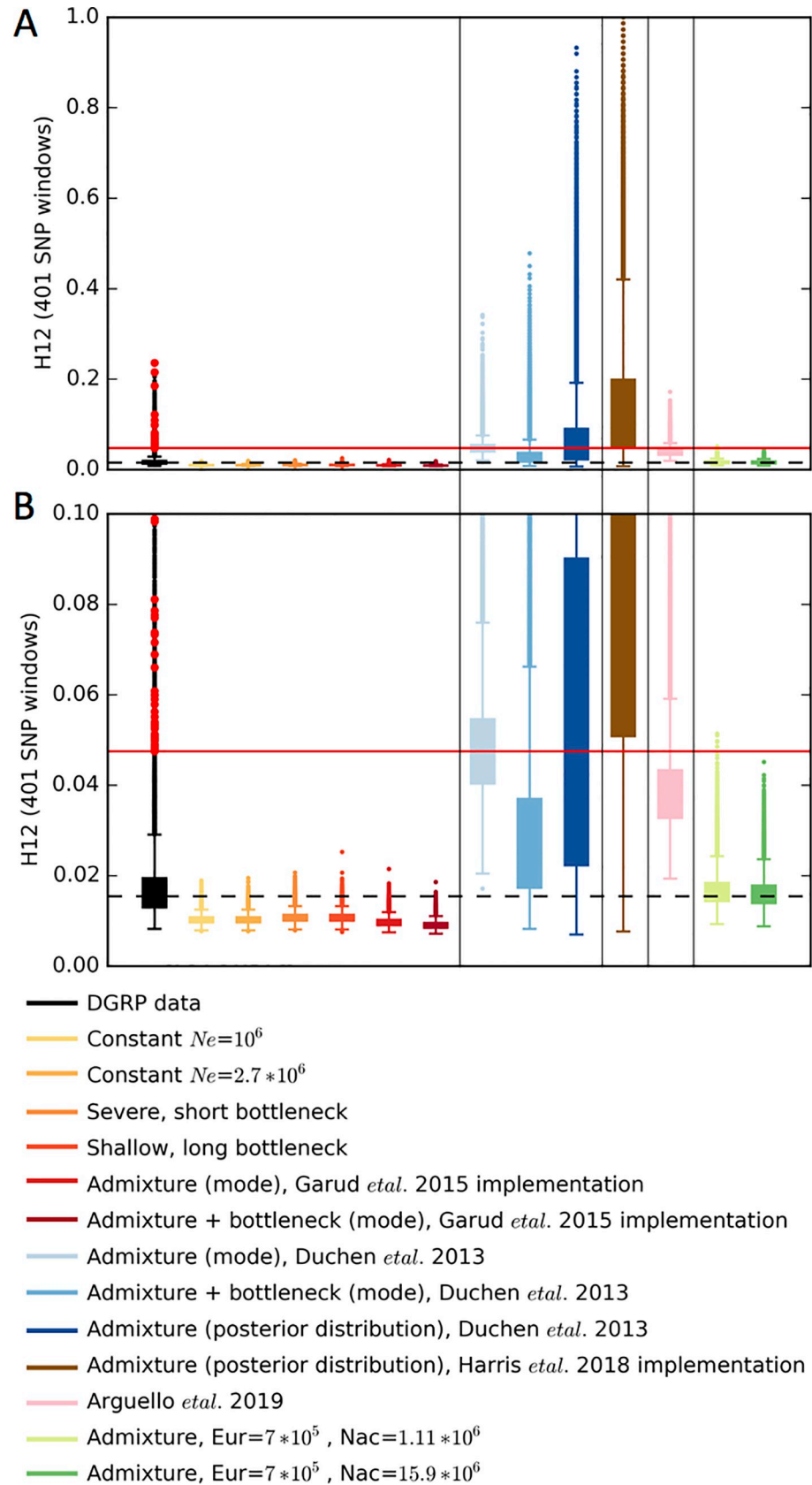Admixture, Eur=$7*10^5$, Nac=$15.9*10^6$

Fig 5. H12 distributions in data and simulations. Distributions of H12 in 401 SNP windows. Shown are the (A) full distribution and (B) truncated y-axis for visual clarity. Simulations were generated with a recombination rate $\rho = 5\times10^{-7}$ cM/bp and H12 was calculated in DGRP data in genomic regions with $\rho \geq 5\times10^{-7}$ cM/bp. The horizontal dashed line indicates the median H12 value in DGRP data and the horizontal red line indicates the lowest H12 value for the top 50 peaks. H12 values from 1.3×10^{5} simulations for each model are plotted. The distribution of genome-wide H12 values measured in DGRP data is shown in black. Overlaid in red points are the H12 values corresponding to the top 50 empirical outliers in the DGRP scan. Histograms and quantile-quantile plots of the full distributions of H12 in data and simulations are shown in **S15–S21 Figs**.

https://doi.org/10.1371/journal.pgen.1009373.g005

deviations away from the median using the Gaussian fit, indicating that these peaks are indeed genome-wide outliers. By contrast, the bulk of the distribution generated by the Duchen et al. 2013 [50] admixture model surpasses the median and bulk of the distribution of H12 values in the data (**Figs 5 and S16 and S19**). This lack of fit of the admixture model to the data is problematic for the inference of selective sweeps: if the tail of the distribution of H12 values from data can be explained by neutrality, then the bulk of the distribution should also be explainable by neutrality. These admixture models do not recapitulate the distribution observed in the data, and instead produce extremely high levels of homozygosity that are incompatible with the data.

One reason for the lack of fit of H12 measured in the data and the simulated admixture model could be that Duchen et al. 2013 [50] initially fit the model to the 242 X-chr fragments of ~500bps using SFS statistics and short-range LD statistics (Kelly ZnS), whereas we analyzed autosomal data. Although Duchen et al. 2013 [50] showed that the model extrapolated to autosomes by fitting to ~50 intronic and intergenic regions on the 3rd chromosome, the models do not fit diversity patterns on short introns, which are putatively the most neutral part of the genome [52]. Additionally, Duchen et al. 2013 [50] did not require that long-range haplotype homozygosity on the scale of ~10kb fit the data, which is the main source of discrepancy between the models and the data.

Models that fail to recapitulate the bulk of the diversity statistics from the data are unlikely to accurately capture the true demographic history of the population. These models are not appropriate for inferring sweeps because they do not set a realistic baseline for the expected diversity pattern in a neutral scenario.

Inference of new demographic models

In this section, we test whether there are variants of the Duchen et al. 2013 [50] and Arguello et al. 2019 [51] admixture models that can achieve a better fit with regard to multiple relevant summary statistics in the DGRP data. Our goal here is to assess whether we can find an admixture model that reasonably fits both SFS and LD-based genome-wide statistics in the data and can also generate tails of elevated H12 values that may explain the outlier peaks observed in the data. We do not claim that other models cannot fit the data equally well or better.

We tested four classes of variants of the Duchen et al. 2013 [50] and Arguello et al. 2019 [51] admixture models (**S22 Fig**). First, we tested models with constant population sizes in North America and Europe (**Figs 3K** and **S22A**), because the Garud et al. 2015 [34] implementation (**Fig 3E**), which had effectively constant population sizes for these two populations, fit the data well in terms of S/bp and Pi/bp. Second, we tested models with varying amounts of growth in North America and Europe (**S22B Fig**). Third, we tested models with varying proportions of admixture (**S22C Fig**), and fourth, we tested models with varying amounts of migration between the continents (**S22D Fig**). For each of these models, we held almost all parameters constant at the mode of the posterior distributions inferred by Duchen et al. 2013 [50]. The only parameters we varied were those relevant to the model being tested (e.g.

proportion of admixture, amount of migration, or rate of growth). Where applicable, the values of these parameters were chosen to span the ranges of the 95CI inferred by Duchen et al. 2013 [50]. These variable parameters are highlighted in red in **S22 Fig**. In sum, we tested a total of 74 admixture model variants. Supplemental **S23**–**S35 Figs** show the distributions of summary statistics S, Pi, H12, short-range LD and long-range LD generated by these models.

The majority of the models tested do not fit the data well, whereby the median values of S/bp, Pi/bp, and H12 measured in the data lie outside the 25th and 75th quantiles measured from simulations (**Figs 4 and 5 and S23**–**S35**). Models that produce extremely high H12 values and low S and Pi values generally have small founding population sizes. Models with depressed H12 values and elevated S and Pi values have larger founding population sizes. Many models fit some summary statistics reasonably well, but no single model fits all five summary statistics.

Only 3 of the 74 models we tested generate distributions overlapping the median genome-wide S, Pi, and H12 values (**S24 Fig**). These models have constant population sizes in North America and Europe (**Figs 3K and S22A**) of magnitudes similar to the one implemented in Garud et al. 2015 [34]. Specifically, the well-fitting models have large North American population sizes ($> = 1.11*10^6$), and intermediate European population sizes ($\sim 0.7*10^6$). The distributions of S, Pi, H12, and LD for two of these models are shown in **Fig 4** as a comparison with all other models considered in this paper so far. Additionally, in **S5**, **S8**, **S11**, **S14**, **S17**, and **S20 Figs**, the full distribution of Pi/bp, S/bp, and H12 in the data vs simulations are plotted and the fit is quantified with a quantile-quantile plot. Generally, the root mean squared errors (RMSE) comparing the observed distribution with the simulated distributions are among the lowest for these new, fitted models. In fact, the admixture model implemented in Garud et al. 2015 (**S3E**, **S6E**, **S9E**, **S12E**, **S15E**, and **S18E Figs)** has comparably low RMSEs for Pi/bp, S/bp and H12 reflecting that the original model we utilized in our analyses fit the data well in terms of multiple summary statistics.

While the three well-fitting models generate Pi/bp, S/bp, and H12 values that overlap the median values measured from genome-wide data, they cannot generate long tails of elevated H12 values. The 1-per-genome FDR values observed in simulations for these models have H12 values that are lower than even the 50th ranking peak in the DGRP H12 scan. This suggests that given a reasonably well-fitting model, the top 50 H12 peaks observed in the DGRP data are still outliers under any of the current models.

Distinguishing hard versus soft sweeps with the H2/H1 statistic

In Garud et al. 2015 [34], we analyzed whether the haplotype patterns observed among the top 50 peaks are more consistent with hard or soft sweeps. First, we visually inspected the haplotype frequency spectra for the top 50 peaks (**Figs 6 and S36**) and observed that multiple haplotypes are present at high frequency for most peaks, including the three positive controls, *Ace*, *Cyp6g1*, and, *CHKov1*. To gain better intution about whether the observed haplotype spectra are expected under hard versus soft sweeps, we simulated recent hard and soft sweeps of varying selection strengths and plotted their haplotype frequency spectra (**Fig 6**). We find that the observed data most closely resembles the frequency spectra generated by soft sweeps and not hard sweeps.

To provide a more quantitative assessment of the likelihood of the observed data being generated by a hard versus soft sweep, in Garud et al. 2015 [34] we introduced a second haplotype homozygosity statistic, which we denoted H2/H1, to distinguish hard from soft sweeps.

H2 is haplotype homozygosity computed excluding the most common haplotype:
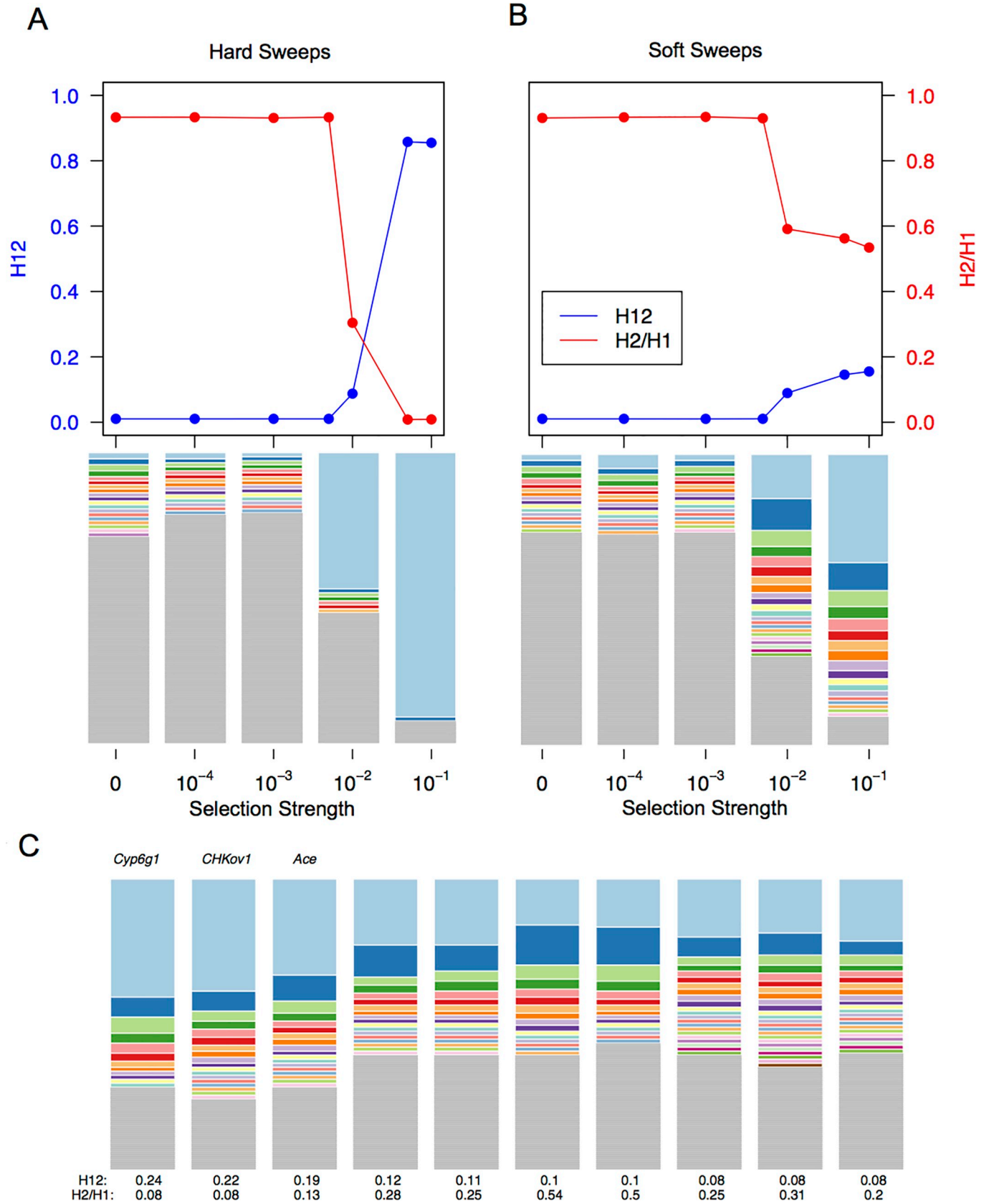
$$H2 = \sum_{2}^{n}(p_i)^2$$

Fig 6. Signatures of hard and soft sweeps in simulations and DGRP data. (A) Top panel: H12 and H2/H1 values associated with hard sweeps simulated with varying selection strengths in a constant $Ne = 2.7*10^6$ model. Each point represents the mean H12 or H2/H1 value for 2000 forward simulations in which selection began $0.0001*Ne$ generations ago. Bottom panel: haplotype frequency spectra for a random simulation for a given selection scenario. (B) same as (A) except for soft sweeps. (C) Haplotype frequency spectra for the top 10 peaks in DGRP data. The analysis window with the highest H12 value for each peak is plotted.

H2 is expected to be small for hard sweeps because the main contributing haplotype to homozygosity is excluded. However, it is expected to be larger for soft sweeps since there should be multiple adaptive haplotypes at high frequency. H2/H1 augments our ability to distinguish hard and soft sweeps since it is even smaller for hard sweeps and larger for soft sweeps than H2.

As **Fig 6** shows, while hard sweeps and neutrality cannot easily generate both high H12 and H2/H1 values, soft sweeps can. Hence, the H2/H1 statistic is powerful for discriminating hard and soft sweeps only when applied to candidate selective sweeps with H12 values exceeding expectations under neutrality. Additionally, H2/H1 is inversely correlated with H12 values [53]. Thus, H12 and H2/H1 must be jointly applied when H12 is sufficiently high to make inferences about the softness of a sweep.

In Garud et al. 2015 [34], we tested whether the H12 and H2/H1 values for the top 50 peaks are more consistent with hard versus soft sweeps. Specifically, we categorized sweeps as hard versus soft by computing Bayes factors: BF = P(H12obs, H2obs/H1obs | soft sweep)/P(H12obs, H2obs/H1obs | hard sweep), whereby H12obs and H2obs/H1obs were computed from data, and hard and soft sweeps were simulated by drawing partial frequencies, selection strengths, and ages from uniform prior distributions (Methods). By using a Bayesian approach, we can then integrate over a wide range of evolutionary scenarios instead of testing a single point hypothesis. In Garud et al. 2015 [34], we found that the top 50 peaks have H12 and H2/H1 values more consistent with soft sweeps than hard sweeps under both constant Ne models and the Garud et al. 2015 implementation of the admixture models. We repeated the BF analysis (Methods) with the new admixture models (**Fig 3K**) inferred in this paper and found that our original findings stand: the majority of the sweeps are classified as soft and ~3–4 are classified as hard (**Fig 7**).

To make their argument that H2/H1 does not have power to distinguish hard and soft sweeps, Harris et al. 2018 [45] assessed whether the top 50 peaks have H2/H1 values consistent with hard versus soft sweeps, even though they do not find evidence for selection at these sites. In their Fig 1D, they conclude that H2/H1 does not have discriminatory power. However, the H2/H1 values for the top 50 peaks lie within the bulk of the distribution generated by soft sweeps and in the tail of the distribution generated by hard sweeps (see their Fig 1D), which appears at odds with their conclusion. Despite their claims that H12 and H2/H1 lack discriminatory power, Harris et al. 2018 [45] also computed Bayes factors (BF) in their S1 Fig and showed that after correctly conditioning on matching H12 and H2/H1 values for the top 50 peaks, the majority of the peaks have values that are consistent with soft sweeps. Thus, Harris et al. 2018 [45] obtain the same result as in Garud et al. 2015 [34].

Discussion

Whether hard or soft sweeps are more common is a topic of much debate. While multiple empirical studies have revealed evidence for soft sweeps in a wide range of organisms including *D. melanogaster*, *P. falciparum* [64,65], viruses [66], humans [42,54,56,67,68], dogs [37], amongst others [21], several articles claim that there is unfounded enthusiasm for soft sweeps and that in fact, they are not as pervasive as the evidence suggests [45,69,70]. Specifically, Harris et al. 2018 [45] suggest that the claim made in Garud et al. 2015 [34]—that there is abundant
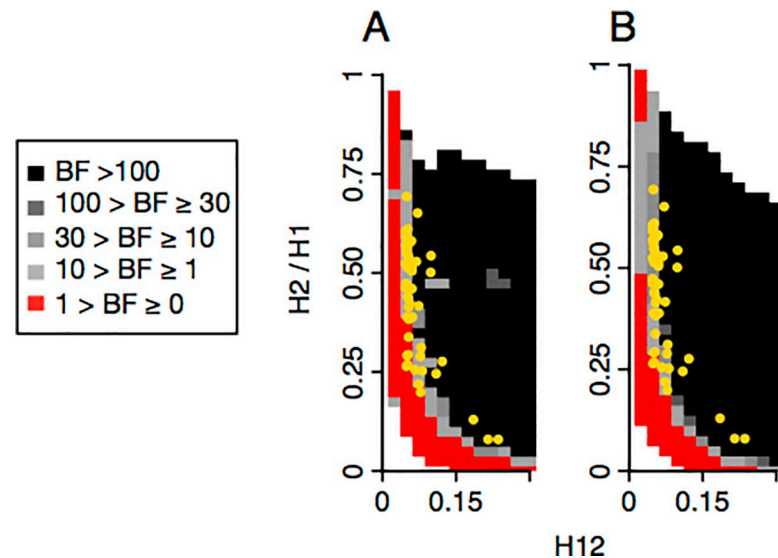
Fig 7. Range of H12 and H2/H1 values expected for hard and soft sweeps under two admixture models. Bayes factors (BFs) were calculated for a grid of H12 and H2/H1 values to demonstrate the range of H12 and H2/H1 values expected under hard versus soft sweeps. Panels A and B show results for variations of the admixture model proposed by Duchen *et al.* 2013, where Africa, North America, and Europe have constant population sizes. In (A), the population sizes for North America and Europe were held constant at 1,110,000 and 700,000 individuals, respectively. In (B), the population sizes for North America and Europe were held fixed at 15,984,500 and 700,000 individuals, respectively. BFs were calculated by computing the ratio of the number of soft sweep versus hard sweep simulations that were within a Euclidean distance of 10% of a given pair of H12 and H2/H1 values. Red portions of the grid represent H12 and H2/H1 values that are more easily generated under hard sweeps, while grey portions represent regions of space more easily generated under soft sweeps. Each panel presents the results from 10^5 hard and soft sweep simulations, respectively. Hard sweeps were generated with $\theta_A = 0.01$ and soft sweeps were generated with $\theta_A = 10$. A recombination rate of $\rho = 5 \times 10^{-7}$ cM/bp was used for all simulations. The H12 and H2/H1 values for the top 50 empirical outliers in the DGRP scan are overlaid in yellow.

https://doi.org/10.1371/journal.pgen.1009373.g007

evidence for many strong and recent soft sweeps in the *D. melanogaster* populations—is not supported after appropriate applications of demographic models. Here we carry out a range of additional analyses and reassert the claims of Garud et al. 2015 [34].

In Garud et al. 2015 [34], we developed the haplotype homozygosity statistics H12 and H2/H1 to systematically detect and differentiate hard and soft sweeps from population genomic data. Our application of these statistics to the DGRP data from North Carolina revealed that soft sweeps are common in this population. Among the top candidates in our scan were the positive controls at *Ace*, *Cyp6g1*, and *CHKov1*. Corroborating our results, we found that approximately half of our sweep candidates were also identified with the popular *iHS* statistic [28], which was designed to detect partial hard sweeps. Finally, we found that soft sweeps are common in a Zambian population as well [17], suggesting that any particular demographic history for a given population is not driving the signal of multiple haplotypes at high frequency. Independently, Sheehan and Song (2018) [44] found that soft sweeps are prevalent in this same population of Zambia.

To ensure low false positive rates, we excluded related individuals and tested for substructure. Additionally, we utilized large analysis window sizes of 401 SNPs (corresponding to ~10kb), since haplotypes of such length are not expected to be at high frequency by chance. Note that we use windows of constant number of SNPs to avoid the issue of H12 co-varying with the number of SNPs in the window. Windows defined in terms of the number of SNPs automatically extends the physical lengths of the windows in regions of low diversity. Such longer windows should then have proportionately higher recombination rates reducing the

expected H12, and thus reducing the probability of false positives (see Box 1). In practice, we also eliminated regions of low recombination (rho <5*10^-7 cM/bp) from the data, as regions of low recombination rate can show elevated false-positive rates in haplotype-bases tests of selection [71]. By contrast, Harris et al. 2018 [45] chose to perform their analyses in 10kb windows, although we note that it is unclear how the H12 and H2/H1 values plotted in their Fig 1 were identified—plausibly they correspond to the 50 identified peaks in Garud et al. 2015 [34] computed in differently defined 401 SNP windows. It is unclear why the plotted values appear to correspond to the values in Garud et al. 2015 [34] despite the application of SNP versus base pair approaches in the respective papers. A more appropriate comparison by Harris et al. would be to overlay the peaks identified in Garud et al. with their S2 Fig since they consider 401 SNP windows in this supplemental figure. It is also unclear whether windows of particularly low nucleotide diversity had been eliminated from the Harris et al.'s 2018 [45] scan or whether any scan was performed at all.

In Garud et al. 2015 [34], we tested several demographic models and found that while they do match Pi and S, they tend to generate values of H12 that are lower than in the data (**Figs 4 and 5**). While we remained agnostic to the cause of this inflation of H12 in the data (misspecification of demography or pervasive draft or both), we chose to focus on empirical outliers as a conservative approach. Our belief was that in general it is not yet possible to ensure that any demographic model is correct and thus the focus on empirical distributions is warranted.

Harris et al. 2018 [45] claim that reasonable demographic models fit the data well, but upon closer inspection of the models they tested, we find that they do not. Specifically, the Duchen et al. 2013 [50] model generates values of S and Pi that are 2-fold lower than the median values in the data, and extremely elevated H12 values that approach 1, suggesting that the North Carolina population should be almost monomorphic while in fact the DGRP data does not have such extreme H12 values. Comparison of empirical values to simulations from a demographic model that does not fit the data is not a valid basis for inference. Thus, is the Duchen et al. 2013 model plausible given that such high homozygosity is never even observed in the data? If the bulk of H12 values in this model are much higher compared to genome-wide levels of H12, then it is impossible to ascertain whether localized regions of high homozygosity in the data are significant departures from neutral expectations. Recently, Arguello et al. 2019 [51] inferred a new admixture model for North American Drosophila, which includes migration between Africa, Europe, and North America. However, this model does not fit H12, S/bp, and Pi/bp in the DGRP data either.

Thus, in the absence of a well-fitting null, we were inspired to look for a more reasonable null model to determine if such a model could in fact generate long H12 tails. We tested over 70 different versions of the admixture models and found that the majority of the tested models did not fit the data (**S22**–**S35 Figs**). When the founding population sizes of Europe and North America were very small, the models predicted a sharp depression of diversity. When the founding population sizes were too large, the models predicted very high diversity unobserved in the data. The models that did fit the data reasonably well in terms of S/bp, Pi/bp, and H12 were the ones with fixed population sizes in North America and Europe, similar to the one implemented in Garud et al. 2015 [34].

We emphasize that these inferred models are not intended to be the 'correct' models, especially since long-range LD and haplotype structure still do not match the data. However, the models proposed in this paper do provide better fits to the data than the ones previously proposed. Thus, they are useful for ascertaining whether a model that can fit multiple summary statistics in the data can also generate a long tail of H12 values. Future work that exhaustively searches the parameter space for a model that fits multiple genome-wide statistics is greatly needed, especially as next generation sequencing of long contiguous genomes becomes even

more ubiquitous [72]. The models that we propose may provide a useful starting point for further demographic inference. For example, models in which rapid oscillations in population size due to seasonal fluctuations may need to be considered [73].

The 50 peaks that we identified in Garud et al. 2015 [34] are all in the extreme tails of the new models that fit Pi/bp, S/bp, and now the bulk of the H12 distribution. These 50 peaks have H12 values that are more than 11 standard deviations away from median H12 value in the data (**S4 Fig**), providing additional evidence that these peaks are outliers given a normal distribution that fits the bulk of the data quite well.

Detecting selective sweeps is only the first goal. The next goal is to distinguish hard and soft sweeps from each other. Comparison of the haplotype frequency spectra for the top sweep candidates versus simulated hard and soft sweeps (**Fig 6**) reveals that the observed data do not resemble classic hard selective sweeps in which a single haplotype has risen to high frequency. Instead, all of the top candidates have multiple haplotypes at high frequency. It is possible that many types of soft sweeps could have given rise to these observed patterns, including soft sweeps from standing genetic variation [19], soft sweeps from multiple de novo mutations [20], spatial soft sweeps [74,75], polygenic sweeps [76], and sweeps arising from periodic selection (such as the oscillating seasonal selection that Drosophila experience [73]). Regardless, all of these scenarios will generate haplotype patterns that differ from both neutrality and classic hard selective sweeps. A potentially exciting future direction will be to quantify the relative occurrences of these complex types of sweeps.

To quantitatively distinguish whether the top peaks are more consistent with hard sweeps versus soft sweeps, in Garud et al. 2015 [34] we developed the statistic H2/H1. Harris et al. 2018 [45] claim that H2/H1 cannot distinguish hard and soft sweeps, even though in their implementation they found that the observed H2/H1 values are in the tail of the values generated by hard sweeps and firmly within the bulk of the distribution generated by soft sweeps.

An additional reason for Harris et al.'s [45] conclusion that H2/H1 does not have sufficient power to distinguish hard versus soft sweeps is that they did not correctly condition on both H12 and H2/H1 in their Fig 1. H2/H1 being high alone is insufficient to determine if a sweep is soft because non-sweeps can easily generate high H2/H1 values (See **Box 1** and **Fig 6**). When H12 values are high and we do have evidence of a sweep, H2/H1 does in fact has high power to distinguish hard and soft sweeps from each other [34,53]. Conditioning on the highest H12 value in a peak can also avoid confounding issues like soft shoulders [69], in which a hard sweep decays due to recombination and mutation events and results in soft sweep-like patterns a short distance away from the sweep center. Indeed, the H12 and H2/H1 statistics have become important discriminating statistics for several recent machine learning methods that detect and differentiate hard and soft sweeps [36,42,43]. Moreover, Figs 1D and S1 from Harris *et al*. 2018 [45] show that after conditioning on H12 values, the top 50 peaks' H2/H1 values are more consistent with soft sweeps, and the strength of this support is quantified in their S1 Fig Bayes factor analysis.

Regardless of the exact statistical methodology used or underlying demographic model, it is a fact that soft sweeps do occur. In *D. melanogaster* alone, there are three well-documented examples of soft sweeps at *Ace*, *Cyp6g1*, and *CHKov1* using direct observations of the same allele on distinct genomic backgrounds [12–18]. More broadly, soft sweeps have been abundantly documented using a variety of methods, data sets and organisms [21]. Our work here is not the final word on the topic as future statistical developments may enable us to better quantify rapid adaptation from population genomic data. Note that despite having tested more than 70 models, none could fit every summary statistic in the data. Thus, it is important to acknowledge that there may not be any purely neutral model that can explain the diversity patterns observed in the data. Factors such as linked selection [77], background selection [78–80],

seasonal adaptation [73], local adaptation [81], and variable recombination rates [82] could all be contributing to diversity patterns in the data. Thus, a combination of demographic and selection forces may be needed to be jointly inferred to be able to fully match diversity patterns in the data. Identifying statistics capable of detecting selection that are robust to the misspecification of demographic and selective models might be one profitable direction for future research given how complex and strong evolutionary forces are known to be.

Methods

Simulations of neutrality and selection

Neutral simulations were generated with the coalescent simulator MS [57], and selection simulations were generated with MSMS [83]. All samples consisted of 145 chromosomes to match the sample depth of the DGRP data analyzed in Garud et al. 2015 [34]. Simulations were generated with a neutral mutation rate of 10^{-9} events/bp/gen [84] and a recombination rate of 5*10^-7 cM/bp.

To simulate hard and soft selective sweeps, we varied the adaptive mutation rate, θ_A = 4*Ne*mu_A. Hard sweeps were simulated with θ_A = 0.01, and soft sweeps were simulated with θ_A = 10, as in Garud et al. 2015. The adaptive mutation was placed in the center of the chromosome. We assumed co-dominance, where a homozygous individual bearing two copies of the advantageous allele has twice the fitness advantage of a heterozygote.

To obtain a minimum of 401 SNPs for computing H12, we simulated chromosomes of length 100,000 bps for neutrality, and 350,000 bps for selection.

Model implementations

For full details and code for model implementations, please refer to the github for this paper (https://github.com/garudlab/Harris_etal_response.git). Specifically, `documentation_Jensen_response_publication.doc` provides a README of the commands run for this paper. The script `generate_MS_commands.py` generates MS commands for all the models previously publishd. The scripts `admixture_parameters_mode_varyGrowth.py`, `admixture_parameters_mode_diffProps.py` and `admixture_parameters_mode_fixedPopSize_MS.py` generate commands for the new models tested in this paper.

We coded the model specified in the Harris et al. supplement as follows: the founding population sizes of North America and Europe were scaled by 4 * African ancestral Ne (EuroNe_anc = $10^{\text{log\_Ne\_Eur\_bn}}$ / (4 * Ne_anc), AmerNe_anc = $10^{\text{log\_Ne\_Ame\_bn}}$ / (4 * Ne_anc)), whereas the present day population sizes were scaled by African ancestral Ne only (scaledNeEuro = Ne_Eur / Ne_anc, scaledNeAmerica = Ne_Ame / Ne_anc). This difference in scaling for the two population sizes resulted in a bottleneck size that was 4 times smaller than reported by Duchen et al. 2013 [50] (**Fig 3I**).

Computation of summary statistics, S, Pi, H12, and LD

S and Pi were computed from putatively neutral SNPs in short introns of the DGRP data, as described in Garud *et al.* 2015 [34] We used the program DaDi [60] to project the DGRP data down to 130 chromosomes to account for missing data. S and Pi was computed from simulations using custom python scripts.

H12 was computed from DGRP data and simulations as described in Garud et al. 2015 [34] using custom python scripts. LD was computed using the R^2 statistic using the same

approach as described in Garud et al. 2015 [34] using custom python scripts. 10^7 R^2 values were averaged over to generate a smooth curve.

Fit of a Gaussian to the distribution of H12 values

We fit a Gaussian distribution to the bulk of the distribution of H12 values. To do so, we first estimated the standard deviation (SD) of data within a 34.1% range of the median, representing roughly 1SD. Then, we simulated a normal distirbution with this inferred SD and a mean corresponding to the median. We then compared data within +/-1 of the inferred SD with the full distribution of a simulated Gaussian using a quantile-quantile plot (**S21 Fig**). We also computed the number of SDs away from the median for the H12 value corresponding to the smallest peak.

Computation of bayes factors

We computed Bayes factors as described in Garud et al. 2015 [34] for two admixture models with constant population sizes in Europe and North America (**Fig 3K**). We approximated BFs using an approximate Bayesian computation approach that integrates out nuisance parameters partial frequency (PF), selection strength (s), and age of a sweep. We stated the hard and soft sweep scenarios as point hypotheses in terms of adaptive mutation rates (θ_A). Specifically, BF = P(H12obs, H2obs/H1obs | soft sweep)/P(H12obs,H2obs/H1obs | hard sweep), whereby H12obs and H2obs/H1obs were computed from data, and hard and soft sweeps were simulated from a range of evolutionary scenarios.

In MSMS, when simulating selection with time-variant demographic models like the admixture model, it is only possible to condition on the time of onset of selection since the simulation runs forward in time. Thus, we assumed a uniform prior distribution of the start time of selection, ~U[0, time of admixture]. The selection coefficient and partial frequency of the sweeps were drawn from uniform priors ranging from 0 to 1.

Supporting information

S1 Fig. Genome wide scan of DGRP data with Pi/bp. Each point represents the mean Pi/bp value in a 10Kb window. Red vertical lines indicate the positions of the positive controls, *Ace*, *Cyp6g1*, and *CHKov1*.
(TIF)

S2 Fig. Genome wide scan of DGRP data with S/bp. Each point represents the mean S/bp value in a 10Kb window. Red vertical lines indicate the positions of the positive controls, *Ace*, *Cyp6g1*, and *CHKov1*.
(TIF)

S3 Fig. Distribution of Pi/bp values in short introns in the DGRP data versus simulations of neutral demographic scenarios tested in Garud et al. 2015 [34]. The distribution of Pi/bp computed in short introns of length 10bps or longer in DGRP data is compared with Pi/bp values computed in a range of simulated neutral demographic models tested in Garud et al. 2015. The models tested are as follows: (A) a constant $N_e = 10^6$ model, (B) a constant $N_e = 2.7 \times 10^6$ model, (C) a severe short bottleneck model, (D) a shallow long bottleneck model, (E) the implemented admixture model in Garud *et al*. 2015 [34], and (F) the implemented admixture + bottleneck model in Garud *et al*. 2015 [34]. Short intron lengths matching those in data were used were used in simulations. Each simulation contains10x the number of short intron fragments as observed in the data. S6 Fig shows the fit of these distributions in quantile-quantile

plots.
(TIF)

S4 Fig. Distribution of Pi/bp values in short introns in the DGRP data versus simulations of neutral demographic scenarios from Duchen et al. [50]**, Harris et al.** [45]**, and Arguello et al** [51]. The distribution of Pi/bp computed in short introns of length 10bps or longer in DGRP data is compared with Pi/bp values computed in a range of simulated neutral demographic models from Duchen et al. [50], Harris et al. [45], and Arguello et al. [51]. The models tested are as follows: (A) The admixture model proposed by Duchen *et al.* 2013 [50], simulated with parameter values corresponding to the mode of the posterior. (B) The admixture + bottleneck model proposed by Duchen *et al.* 2013 [50], simulated with parameter values corresponding to the mode of the posterior. (C) The admixture model proposed by Duchen *et al.* 2013 [50], simulated with parameter values drawn from the posterior distribution. (D) The implemented admixture model in Harris *et al.* 2018 [45], simulated with parameter values drawn from the posterior distribution. (E) The admixture model proposed by Arguello *et al.* 2019 [51]. Short intron lengths matching those in data were used were used in simulations. Each simulation contains10x the number of short intron fragments as observed in the data. S7 Fig shows the fit of these distributions in quantile-quantile plots.
(TIF)

S5 Fig. Distribution of Pi/bp values in short introns in the DGRP data versus simulations of neutral demographic scenarios inferred from this paper. The distribution of Pi/bp computed in short introns of length 10bps or longer in DGRP data is compared with Pi/bp values computed in two demographic models inferred in this paper to fit the DGRP (Models presented in **Fig 3J and 3K**) (A) A variant of the Duchen *et al.* 2013 [50] admixture model where North America, Europe, and Africa have fixed population sizes. North American population size = 1.11x10^6, European population size = .7x10^6, and African population size held constant at the value inferred in Duchen et al. 2013 [50]. (B) A variant of the Duchen *et al.* 2013 [50] admixture model where North America, Europe, and Africa have fixed population sizes. North American population size = 1.6x10^6, European population size = .7x10^6, and African population size held constant at the value inferred in Duchen et al. 2013 [50]. Short intron lengths matching those in data were used were used in simulations. Each simulation contains10x the number of short intron fragments as observed in the data. S8 Fig shows the fit of these distributions in quantile-quantile plots.
(TIF)

S6 Fig. QQ plot comparing the distribution of Pi/bp values in short introns in the DGRP data versus simulations of neutral demographic scenarios tested in Garud et al. 2015 [34]. These plots quantify the fit of the distributions plotted in S3 Fig. The distribution of Pi/bp computed in short introns of length 10bps or longer in DGRP data is compared with Pi values computed in a range of simulated neutral demographic models tested in Garud et al. 2015 [34]. The models tested are as follows: (A) a constant $N_e = 10^6$ model, (B) a constant $N_e = 2.7x10^6$ model, (C) a severe short bottleneck model, (D) a shallow long bottleneck model, (E) the implemented admixture model in Garud *et al.* 2015 [34], and (F) the implemented admixture + bottleneck model in Garud *et al.* 2015 [34]. The root mean square error (RMSE) of the fit is reported in the top left corner of each plot.
(TIF)

S7 Fig. QQ plot comparing the distribution of Pi/bp values in short introns in the DGRP data versus simulations of neutral demographic scenarios from Duchen et al. [50]**, Harris et al.** [45]**, and Arguello et al** [51]. These plots quantify the fit of the distributions plotted in

S4 Fig. The distribution of Pi/bp computed in short introns of length 10bps or longer in DGRP data is compared with Pi/bp values computed in a range of simulated neutral demographic models from Duchen et al. [50], Harris et al. [45], and Arguello et al. [51]. The models tested are as follows: (A) The admixture model proposed by Duchen *et al.* 2013 [50], simulated with parameter values corresponding to the mode of the posterior. (B) The admixture + bottleneck model proposed by Duchen *et al.* 2013 [50], simulated with parameter values corresponding to the mode of the posterior. (C) The admixture model proposed by Duchen *et al.* 2013 [50], simulated with parameter values drawn from the posterior distribution. (D) The implemented admixture model in Harris *et al.* 2018 [45], simulated with parameter values drawn from the posterior distribution. (E) The admixture model proposed by Arguello *et al.* 2019 [51]. The root mean square error (RMSE) of the fit is reported in the top left corner of each plot.
(TIF)

S8 Fig. QQ plot comparing the distribution of Pi/bp values in short introns in the DGRP data versus simulations of neutral demographic scenarios inferred from this paper. These plots quantify the fit of the distributions plotted in S6 Fig. The distribution of Pi/bp computed in short introns of length 10bps or longer in DGRP data is compared with Pi/bp values computed in two demographic models inferred in this paper to fit the DGRP (Models presented in **Fig 3J and 3K**) (A) A variant of the Duchen *et al.* 2013 [50] admixture model where North America, Europe, and Africa have fixed population sizes. North American population size = 1.11×10^6, European population size = $.7 \times 10^6$, and African population size held constant at the value inferred in Duchen et al. 2013 [50]. (B) A variant of the Duchen *et al.* 2013 [50] admixture model where North America, Europe, and Africa have fixed population sizes. North American population size = 1.6×10^6, European population size = $.7 \times 10^6$, and African population size held constant at the value inferred in Duchen et al. 2013 [50]. The root mean square error (RMSE) of the fit is reported in the top left corner of each plot.
(TIF)

S9 Fig. Distribution of S/bp values in short introns in the DGRP data versus simulations of neutral demographic scenarios tested in Garud et al. 2015 [34]. The distribution of S/bp computed in short introns of length 10bps or longer in DGRP data is compared with S/bp values computed in a range of simulated neutral demographic models tested in Garud et al. 2015. The models tested are as follows: (A) a constant $N_e = 10^6$ model, (B) a constant $N_e = 2.7 \times 10^6$ model, (C) a severe short bottleneck model, (D) a shallow long bottleneck model, (E) the implemented admixture model in Garud *et al.* 2015, and (F) the implemented admixture + bottleneck model in Garud *et al.* 2015. Short intron lengths matching those in data were used were used in simulations. Each simulation contains10x the number of short intron fragments as observed in the data. S12 Fig shows the fit of these distributions in quantile-quantile plots.
(TIF)

S10 Fig. Distribution of S/bp values in short introns in the DGRP data versus simulations of neutral demographic scenarios from Duchen et al. [50]**, Harris et al.** [45]**, and Arguello et al** [51]. The distribution of S/bp computed in short introns of length 10bps or longer in DGRP data is compared with S/bp values computed in a range of simulated neutral demographic models from Duchen et al. [50], Harris et al. [45], and Arguello et al. [51]. The models tested are as follows: (A) The admixture model proposed by Duchen *et al.* 2013 [50], simulated with parameter values corresponding to the mode of the posterior. (B) The admixture + bottleneck model proposed by Duchen *et al.* 2013 [50], simulated with parameter values corresponding to the mode of the posterior. (C) The admixture model proposed by Duchen *et al.* 2013

[50], simulated with parameter values drawn from the posterior distribution. (D) The implemented admixture model in Harris *et al.* 2018 [45], simulated with parameter values drawn from the posterior distribution. (E) The admixture model proposed by Arguello *et al.* 2019 [51]. Short intron lengths matching those in data were used were used in simulations. Each simulation contains10x the number of short intron fragments as observed in the data. S13 Fig shows the fit of these distributions in quantile-quantile plots.
(TIF)

S11 Fig. Distribution of S/bp values in short introns in the DGRP data versus simulations of neutral demographic scenarios inferred from this paper. The distribution of S/bp computed in short introns of length 10bps or longer in DGRP data is compared with S/bp values computed in two demographic models inferred in this paper to fit the DGRP (Models presented in **Fig 3J and 3K**) (A) A variant of the Duchen *et al.* 2013 [50] admixture model where North America, Europe, and Africa have fixed population sizes. North American population size = 1.11x10^6, European population size = .7x10^6, and African population size held constant at the value inferred in Duchen et al. 2013 [50]. (B) A variant of the Duchen *et al.* 2013 [50] admixture model where North America, Europe, and Africa have fixed population sizes. North American population size = 1.6x10^6, European population size = .7x10^6, and African population size held constant at the value inferred in Duchen et al. 2013 [50]. Short intron lengths matching those in data were used were used in simulations. Each simulation contains10x the number of short intron fragments as observed in the data. S8 Fig shows the fit of these distributions in quantile-quantile plots.
(TIF)

S12 Fig. QQ plot comparing the distribution of S/bp values in short introns in the DGRP data versus simulations of neutral demographic scenarios tested in Garud et al. 2015 [34]. These plots quantify the fit of the distributions plotted in S9 Fig. The distribution of S/bp computed in short introns of length 10bps or longer in DGRP data is compared with S/bp values computed in a range of simulated neutral demographic models tested in Garud et al. 2015 [34]. The models tested are as follows: (A) a constant $N_e = 10^6$ model, (B) a constant $N_e = 2.7x10^6$ model, (C) a severe short bottleneck model, (D) a shallow long bottleneck model, (E) the implemented admixture model in Garud *et al.* 2015 [34], and (F) the implemented admixture + bottleneck model in Garud *et al.* 2015 [34]. The root mean square error (RMSE) of the fit is reported in the top left corner of each plot.
(TIF)

S13 Fig. QQ plot comparing the distribution of S/bp values in short introns in the DGRP data versus simulations of neutral demographic scenarios from Duchen et al. [50]**, Harris et al.** [45]**, and Arguello et al** [51]. These plots quantify the fit of the distributions plotted in **S10 Fig**. The distribution of S/bp computed in short introns of length 10bps or longer in DGRP data is compared with S/bp values computed in a range of simulated neutral demographic models from Duchen et al. [50], Harris et al. [45], and Arguello et al. [51]. The models tested are as follows: (A) The admixture model proposed by Duchen *et al.* 2013 [50], simulated with parameter values corresponding to the mode of the posterior. (B) The admixture + bottleneck model proposed by Duchen *et al.* 2013 [50], simulated with parameter values corresponding to the mode of the posterior. (C) The admixture model proposed by Duchen *et al.* 2013 [50], simulated with parameter values drawn from the posterior distribution. (D) The implemented admixture model in Harris *et al.* 2018 [45], simulated with parameter values drawn from the posterior distribution. (E) The admixture model proposed by Arguello *et al.* 2019 [51]. The root mean square error (RMSE) of the fit is reported in the top left corner of each

plot.
(TIF)

S14 Fig. QQ plot comparing the distribution of S/bp values in short introns in the DGRP data versus simulations of neutral demographic scenarios inferred from this paper. These plots quantify the fit of the distributions plotted in S11 Fig. The distribution of S/bp computed in short introns of length 10bps or longer in DGRP data is compared with S/bp values computed in two demographic models inferred in this paper to fit the DGRP (Models presented in **Fig 3J and 3K**) (A) A variant of the Duchen *et al*. 2013 [50] admixture model where North America, Europe, and Africa have fixed population sizes. North American population size = 1.11x10^6, European population size = .7x10^6, and African population size held constant at the value inferred in Duchen et al. 2013 [50]. (B) A variant of the Duchen *et al*. 2013 [50] admixture model where North America, Europe, and Africa have fixed population sizes. North American population size = 1.6x10^6, European population size = .7x10^6, and African population size held constant at the value inferred in Duchen et al. 2013 [50]. The root mean square error (RMSE) of the fit is reported in the top left corner of each plot.
(TIF)

S15 Fig. H12 values in DGRP data and in simulations of neutral demographic scenarios tested in Garud et al. 2015 [34]. The DGRP H12 values are compared with H12 values computed in a range of simulated neutral demographic models from Garud et al. 2015. The models tested are as follows: (A) a constant $N_e = 10^6$ model, (B) a constant $N_e = 2.7x10^6$ model, (C) a severe short bottleneck model, (D) a shallow long bottleneck model, (E) the implemented admixture model in Garud *et al*. 2015 [34], and (F) the implemented admixture + bottleneck model in Garud *et al*. 2015. The number of analysis windows generated for the simulated models (n = 69,113) equals the number of analysis windows for the DGRP data, after excluding regions of low recombination rates. The red points indicate the H12 values for the top 50 peaks in the DGRP data.
(TIF)

S16 Fig. H12 values in DGRP data compared to values measured in simulations of neutral demographic scenarios from Duchen et al. [50]**, Harris et al.** [45]**, and Arguello et al** [51]. The DGRP H12 values are compared with H12 values computed in a range of simulated neutral demographic models from Duchen et al. [50], Harris et al. [45], and Arguello et al. [51]. The models tested are as follows: (A) The admixture model proposed by Duchen *et al*. 2013 [50], simulated with parameter values corresponding to the mode of the posterior. (B) The admixture + bottleneck model proposed by Duchen *et al*. 2013 [50], simulated with parameter values corresponding to the mode of the posterior. (C) The admixture model proposed by Duchen *et al*. 2013 [50], simulated with parameter values drawn from the posterior distribution. (D) The implemented admixture model in Harris *et al*. 2018 [45], simulated with parameter values drawn from the posterior distribution. (E) The admixture model proposed by Arguello *et al*. 2019 [51]. The number of analysis windows generated for the simulated models (n = 69,113) equals the number of analysis windows for the DGRP data, after excluding regions of low recombination rates. The red points indicate the H12 values for the top 50 peaks in the DGRP data.
(TIF)

S17 Fig. H12 values in DGRP data compared to values measured in simulations of neutral demographic scenarios inferred from this paper. The two models are depicted in **Fig 3J and 3K** and are (A) A variant of the Duchen *et al*. 2013 [50] admixture model where North America, Europe, and Africa have fixed population sizes. North American population

size = 1.11x10^6, European population size = .7x10^6, and African population size held constant at the value inferred in Duchen et al. 2013 [50]. (B) A variant of the Duchen *et al.* 2013 [50] admixture model where North America, Europe, and Africa have fixed population sizes. North American population size = 1.6x10^6, European population size = .7x10^6, and African population size held constant at the value inferred in Duchen et al. 2013 [50]. The number of analysis windows generated for the simulated models (n = 69,113) equals the number of analysis windows for the DGRP data, after excluding regions of low recombination rates. The red points indicate the H12 values for the top 50 peaks in the DGRP data.
(TIF)

S18 Fig. QQ plot comparing the distribution of H12 values in the DGRP data versus simulations of neutral demographic scenarios tested in Garud et al. 2015 [34]**.** These plots quantify the fit of the distributions plotted in S15 Fig. The models tested are as follows: (A) a constant $N_e = 10^6$ model, (B) a constant $N_e = 2.7x10^6$ model, (C) a severe short bottleneck model, (D) a shallow long bottleneck model, (E) the implemented admixture model in Garud *et al.* 2015 [34], and (F) the implemented admixture + bottleneck model in Garud *et al.* 2015 [34]. The root mean square error (RMSE) of the fit is reported in the top left corner of each plot.
(TIF)

S19 Fig. QQ plot comparing the distribution of H12 values in the DGRP data versus simulations of neutral demographic scenarios from Duchen et al. [50]**, Harris et al.** [45]**, and Arguello et al** [51]**.** These plots quantify the fit of the distributions plotted in S16 Fig. The models tested are as follows: (A) The admixture model proposed by Duchen *et al.* 2013 [50], simulated with parameter values corresponding to the mode of the posterior. (B) The admixture + bottleneck model proposed by Duchen *et al.* 2013 [50], simulated with parameter values corresponding to the mode of the posterior. (C) The admixture model proposed by Duchen *et al.* 2013 [50], simulated with parameter values drawn from the posterior distribution. (D) The implemented admixture model in Harris *et al.* 2018 [45], simulated with parameter values drawn from the posterior distribution. (E) The admixture model proposed by Arguello *et al.* 2019 [51]. The root mean square error (RMSE) of the fit is reported in the top left corner of each plot.
(TIF)

S20 Fig. QQ plot comparing the distribution of H12 values in short the DGRP data versus simulations of neutral demographic scenarios inferred from this paper. These plots quantify the fit of the distributions plotted in S17 Fig. The models tested are depicted in **Fig 3J and 3K** (A) A variant of the Duchen *et al.* 2013 [50] admixture model where North America, Europe, and Africa have fixed population sizes. North American population size = 1.11x10^6, European population size = .7x10^6, and African population size held constant at the value inferred in Duchen et al. 2013 [50]. (B) A variant of the Duchen *et al.* 2013 [50] admixture model where North America, Europe, and Africa have fixed population sizes. North American population size = 1.6x10^6, European population size = .7x10^6, and African population size held constant at the value inferred in Duchen et al. 2013 [50]. The root mean square error (RMSE) of the fit is reported in the top left corner of each plot.
(TIF)

S21 Fig. H12 values in the bulk of the DGRP data compared to a fitted Gaussian. (A) Quantile-quantile plot of H12 values within +/-1 SD of the median value in the DGRP data are compared with a random sample from the fitted Gaussian. The Gaussian was simulated with the mean equalling the median value of H12 in the DGRP data, and the standard deviation

estimated from points within 1 standard deviation around the median (Methods). (B) Comparison of distribution of H12 values in DGRP data with that of a simulated Gaussian with a mean and standard deviation from (A). The vertical blue line indicates 11 standard deviations away from the mean of the simulated Gaussian distribution. The red points indicate the H12 values for the top 50 peaks in the DGRP data. (C) QQ-plot of H12 from the entire distribution of DGRP values compared with the same simulated Gaussian from (A). The distribution of H12 values in DGRP data has an extreme elevated tail compared to expectations under a Gaussian (D) QQ-plot comparing H12 values from a constant $Ne = 2.7*10^6$ model with a Gaussian fitted to its bulk (Methods). Neutral simulations lack the elevated tail present in the data.
(TIF)

S22 Fig. Variants of Duchen et al. 2013 [50] admixture models tested in this paper. We computed Pi, S, and H12 in variants of the admixture model proposed by Duchen et al. 2013 [50]. The admixture models include: (A) constant population sizes for North America and Europe, (B) different growth rates for North America and Europe, (C) different proportions of admixture, (D) different migration rates. In all cases, the 11 parameters originally inferred by Duchen et al. 2013 [50] were kept constant at the mode of the parameters' posterior distributions, unless highlighted in red. The parameters highlighted in red were varied.
(TIF)

S23 Figs. Pi, S, H12, and linkage disequilibrium measured in simulated admixture models with fixed population sizes in North America and Europe. Summary statistics S, Pi, H12, and LD were measured in admixture models with constant population sizes in Europe and North America (S5A Fig). In S6 through S10 Figs, the European population size was held constant at the values 16,982, 67,608, 700,000, 2,000,000, and 9,550,000, respectively. Along the x-axis of each figure, the North American population sizes were held constant at the values 2,500, 61,659, 1,110,000, 15,984,500, and 28,800,000. These population sizes span the ranges of the 95 CI for the posterior distributions of the European and North American population sizes in Duchen et al. 2013. All other parameters in the admixture model were held constant at the mode of the posterior distribution inferred by Duchen et al. 2013 [50]. Each boxplot is comprised of 3,000 simulations.
(TIF)

S24 Figs. Pi, S, H12, and linkage disequilibrium measured in simulated admixture models with fixed population sizes in North America and Europe. Summary statistics S, Pi, H12, and LD were measured in admixture models with constant population sizes in Europe and North America (S5A Fig). In S6 through S10 Figs, the European population size was held constant at the values 16,982, 67,608, 700,000, 2,000,000, and 9,550,000, respectively. Along the x-axis of each figure, the North American population sizes were held constant at the values 2,500, 61,659, 1,110,000, 15,984,500, and 28,800,000. These population sizes span the ranges of the 95 CI for the posterior distributions of the European and North American population sizes in Duchen et al. 2013. All other parameters in the admixture model were held constant at the mode of the posterior distribution inferred by Duchen et al. 2013 [50]. Each boxplot is comprised of 3,000 simulations.
(TIF)

S25 Figs. Pi, S, H12, and linkage disequilibrium measured in simulated admixture models with fixed population sizes in North America and Europe. Summary statistics S, Pi, H12, and LD were measured in admixture models with constant population sizes in Europe and North America (S5A Fig). In S6 through S10 Figs, the European population size was held

constant at the values 16,982, 67,608, 700,000, 2,000,000, and 9,550,000, respectively. Along the x-axis of each figure, the North American population sizes were held constant at the values 2,500, 61,659, 1,110,000, 15,984,500, and 28,800,000. These population sizes span the ranges of the 95 CI for the posterior distributions of the European and North American population sizes in Duchen et al. 2013. All other parameters in the admixture model were held constant at the mode of the posterior distribution inferred by Duchen et al. 2013 [50]. Each boxplot is comprised of 3,000 simulations.
(TIF)

S26 Figs. Pi, S, H12, and linkage disequilibrium measured in simulated admixture models with fixed population sizes in North America and Europe. Summary statistics S, Pi, H12, and LD were measured in admixture models with constant population sizes in Europe and North America (S5A Fig). In S6 through S10 Figs, the European population size was held constant at the values 16,982, 67,608, 700,000, 2,000,000, and 9,550,000, respectively. Along the x-axis of each figure, the North American population sizes were held constant at the values 2,500, 61,659, 1,110,000, 15,984,500, and 28,800,000. These population sizes span the ranges of the 95 CI for the posterior distributions of the European and North American population sizes in Duchen et al. 2013. All other parameters in the admixture model were held constant at the mode of the posterior distribution inferred by Duchen et al. 2013 [50]. Each boxplot is comprised of 3,000 simulations.
(TIF)

S27 Figs. Pi, S, H12, and linkage disequilibrium measured in simulated admixture models with fixed population sizes in North America and Europe. Summary statistics S, Pi, H12, and LD were measured in admixture models with constant population sizes in Europe and North America (S5A Fig). In S6 through S10 Figs, the European population size was held constant at the values 16,982, 67,608, 700,000, 2,000,000, and 9,550,000, respectively. Along the x-axis of each figure, the North American population sizes were held constant at the values 2,500, 61,659, 1,110,000, 15,984,500, and 28,800,000. These population sizes span the ranges of the 95 CI for the posterior distributions of the European and North American population sizes in Duchen et al. 2013. All other parameters in the admixture model were held constant at the mode of the posterior distribution inferred by Duchen et al. 2013 [50]. Each boxplot is comprised of 3,000 simulations.
(TIF)

S28 Figs. Pi, S, H12, and linkage disequilibrium measured in simulated admixture models with differing growth rates in North America and Europe. Summary statistics S, Pi, H12, and LD were measured in admixture models with varying growth rates in Europe and North America (S5B Fig). In S11– S13 Figs, the starting population size for Europe was 16,982, and ending population sizes were 700,000, 2,000,000, and 9,550,000, respectively. In S14 through S16 Figs, the starting population size for Europe was 67,608, and ending population sizes were 700,000, 2,000,000, and 9,550,000, respectively. Along the x-axis of each figure, the North American starting population sizes were either 2,500 or 61,659, and ending population sizes were either 1,110,000, 15,984,500, or 28,800,000. These population sizes span the ranges of the 95 CI for the posterior distributions of the European and North American population sizes in Duchen et al. 2013 [50]. All other parameters in the admixture model were held constant at the mode of the posterior distribution inferred by Duchen et al. 2013. Each boxplot is comprised of 3,000 simulations.
(TIF)

S29 Figs. Pi, S, H12, and linkage disequilibrium measured in simulated admixture models with differing growth rates in North America and Europe. Summary statistics S, Pi, H12, and LD were measured in admixture models with varying growth rates in Europe and North America (S5B Fig). In S11– S13 Figs, the starting population size for Europe was 16,982, and ending population sizes were 700,000, 2,000,000, and 9,550,000, respectively. In S14 through S16 Figs, the starting population size for Europe was 67,608, and ending population sizes were 700,000, 2,000,000, and 9,550,000, respectively. Along the x-axis of each figure, the North American starting population sizes were either 2,500 or 61,659, and ending population sizes were either 1,110,000, 15,984,500, or 28,800,000. These population sizes span the ranges of the 95 CI for the posterior distributions of the European and North American population sizes in Duchen et al. 2013 [50]. All other parameters in the admixture model were held constant at the mode of the posterior distribution inferred by Duchen et al. 2013. Each boxplot is comprised of 3,000 simulations.
(TIF)

S30 Figs. Pi, S, H12, and linkage disequilibrium measured in simulated admixture models with differing growth rates in North America and Europe. Summary statistics S, Pi, H12, and LD were measured in admixture models with varying growth rates in Europe and North America (S5B Fig). In S11– S13 Figs, the starting population size for Europe was 16,982, and ending population sizes were 700,000, 2,000,000, and 9,550,000, respectively. In S14 through S16 Figs, the starting population size for Europe was 67,608, and ending population sizes were 700,000, 2,000,000, and 9,550,000, respectively. Along the x-axis of each figure, the North American starting population sizes were either 2,500 or 61,659, and ending population sizes were either 1,110,000, 15,984,500, or 28,800,000. These population sizes span the ranges of the 95 CI for the posterior distributions of the European and North American population sizes in Duchen et al. 2013 [50]. All other parameters in the admixture model were held constant at the mode of the posterior distribution inferred by Duchen et al. 2013. Each boxplot is comprised of 3,000 simulations.
(TIF)

S31 Figs. Pi, S, H12, and linkage disequilibrium measured in simulated admixture models with differing growth rates in North America and Europe. Summary statistics S, Pi, H12, and LD were measured in admixture models with varying growth rates in Europe and North America (S5B Fig). In S11– S13 Figs, the starting population size for Europe was 16,982, and ending population sizes were 700,000, 2,000,000, and 9,550,000, respectively. In S14 through S16 Figs, the starting population size for Europe was 67,608, and ending population sizes were 700,000, 2,000,000, and 9,550,000, respectively. Along the x-axis of each figure, the North American starting population sizes were either 2,500 or 61,659, and ending population sizes were either 1,110,000, 15,984,500, or 28,800,000. These population sizes span the ranges of the 95 CI for the posterior distributions of the European and North American population sizes in Duchen et al. 2013 [50]. All other parameters in the admixture model were held constant at the mode of the posterior distribution inferred by Duchen et al. 2013. Each boxplot is comprised of 3,000 simulations.
(TIF)

S32 Figs. Pi, S, H12, and linkage disequilibrium measured in simulated admixture models with differing growth rates in North America and Europe. Summary statistics S, Pi, H12, and LD were measured in admixture models with varying growth rates in Europe and North America (S5B Fig). In S11– S13 Figs, the starting population size for Europe was 16,982, and ending population sizes were 700,000, 2,000,000, and 9,550,000, respectively. In S14 through

S16 Figs, the starting population size for Europe was 67,608, and ending population sizes were 700,000, 2,000,000, and 9,550,000, respectively. Along the x-axis of each figure, the North American starting population sizes were either 2,500 or 61,659, and ending population sizes were either 1,110,000, 15,984,500, or 28,800,000. These population sizes span the ranges of the 95 CI for the posterior distributions of the European and North American population sizes in Duchen et al. 2013 [50]. All other parameters in the admixture model were held constant at the mode of the posterior distribution inferred by Duchen et al. 2013. Each boxplot is comprised of 3,000 simulations.
(TIF)

S33 Figs. Pi, S, H12, and linkage disequilibrium measured in simulated admixture models with differing growth rates in North America and Europe. Summary statistics S, Pi, H12, and LD were measured in admixture models with varying growth rates in Europe and North America (S5B Fig). In S11– S13 Figs, the starting population size for Europe was 16,982, and ending population sizes were 700,000, 2,000,000, and 9,550,000, respectively. In S14 through S16 Figs, the starting population size for Europe was 67,608, and ending population sizes were 700,000, 2,000,000, and 9,550,000, respectively. Along the x-axis of each figure, the North American starting population sizes were either 2,500 or 61,659, and ending population sizes were either 1,110,000, 15,984,500, or 28,800,000. These population sizes span the ranges of the 95 CI for the posterior distributions of the European and North American population sizes in Duchen et al. 2013 [50]. All other parameters in the admixture model were held constant at the mode of the posterior distribution inferred by Duchen et al. 2013. Each boxplot is comprised of 3,000 simulations.
(TIF)

S34 Fig. Pi, S, H12, and linkage disequilibrium measured in simulated admixture models with differing admixture proportions. Summary statistics S, Pi, H12, and LD were measured in admixture models with varying admixture proportions between Europe and North America (S5C Fig). Admixture proportions varied from 0 to 0.9. All other parameters in the admixture model were held constant at the mode of the posterior distribution inferred by Duchen et al. 2013 [50]. Each boxplot is comprised of 3,000 simulations.
(TIF)

S35 Fig. Pi, S, H12, and linkage disequilibrium measured in simulated admixture models with differing migration rates. Summary statistics S, Pi, H12, and LD were measured in admixture models with varying amounts of migration between Europe and North America (S5D Fig). Migration rates varied from 0 to 0.75. All other parameters in the admixture model were held constant at the mode of the posterior distribution inferred by Duchen et al. 2013 [50]. Each boxplot is comprised of 3,000 simulations.
(TIF)

S36 Fig. Haplotype frequency spectra for the 11th-50th peaks. Same as Fig 7, except plotted are haplotype frequency spectra for the (A)11th-30th and the (B) 31st—50th peaks in the DGRP scan.
(TIFF)

Acknowledgments

Author Contributions

Conceptualization: Nandita R. Garud, Philipp W. Messer, Dmitri A. Petrov.

Data curation: Nandita R. Garud.

Formal analysis: Nandita R. Garud.

Methodology: Nandita R. Garud, Philipp W. Messer, Dmitri A. Petrov.

Software: Nandita R. Garud.

Visualization: Nandita R. Garud, Dmitri A. Petrov.

Writing – original draft: Nandita R. Garud.

Writing – review & editing: Nandita R. Garud, Philipp W. Messer, Dmitri A. Petrov.

References

1. Andolfatto P, Przeworski M. A Genome-Wide Departure From the Standard Neutral Model in Natural Populations of Drosophila. Genetics. 2000; 156(1):257–68. PMID: 10978290

2. Fay JC, Wu CI. Hitchhiking under positive Darwinian selection. Genetics [Internet]. 2000; 155(3):1405–13. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10880498. PMID: 10880498

3. Smith NG, Eyre-Walker A. Adaptive protein evolution in Drosophila. Nature [Internet]. 2002; 415 (6875):1022–4. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11875568. https://doi.org/10.1038/4151022a PMID: 11875568

4. Bierne N, Eyre-Walker A. The genomic rate of adaptive amino acid substitution in Drosophila. Mol Biol Evol [Internet]. 2004; 21(7):1350–60. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15044594. https://doi.org/10.1093/molbev/msh134 PMID: 15044594

5. Andolfatto P. Adaptive evolution of non-coding DNA in Drosophila. Nature [Internet]. 2005; 437 (7062):1149–52. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16237443. https://doi.org/10.1038/nature04107 PMID: 16237443

6. Eyre-Walker A, Keightley PD. The distribution of fitness effects of new mutations. Vol. 8, Nature Reviews Genetics. 2007. p. 610–8. https://doi.org/10.1038/nrg2146 PMID: 17637733

7. Macpherson JM, Sella G, Davis JC, Petrov DA. Genomewide spatial correspondence between nonsynonymous divergence and neutral polymorphism reveals extensive adaptation in Drosophila. Genetics [Internet]. 2007; 177(4):2083–99. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18073425. https://doi.org/10.1534/genetics.107.080226 PMID: 18073425

8. Shapiro JA, Huang W, Zhang C, Hubisz MJ, Lu J, Turissini DA, et al. Adaptive genic evolution in the Drosophila genomes. Proc Natl Acad Sci U S A [Internet]. 2007; 104(7):2271–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17284599. https://doi.org/10.1073/pnas.0610385104 PMID: 17284599

9. Sella G, Petrov DA, Przeworski M, Andolfatto P. Pervasive natural selection in the Drosophila genome? PLoS Genet [Internet]. 2009; 5(6):e1000495. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19503600. https://doi.org/10.1371/journal.pgen.1000495 PMID: 19503600

10. Schneider A, Charlesworth B, Eyre-Walker A, Keightley PD. A method for inferring the rate of occurrence and fitness effects of advantageous mutations. Genetics [Internet]. 2011; 189(4):1427–37. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21954160. https://doi.org/10.1534/genetics.111.131730 PMID: 21954160

11. Kousathanas A, Keightley PD. A comparison of models to infer the distribution of fitness effects of new mutations. Genetics [Internet]. 2013; 193(4):1197–208. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23341416. https://doi.org/10.1534/genetics.112.148023 PMID: 23341416

12. Daborn P, Boundy S, Yen J, Pittendrigh B, Ffrench-Constant R. DDT resistance in Drosophila correlates with Cyp6g1 over-expression and confers cross-resistance to the neonicotinoid imidacloprid. Mol Genet Genomics [Internet]. 2001; 266(4):556–63. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11810226. https://doi.org/10.1007/s004380100531 PMID: 11810226

13. Schmidt JM, Good RT, Appleton B, Sherrard J, Raymant GC, Bogwitz MR, et al. Copy number variation and transposable elements feature in recent, ongoing adaptation at the Cyp6g1 locus. PLoS Genet [Internet]. 2010; 6(6):e1000998. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20585622. https://doi.org/10.1371/journal.pgen.1000998 PMID: 20585622

14. Aminetzach YT, Macpherson JM, Petrov DA. Pesticide resistance via transposition-mediated adaptive gene truncation in Drosophila. Science (80-) [Internet]. 2005; 309(5735):764–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16051794. https://doi.org/10.1126/science.1112699 PMID: 16051794

15. Magwire MM, Bayer F, Webster CL, Cao C, Jiggins FM. Successive increases in the resistance of Drosophila to viral infection through a transposon insertion followed by a Duplication. PLoS Genet [Internet]. 2011; 7(10):e1002337. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22028673. https://doi.org/10.1371/journal.pgen.1002337 PMID: 22028673

16. Karasov T, Messer PW, Petrov DA. Evidence that adaptation in Drosophila is not limited by mutation at single sites. PLoS Genet [Internet]. 2010; 6(6):1–10. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20585551. https://doi.org/10.1371/journal.pgen.1000905 PMID: 20386746

17. Garud NR, Petrov DA. Elevated linkage disequilibrium and signatures of soft sweeps are common in drosophila melanogaster. Genetics. 2016; https://doi.org/10.1534/genetics.115.184002 PMID: 27098909

18. Vy HMT, Won YJ, Kim Y. Multiple Modes of Positive Selection Shaping the Patterns of Incomplete Selective Sweeps over African Populations of Drosophila melanogaster. Mol Biol Evol [Internet]. 2017; 34(11):2792–807. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28981697. https://doi.org/10.1093/molbev/msx207 PMID: 28981697

19. Hermisson J, Pennings PS. Soft sweeps: molecular population genetics of adaptation from standing genetic variation. Genetics [Internet]. 2005; 169(4):2335–52. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15716498. https://doi.org/10.1534/genetics.104.036947 PMID: 15716498

20. Pennings PS, Hermisson J. Soft sweeps II—molecular population genetics of adaptation from recurrent mutation or migration. Mol Biol Evol [Internet]. 2006; 23(5):1076–84. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16520336. https://doi.org/10.1093/molbev/msj117 PMID: 16520336

21. Messer PW, Petrov DA. Population genomics of rapid adaptation by soft selective sweeps. Trends Ecol Evol [Internet]. 2013; 28(11):659–69. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24075201. https://doi.org/10.1016/j.tree.2013.08.003 PMID: 24075201

22. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics [Internet]. 1989; 123(3):585–95. Available from: http://www.ncbi.nlm.nih.gov/pubmed/2513255. PMID: 2513255

23. Guttman DS, Dykhuizen DE. Detecting selective sweeps in naturally occurring Escherichia coli. Genetics. 1994; 138(4):993–1003. PMID: 7896119

24. Braverman JM, Hudson RR, Kaplan NL, Langley CH, Stephan W. The hitchhiking effect on the site frequency spectrum of DNA polymorphisms. Genetics [Internet]. 1995; 140(2):783–96. Available from: http://www.ncbi.nlm.nih.gov/pubmed/7498754. PMID: 7498754

25. Depaulis F, Veuille M. Neutrality tests based on the distribution of haplotypes under an infinite-site model. Mol Biol Evol [Internet]. 1998; 15(12):1788–90. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9917213. https://doi.org/10.1093/oxfordjournals.molbev.a025905 PMID: 9917213

26. Sabeti PC, Reich DE, Higgins JM, Levine HZ, Richter DJ, Schaffner SF, et al. Detecting recent positive selection in the human genome from haplotype structure. Nature [Internet]. 2002; 419(6909):832–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12397357. https://doi.org/10.1038/nature01140 PMID: 12397357

27. Nielsen R, Williamson S, Kim Y, Hubisz MJ, Clark AG, Bustamante C. Genomic scans for selective sweeps using SNP data. Genome Res [Internet]. 2005; 15(11):1566–75. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16251466. https://doi.org/10.1101/gr.4252305 PMID: 16251466

28. Voight BF, Kudaravalli S, Wen X, Pritchard JK. A map of recent positive selection in the human genome. PLoS Biol [Internet]. 2006; 4(3):e72. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16494531. https://doi.org/10.1371/journal.pbio.0040072 PMID: 16494531

29. Vitti JJ, Grossman SR, Sabeti PC. Detecting natural selection in genomic data. Annu Rev Genet [Internet]. 2013; 47:97–120. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24274750. https://doi.org/10.1146/annurev-genet-111212-133526 PMID: 24274750

30. Przeworski M, Coop G, Wall JD. The signature of positive selection on standing genetic variation. Evolution (N Y) [Internet]. 2005; 59(11):2312–23. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16396172. PMID: 16396172

31. Pennings PS, Hermisson J. Soft sweeps III: the signature of positive selection from recurrent mutation. PLoS Genet [Internet]. 2006; 2(12):e186. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17173482. https://doi.org/10.1371/journal.pgen.0020186 PMID: 17173482

32. Teshima KM, Coop G, Przeworski M. How reliable are empirical genomic scans for selective sweeps? Genome Res [Internet]. 2006; 16(6):702–12. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16687733. https://doi.org/10.1101/gr.5105206 PMID: 16687733

33. Cutter AD, Payseur BA. Genomic signatures of selection at linked sites: unifying the disparity among species. Nat Rev Genet [Internet]. 2013; 14(4):262–74. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23478346. https://doi.org/10.1038/nrg3425 PMID: 23478346

34. Garud NR, Messer PW, Buzbas EO, Petrov DA. Recent Selective Sweeps in North American Drosophila melanogaster Show Signatures of Soft Sweeps. PLoS Genet. 2015; https://doi.org/10.1371/journal.pgen.1005004 PMID: 25706129

35. Mackay TF, Richards S, Stone EA, Barbadilla A, Ayroles JF, Zhu D, et al. The Drosophila melanogaster Genetic Reference Panel. Nature [Internet]. 2012; 482(7384):173–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22318601. https://doi.org/10.1038/nature10811 PMID: 22318601

36. Sheehan S, Song YS. Deep Learning for Population Genetic Inference. PLoS Comput Biol [Internet]. 2016; 12(3):e1004845. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27018908. https://doi.org/10.1371/journal.pcbi.1004845 PMID: 27018908

37. Schlamp F, van der Made J, Stambler R, Chesebrough L, Boyko AR, Messer PW. Evaluating the performance of selection scans to detect selective sweeps in domestic dogs. Mol Ecol [Internet]. 2016; 25(1):342–56. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26589239. https://doi.org/10.1111/mec.13485 PMID: 26589239

38. Zhao Y, Wang J, Zou H, Chen L, Long X, Lan J, et al. Convergent and divergent genetic changes in the genome of Chinese and European pigs. Sci Rep. 2017; https://doi.org/10.1038/s41598-017-09061-z PMID: 28819228

39. González-Rodríguez A, Munilla S, Mouresan EF, Cañas-Álvarez JJ, Díaz C, Piedrafita J, et al. On the performance of tests for the detection of signatures of selection: a case study with the Spanish autochthonous beef cattle populations. Genet Sel Evol. 2016; https://doi.org/10.1186/s12711-016-0258-1 PMID: 27793093

40. Yurchenko AA, Daetwyler HD, Yudin N, Schnabel RD, Vander Jagt CJ, Soloshenko V, et al. Scans for signatures of selection in Russian cattle breed genomes reveal new candidate genes for environmental adaptation and acclimation. Sci Rep. 2018; https://doi.org/10.1038/s41598-018-31304-w PMID: 30154520

41. Zhong L, Yang Q, Yan X, Yu C, Su L, Zhang X, et al. Signatures of soft sweeps across the Dt1 locus underlying determinate growth habit in soya bean [Glycine max (L.) Merr.]. Mol Ecol. 2017; https://doi.org/10.1111/mec.14209 PMID: 28627128

42. Schrider DR, Kern AD. Soft Sweeps Are the Dominant Mode of Adaptation in the Human Genome. Mol Biol Evol [Internet]. 2017; 34(8):1863–77. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28482049. https://doi.org/10.1093/molbev/msx154 PMID: 28482049

43. Schrider DR, Kern AD. S/HIC: Robust Identification of Soft and Hard Sweeps Using Machine Learning. PLoS Genet. 2016; https://doi.org/10.1371/journal.pgen.1005928 PMID: 26977894

44. Sheehan S, Song YS. Deep Learning for Population Genetic Inference. PLoS Comput Biol. 2016; https://doi.org/10.1371/journal.pcbi.1004845 PMID: 27018908

45. Harris RB, Sackman A, Jensen JD. On the unfounded enthusiasm for soft selective sweeps II: Examining recent evidence from humans, flies, and viruses. PLoS Genet [Internet]. 2018; 14(12):e1007859. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30592709. https://doi.org/10.1371/journal.pgen.1007859 PMID: 30592709

46. Maynard Smith J, Haigh J. The hitch-hiking effect of a favourable gene. Genet Res. 1974; 23(1):23–35. PMID: 4407212

47. Kaplan NL, Hudson RR, Langley CH. The "hitchhiking effect" revisited. Genetics [Internet]. 1989; 123(4):887–99. Available from: http://www.ncbi.nlm.nih.gov/pubmed/2612899. PMID: 2612899

48. Kim Y, Stephan W. Detecting a local signature of genetic hitchhiking along a recombining chromosome. Genetics [Internet]. 2002; 160(2):765–77. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11861577. PMID: 11861577

49. Hudson RR, Bailey K, Skarecky D, Kwiatowski J, Ayala FJ. Evidence for positive selection in the superoxide dismutase (Sod) region of Drosophila melanogaster. Genetics [Internet]. 1994; 136(4):1329–40. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8013910. PMID: 8013910

50. Duchen P, Zivkovic D, Hutter S, Stephan W, Laurent S. Demographic inference reveals African and European admixture in the North American Drosophila melanogaster population. Genetics [Internet]. 2013; 193(1):291–301. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23150605. https://doi.org/10.1534/genetics.112.145912 PMID: 23150605

51. Arguello JR, Laurent S, Clark AG, Gaut B. Demographic History of the Human Commensal Drosophila melanogaster. Genome Biol Evol. 2019; https://doi.org/10.1093/gbe/evz022 PMID: 30715331

52. Lawrie DS, Messer PW, Hershberg R, Petrov DA. Strong purifying selection at synonymous sites in D. melanogaster. PLoS Genet [Internet]. 2013; 9(5):e1003527. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23737754. https://doi.org/10.1371/journal.pgen.1003527 PMID: 23737754

53. Garud NR, Rosenberg NA. Enhancing the mathematical properties of new haplotype homozygosity statistics for the detection of selective sweeps. Theor Popul Biol. 2015; https://doi.org/10.1016/j.tpb.2015.04.001 PMID: 25891325

54. Harris AM, Garud NR, DeGiorgio M. Detection and Classification of Hard and Soft Sweeps from Unphased Genotypes by Multilocus Genotype Identity. Genetics [Internet]. 2018; 210(4):1429–52. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30315068. https://doi.org/10.1534/genetics.118.301502 PMID: 30315068

55. Harpak A, Garud N, Rosenberg NA, Petrov DA, Combs M, Pennings PS, et al. Genetic Adaptation in New York City Rats. Genome Biol Evol. 2020;

56. Ferrer-Admetlla A, Liang M, Korneliussen T, Nielsen R. On detecting incomplete soft or hard selective sweeps using haplotype structure. Mol Biol Evol [Internet]. 2014; 31(5):1275–91. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24554778. https://doi.org/10.1093/molbev/msu077 PMID: 24554778

57. Hudson RR. Generating samples under a Wright-Fisher neutral model of genetic variation. Bioinformatics [Internet]. 2002; 18(2):337–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11847089. https://doi.org/10.1093/bioinformatics/18.2.337 PMID: 11847089

58. Messer PW. SLiM: simulating evolution with selection and linkage. Genetics [Internet]. 2013; 194 (4):1037–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23709637. https://doi.org/10.1534/genetics.113.152181 PMID: 23709637

59. Begun DJ, Aquadro CF. Levels of naturally occur.ring DNA polymorphism correlate with recombination rates in D. melanogaster. Nature [Internet]. 1992; 356(6369):519–20. Available from: http://www.ncbi.nlm.nih.gov/pubmed/1560824. https://doi.org/10.1038/356519a0 PMID: 1560824

60. Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD. Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. PLoS Genet [Internet]. 2009; 5 (10):e1000695. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19851460. https://doi.org/10.1371/journal.pgen.1000695 PMID: 19851460

61. Li H, Durbin R. Inference of human population history from individual whole-genome sequences. Nature. 2011; https://doi.org/10.1038/nature10231 PMID: 21753753

62. Enard D, Cai L, Gwennap C, Petrov DA. Viruses are a dominant driver of protein adaptation in mammals. Elife. 2016; https://doi.org/10.7554/eLife.12469 PMID: 27187613

63. Enard D, Petrov DA. Evidence that RNA Viruses Drove Adaptive Introgression between Neanderthals and Modern Humans. Cell. 2018; https://doi.org/10.1016/j.cell.2018.08.034 PMID: 30290142

64. Cheeseman IH, Miller BA, Nair S, Nkhoma S, Tan A, Tan JC, et al. A major genome region underlying artemisinin resistance in malaria. Science (80-) [Internet]. 2012; 336(6077):79–82. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22491853.

65. Cheeseman IH, Miller B, Tan JC, Tan A, Nair S, Nkhoma SC, et al. Population structure shapes copy number variation in malaria parasites. Mol Biol Evol. 2016; 33(3):603–20. https://doi.org/10.1093/molbev/msv282 PMID: 26613787

66. Feder AF, Rhee SY, Holmes SP, Shafer RW, Petrov DA, Pennings PS. More effective drugs lead to harder selective sweeps in the evolution of drug resistance in HIV-1. Elife. 2016; 5(FEBRUARY2016). https://doi.org/10.7554/eLife.10670 PMID: 26882502

67. Tishkoff SA, Reed FA, Ranciaro A, Voight BF, Babbitt CC, Silverman JS, et al. Convergent adaptation of human lactase persistence in Africa and Europe. Nat Genet [Internet]. 2007; 39(1):31–40. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17159977. https://doi.org/10.1038/ng1946 PMID: 17159977

68. Peter BM, Huerta-Sanchez E, Nielsen R. Distinguishing between selective sweeps from standing variation and from a de novo mutation. PLoS Genet [Internet]. 2012; 8(10):e1003011. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23071458 https://doi.org/10.1371/journal.pgen.1003011 PMID: 23071458

69. Schrider DR, Mendes FK, Hahn MW, Kern AD. Soft shoulders ahead: spurious signatures of soft and partial selective sweeps result from linked hard sweeps. Genetics [Internet]. 2015; 200(1):267–84. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25716978. https://doi.org/10.1534/genetics.115.174912 PMID: 25716978

70. Jensen JD. On the unfounded enthusiasm for soft selective sweeps. Nature Communications. 2014. https://doi.org/10.1038/ncomms6281 PMID: 25345443

71. O'Reilly PF, Birney E, Balding DJ. Confounding between recombination and selection, and the Ped/Pop method for detecting selection. Genome Res. 2008;

72. Kim BY, Wang J, Miller DE, Barmina O, Delaney EK, Thompson A, et al. Highly contiguous assemblies of 101 drosophilid genomes. bioRxiv. 2020;

73. Bergland AO, Behrman EL, O'Brien KR, Schmidt PS, Petrov DA. Genomic Evidence of Rapid and Stable Adaptive Oscillations over Seasonal Time Scales in Drosophila. PLoS Genet. 2014; https://doi.org/10.1371/journal.pgen.1004775 PMID: 25375361

74. Ralph P, Coop G. Parallel adaptation: One or many waves of advance of an advantageous allele? Genetics. 2010; https://doi.org/10.1534/genetics.110.119594 PMID: 20660645

75. Ralph PL, Coop G. The role of standing variation in geographic convergent adaptation. Am Nat. 2015; https://doi.org/10.1086/682948 PMID: 26656217

76. Pritchard JK, Pickrell JK, Coop G. The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. Curr Biol [Internet]. 2010; 20(4):R208–15. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20178769. https://doi.org/10.1016/j.cub.2009.11.055 PMID: 20178769

77. Elyashiv E, Sattath S, Hu TT, Strutsovsky A, McVicker G, Andolfatto P, et al. A Genomic Map of the Effects of Linked Selection in Drosophila. PLoS Genet. 2016; 12(8). https://doi.org/10.1371/journal.pgen.1006130 PMID: 27536991

78. Comeron JM. Background selection as baseline for nucleotide variation across the Drosophila genome. PLoS Genet [Internet]. 2014; 10(6):e1004434. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24968283. https://doi.org/10.1371/journal.pgen.1004434 PMID: 24968283

79. Comeron JM. Background selection as null hypothesis in population genomics: Insights and challenges from drosophila studies. Philos Trans R Soc B Biol Sci. 2017; https://doi.org/10.1098/rstb.2016.0471 PMID: 29109230

80. Charlesworth B. Background selection and patterns of genetic diversity in Drosophila melanogaster. Genet Res. 1996; https://doi.org/10.1017/s0016672300034029 PMID: 8940902

81. Bergland AO, Tobler R, Gonzalez J, Schmidt PS, Petrov DA. Secondary contact and local adaptation contribute to genome-wide patterns of clinal variation in Drosophila melanogaster. bioRxiv. 2015;

82. Comeron JM, Ratnappan R, Bailin S. The many landscapes of recombination in Drosophila melanogaster. PLoS Genet [Internet]. 2012; 8(10):e1002905. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23071443. https://doi.org/10.1371/journal.pgen.1002905 PMID: 23071443

83. Ewing G, Hermisson J. MSMS: a coalescent simulation program including recombination, demographic structure and selection at a single locus. Bioinformatics [Internet]. 2010; 26(16):2064–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20591904. https://doi.org/10.1093/bioinformatics/btq322 PMID: 20591904

84. Keightley PD, Trivedi U, Thomson M, Oliver F, Kumar S, Blaxter ML. Analysis of the genome sequences of three Drosophila melanogaster spontaneous mutation accumulation lines. Genome Res [Internet]. 2009; 19(7):1195–201. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19439516. https://doi.org/10.1101/gr.091231.109 PMID: 19439516