bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

1 **Title:**

2 **SNP-level F_{ST} outperforms window statistics for detecting soft sweeps in local adaptation**

3

4 **Authors and affiliations:**

5 Tiago da Silva Ribeiro[1,2,@], José A. Galván[3], John E. Pool[1,2,@]

6 [1] Department of Integrative Biology, University of Wisconsin-Madison, Madison, WI, 53706, USA

7 [2] Laboratory of Genetics, University of Wisconsin-Madison, Madison, WI, 53706, USA

8 [3] John Jay College of Criminal Justice, New York, NY, 10019, USA

9

10 [@] Corresponding authors:

11 Tiago da Silva Ribeiro, tribeiro@wisc.edu

12 John E. Pool, jpool@wisc.edu

13

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Abstract

Local adaptation can lead to elevated genetic differentiation at the targeted genetic variant and nearby sites. Selective sweeps come in different forms, and depending on the initial and final frequencies of a favored variant, very different patterns of genetic variation may be produced. If local selection favors an existing variant that had already recombined onto multiple genetic backgrounds, then the width of elevated genetic differentiation (high F_{ST}) may be too narrow to detect using a typical windowed genome scan, even if the targeted variant becomes highly differentiated. We therefore used a simulation approach to investigate the power of SNP-level F_{ST} (specifically, the maximum SNP F_{ST} value within a window) to detect diverse scenarios of local adaptation, and compared it against whole-window F_{ST} and the Comparative Haplotype Identity statistic. We found that SNP F_{ST} had superior power to detect complete or mostly complete soft sweeps, but lesser power than window-wide statistics to detect partial hard sweeps. To investigate the relative enrichment and nature of SNP F_{ST} outliers from real data, we applied the two F_{ST} statistics to a panel of *Drosophila melanogaster* populations. We found that SNP F_{ST} had a genome-wide enrichment of outliers compared to demographic expectations, and though it yielded a lesser enrichment than window F_{ST}, it detected mostly unique outlier genes and functional categories. Our results suggest that SNP F_{ST} is highly complementary to typical window-based approaches for detecting local adaptation, and merits inclusion in future genome scans and methodologies.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Key words

Local adaptation, soft sweeps, partial sweeps, population genomics, *Drosophila melanogaster*

Significance statement

Studies that use genetic variation to search for genes evolving under population-specific natural selection tend to analyze data at the level of genomic windows that may each contain hundreds of variable sites. However, some models of natural selection (*e.g.* favoring an existing genetic variant) may result in genetic signals of local adaptation that are too narrow to be detected by such approaches. Here we use both simulations and empirical data analysis to show that searching for a site-specific signal of elevated genetic differentiation can find instances of local adaptation that other approaches miss, and therefore the integration of this signal into future studies may significantly improve our understanding of adaptive evolution and its genetic targets.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Introduction

Geographically distinct populations are exposed to different selective pressures, which may result in local adaptation. The detection of genomic regions under positive selection specific to one population is essential to uncovering the genetic basis of locally adaptive trait variation. Local adaptation can exist between populations with low genome-wide genetic differentiation, and comparing genetic variation between these closely-related populations can allow for much more powerful detection of positive selection than is possible from a single population. In light of that advantage, as well as the potential applicability of genetic mapping and functional approaches to locally adaptive traits, local adaptation has played a key role in our increasing understanding of adaptive evolution at the genetic level (Kawecki and Ebert 2004; Yeaman 2015; Tigano and Friesen 2016). In addition to its importance for evolutionary biology and ecology, the identification of regions under selection has implications for applied fields such as health sciences and agriculture because it can also pinpoint regions of the genome that hold functional diversity (Bamshad and Wooding 2003; Ross-Ibarra *et al*. 2007). There has also been increasing recognition of the importance of local adaptation for a species' future adaptive potential, with implications for conservation genetics and adaptation to climate change (Funk *et al.* 2012; Aitken and Whitlock 2013; Fitzpatrick and Keller 2015).

Population genomic scans for local adaptation compare genetic variation between two populations, often searching for specific genomic windows that depart from genome-wide patterns of differentiation in a manner consistent with population-specific natural selection. Positive selection has traditionally been conceptualized and modeled as a selective sweep,

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

which traditionally involves a new beneficial mutation rising to fixation, with strong effects on genetic variation at linked sites (Maynard Smith and Haigh 1974; Kaplan *et al*. 1989). However, there are different kinds of selective sweeps, depending on the initial and final frequencies of the favored variant, and different statistical tests for deviations from neutrality vary in their power to detect them.

First, selective sweeps can be classified as hard or soft sweeps. In a hard sweep, only a single original haplotype carrying the advantageous allele is boosted by natural selection. This situation might be expected if selection favors either a newly occurring mutation or else a variant at low enough frequency that only one copy contributes to the sweep by chance. In a soft sweep, two or more distinct haplotypes carrying the beneficial variant increase in frequency. In some cases, soft sweeps occur because the advantageous allele was present in the population, segregating neutrally, prior to the onset of selection (Hermisson and Pennings 2005). But they can also be the result of recurrent mutations or influx of new alleles through migration (Pennings and Hermisson 2006a, 2006b).

Selective sweeps can also be classified as complete or partial sweeps. In a complete sweep, the advantageous allele reaches fixation in the population. In a partial sweep, the advantageous allele is at an intermediary frequency. This may occur either because the sweep is still ongoing or because positive selection ended prior to fixation. Situations in which a sweep might terminate prematurely include an environmental change, a polygenic trait reaching its new optimum or threshold value, or an allele reaching a balanced equilibrium in a scenario such as heterozygote advantage.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Different kinds of selective sweeps leave different signatures of local adaptation and our power to detect them will differ depending on which methods we use (Lange and Pool 2016). Some common approaches to scanning the genome for population-specific selective sweeps use F_{ST} (or F_{ST}-based) statistics to quantify genetic differentiation between populations. Local adaptation is expected to create genomic regions with higher differentiation than what would be expected under neutrality, since allele frequencies in these regions will change faster as the beneficial allele increases in frequency (Lewontin and Krakauer 1973). Neutral expectations can be inferred either with demographic simulations or an outlier approach. Demographic simulations, based on a previously estimated model of population history, can be used to mimic the history of the populations being studied in the absence of natural selection. Outlier approaches rely on the genome-wide distribution of F_{ST} as a proxy for the neutral distribution, since neutral forces (including those due to demographic history) can broadly be expected to affect the whole genome similarly. Genome scans for regions under selection have typically focused on measuring F_{ST} or other statistics in windows of the genome of some predefined size to search for highly differentiated genomic regions.

A motivating empirical example for the present study comes from an investigation of the genetic basis of locally adaptive melanism in high altitude *Drosophila melanogaster* populations. Here, the authors used QTL mapping to identify genomic regions associated with derived dark pigmentation traits, and then used F_{ST} to scan these regions for signatures of selection (Bastide *et al.* 2016). One very narrow and strong QTL for highland Ethiopian melanism contained the well-known pigmentation gene *ebony*, which also contributed to melanic evolution in a Uganda

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

111 population (Pool and Aquadro 2007; Rebeiz *et al.* 2009). Assessing genetic differentiation
112 between the Ethiopia and Zambia populations for the window containing *ebony*, although
113 window-wide F_{ST} was only marginally elevated, it had a SNP with extremely high F_{ST} (0.85).
114 Compared to demographic simulations, this window's maximum SNP F_{ST} value was among the
115 top 1% of all windows, while its window-wide F_{ST} was only among the 7% highest (Bastide *et al.*
116 2016). Simulated scenarios of soft sweeps from standing variation replicated this pattern of
117 extremely high SNP F_{ST} and only moderately high window F_{ST}, suggesting that some kinds of
118 selective sweeps that may not be detected using window-wide F_{ST} could potentially be detected
119 with a SNP-level F_{ST} approach. Further potential support for the use of SNP F_{ST} to detect
120 adaptive events in this same species is demonstrated by much stronger parallel signatures of
121 selection seen at the SNP level compared to the window level in populations that independently
122 adapted to cold environments (Pool *et al*. 2017).

123 Challenges of using SNP F_{ST} values include their variability due to random sampling
124 effects (Weir *et al.* 2005) and the large number of tests that need to be made against a null
125 distribution. Therefore, larger sample sizes are needed than for window F_{ST}. By using only the
126 highest SNP F_{ST} value within a window, and comparing against null simulations with demography
127 and recombination, we may somewhat improve the multiple testing issue, since here we are not
128 treating all tightly linked SNPs as fully independent tests. Another advantage of this approach is
129 to make a SNP F_{ST} genome scan easier to compare to window-wide statistics. If these metrics
130 are able to detect different types of selective events, then a more comprehensive scan for
131 signatures of selection could benefit from using both window and SNP-based methods. The

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

genome-wide distribution of these statistics in natural populations, compared to their neutral expectations, might also shed light on the contribution of different kinds of selective sweeps to local adaptation.

To understand the utility of using the highest F_{ST} value of any SNP within a window (herein F_{ST_MaxSNP}) as a local adaptation summary statistic, we performed power analyses based on extensive simulations, and then applied these results to empirical data from natural populations of *D. melanogaster*. We calculated the power of F_{ST_MaxSNP} to detect signatures of local adaptation under a wide range of different selective scenarios and demographic histories. We performed demographic simulations and compared the power of F_{ST_MaxSNP} to both window-wide F_{ST} (herein, F_{ST_Window}) and a comparative haplotype-based statistic (χ_{MD}). Then, we investigated the genome-wide distribution of F_{ST_MaxSNP} and F_{ST_Window} among several natural populations of *D. melanogaster*, to determine whether either statistic was enriched genome-wide in empirical data compared to neutral expectations. Finally, we used an outlier approach to perform a genome scan for regions potentially under local adaptation between the Ethiopia and Zambia populations mentioned above, using F_{ST_MaxSNP}, F_{ST_Window}, and χ_{MD}, and we determined the extent of overlap between candidate regions identified according to these different methods. These analyses allowed us to both identify the parameter space in which F_{ST_MaxSNP} outperforms other statistics, and to assess the utility and complementarity of applying these approaches to real data.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Results

SNP-level F_{ST} and window-wide summaries have complementary power to detect local adaptation

We performed power analyses of F_{ST_MaxSNP}, F_{ST_Window}, and χ_{MD} using population genetic simulations with and without natural selection. We used *msms* (Ewing and Hermisson 2010) to simulate population-specific selective sweeps with constrained initial and final allele frequencies, as well as scenarios with population size bottlenecks or migration (simulation commands in Table S1). For each scenario, we simulated populations with high effective population size (N_e) using a set of parameters based on *D. melanogaster* and populations with low N_e using parameters based on humans, following the design of a previous power analysis study that did not include F_{ST_MaxSNP} (Lange and Pool 2016). Power was defined in a locus-specific context, based on the proportion of selection simulations giving a more extreme value of the summary statistic than the 95th quantile of its distribution from neutral simulations.

Unsurprisingly, all three statistics were found to have high power for the case of complete hard sweeps (Figure 1; Table S1). These simulations were conditioned on fixation of a beneficial new mutation in one population that had not occurred in the other population. In light of this fixed difference, F_{ST_MaxSNP} in all replicates had its maximum value ($F_{ST\_MaxSNP} = 1$). In such cases, the power of F_{ST_MaxSNP} was binary, either zero or one, depending on whether or not 5% of the corresponding neutral replicates had an allele that reached fixation. In our simple isolation model, the likelihood that a neutral allele can reach fixation increases with the split time (Table S1; Figure S1). Stronger bottlenecks also boost the likelihood of having neutral

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

174 alleles reach fixation (Table S1; Figure S2, Figure S3). Hence, power for F_{ST_MaxSNP} to detect
175 complete hard sweeps goes from high, for recent splits and weaker bottlenecks, to zero for
176 histories in which more than 5% of neutral replicates contain a fixed difference. Similarly,
177 F_{ST_Window} and χ_{MD} had higher power to detect signatures of local adaptation following recent
178 splits and in weaker bottlenecks, but their change in power was gradual and continuous instead
179 of binary.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

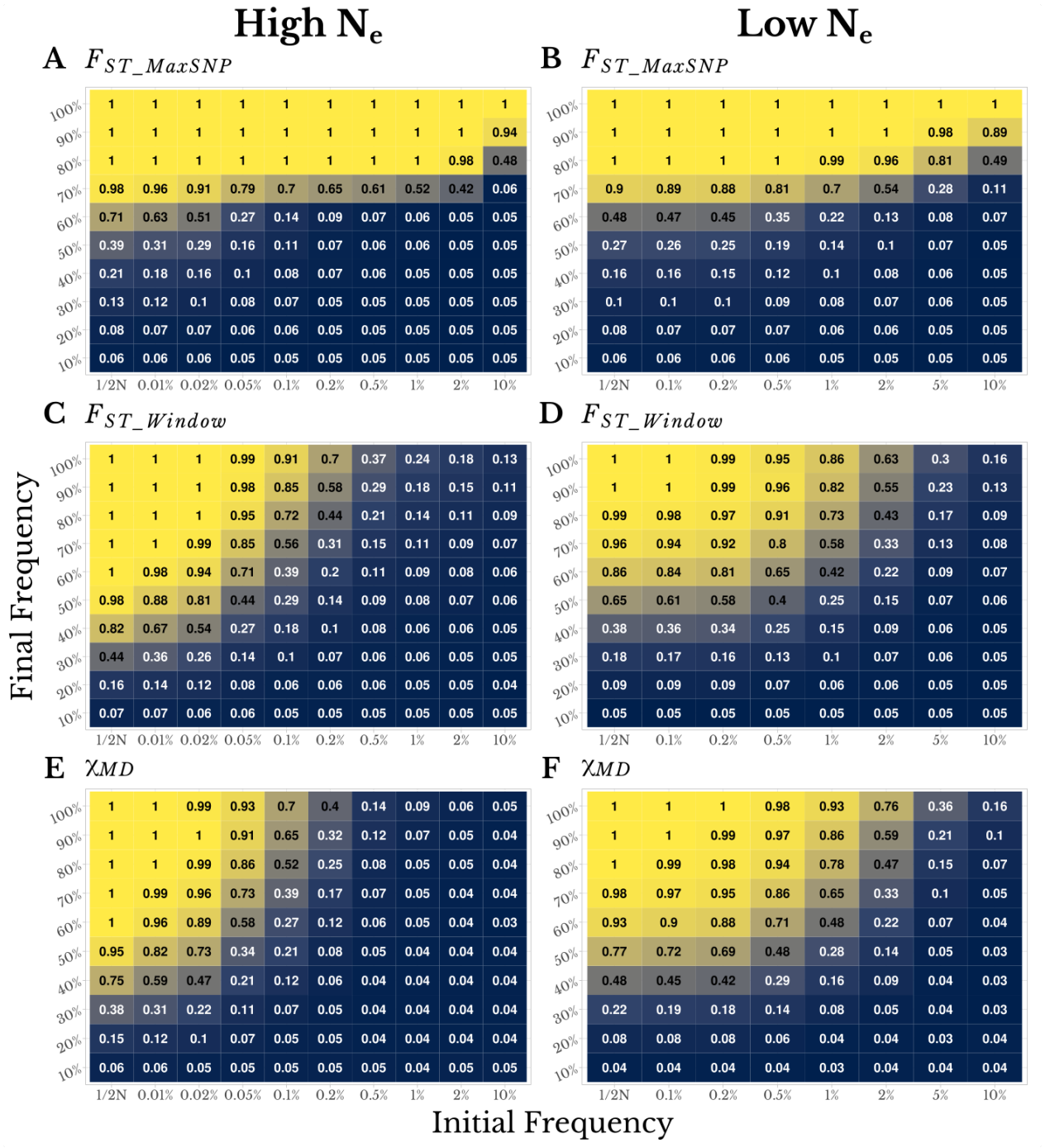


Figure 1. SNP-level F_{ST} and window-wide statistics show complementary power to detect local adaptation, depending on the type of selective sweep simulated. Numbers and colors in each panel both depict statistical power to detect local adaptation, in high N_e populations (s=0.001,

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

left column) and low N_e populations (s=0.01, right column). In each panel, the x-axis illustrates the pre-selection frequency of a favored variant (with the left column indicating selection on newly-occurring mutations) and the y-axis illustrates the final frequency of the sweep (with the top row showing complete sweeps). Detection power is shown for (A and D) F_{ST_MaxSNP}, (B and E) F_{ST_Window}, and (C and F) χ_{MD}. These results are based on a demographic history of simple isolation between two populations without change in population size, with a split time of $0.2N_e$ generations.

In the case of complete or nearly complete soft sweeps, F_{ST_MaxSNP} showed a clear power advantage over F_{ST_Window} and χ_{MD}. Notably, for sweeps ending between 80% and 100% frequency, F_{ST_MaxSNP} had high power to detect local adaptation, even for cases with rather high initial frequencies of the beneficial allele (e.g. 10%; Figure 1; Figure 2). In contrast, F_{ST_Window} and χ_{MD} showed rapidly diminishing performance as sweeps became softer (Figure 1; Figure 2). These results make sense, in that beneficial alleles that drift to higher pre-selection frequencies have more time to recombine onto multiple haplotypes, and recombination events will have happened closer to the selected site on average. Therefore, soft sweeps are generally narrower in width and may not substantially alter window-wide statistics (Catania *et al.* 2004; Schlenke and Begun 2004; Hermisson and Pennings 2005). Although the two window-wide statistics maintained good power for lower initial frequencies, some of the replicates of those scenarios are actually generating hard sweeps due to the chance survival of a single haplotype carrying the favored variant (Jensen 2014), as shown by an average number of beneficial haplotypes

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

lower than two in these simulations (Figure 2). Moreover, as the average number of haplotypes carrying the favored variant increased, the power of the window-wide statistics decreased (Figure 2), while the power of F_{ST_MaxSNP} was unchanged.

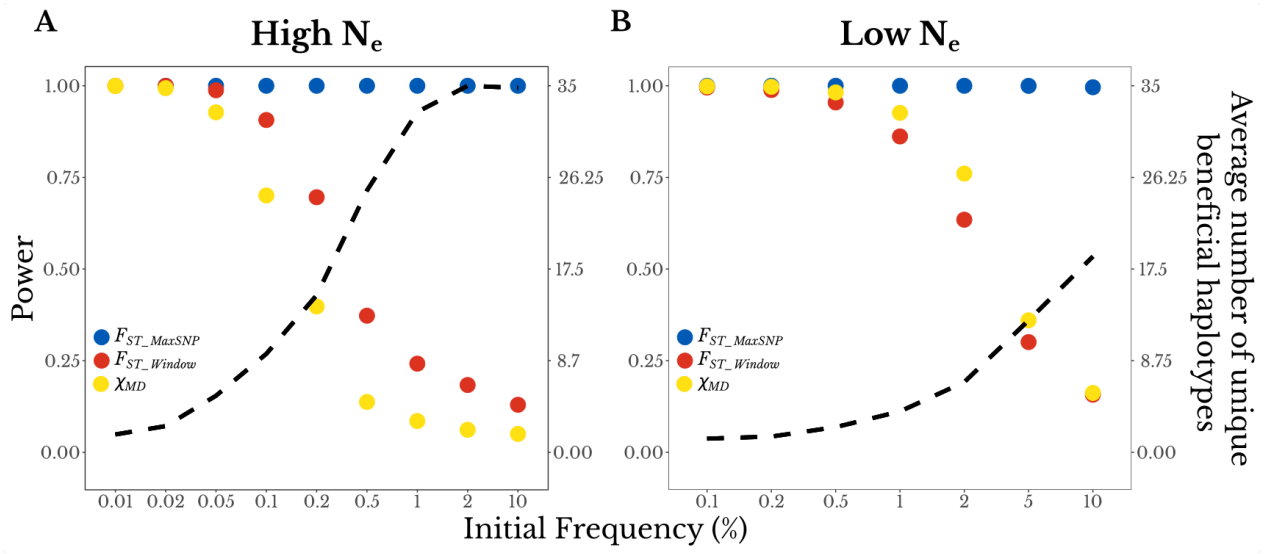


Figure 2. F_{ST_MaxSNP} shows an increasing power advantage as sweeps become softer. For complete sweeps with a range of initial frequencies (x-axis), the two y-axes show detection power for each statistic (left axis, dots) and the average number of unique beneficial haplotypes present at the end of the simulation (right axis, dashed line). Results are shown for (A) high N_e populations (s=0.001) and (B) low N_e populations (s=0.01), for the same demographic history as in Figure 1.

Contrasting results were obtained for partial, harder sweep scenarios. In cases where new mutations or rare standing variants were only boosted to intermediate frequencies, F_{ST_Window} and χ_{MD} had fairly strong power, whereas F_{ST_MaxSNP} declined sharply in effectiveness at

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

around 60% final frequency for hard sweeps (Figure 1). These results are also intuitive, in that partial hard sweeps can meaningfully alter allele frequencies across a whole window and generate a class of identical haplotypes, even though no single SNP traverses an extreme range of frequencies. The broadly similar power profiles of F_{ST_Window} and χ_{MD} are somewhat surprising in light of their distinct basis (albeit consistent with Lange and Pool, 2016). Less surprising is that for the challenging scenario of partial soft sweeps, none of the three statistics showed strong power in the scenarios examined (Figure 1).

Whereas the above simulations had no migration, we also wondered if F_{ST_MaxSNP} might prove useful in detecting targets of local adaptation for which genetic differentiation had been whittled down in width by recombination with migrant alleles over time. We therefore simulated scenarios with varying combinations of migration rate and population split time, while assuming symmetric migration rates and equal but opposing selective pressures. Overall, F_{ST_MaxSNP} and F_{ST_Window} performed very similarly to each other and better than χ_{MD}. Particularly in the high N_e scenarios (which feature a higher ratio of recombination to mutation events) with intermediate migration rates, there was a narrow space of parameters in which F_{ST_MaxSNP} performed slightly better than F_{ST_Window} (Figure S4). The split time between the populations greatly affected the power of χ_{MD}, which performed better on recent splits. The power of the F_{ST} statistics showed a small improvement for more recent splits and intermediate migration rates. Although small, the effect of split time also seemed more pronounced on F_{ST_Window} than F_{ST_MaxSNP} (Figure S4). Overall, these analyses provide only modest support for the notion that

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

240 F_{ST_MaxSNP} could help detect peaks of genetic differentiation driven by local adaptation that have
241 been narrowed by migration and recombination.

242 In the above simulations, we used a sample size of 50 chromosomes per population. We
243 generally expect statistical power to be correlated with sample size and understanding the
244 effect of sample size on the power of each statistic is relevant when designing an experiment or
245 choosing which statistics to use. We analyzed the power of F_{ST_MaxSNP}, F_{ST_Window}, and χ_{MD} in
246 three scenarios for high N_e and three for low N_e. We chose scenarios in which F_{ST_MaxSNP} and the
247 window wide statistics performed differently: a mostly complete soft sweep, a complete soft
248 sweep with a bottleneck, and a partial hard sweep. We found that sample size had a stronger
249 effect on F_{ST_MaxSNP} than on the window wide statistics (Figure 3). F_{ST_MaxSNP} is based on allele
250 frequencies at a single site, so it is more sensitive to the increased sampling variance at lower
251 sample sizes than window wide statistics. The sampling variance in each SNP in a window
252 should fluctuate around the mean, so when information from each SNP is combined the
253 window-wide statistic suffers less from the reduced sample size. Demographic history also
254 affected the effect of sample size on each statistic: in scenarios with a population bottleneck,
255 which also increases sampling variance, the power of F_{ST_MaxSNP} changed from near 1 at sample
256 size 50 or higher to 0 at sample sizes smaller than 50 (Figure 3C, 3D). More generally, F_{ST_MaxSNP}
257 was found to perform much better with 50 chromosomes than with 20, but showed relatively
258 less improvement for sample sizes larger than 50.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

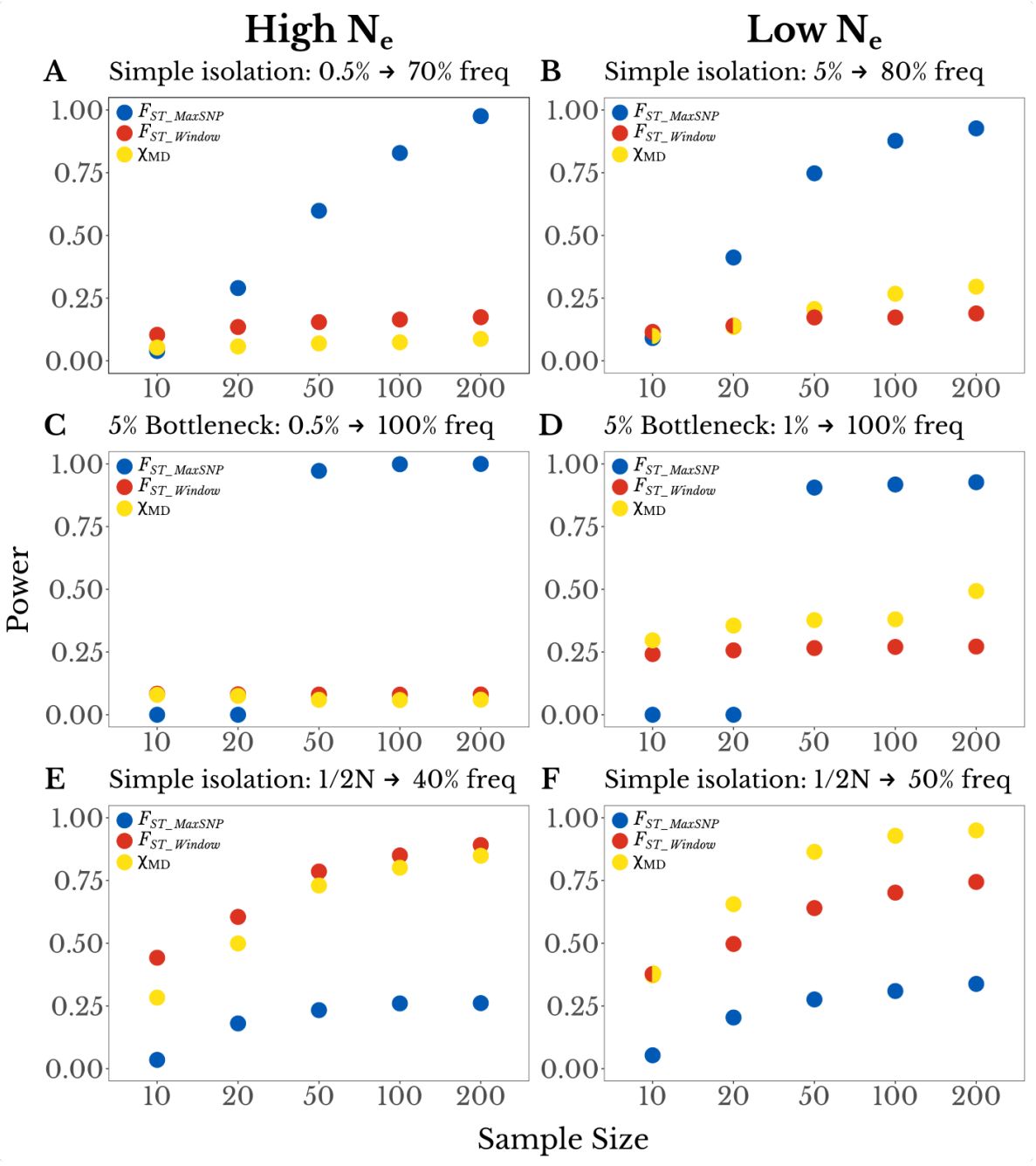


Figure 3. The power of F_{ST_MaxSNP} is particularly sensitive to sample size. Here, the power of each statistic (y-axis) is plotted as a function of sample size (x-axis; number of chromosomes per population). We found that depending on sample size, F_{ST_MaxSNP} outperforms F_{ST_Window} and χ_{MD} for a simple isolation model, for: (A) a high N_e population with initial beneficial allele frequency

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

of 0.005 and final frequency of 0.70,and (B) a low N_e population with initial frequency 0.05 and final frequency of 0.80. Similar results were observed for a complete soft sweep with a population bottleneck of 0.05, except that the loss of power for F_{ST_MaxSNP} was more immediate at lower sample sizes, for: (C) a high N_e population with initial frequency 0.05, (D) a low N_e population with initial frequency 0.01. For partial hard sweep scenarios where F_{ST_Window} and χ_{MD} outperform F_{ST_MaxSNP}, all three statistics show more gradual sample size effects, specifically for new mutations and: (E) a final frequency of 0.40 in a high N_e population, and (F) a final frequency of 0.50 in a low N_e population.

We also analyzed the effect of window size on the power of each statistic, with the aim of determining whether there would be a window size for which a single statistic would perform well in contrasting scenarios. For example, one might hope that F_{ST_Window} for a narrower window might retain good performance for partial hard sweeps, while also capturing the advantages of F_{ST_MaxSNP} for complete soft sweeps. We explored four scenarios of partial sweeps, two for the high N_e and two for the low N_e. For each population size, we chose one scenario in which the power of F_{ST_MaxSNP} outperformed F_{ST_Window} and χ_{MD}, and one in which it underperformed. In practice, a reduction in window size would result in an increase in the number of tests performed in a genome scan. Therefore, we applied a Bonferroni correction to the p-value proportional to the reduction in size. Our results showed that, for the two scenarios in which F_{ST_MaxSNP} outperformed F_{ST_Window} and χ_{MD}, the power of each statistic remained mostly constant (Figure 4). For the scenarios in which F_{ST_Window} and χ_{MD} had an advantage, the power

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

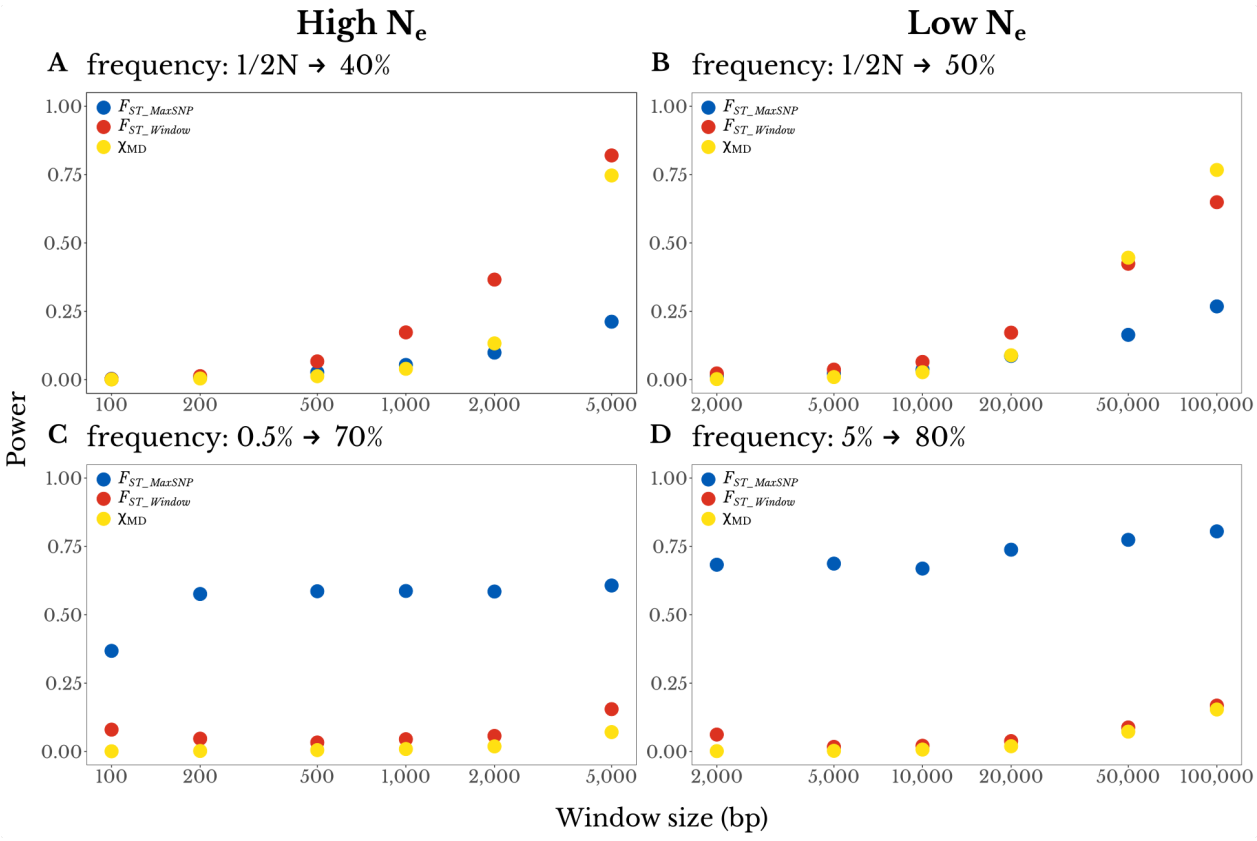
285 of each statistic, as well as the difference among them, declined with smaller window sizes.

286 Overall, there was no window size in which a single statistic performed well for all scenarios,

287 and hence it may be preferable to apply F_{ST_MaxSNP} and window-wide statistics separately to

288 empirical data.

289



290

291 **Figure 4.** Varying window size does not reveal a single statistic with broad detection power.

292 The top panels show partial hard sweeps for which F_{ST_Window} and χ_{MD} outperform F_{ST_MaxSNP}: (A)

293 a high N_e population with a final beneficial allele frequency of 0.40, And (B) a low N_e population

294 with a final frequency of 0.50. The bottom panels show mostly complete soft sweeps for which

295 F_{ST_MaxSNP} outperforms F_{ST_Window} and χ_{MD}: (C) a high N_e population with an initial beneficial allele

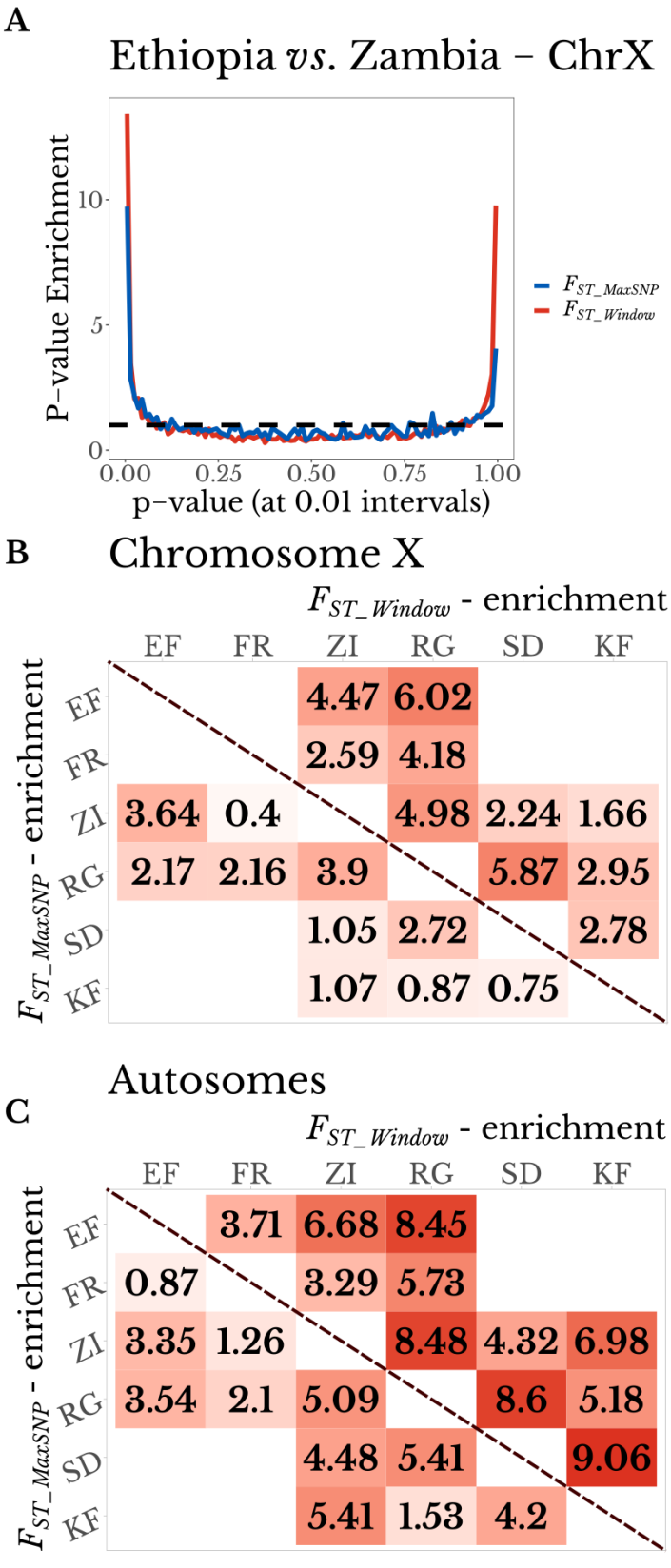
bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

296 frequency of 0.005 and final frequency of 0.70, and (D) a low N_e population with initial
297 frequency 0.05 and final frequency 0.80. These power values reflect a Bonferroni-corrected
298 significance threshold to reflect the relatively larger number of smaller windows needed.
299 Results do not suggest that any statistic in a smaller window size captures the advantages of
300 both F_{ST_MaxSNP} and the window-wide statistics.
301
302 **Outliers for F_{ST_MaxSNP} and F_{ST_Window} are enriched in empirical data**
303 In light of the above results, we were interested in applying both F_{ST_MaxSNP} and F_{ST_Window} to an
304 empirical data set, in part with an interest in quantifying the relative enrichment of outliers for
305 each statistic, and what that might suggest about the modes of selection active in these
306 populations. We chose to focus on data from the *Drosophila* Genome Nexus (Lack *et al*. 2015,
307 2016), because it contained several populations of *D. melanogaster* that were linked by an
308 estimated model of population history (Sprengelmeyer *et al*. 2020) and had at least minimal
309 sample sizes available for studying genome-wide patterns of F_{ST} (Table S2). We included six
310 natural populations of flies. From the ancestral range in Zambia, we included one town
311 population (Siavonga) and one wilderness population (Kafue). We also included four additional
312 town populations: from Rwanda, South Africa, Ethiopia, and France (the latter three having
313 independently colonized colder environments; Pool *et al*. 2017).
314 We calculated a p-value for each empirical window in each pairwise population
315 comparison, based on neutral distributions of F_{ST_MaxSNP} or F_{ST_Window} generated using coalescent
316 simulations of the estimated demographic history (Sprengelmeyer *et al.* 2020; simulation

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

commands in Table S2). Under neutrality, a uniform distribution of p-values is expected. In general, for most population pairs, the distribution of p-values for F_{ST_MaxSNP} and F_{ST_Window} showed a U-shape instead of an uniform distribution (e.g. Figure 5A). Nonetheless, the enrichment of high F_{ST} (defined as p-values from 0 to 0.05) and low F_{ST} (p-values from 0.95 to 1) varied for each statistic and across the population comparisons (Figure 5B-C).

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.



322

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

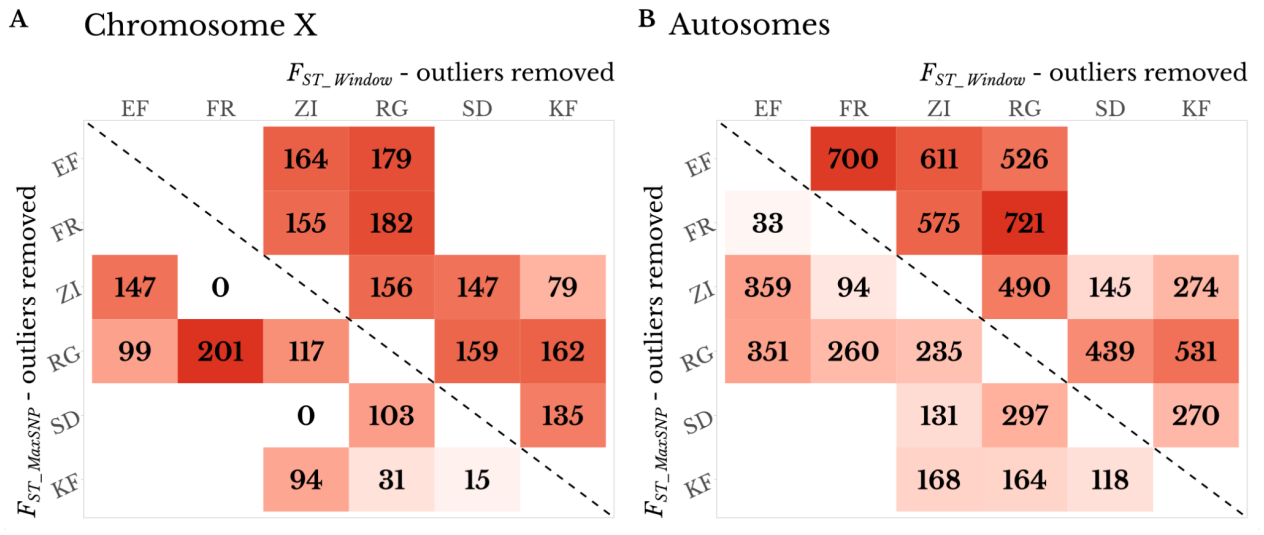
Figure 5. F_{ST_MaxSNP} and F_{ST_Window} both show outlier enrichment between natural populations of *D. melanogaster*. (A) Ethiopia-Zambia F_{ST_MaxSNP} and F_{ST_Window} values on chromosome X show enrichment of low (right) and especially high values (left), based on the distribution of p-values obtained from neutral demographic simulations. (B and C) F_{ST_MaxSNP} (lower diagonal) and F_{ST_Window} (upper diagonal) both show enrichment of high outliers on (B) chromosome X and (C) combined autosome arms. F_{ST_Window} shows a greater enrichment in nearly all cases. Populations: SD = South Africa. ZI = Zambia. KF = Kafue, Zambia. RG = Rwanda. EF = Ethiopia. Population pairs not present in the same demographic model were not evaluated. Color scale ranges from the minimum to maximum value within each panel.

All population pair comparisons showed an enrichment for windows with high F_{ST_Window}. The smallest enrichment was found between the Zambia (town) and France populations, for which there were 3.29 more windows with high F_{ST_Window} than expected by chance. The highest enrichment was found in the comparison between the South Africa and Kafue (Zambia wilderness) populations, with an enrichment factor of 9.06. For F_{ST_MaxSNP}, eight population pairs had an enrichment value > 2, the highest being 5.41 (between the Zambian town and wilderness populations, and between South Africa and Rwanda). On the other hand, one population pair was slightly depleted of windows with high F_{ST_MaxSNP} (enrichment to 0.87 between France and Ethiopia). In most comparisons, F_{ST_Window} showed higher enrichment than F_{ST_MaxSNP}. The only exception was the comparison between South Africa and Zambia (town population), in which both enrichments were very similar: F_{ST_MaxSNP} enrichment was 4.48 and F_{ST_Window} enrichment

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

4.32 (Figure 5). This large variation in enrichment between populations suggests that the kind and prevalence of selective sweeps unique to each population may vary among populations.

The almost universally greater enrichment of F_{ST_Window} relative to F_{ST_MaxSNP} could hint that sweeps in the unique detection parameter space of F_{ST_Window} (*i.e.* partial harder sweeps) are more common among these populations than sweeps in the unique space of F_{ST_MaxSNP} (*i.e.* more complete softer sweeps). However, the above enrichments may be influenced by locally adaptive sweeps that create multiple linked outlier windows. We therefore pursued a complementary analysis in which nearby outlier windows were merged into "outlier regions", which were then removed one at a time until the observed enrichment was erased (see Materials and Methods). For almost every population pair, we had to remove a larger number of regions to erase the signal of enrichment of F_{ST_Window} than the signal of F_{ST_MaxSNP} (Figure 6). Hence, the greater enrichment of F_{ST_Window} relative to F_{ST_MaxSNP} does not appear to be a product of broader linkage signals of F_{ST_Window} outliers alone.



bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

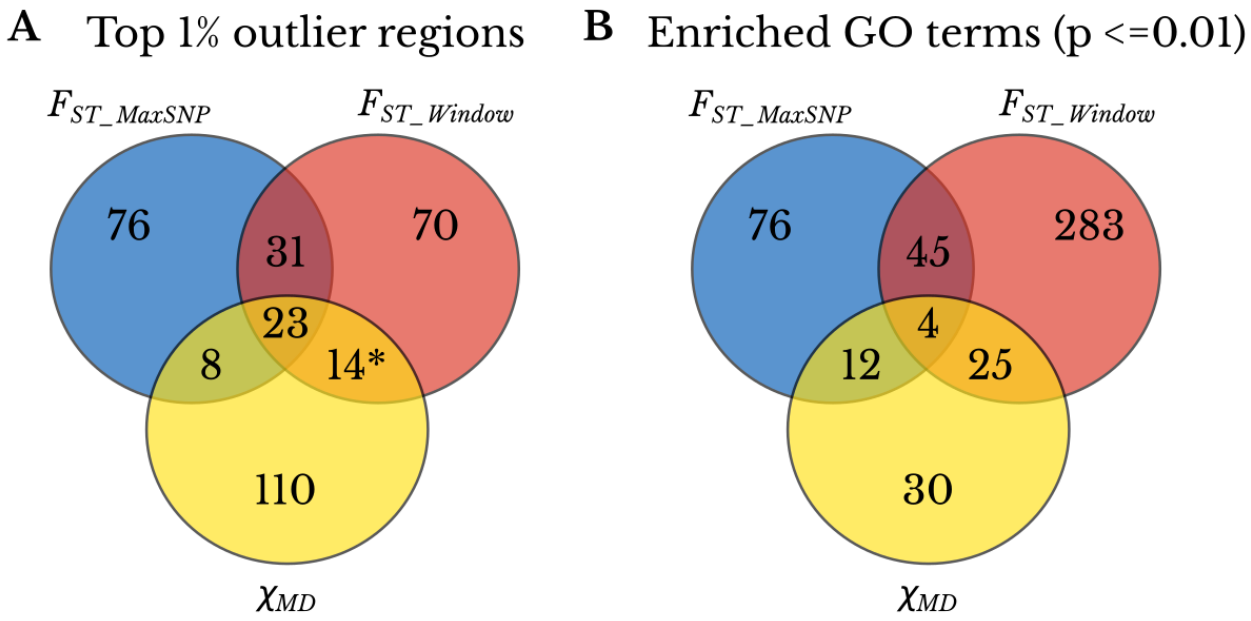
Figure 6. Number of outlier regions that were removed to erase the signature of enrichment for high F_{ST_MaxSNP} (lower diagonal) and F_{ST_Window} (upper diagonal) for each population on (A) chromosome X and (B) the combined autosome arms. F_{ST_Window} was associated with a greater outlier region enrichment for most population pairs, reinforcing the window-level patterns shown in Figure 5. Populations: SD = South Africa. ZI = Zambia. KF = Kafue, Zambia. RG = Rwanda. EF = Ethiopia. Population pairs not present in the same demographic model were not evaluated. Color scale ranges from the minimum to maximum value within each panel.

Genome Scan for Signatures of Selection

We chose to complement the above multi-population analysis of genome-wide patterns with a closer analysis of a single population pair. We chose to compare the Ethiopia and Zambia town populations because (1) Their relatively large sample sizes of 129-181 and 60-76 respectively for each chromosome arm (Table S2) are more conducive to the analysis of specific F_{ST_MaxSNP} outliers, (2) These populations showed enrichments of both F_{ST_MaxSNP} and F_{ST_Window} (Figure 4), and (3) Past results from these populations helped motivate the present study (*e.g.* Bastide *et al.* 2016). We performed genome scans for regions potentially under population-specific selection between these populations using F_{ST_MaxSNP}, F_{ST_Window}, and χ_{MD}. For each statistic, we obtained a list of outlier windows (top 1%), and as above, we merged nearby outlier windows into regions (Materials and Methods). We obtained 138 outlier regions for F_{ST_MaxSNP}, 138 for F_{ST_Window} , and 155 for χ_{MD}. Our results showed an overlap of just 39% between the outlier regions detected with F_{ST_MaxSNP} and F_{ST_Window}. Perhaps surprisingly in light of the above power

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

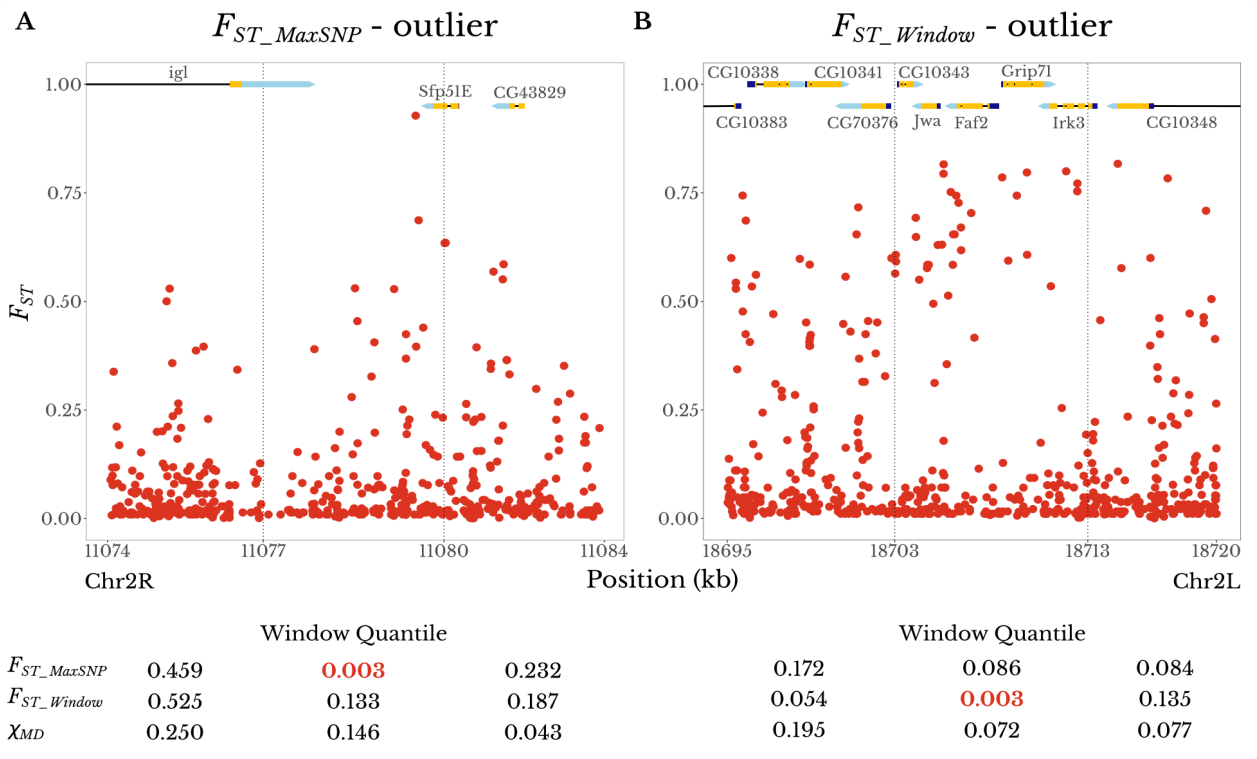
379 results, there was a smaller overlap of either F_{ST} metric with χ_{MD} (Figure 7A), although the
380 overlap of the haplotype statistic with F_{ST_Window} was indeed slightly greater. In regions that
381 were outliers for F_{ST_MaxSNP} but not F_{ST_Window}, the distribution of individual SNP F_{ST} values often
382 had a narrow sharp F_{ST} peak, with most of the other SNPs having low F_{ST} values. On the contrary,
383 in regions there were outliers for F_{ST_Window} but not F_{ST_MaxSNP}, often no single SNP had a large F_{ST}
384 value, but there was a broad moderate F_{ST} plateau with many SNPs showing intermediate F_{ST}
385 values (Figure 8).



386

387 **Figure 7.** The three statistics detect mostly unique genomic regions and functional categories.
388 (A) Overlap between the top 1% outlier regions detected with F_{ST_MaxSNP}, F_{ST_Window}, and χ_{MD}. *
389 indicates the average number of outlier regions between the two statistics: 15 F_{ST_Window} outlier
390 regions exclusively overlap χ_{MD} outliers and 13 χ_{MD} outlier regions exclusively overlap F_{ST_Window}
391 outliers. (B) Overlap between enriched GO terms with raw p-value <= 0.01 based on the outlier
392 regions detected with F_{ST_MaxSNP}, F_{ST_Window}, and χ_{MD}.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.



	Window Quantile			Window Quantile		
F_{ST_MaxSNP}	0.459	**0.003**	0.232	0.172	0.086	0.084
F_{ST_Window}	0.525	0.133	0.187	0.054	**0.003**	0.135
χ_{MD}	0.250	0.146	0.043	0.195	0.072	0.077

Figure 8. Examples of the distinct SNP-level F_{ST} landscapes associated with F_{ST_MaxSNP} versus F_{ST_Window} outliers. Each plot shows an outlier window for an Ethiopia-Zambia F_{ST} statistic, plus its adjacent windows. Dashed vertical lines delimit the boundaries of the windows. Numbers under each window are the empirical quantiles of that window's statistic (F_{ST_MaxSNP}, F_{ST_Window}, and χ_{MD}) in relation to the chromosome arm-wide distribution of the same statistic, with the outlier (quantile < 0.01) value in red. (A) An outlier window for F_{ST_MaxSNP} (center) shows a peak-like F_{ST} landscape with one particularly differentiated SNP. (B) An outlier window for F_{ST_Window} (center) shows a broad plateau of fairly high F_{ST} values. Gene names and structures are shown at the top of each plot. Protein-coding exons are in yellow, while 5' and 3' untranslated regions are in dark blue and light blue, respectively.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

We then performed GO term enrichment analysis separately for each statistic's list of outlier regions. Considering only GO terms with raw p-value < 0.01 from each list, we found mostly lower overlaps between enriched GO terms compared to the spatial overlap between outlier regions (Figure 7B; Table S3). The three statistics differed substantially in the number of enriched GO terms by this criterion: 357 for F_{ST_Window}, 133 for F_{ST_MaxSNP}, and 71 for χ_{MD} (although we emphasize that these terms are not independent and any given list of enriched GO terms will contain overlapping categories). The relative overlap between GO terms enriched for each statistic largely followed the relative numbers of enriched GO terms for each (Figure 7B). Mirroring the outlier region results, most enriched GO terms were detected for only one of the three statistics, consistent with their complementary detection powers described above.

Discussion

F_{ST_MaxSNP} complements other statistics by detecting soft sweeps

Identifying regions under selection can help us answer further questions about the evolution of local adaptation, such as which biological functions are under selective pressure, the number of loci underlying adaptive events, the source of the adaptive variation, and the kinds of genetic changes that might be under selection. Our results underscore the importance of deploying methods capable of capturing different kinds of selective sweeps when the aim of the study is to identify as many genes potentially under local adaptation as possible.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

425 F_{ST_MaxSNP} in particular, seems to be especially useful to detect soft sweeps with relatively
426 large initial and final frequencies of the beneficial allele. Instances of mostly complete soft
427 sweeps, as simulated here, represent regions in which a beneficial allele was present in several
428 different haplotypes that might have increased in frequency along with the beneficial allele.
429 While the selected SNP itself changed in frequency drastically, resulting in a large F_{ST_MaxSNP}, the
430 alleles around it must have changed in frequency to a lesser degree because many background
431 haplotypes were hitchhiking along with the beneficial allele. Therefore, while the beneficial
432 variant can have an extreme F_{ST} value, the lower allele frequency changes in the other SNPs in
433 that window would result in a F_{ST_Window} that is not statistically significant, and thus a low power
434 to detect a selective sweep under these conditions.
435 The window-wide metrics, F_{ST_Window} and χ_{MD}, had greater power than F_{ST_MaxSNP} to detect
436 relatively harder, partial sweeps that had intermediate final allele frequencies. In these sweeps,
437 no individual SNP changed dramatically in frequency, so none have F_{ST} values higher than what
438 could be obtained randomly in the genome. However, the increase in frequency of one or a few
439 haplotypes resulted in many SNPs in the same region with intermediate F_{ST}, producing a
440 window-wide pattern that is too extreme to be generated by chance - even if each single marker
441 individually did not have an extreme F_{ST} value.
442 There was little difference in the power of F_{ST_MaxSNP} and F_{ST_Window} to detect regions
443 under selection in scenarios with varying migration rates. We had expected that F_{ST_MaxSNP}
444 would outperform F_{ST_Window} in scenarios with older splits, as selection might only maintain a
445 narrow window of differentiation between the two populations in the presence of long-term

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

recombination with migrant haplotypes. Nonetheless, differences in split time between the two populations only had a small effect in a very narrow space of parameters (intermediate migration rates for high N_e populations, Figure S1), suggesting that even in scenarios with recent divergence, the populations had already reached a state of equilibrium and the balance between migration, selection, and recombination did not result in distinguishable signatures of selection between F_{ST_MaxSNP} and F_{ST_Window}. However, both metrics outperformed χ_{MD} on the simulated scenarios, indicating that selection could not maintain long shared haplotypes in the presence of migration.

In light of the complementary performance of F_{ST_MaxSNP} and F_{ST_Window} for the non-migration cases, we tested whether F_{ST_Window} across shorter windows could yield a balance of reasonable power to detect both complete soft sweeps and partial hard sweeps. However, the relationship between window size and the power - while accounting for the increase in the number of tests in smaller windows - did not follow this prediction. Our results suggest that applying both F_{ST_MaxSNP} and F_{ST_Window} to conventionally-sized windows is preferable to shrinking the window size in an effort to identify narrower soft sweeps. More generally, we suggest that genetic differentiation on both SNP and broader scales should be incorporated into scans for local adaptation, whether using the specific summary statistics described here, or attempting to develop a single statistic or integrated analysis framework that encompasses the advantages of both.

In this study, we have used neutral demographic simulations to estimate statistical power at the single window level, only penalizing multiple tests when comparing between

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

467 window sizes. Clearly, our results do not imply the power to identify genome-wide significant
468 loci, which is only rarely attainable for population genomic scans. Instead, most genome scans
469 aim to identify good candidates for downstream study, and our results are best interpreted in
470 terms of the relative utility of these summary statistics to identify local adaptation candidates.
471 An important caveat of using F_{ST_MaxSNP} is that it requires a greater sample size than
472 F_{ST_Window}. With smaller samples, it is easy to get a large F_{ST_MaxSNP} at one of the many analyzed
473 SNPs through sampling variance alone, whereas an extreme F_{ST_Window} value is less likely in this
474 scenario. It is difficult to provide any universal advice regarding sample size, because the
475 neutral variance of F_{ST_MaxSNP} also depends strongly on demographic history, as shown above.
476 Nonetheless, we have shown that in two scenarios in which F_{ST_MaxSNP} outperformed F_{ST_Window} its
477 power declined considerably when we decreased the sample size from 50 to 20 chromosomes.
478 Although the relationship between sample size and power will depend on the specific
479 populations being studied, the utility of F_{ST_MaxSNP} seems most promising when sample sizes are
480 around 100 alleles per population or more. However, it would be advisable to conduct neutral
481 simulations based on estimated or suspected demography, in order to identify sample sizes for
482 which it is very unlikely to get extreme SNP F_{ST} values in the absence of local adaptation.
483
484 **Both F_{ST_Window} and F_{ST_MaxSNP} outliers are enriched among *Drosophila* populations**
485 When we applied F_{ST_Window} and F_{ST_MaxSNP} to empirical data from *D. melanogaster* populations,
486 we found that enrichment patterns of F_{ST_Window} and F_{ST_MaxSNP} varied among population pairs,
487 both for high and low F_{ST} values. The excess of windows with high F_{ST} observed could be

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

488 explained by local adaptation: unique selective sweeps in one population increase the
489 differentiation between two populations in that region. Not all population pairs showed the
490 same degree of enrichment for high F_{ST}. A larger enrichment could be due to a higher number
491 of selective sweeps between two populations, stronger selective events that impacted a larger
492 region of the genome, or a neutral history more conducive to outlier detection. The populations
493 we studied cover a large geographical scale, most are located in sub-Saharan Africa and one in
494 Europe. These populations are exposed to a variety of environments, ranging from warm
495 tropical lowlands to cool high latitude and high altitude regions, in addition to commensal
496 versus wilderness settings (Sprengelmeyer *et al.* 2020). Hence, they are most likely exposed to
497 several unique selective pressures that could be underlying local adaptation and an enrichment
498 of high F_{ST} values.

499 Alternatively, enrichment for high F_{ST} could also be explained by background selection,
500 which is expected to reduce genetic diversity and therefore result in lower effective population
501 sizes in that genomic region. Genetic drift is stronger in regions of low N_e, which could increase
502 the differentiation between two populations and produce high F_{ST} (Charlesworth *et al*. 1993).
503 However, a simulation study of background selection targeting stickleback exons found no
504 evidence for background selection increasing F_{ST} outliers (Matthey-Doret and Whitlock 2019).

505 On the other extreme, the existence of enrichment for low values of F_{ST} suggests that
506 many regions of the genome maintained unexpectedly similar allele frequencies between two
507 populations. Following a population split, neutral evolutionary forces such as genetic drift are
508 expected to increase the genetic differences between two populations. The fact that many

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

509 regions seemed to have changed less than what was expected due to neutral forces could also
510 be explained by the action of natural selection. This could be the product of shared selective
511 sweeps (i.e. similar selective pressures) taking place in both populations, instead of local
512 adaptation. Shared balancing selection could also be acting at some loci to maintain allele
513 frequencies constant between two populations, perhaps even from before their split time.
514 We should also acknowledge that the demographic models applied here are simply the
515 best available estimates of population history, and no demographic model fully accounts for the
516 complexity of natural populations. Demographic model misspecification could result in some
517 enrichment of high and/or low F_{ST} values. One potential source of error in demographic
518 estimation is natural selection. The demographic models were estimated based on tentatively
519 neutral regions of the genome (Sprengelmeyer *et al.* 2020). However, these regions could be
520 under the influence of linked positive and negative selection, with the potential to bias
521 demographic estimation. For example, if the presumed neutral data was substantially affected
522 by either local adaptation or shared sweeps, it could bias the neutral distribution of F_{ST} towards
523 higher or lower values, respectively, making it more difficult to detect F_{ST} outliers in that
524 direction. Nonetheless, previous work suggests that this effect might be weak on demographic
525 inference in *D. melanogaster* (Lange and Pool 2018).
526 In nearly all population pairs, F_{ST_Window} showed a larger enrichment than F_{ST_MaxSNP}. The
527 greater enrichment of F_{ST_Window} persisted when we instead pursued an outlier region removal
528 strategy. In light of the complementary zones of power shown in Figure 1, these results suggest
529 that roughly speaking,there might be a larger contribution of partial hard sweeps than complete

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

soft sweeps to local adaptation among these populations. Furthermore, the fairly low levels of outlier overlap between F_{ST_Window} and F_{ST_MaxSNP} may suggest that the sweeps both statistics can reliably detect (*i.e.* more complete harder sweeps) are not the primary drivers of local adaptation in this data set. Overall, these results suggest that partial sweeps might have played a large role in the adaptation of fly populations to diverse environments. The importance of partial sweeps in populations of *D. melanogaster* has been proposed previously, including for some of the populations studied here (Pool and Aquadro 2007; Bastide *et al.* 2016; Garud and Petrov 2016; Vy *et al*. 2017).

Here, we have shown that SNP-level F_{ST} (F_{ST_MaxSNP}) offers strong power to detect soft sweeps, and is highly complementary to window-wide frequency and haplotype statistics for detecting local adaptation. These results stress the importance of taking into account the different signatures left by different kinds of selective sweeps in the genome when deciding how to perform a genome scan. The raw summary statistics evaluated here can either be applied in parallel, or their signals can be integrated into frameworks such as approximate Bayesian computation and machine learning. Thus far, the latter methodologies have been used more extensively to detect and classify selective sweeps within a single population (Peter *et al*. 2012; Sheehan and Song 2016; Schrider and Kern 2016, 2017). However, such approaches are equally applicable to the study of local adaptation (Key *et al.* 2014). Future work could investigate whether methods that combine multiple statistics would benefit from including F_{ST_MaxSNP}, potentially increasing their power to detect soft sweeps and their accuracy in classifying different types of sweeps. Because studies of genetic differentiation between

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

populations inherently control for evolutionary variance in the shared ancestral population, local adaptation may offer a better "signal to noise ratio" regarding the types of positive selection acting in natural populations, compared to single population studies. Hence, our results may contribute toward not only an improved ability to detect local adaptation, but also a clearer understanding of adaptation in nature more generally.

Methods

Simulation Power Analysis

To generate adaptive and neutral distributions of genetic diversity, we performed simulations of demographic history scenarios with and without natural selection using *msms* (Ewing and Hermisson 2010). For each model, we obtained 10,000 replicates from which we calculated the statistics of interest. Power was calculated as the proportion of replicates under selection with a statistical value larger than 95% of the values obtained in its corresponding replicates without selection. We investigated the power of three different statistics: F_{ST_MaxSNP}, F_{ST_Window} and χ_{MD} (Lange and Pool 2016), which were calculated on windows of fixed size. F_{ST_MaxSNP} is based on the SNP within a window with the highest F_{ST} value. F_{ST_Window} was calculated as the weighted average of all SNPs in a window (Reynolds *et al*. 1983). χ_{MD} stands for Comparative Haplotype Identity; it compares the average length of identical haplotypes in a window between two populations, and was calculated following Lange and Pool (2016). The window size used was 5,000 bp for simulations of populations with high effective population size (N_e) and 100,000 bp for simulations of populations with low N_e. Except where otherwise stated, the sample size was

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

572 50 chromosomes. The high N_e simulations used parameters similar to those from flies
573 (*Drosophila melanogaster*) while the low N_e had parameters similar to humans (simulation
574 parameters followed Lange and Pool, 2016).

575 We initially used scenarios of constant population size and a simple population split to
576 simulate scenarios of selective sweeps with varying initial and final allele frequencies,
577 representing hard and soft sweeps as well as complete and partial sweeps. We also simulated
578 scenarios of population bottlenecks and population splits for complete selective sweeps, and for
579 scenarios with varying migration rates for hard sweeps (not constrained by ending allele
580 frequency). For bottlenecks, the population that will experience local adaptation underwent a
581 period of reduced population size for the first 0.01 coalescent units after the population split
582 (which in most scenarios including these, occurred 0.05 coalescent units ago; Table S1).

583 The simulations of populations with high N_e were done for two different selection
584 coefficients (s = 0.01 and s = 0.001) and simulations of populations with low N_e only included s =
585 0.01 (Table S1). Simulations of complete sweeps only used replicates in which the beneficial
586 allele went to fixation. Simulations of partial sweeps only accepted replicates in which the
587 beneficial allele stayed within 4% of the targeted ending frequency. Selection initiation time
588 was adjusted in each case to maximize the proportion of accepted replicates. Moreover, in the
589 scenarios with initial allele frequencies larger than $1/2N_e$, both the selected and non-selected
590 populations had the same initial frequency.

591 For models that included migration (gene flow), selection of equal magnitudes but in
592 opposite directions was imposed on each population. Per generation migration rates varied

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

593 from 0.0004 to 0.004 in simulations with high N_e populations and from 0.01 to 0.10 in
594 simulations with low N_e populations. For each migration rate, split times varied from 0.1 to 1
595 coalescent unit.

596 We calculated the effect of sample size on the power of each statistic in six different
597 scenarios: four models with demographic history of a simple isolation between two populations
598 and two models with population size bottleneck. Of the simple isolation models, two models for
599 high N_e populations were considered: one in which F_{ST_Window} outperformed F_{ST_MaxSNP} (initial
600 allele frequency of $1/2N_e$ and final allele frequency of 0.4) and another where F_{ST_MaxSNP}
601 outperformed F_{ST_Window} (initial frequency of 0.005 and final frequency of 0.7). Two scenarios for
602 low N_e populations were also considered: one in which F_{ST_Window} outperformed F_{ST_MaxSNP} (initial
603 allele frequency of $1/2N_e$ and final allele frequency of 0.5) and another where F_{ST_MaxSNP}
604 outperformed F_{ST_Window} (initial frequency of 0.05 and final frequency of 0.8). For the bottleneck
605 models, we used models with a bottleneck of 5% (*i.e.* a reduction to 5% of the prior N_e for 0.01
606 coalescent units in the adapting population immediately following the population split) and only
607 models in which F_{ST_MaxSNP} outperformed the window wide statistics were considered: one
608 model for high N_e population (initial allele frequency from 0.5% to 100%) and one for low N_e
609 populations (initial allele frequency from 1% to 100%). For all the six scenarios, we used sample
610 sizes of 10, 20, 50 (original sample size), 100, and 200 chromosomes.

611 We calculated the effect of window sizes on the power of each statistic in four different
612 scenarios, the same scenarios of simple isolation used to calculate the power of sample sizes
613 above. For the high N_e scenarios, we used window sizes of 5 kb (original size), 2 kb, 1 kb, 0.5 kb,

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

614 0.2 kb, and 0.1 kb. For the low N_e scenarios, we used window sizes of 100 kb (original size), 50
615 kb, 20 kb, 10 kb, 5 kb, and 1 kb. To calculate χ_{MD}, we used a minimum haplotype threshold of
616 10% of the window size (as was used for the original analyses). For each window size smaller
617 than the original, we applied a p-value Bonferroni multiple testing correction proportional to the
618 reduction in size (or equivalently, the increased number of windows needed to cover a given
619 genomic region) to calculate power. That is, while for the standard window size power is the
620 number of replicates with a p-value of 0.05 or lower, for a window half the size of the original
621 the p-value would need to be 0.025 or lower.
622
623 **Empirical Enrichment of F_{ST_MaxSNP} and F_{ST_Window} - data and simulations**
624 Our data set consists of individual fly strain genomes from six natural populations of *D.*
625 *melanogaster*: one non-human commensal population from Kafue, Zambia (KF) and five human
626 commensal populations from different countries: Zambia (ZI), South Africa (SD), Rwanda (RG),
627 Ethiopia (EF) and France (FR), using data from Lack *et al.* (2016) and Sprengelmeyer *et al.* (2020).
628 From each population, for each chromosome arm (ChrX, Chr2L, Chr2R, Chr3L, Chr3R), we
629 excluded genomes from lines with a known inversion for that arm. To boost the sample size of
630 two populations with genomes from partially inbred lines (Ethiopia and France), instead of only
631 using homozygous regions of the genome (as in the original filtering of the published data set)
632 we also included heterozygous regions identified by Lack *et al*. (2016), and therefore counted
633 two alleles at each site from these regions. For any pair of lines with excess identity by descent
634 (IBD) between them (defined as more than 10 megabases of IBD outside previously defined

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

regions of low recombination; Lack *et al.*, 2016), we excluded one member of the pair from this data set. For each population sample and each chromosome arm, we chose a sample size to jointly maximize the number of analyzable sites and the sample size itself. Our resulting sample sizes are shown on Table S2. For sites with more than that number of alleles called, we downsampled to match the chosen sample size.

We calculated pairwise F_{ST_Window} and F_{ST_MaxSNP} for all populations using diversity-scaled window sizes designed to contain 250 non-singleton SNPs in the ZI sample. To compare empirical and null distributions for similar recombination rates, each window was assigned to one of five recombination rates bins based on estimates from Comeron *et al*. (2012); the bins corresponded to recombination rates from 0.5-1, 1-1.5, 1.5-2, 2-3, and greater than 3. Windows with recombination rates lower than 0.5 were not used due to low spatial resolution for localizing signatures of selection in low recombination regions. We obtained p-values for each window using neutral demographic simulations performed using *ms* (Hudson 2002). Demographic simulations were performed using parameters estimated for the evolutionary history of nine populations of *D. melanogaster*, including all the populations we analyzed (Sprengelmeyer *et al.* 2020). The other three populations were lowland Ethiopia (EA), Cameroon (CO), and Egypt (EG). We did not use those three populations in our empirical analyses due to their lower sample sizes. Nonetheless, they were included in the simulations in order to accurately reflect the estimated patterns of migration.

Each demographic model had been estimated based on tentatively neutral genetic markers (short introns and 4-fold synonymous sites from regions with sex-averaged

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

recombination rates of at least 1 cM/Mb) from inversion-free chromosome arms (Sprengelmeyer *et al.* 2020). A model was estimated for each of three chromosome arms that had lower inversion frequencies (X, 2R, and 3L), and the history was inferred iteratively, such that not all population samples were present in the same model. To better approximate genetic diversity in all populations, we used two sets of demographic models: Northern model (containing ZI, RG, CO, EF, FR, EG, EA) and Southern model (containing ZI, RG, CO, SD, and KF). The Northern model for the chromosome X was subdivided into two sub-models (one with ZI, RG, CO, EF, EA and another with ZI, RG, CO, FR, EG). Hence, we simulated four Northern models and three Southern models (command lines in Table S2). The models for the autosomal chromosome arms (2R and 3L) were simulated using the highest sample sizes for any autosomal arm of each population (Table S2). Simulated sample sizes were downsampled to match the sample sizes of each specific arm when comparing empirical and simulated F_{ST} patterns for any given arm. The window size and crossing over rate used in each replicate were based on a random sampling with replacement from the empirical windows, and the single gene conversion rate and mean tract length were based on the estimates of Comeron *et al*. (2012). Therefore, a null distribution was generated for each model and each recombination bin (described above). For each model and each recombination bin, 50,000 replicates were simulated.

Enrichment calculation

F_{ST_Window} and F_{ST_MaxSNP} were calculated for each population pair and each chromosome arm. F_{ST} was calculated for the simulated data using the same sample sizes as the empirical data (Table

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

S2). For sites with more than two alleles, only the two most common alleles were kept. Sites with minor allele counts lower than two were discarded from empirical and simulated analyses.

P-values were calculated for each window based on the neutral distribution of its corresponding recombination group. Windows from chromosome X were compared to neutral distributions based on the model for chromosome X. For autosomal loci, we determined that simulations from the 3L model yielded somewhat milder outlier enrichments than the 2R model, and therefore we conservatively focused on results from the 3L model.

We calculated p-value enrichments for F_{ST_Window} and F_{ST_MaxSNP} using p-value bins of width equal to 0.05, resulting in 20 bins of p-value 0 to 1. We counted how many windows had a given p-value for each bin and divided the observed number by how many windows we expected to have with a p-value in that bin based on simulated data.

Neighboring windows with low p-value could be showing the effect of a single selective sweep. Therefore, we complemented this outlier window enrichment analysis with one based on "outlier regions". We intentionally defined outlier regions generously, preferring to falsely lump two sweeps versus splitting a single sweep into two or more regions. Formally, starting with the window containing the lowest p-values, we extended the region surrounding it until we reached a stretch of five consecutive windows with $p > 0.1$ to create an outlier region. We removed the outlier regions from our analysis and repeated the process until the signal of enrichment was erased (defined as the $p < 0.05$ bin having no more enrichment than the $0.05 < p < 0.1$ bin). For each of F_{ST_MaxSNP} and F_{ST_Window}, we recorded the total number of outlier regions that had to be removed for a given population pair.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Genome scan for regions under selection - Ethiopia vs. Zambia

We performed a genome scan for candidate regions under selection between the Ethiopia (EF) and Zambia (ZI) populations. We calculated F_{ST_Window}, F_{ST_MaxSNP}, and χ_{MD} for each window of the genome. We used an outlier approach and considered windows in the top 1% of each statistic to be the candidate regions under selection. Here, we combined multiple outlier windows into the same outlier region if they were separated by no more than five windows with p-value > 0.01. To investigate whether the candidate regions detected with each statistic were the same or unique, we calculated how many regions overlapped between the different statistics. We considered that two regions were overlapping if at least 50% of the smaller region overlapped the larger one.

For each list of candidate regions under selection, we performed a GO term enrichment analysis using a method initially described by Pool *et al.* 2012. For each gene within a candidate region, we obtained GO term annotations from FlyBase. The GO terms for each gene also included all the parents of each term. GO terms that appeared repeatedly in a candidate region were counted only once for that region. We calculated the p-values for each GO term based on 10,000 permutations of the genomic locations of the outlier regions. This procedure allows genes to have different null probabilities of being outliers, particularly based on their length. We obtained a list of enriched GO terms for each statistic defined as the GO terms with raw p-values less than or equal or to 0.01. We then determined the overlap between the three lists of enriched GO terms.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

719

720 **Data Availability Statement**

721 No new empirical data were generated for this research. Scripts used in the analyses presented

722 can be found at https://github.com/ribeirots/fst_maxsnp.git.

723

724 **Acknowledgments**

725 We appreciate comments from multiple Pool lab members on this manuscript. This research

726 was funded by NIH grants R01 GM127480 and R35 GM13630, and by NSF grant DEB 1754745.

727

728 **References**

729 Aitken SN, Whitlock MC. 2013. Assisted gene flow to facilitate local adaptation to climate

730 change. *Annu Rev Ecol Evol Syst*. 44(1):367–388. doi:10.1146/annurev-ecolsys-110512-135747.

731 Bamshad M, Wooding SP. 2003. Signatures of natural selection in the human genome. *Nat Rev*

732 *Genet*. 4(2):99–110. doi:10.1038/nrg999.

733 Bastide H, Lange JD, Lack JB, Yassin A, Pool JE. 2016. A variable genetic architecture of melanic

734 evolution in *Drosophila melanogaster*. *Genetics*. 204(3):1307–1319.

735 doi:10.1534/genetics.116.192492.

736 Catania F, *et al.* 2004. World-wide survey of an *Accord* insertion and its association with DDT

737 resistance in *Drosophila melanogaster*. *Mol Ecol*. 13(8):2491–2504. doi:10.1111/j.1365-

738 294X.2004.02263.x.

739 Charlesworth B, Morgan MT, Charlesworth D. 1993. The effect of deleterious mutations on

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

740 neutral molecular variation. *Genetics*. 134(4):1289–1303.

741 Comeron JM, Ratnappan R, Bailin S. 2012. The many landscapes of recombination in *Drosophila*

742 *melanogaster*. *PLoS Genet.* 8(10):e1002905. doi:10.1371/journal.pgen.1002905.

743 Ewing G, Hermisson J. 2010. MSMS: a coalescent simulation program including recombination,

744 demographic structure and selection at a single locus. *Bioinformatics*. 26(16):2064–2065.

745 doi:10.1093/bioinformatics/btq322.

746 Fitzpatrick MC, Keller SR. 2015. Ecological genomics meets community-level modelling of

747 biodiversity: mapping the genomic landscape of current and future environmental adaptation.

748 *Ecol Lett*. 18(1):1–16. doi:10.1111/ele.12376.

749 Funk WC, McKay JK, Hohenlohe PA, Allendorf FW. 2012. Harnessing genomics for delineating

750 conservation units. *Trends Ecol Evol*. 27(9):489–496. doi:10.1016/j.tree.2012.05.012.

751 Garud NR, Petrov DA. 2016. Elevated linkage disequilibrium and signatures of soft sweeps are

752 common in *Drosophila melanogaster*. *Genetics*. 203(2):863–880.

753 doi:10.1534/genetics.115.184002.

754 Hermisson J, Pennings PS. 2005. Soft sweeps: molecular population genetics of adaptation from

755 standing genetic variation. *Genetics.* 169(4):2335–2352. doi:10.1534/genetics.104.036947.

756 Hudson RR. 2002. Generating samples under a Wright–Fisher neutral model of genetic variation.

757 *Bioinformatics*. 18(2):337–338. doi:10.1093/bioinformatics/18.2.337.

758 Jensen JD. 2014. On the unfounded enthusiasm for soft selective sweeps. *Nat Commun*.

759 5(1):5281. doi:10.1038/ncomms6281.

760 Kaplan NL, Hudson RR, Langley CH. 1989. The “hitchhiking effect” revisited. *Genetics.* 123(4):

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

887–899. doi:10.1093/genetics/123.4.887.

Kawecki TJ, Ebert D. 2004. Conceptual issues in local adaptation. *Ecol Lett*. 7(12):1225–1241. doi:10.1111/j.1461-0248.2004.00684.x.

Key FM, *et al.* 2014. Selection on a variant associated with improved viral clearance drives local, adaptive pseudogenization of interferon lambda 4 (IFNL4). *PLoS Genet*. 10(10):e1004681. doi:10.1371/journal.pgen.1004681.

Lack JB, *et al.* 2015. The *Drosophila* Genome Nexus: a population genomic resource of 623 *Drosophila melanogaster* genomes, including 197 from a single ancestral range population. *Genetics*. 199(4):1229–1241. doi:10.1534/genetics.115.174664.

Lack JB, Lange JD, Tang AD, Corbett-Detig RB, Pool JE. 2016. A thousand fly genomes: an expanded *Drosophila* Genome Nexus. *Mol Biol Evol*. 33:3308-3313. doi:10.1093/molbev/msw195.

Lange JD, Pool JE. 2016. A haplotype method detects diverse scenarios of local adaptation from genomic sequence variation. *Mol Ecol*. 25(13):3081–3100. doi:10.1111/mec.13671.

Lange JD. Pool JE. 2018. Impacts of recurrent hitchhiking on divergence and demographic inference in *Drosophila*. *Genome Biol Evol*. 10(8):1882–1891. doi:10.1093/gbe/evy142.

Lewontin RC, Krakauer J. 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics*. 74(1):175–195. doi:10.1093/genetics/74.1.175.

Matthey-Doret R, Whitlock MC. 2019. Background selection and F_{ST}: consequences for detecting local adaptation. *Mol Ecol*. 28(17):3902–3914. doi:10.1111/mec.15197.

Maynard Smith J, Haigh J. 1974. The hitch-hiking effect of a favourable gene. *Genet Res*.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

782 23(01):23–35. doi:10.1017/S0016672300014634.

783 Pennings PS, Hermisson J. 2006a. Soft sweeps II—molecular population genetics of adaptation

784 from recurrent mutation or migration. *Mol Biol Evol.* 23(5):1076–1084.

785 doi:10.1093/molbev/msj117.

786 Pennings PS, Hermisson J. 2006b. Soft sweeps III: the signature of positive selection from

787 recurrent mutation. *PLoS Genet.* 2(12):e186. doi:10.1371/journal.pgen.0020186.

788 Peter BM, Huerta-Sanchez E, Nielsen R. 2012. Distinguishing between selective sweeps from

789 standing variation and from a de novo mutation. *PLoS Genet*. 8(10):e1003011.

790 doi:10.1371/journal.pgen.1003011.

791 Pool JE, *et al*. 2012. Population genomics of sub-Saharan *Drosophila melanogaster*: African

792 diversity and non-African admixture. PLoS Genet. 8:e1003080.

793 doi:10.1371/journal.pgen.1003080.

794 Pool JE, Aquadro CF. 2007. The genetic basis of adaptive pigmentation variation in *Drosophila*

795 *melanogaster*. Mol Ecol. 16(14):2844–2851. doi:10.1111/j.1365-294X.2007.03324.x.

796 Pool JE, Braun DT, Lack JB. 2017. Parallel evolution of cold tolerance within *Drosophila*

797 *melanogaster*. Mol Biol Evol. 34(2):349–360. doi:10.1093/molbev/msw232.

798 Rebeiz M, Pool JE, Kassner VA, Aquadro CF, Carroll SB. 2009. Stepwise modification of a modular

799 enhancer underlies adaptation in a *Drosophila* population. Science. 326(5960):1663–1667.

800 doi:10.1126/science.1178357.

801 Reynolds J, Weir BS, Cockerham CC. 1983. Estimation of the coancestry coefficient: basis for a

802 short-term genetic distance. Genetics. 105(3):767–779. doi:10.1093/genetics/105.3.767.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

803 Ross-Ibarra J, Morrell PL, Gaut BS. 2007. Plant domestication, a unique opportunity to identify
804 the genetic basis of adaptation. Proc Nat Acad Sci. 104(Suppl 1):8641–8648.
805 doi:10.1073/pnas.0700643104.
806 Schlenke TA, Begun DJ. 2004. Strong selective sweep associated with a transposon insertion in
807 *Drosophila simulans*. Proc Nat Acad Sci. 101(6):1626–1631. doi:10.1073/pnas.0303793101.
808 Schrider DR, Kern AD. 2016. S/HIC: Robust identification of soft and hard sweeps using machine
809 learning. PLoS Genet. 12(3):e1005928. doi:10.1371/journal.pgen.1005928.
810 Schrider DR, Kern AD. 2017. Soft sweeps are the dominant mode of adaptation in the human
811 genome. Mol Biol Evol. 34(8):1863–1877. doi:10.1093/molbev/msx154.
812 Sheehan S, Song YS. 2016. Deep learning for population genetic inference. PLoS Comput Biol.
813 12(3):e1004845. doi:10.1371/journal.pcbi.1004845.
814 Sprengelmeyer QD, *et al.* 2020. Recurrent collection of *Drosophila melanogaster* from wild
815 African environments and genomic insights into species history. *Mol Biol Evol*. 37(3):627–638.
816 doi:10.1093/molbev/msz271.
817 Tigano A, Friesen VL. 2016. Genomics of local adaptation with gene flow. *Mol Ecol*. 25(10):2144–
818 2164. doi:10.1111/mec.13606.
819 Vy HMT, Won YJ, Kim Y. 2017. Multiple modes of positive selection shaping the patterns of
820 incomplete selective sweeps over African populations of *Drosophila melanogaster*. *Mol Biol*
821 *Evol*. 34(11):2792–2807. doi:10.1093/molbev/msx207.
822 Weir BS, Cardon LR, Anderson AD, Nielsen DM, Hill WG. 2005. Measures of human population
823 structure show heterogeneity among genomic regions. *Genome Res*. 15(11):1468–1476.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.12.476036; this version posted January 12, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

824 doi:10.1101/gr.4398405.

825 Yeaman S. 2015. Local adaptation by alleles of small effect. *Am Nat*. 186(S1):S74–S89.

826 doi:10.1086/682405.