1 *The Precision and Power of Population Branch Statistics in Identifying the Genomic*

2 *Signatures of Local Adaptation*

3

4 Max Shpak*, Kadee N. Lawrence, John E. Pool*

5 Laboratory of Genetics, University of Wisconsin–Madison, Madison, WI, USA

6

7 * Authors for Correspondence:

8 425-G Henry Mall, Madison, WI, 53705, USA, (608)263-6804

9 shpak.max@gmail.com; jpool@wisc.edu

10

11 <u>**Keywords**</u>:

12 Population branch statistics, fixation index, selective sweeps, local adaptation, population

13 genetic simulation

14

15 <u>**Running Title**</u>:

16 Precision and Power of Population Branch Statistics

17

18 <u>**Significance Statement**</u>:

19 Population branch statistics are widely used in genome-wide scans to identify loci

20 associated with local adaptation. This study finds that branch statistics are more accurate

21 than F_{ST} at identifying local selective sweeps under a wide range of demographic

22 parameters and models of evolution. It also demonstrates that certain branch statistics

23 have improved ability to distinguish local adaptation from other models of natural

24 selection.

25

Abstract:

Population branch statistics, which estimate the branch lengths of focal populations with respect to two outgroups, have been used as an alternative to F_{ST}-based genome-wide scans for identifying loci associated with local selective sweeps. In addition to the original population branch statistic (PBS), there are subsequently proposed branch rescalings: normalized population branch statistic (PBSn1), which adjusts focal branch length with respect to outgroup branch lengths at the same locus, and population branch excess (PBE), which also incorporates median branch lengths at other loci. PBSn1 and PBE have been proposed to be less sensitive to allele frequency divergence generated by background selection or geographically ubiquitous positive selection rather than local selective sweeps. However, the accuracy and statistical power of branch statistics have not been systematically assessed. To do so, we simulate genomes in representative large and small populations with varying proportions of sites evolving under genetic drift or background selection (approximated using variable N_e), local selective sweeps, and geographically parallel selective sweeps. We then assess the probability that local selective sweep loci are correctly identified as outliers by F_{ST} and by each of the branch statistics. We find that branch statistics consistently outperform F_{ST} at identifying local sweeps. When background selection and/or parallel sweeps are introduced, PBSn1 and especially PBE correctly identify local sweeps among their top outliers at a higher frequency than PBS. These results validate the greater specificity of rescaled branch statistics such as PBE to detect population-specific positive selection, supporting their use in genomic studies focused on local adaptation.

49 **Introduction**:

50 One of the cornerstones of evolutionary genetics is understanding the processes that drive

51 genetic differences within and among populations. A research area of particular interest is

52 determining which observed differences in allele frequency between populations are driven

53 by local adaptation versus genetic drift or background selection. Local adaptation is of

54 special importance in part because population-specific selective sweeps can contribute to

55 the initial stages of allopatric or parapatric speciation by generating genetic

56 incompatibilities between regional variants. Furthermore, the genomic signatures of local

57 selective sweeps can provide insights into the molecular basis and genetic architecture of

58 adaptive phenotypic traits as they relate to regional variation.

59

60 Among the most notable and well-documented examples of local adaptation include

61 grasses that tolerate high heavy metal concentrations in contaminated soils (Antonovics

62 and Bradshaw 1970, Macnair 1987), mouse coat color variants adapted to different

63 substrates by selection for camouflage (Nachman et al. 2003, Hoekstra et al 2006), the

64 evolution of body shape and coloration of sticklebacks in response to different species of

65 predator (Colosimo et al. 2005, Miller et al. 2015, Gygax et al. 2018), the evolution of toxin

66 resistance in garter snakes that co-occur in regions with highly toxic newts as prey (Brodie

67 et al. 2002), and the independent adaptation to hypoxic environments for humans and

68 other large mammals living at high altitudes (Yi et al. 2010, Julian and Moore 2019, Witt

69 and Huerta-Sanchez 2019). *Drosophila melanogaster* and its close relatives have also

70 provided important model systems for genetic studies of local adaptation, including for

71 studies focused on adaptation to high altitudes (Lack et al 2016, Sprengelmeyer and Pool

72 2021, Sprengelmeyer et al 2022), upper latitudes (Adrion et al. 2015; Svetec et al. 2015;

73 Siddiq et al. 2019), and novel food sources (Yassin et al 2016).

74

75 One general approach to identifying and characterizing local adaptations is based on

76 quantifying the differences in allele frequencies among populations at different loci, on the

77 assumption that highly divergent loci are potential genomic signatures of local adaptation.

78 A variety of statistical approaches have been proposed and used to leverage this

79 information.

80

81   *Fixation Index and Population Branch Statistics*:

82   Genome-wide scans for selection identify outliers in the distribution of differences in allele

83   frequency between a focal population and its outgroup(s) using various measures of

84   genetic distance. Among the most widely used is Wright's fixation index F_{ST}, defined as the

85   ratio of among population variance to the total variance across populations. The quantity

86   was originally derived for single loci and sites, but can be generalized as a statistic for

87   genomic regions (e.g. Reynolds et al 1983). Many studies have used high F_{ST} between two

88   populations at specific loci in comparison to the rest of the genome or chromosome region

89   to infer local adaptation (*e.g.* Akey et al. 2009, Amato et al. 2009, Kapun et al. 2020).

90

91   There are several drawbacks to using F_{ST} to measure population divergence, such as its

92   non-linearity and non-additivity (making comparisons across multiple populations

93   difficult). Importantly, when comparing very large numbers of sites in a genome-wide scan,

94   it is inevitable that a significant fraction will be divergent among populations due to genetic

95   drift, resulting in some fraction of loci that are falsely identified as having evolved under

96   positive directional selection. Further, the use of F_{ST} leaves the problem of determining

97   which population in a divergent pair has experienced local adaptation.

98

99   Particularly to address the latter problem, several statistics have been proposed that use a

100   rescaled F_{ST}-based value across population triplets (a focal population and two outgroup

101   populations) to estimate population-specific differentiation. Among these is the Population

102   Branch Statistic (PBS), first proposed in Yi et al (2010) as an estimate of the branch length

103   of the focal population. The distance metric for population pairs is a log-transformed

104   (Cavalli-Sforza 1969) function of F_{ST},

105

106   T = -log(1- F_{ST})   (Eq 1)

107

108   Defining a three population tree (Figure 1A) with focal population A, a relatively closely

109   related population B, and a 'population outgroup' C, we can estimate the corresponding

110   branch lengths of *a,b,c*. Because the pairwise distances T are approximately additive, the

111   branch length *a* from the three pairwise T values is estimated as:

112

113    $\text{PBS}_A = \frac{T_{AB} + T_{AC} - T_{BC}}{2}$ (Eq 2)

114

115   PBS as an estimator of focal branch length has been used in a number of recent studies to

116   identify sites or regions of the genome evolving under selective sweeps (e.g. Jiang and Assis

117   2020), with the logic that loci with high PBS compared to other loci in the genome are

118   strong candidates for selective sweeps since population divergence.

119

120   To the extent that studies aim to identify population-specific positive selection in particular,

121   a potential disadvantage to PBS is its inability to distinguish between cases where only the

122   focal branch is especially long from cases where all population branches are long. While

123   sites under strong local directional selection will have high PBS, the same is true at sites

124   undergoing similar, parallel selective sweeps in all three populations. Depending on the

125   goals of the study, such loci may or may not be of interest, as they may often reflect

126   instances of positive selection unrelated to adaptation to local environments, such as arms

127   races driven by meiotic drive or reproductive competition. Furthermore, background

128   selection (Charlesworth et al 1993, 1995) may have a relatively widespread influence on

129   genetic diversity in at least some regions of the genome (e.g. regions with low

130   recombination rates and/or a high density of functional sites), and frequency changes

131   driven by background selection may also lead all populations to have unusually long

132   branches at certain loci. Consequently, PBS may be effective at identifying instances of

133   recent natural selection generally while being less effective at discriminating between

134   population-specific versus species-wide selection pressures (although such predictions

135   remain to be tested).

136

137   To address the above concerns regarding PBS, two branch statistic rescalings have been

138   proposed with the specific purpose of identifying loci with population-specific elevations in

139   genetic differentiation. The normalized population branch statistic, *i.e.* PBSn1, was

5

140   introduced by Malaspinas et al. (2016; see supplemental section S16) and also featured in

141   subsequent studies (Crawford et al 2017, Vicuna et al 2019). This metric rescales PBS_A with

142   respect to the total tree length, e.g.

143

144    $$PBSn1_A = \frac{PBS_A}{1 + PBS_A + PBS_B + PBS_C} \qquad \text{(Eq 3)}$$

145

146   This rescaling has the potential to discriminate between cases where PBS_A is high in

147   comparison to the lineages of populations B and/or C from cases where B or C also have

148   long branches. Consequently, PBSn1 is expected to give a lower false positive rate in

149   identifying local sweeps at loci that are evolving under parallel (*i.e.* separately occurring)

150   selective sweeps in multiple populations. The addition of 1 in the denominator prevents

151   division by zero and provides a baseline (although arbitrary) normalization term so that

152   cases where all three branches are extremely short but the focal population is several times

153   longer is not scored as significant in the same way that an especially long focal branch in

154   particular would be (*e.g.* without the 1, a case of branch lengths *a,b,c* = 0.2,0.1,0.1 and *a,b,c*

155   = 20,10,10 would be treated equivalently).

156

157   Unlike the above statistics, another elaboration on PBS, known as Population Branch Excess

158   (PBE; Yassin et al. 2016), incorporates branch length information from additional genomic

159   loci. This method posits that in the absence of local adaptation at a given locus, the

160   relationship between the focal population's branch length at this locus (*i.e.* PBS_A) and the

161   combined lengths of the two non-focal population branches (*i.e. T_{BC}*) should be similar to

162   the relationship between these two quantities observed at most other genomic loci. PBE

163   employs this logic to obtain an expected value for PBS at this locus, based on T_{BC} at this

164   locus and the median values of both PBS and T_{BC} observed at all other analyzed loci in the

165   genome (or else a component of it such as the same chromosome arm). PBE then quantifies

166   the degree to which the observed PBS value for this locus exceeds its expected value, as

167   follows:

168

169    $\mathrm{PBE_A} = \mathrm{PBS_A} - \mathrm{PBS_{A\_exp}} = \mathrm{PBS_A} - \dfrac{T_{BC} \times \mathrm{PBS_{A\_med}}}{T_{BC\_med}}$ (Eq 4).

170

171   The comparison to an expected PBS value may make PBE values somewhat more

172   comparable across studies that involve populations with very difficult demographic

173   histories. For example,

174   if all loci are highly differentiated due to a recent bottleneck in the focal population's

175   history, PBE will adjust all values based on the relatively high expected PBS (although an

176   elevated neutral variance may still be expected). The ratio of branch lengths T_{BC}/T_{BC_med}

177   also corrects for cases where the non-focal branches are also long due to selection in those

178   populations as well (whether sweeps or background selection). Or viewed another way, a

179   larger-than-usual value of T_{BC} at this locus increases the value expected for PBS, and thus

180   will tend to yield lower PBE values, which is desirable if the goal is a specific focus on

181   population-specific positive selection.

182

183   We also note that multiple branch-oriented statistics have been developed to detect local

184   adaptation from data sets that include greater than three populations (Schmidt et al. 2019;

185   Schlebush et al. 2020; Cheng et al. 2022). However, this study's focus is restricted to

186   frequency-based statistics that incorporate at most three populations, in part for the sake

187   of clarity, and in part because of the relatively larger number of empirical studies that can

188   generate adequate population genomic data for these methods.

189

190   The heuristic considerations that prompted the proposal of PBS, PBSn1, and PBE as

191   measures of population divergence suggest that all branch statistics may have greater

192   power and accuracy in identifying instances of local selective sweeps against a backdrop of

193   neutral evolution than F_{ST}, and that PBSn1 and PBE may outperform PBS when some loci

194   evolve under parallel sweeps while the evolution of other loci is driven by local adaptation.

195   However, their relative efficacy and statistical power have not been systematically

196   examined. Indeed, all three of the branch statistics highlighted here were introduced in the

197   context of empirical genome scan studies which did not evaluate the performance of these

198   statistics via simulation.

199

200   In this study, we analyze data from population genetic simulations representing a wide
201   range of scenarios, in order to assess the above predictions regarding the performance of
202   the above-mentioned frequency-based statistics to detect population-specific positive
203   selection. We include simulations with large and small population sizes (motivated by
204   *Drosophila* and humans, respectively), and we investigate models with and without gene
205   flow after population divergence. We combine simulated replicates into simulated
206   genomes, and assess which scenarios are placed in the top (upper 1%) outlier quantile for
207   each statistic, emulating an empirical outlier genome scan. This approach allows us to
208   differentiation local adaptation not only from neutral evolution but also from background
209   selection (modeled as reduced population size) and parallel selective sweeps. Results of
210   these analyses will allow researchers to make more informed choices regarding which
211   statistics to use in future population genomic scans for local adaptation.
212

213   **Results:**
214   We performed population genetic simulations to test the ability of four statistics (F_{ST}, PBS,
215   PBSn1, and PBE) to detect population-specific positive selection, examining three
216   population models in small and large N_e cases motivated by human and *Drosophila* data,
217   respectively. The demographic and genomic parameters used in these models are
218   summarized in Figures 1B-C. We aggregated simulations from distinct neutral/selection
219   scenarios into "model genomes" (Figure 2A) to test how reliably each statistic placed the
220   1% of true local sweep loci in its the upper 1% tail, with or without an additional 1% of loci
221   subject to parallel sweeps in all three populations, and with the remaining loci subject to
222   either neutral evolution or background selection (BGS; modeled as variably reduced N_e).
223

224   We simulated local complete hard sweeps with selection coefficients $s = 0.025$ and 0.001 in
225   the small and large populations, respectively. When these local sweeps occur amongst
226   neutral loci, the loci under selection can be identified with a high level of accuracy by all
227   statistics, with only slight advantages of the three population statistics over pairwise F_{ST}. As
228   we deviated from this simple scenario, the statistics began to diverge from one another in
229   both their precision based on the model genomes defined above (Figure 3) and their more

230 traditional statistical power based on comparing individual local sweep replicates to
231 neutral or BGS distributions (Table S1).
232
233 Precision was reduced in the cases of partial and soft local sweeps in otherwise neutral
234 genomes, with PBS, PBSn1, and PBE often performing similarly, and F_{ST} consistently
235 showing the lowest precision (Figure 3). In contrast, in model genomes where most loci
236 evolve under BGS rather than genetic drift, the relative performances of the branch
237 statistics became more distinct: here, the precision of PBSn1 and PBE was invariably larger
238 than for PBS in both the large and small populations (Figure 3). PBE also outperformed
239 PBSn1 in each of these cases, although the difference was more notable in the large
240 populations. In the scenario that yielded the most disparate outcomes, that of a soft sweep
241 from 0.5% initial frequency in a large population with BGS, precision was 1.6% for F_{ST},
242 10.8% for PBS, 16.7% for PBSn1, and 31.6% for PBE. In a few of the most challenging cases,
243 the precision of some or all statistics fell considerably. In the large population case, strong
244 effects of the examined BGS model even led some precision estimates to fall below the null
245 expectation of 0.01, e.g. for a soft sweep with $p_0 = 0.01$ (perhaps partly because the sweep
246 loci themselves were not subjected to BGS). Under these scenarios, BGS tends to generate
247 longer branch lengths than genetic drift (especially in the large populations, which have a
248 higher fraction of loci with $B < 0.5$; Figure 2B). Here, PBE's rescaling of branch lengths with
249 respect to all loci in the "genome" may improve its performance in comparison to PBS and
250 even with respect to PBSn1, which rescales the focal branch length with respect to total tree
251 length at the focal locus but not with respect to an overall genome-wide baseline.
252
253 In simulations that included both parallel and local sweeps at 1% of the loci, in otherwise
254 neutrally evolving genomes, the overall frequency of false positives increased for all of the
255 statistics when compared to otherwise equivalent scenarios in which all sweeps were local
256 (Figure 3). Here, the advantages of PBE and PBSn1 over PBS became more pronounced,
257 with PBE again showing a larger advantage for the large population scenarios. Whereas, the
258 performance of F_{ST} became especially poor with parallel sweeps. In these scenarios, we
259 expect that parallel sweeps may generate higher F_{ST} values than local sweeps, so parallel
260 sweeps will disproportionately contribute to the upper quantiles of F_{ST} and potentially PBE

261   (which simply looks for a long estimated focal population branch length). The enhanced

262   differentiation of parallel sweeps (and thus their contribution to false positives) can be

263   greater for hard complete sweeps than for partial or soft sweeps (Figure 4), which may

264   contribute to the otherwise surprising observation that in some cases, F_{ST} and to lesser

265   extent PBS had lower precision and power with hard complete sweeps than with certain

266   soft or partial sweeps. This phenomenon is less apparent in the large populations, where

267   the modeled selection coefficients are weaker and the rate of recombination relative to

268   mutation is somewhat higher (see Materials and Methods).

269

270   With hard complete local and parallel sweeps (as well as partial sweeps to 80% terminal

271   frequency), most false positives were contributed by parallel sweep loci (Fig. 4, Table S2),

272   while for soft sweeps at high initial frequency and for partial sweeps to 50%, there was a

273   greater contribution from the neutral and BGS loci (especially for the case of BGS in the

274   large populations – e.g. 46-94% of parallel sweep loci contributed to upper quantiles under

275   80% partial sweeps whereas only about 1% of parallel sweep loci were in the upper

276   quantile in soft sweeps with $p_0 = 0.01$). The proportional contribution of parallel sweeps

277   vs. neutral/BGS loci to false positives impacts the relative ability of the population statistics

278   to detect local sweeps. For example, in a small population with both local and parallel

279   sweeps, the precisions for some of the statistics, especially F_{ST} and PBS, are higher for soft

280   sweeps (with $p_0 = 0.005$) and partial sweeps ($p_t = 0.80$) than for complete hard sweeps.

281   This counter-intuitive pattern is not seen in the large population simulations, which have

282   lower false positive rates from the parallel sweeps than in the small populations. In

283   contrast, with very soft sweeps ($p_0 = 0.01$) or partial sweeps to only 50% final frequency,

284   there is also a significant relative contribution of neutral or BGS sites to the false positives.

285

286   When populations are subject to both parallel sweeps and background selection, similar

287   trends hold, with slightly to moderately reduced precision for each statistic to detect local

288   sweeps compared to cases with parallel sweeps without BGS. In all cases, F_{ST} showed the

289   poorest performance, while PBS lagged behind the other two branch statistics. For small

290   populations, PBSn1 and PBE show similar advantages, while for large populations, PBE

291   displayed consistently greater precision, generally showing similar or greater advantages

292 over PBSn1 and the other statistics than observed with either parallel sweeps or BGS alone.

293 For example, in the case of soft sweeps with $p_0 = 0.005$, PBE at 0.295 gave a precision

294 nearly two-fold and three-fold higher than PBSn1 and PBS, and twenty-fold higher than F_{ST}.

295

296 While the above-mentioned results all pertain to simulations with genetic drift but no

297 migration, we also explored cases in which all pairs of populations have exchanged one

298 migrant per generation (*i.e.* $N_e m = 1$) since their divergence. Adding this level of migration

299 to the previously studied population history had a relatively weak effect on the precision

300 and power of the statistics to detect hard local sweeps under scenarios lacking any other

301 form of selection (Figure 3; Table S1; "match tree"). Here, while migration may have

302 reduced the frequency differences at locally adaptive loci, it also resulted in populations

303 with lower neutral genetic differentiation than in migration-free base scenario. The effect of

304 migration on precision and power is more perceptible under the scenarios in which we

305 extended the divergence times of our populations to emulate the neutral F_{ST} values

306 observed among populations without migration (Figure 1B), since here allele frequency

307 differences should be reduced at local sweep but not neutral loci compared to the base

308 history (Figure 3; Table S1; "match F_{ST}"). While the effects of migration on its own were not

309 dramatic for either population size or population tree length, its effects synergized with

310 BGS to reduce precision, particularly for the large population size with its more powerful

311 modeled BGS effects (Figure 3). With parallel sweeps, the performance of F_{ST} and in some

312 cases PBS actually improved with the introduction of migration. To some extent, migration

313 may be blunting the propensity of parallel sweeps to generate false positives (Figure 4;

314 Table S2), analogous to the outcomes of some partial or soft sweep models discussed above.

315 In general with migration, as observed otherwise, PBE and PBSn1 showed the greatest

316 precision when other types of selection were present, and PBE in particular showed greater

317 advantages in the large populations.

318

319 **Discussion:**

320 Population branch statistics were introduced as alternatives to F_{ST} in genome-wide scans

321 for local selective sweeps on heuristic grounds, without explicit assessment of their

322 performance. The present study provides simulation-based evidence favoring their use

11

323 over F_{ST} as well as supporting the use of rescaled branch statistics PBSn1 and PBE rather
324 than PBS for a wide range of evolutionary models. Generally, only in the most trivial
325 scenario of a small number of loci experiencing local selective sweeps while all others
326 evolve neutrally do F_{ST}-based scans or PBS have precision and power comparable to the
327 rescaled population branch statistics PBSn1 or PBE. Otherwise, the rescaled branch
328 statistics were consistently more robust in contributing to the upper quantiles when
329 potentially confounding selective processes such as BGS or parallel sweeps were
330 introduced.
331
332 This study also provides qualified support for the use of PBE over PBSn1 under many
333 selection regimes, including the potentially more biologically realistic models of partial and
334 soft sweeps, and in genomes where background selection plays an important role in
335 shaping genetic differentiation. On the other hand, there are certain scenarios, such as
336 those where there are strong parallel as well as local sweeps, where PBSn1 performs as
337 well as or slightly better than PBE, though it is difficult to generalize from these cases
338 because the differences are typically modest and contingent on multiple parameters.
339 Furthermore, complete, hard selective sweeps are probably quite rare, at least in human
340 populations (Hernandez et al 2011), which may limit the importance of the only case in
341 which PBSn1 had a non-trivial advantage over PBE (*i.e.* both complete hard local and
342 parallel sweeps in small populations). One factor that may modulate the relative precision
343 PBSn1 compared to other statistics is the arbitrary scaling factor of 1.0 in the denominator
344 (Eq 2), which may serve to dampen noise from excessively small denominators.
345 Conceivably, the precision of PBSn1 relative to PBE might vary if other values of this scaling
346 factor were used. On the other hand, it is worth emphasizing that PBSn1 (like PBS and F_{ST})
347 can be calculated using only genetic variation at a specific locus, whereas each value of PBE
348 necessarily draws on information from other comparable loci.
349
350 Given the lack of prior results, our study represents a useful starting point in evaluating the
351 relative performance of the statistics examined. Future studies could use demographically
352 expanded simulations to incorporate methods that accommodate more than three
353 populations (e.g. Cheng et al. 2022). It would also be instructive to investigate a wider array

354   of demographic models, such as strong population bottlenecks, additional migration rates,

355   or admixture models, to further assess the context-dependency of the differences in

356   precision we observed among statistics.

357

358   Lastly, while our use of reduced population size as a proxy for the effects of BGS on allele

359   frequency differentiation is reasonable in the context of this study, it will be worthwhile to

360   ultimately supplement our analyses with explicit simulations of negative selection at linked

361   sites (Ewing & Jensen 2016). Such simulations may be especially computationally

362   demanding for *Drosophila*-like cases, in light of the chromosomal scale that may be relevant

363   (*e.g.* Comeron *et al.* 2014) and the unclear effects of population size rescaling (as often

364   conducted for simulations with *Drosophila*-like N_e) on the outcomes of BGS simulations.

365   Relatedly, while our simulation results suggested a strong potential of background selection

366   to generate outliers in F_{ST}-based statistics in the *Drosophila*-like simulations, it is worth

367   noting that Booker et al. (2020) found that the empirical relationship between crossing-

368   over rate and F_{ST} between *D. melanogaster* populations was roughly similar to that

369   expected under a simulated model without background selection, particularly above the

370   female crossing-over rate threshold of 1 cM/Mb that our B-value cutoff approximates.

371

372   Even in an era featuring increasingly complex population genetic methods for detecting

373   local adaptation, we suggest that there is value in simpler statistics that have relatively

374   straightforward interpretation with regard to genetic variation. In this context, we suggest

375   that rescaled population branch statistics such as PBE and PBSn1 are relatively powerful

376   and efficient tools for identifying the genetic signatures of local adaptation in genome-wide

377   scans. The results of this study argue in favor for their expanded use in the toolkit of those

378   seeking evidence of local adaptation in natural populations, particularly in cases where

379   other forms of positive or negative selection may impact genetic differentiation.

380   Furthermore, these statistics and their underlying logic may continue to provide useful

381   building blocks for more advanced methodologies in the future.

382

383   **Methods:**

384   We simulated evolution in a three-population model where the evolutionary dynamics of

385   genotype frequencies are driven by (at different loci in various combinations) neutral

386   genetic drift, background selection against deleterious mutations, local selective sweeps,

387   and parallel selective sweeps in all three populations. Unless otherwise indicated,

388   simulations were run using Python 3.8.8 scripts and all data analyses were performed using

389   R 4.1.1.

390

391   *Population Demographic History*:

392   We chose two base demographic models to approximate the population histories of

393   humans and *Drosophila melanogaster* as representatives of large and small population

394   species with respective effective sizes of $N_e \sim 10^4$ and 10^6. We set the outer and inner

395   population split times at 2×10^3 and 10^3 generations before present for the small

396   population case and 2×10^5 and 10^5 generations for the large population case (as shown in

397   Figure 1B). These time intervals roughly approximate the split times between Africa vs.

398   Eurasia and Eurasia vs. the Americas for humans and between sub-Saharan Africa vs. the

399   Mediterranean and the Mediterranean vs. north-central Europe for the flies.

400

401   We applied a reduced population size for the internal branch representing the common

402   ancestor of populations A and B (labeled in Figure 1B as population D, which is half the size

403   of its descendant populations), in order to roughly recapitulate the genetic drift involved in

404   both of the motivating species' founder event expansions from Africa into Eurasia (note

405   that this bottleneck was much stronger in human evolutionary history). Note that we are

406   not attempting to model the precise demographic details of either species' history, rather,

407   we use these approximate and simplified topologies to capture the qualitative differences

408   between large and small population models.

409

410   To make large population forward simulations computationally feasible in light of the very

411   high number of generations needed, we modeled evolution using a rescaling approximation

412   where a 50-fold smaller N_e population was simulated in combination with 50-fold higher

413   rescaling of selection coefficients, mutations rates, and other populations parameters, an

414   approach similar to the one used by Lange et al. (2018). The demographic and genetic

14

415   parameters are listed in Figure 1C for both the small and large populations (based on

416   ancestral $N_e = 2$ x 10^4 and 2×10^6, respectively, simulated with a corresponding 50x

417   rescaling in the large population). We used this rescaling in both the forward time

418   (selection) and backward time (coalescent drift) simulations for consistency when

419   simulating evolution in the large populations.

420

421   *Parameterizing the Model Genomes*:

422   We generated genomic windows so that in our simulations a "locus" is defined not as a

423   single site, but as a chromosome region. We simulated 100 kb and 5 kb regions for the small

424   and large populations, respectively, following previous simulation studies (*e.g.* Lange and

425   Pool 2016; da Silva Ribeiro *et al.* 2022), and equal or close to window sizes from empirical

426   genome scan studies (*e.g.* Voight *et al.* 2006; Granka *et al.* 2012; Lack *et al.* 2015; Pool *et al.*

427   2017). The 20x difference in window length also represents a compromise between the

428   approximately 10x higher population mutation rate and the roughly 40x higher population

429   recombination rate estimated for fruit fly versus human populations. Per-site mutation,

430   recombination, and gene-conversion rates were based on the mean autosomal estimates in

431   *D. melanogaster* from Comeron et al (2012) and for humans from Schiffels and Durbin

432   (2014), Williams et al (2015), and Jonsson et al (2017). Figure 1C summarizes the genomic

433   parameters used in the simulations.

434

435   *Simulating Genetic Drift*:

436   In purely neutral simulations, genetic drift was modeled as a backward time coalescent

437   process using the msprime package 1.2.0 (Baumdicker et al. 2022) for the demographic

438   models and parameters summarized in Figure 1B, 1C. The msprime commands were

439   embedded in python script "wrappers" that calculated allele frequencies and the

440   population statistics.

441

442   At the termination of each simulation run, we calculated F_{ST} among subpopulations based

443   on the Reynolds et al. (1983) estimate of the fixation index as well as PBS and PBSn1. The

444   branch length T_{BC} was also output for each simulated replicate, since this is needed

445   alongside PBS values for the subsequent calculation of PBE, which is calculated with

15

446 respect to an aggregate of loci in a genome rather (Eq. 4). Additionally, sequences of

447 segregating sites generated by the simulations were saved as output.

448

449 *Background Selection*:

450 The effects of background selection (Charlesworth et al. 1993, 1995) were simulated

451 without explicitly modeling deleterious mutations or negative selection as such. Rather, we

452 modeled BGS implicitly by simulating the effects of negative selection on the genetic

453 diversity of linked neutral variation. Charlesworth (2012) demonstrated that the reduction

454 in linked neutral genetic variation due to BGS is equivalent to that under a reduction of

455 effective population size by a factor *B* which depends on the recombination rate in a

456 genomic window (i.e. regions of low recombination correspond to B-values much less than

457 one). As examples, rescaling N_e (and therefore expected coalescent times) with B-values to

458 model the effects of BGS on neutral genetic variation was applied by Jiang and Assis (2002)

459 to estimates of F_{ST} among populations evolving with BGS, and in the calculations of Huber

460 et al. (2015) of allele frequency spectra under BGS. There is evidence that modeling BGS as

461 drift with reduced *Ne* is inadequate for demographic inference (Ewing and Jensen 2016),

462 and it would seem similarly inappropriate for studies making quantitative inferences about

463 the background selection process itself. In contrast, because the B-value approximation

464 does capture the reduction in nucleotide diversity due to BGS, it should recapitulate the

465 resulting increases in allele frequency differentiation between populations to a sufficient

466 degree for a study not focused on parameter estimation.

467

468 Specifically, we model the effects of BGS on neutral genetic variation by using the

469 distribution of B-values estimated for chromosome regions in *D. melanogaster* in Comeron

470 (2014) and for the human genome in McVicker et al. (2009). In each simulation, a value of

471 *B* is randomly drawn from the frequency distribution (from the *Drosophila* autosome

472 distribution for large population simulations, the human autosome distribution for small

473 populations) and used as a multiplier to rescale the effective population size. The evolution

474 of genetic diversity in a region of genome with this value of *B* is simulated as genetic drift

475 using msprime with $N_e^* = BN_e$ in place of N_e.

476

477 Certain chromosome regions, particularly those near centromeres in *Drosophila*, have very
478 low B, i.e. < 0.01 (Figure 2B), resulting in very small rescaled effective population size and
479 greatly inflated F_{ST} for simulated loci in those intervals. We corrected this potential artifact
480 by removing B-values corresponding to regions where the sex-averaged crossing-over rate
481 $r < 0.5$ cM/Mb, analogous to the removal of low recombination regions such as those
482 around autosomal centromeres in many genome-wide scans of *D. melanogaster* (*e.g.* Pool et
483 al. 2017). Low recombination regions are often excluded from genome scans because of (1)
484 the greater variance in diversity patterns expected under neutrality (Booker et al. 2020),
485 (2) stronger effects of background selection, and (3) the difficulty in localizing outliers due
486 to positive selection in low recombination regions.
487
488 To find an appropriate cut-off, we performed a LOESS regression of *D. melanogaster* B-
489 values against recombination rate and identify the value of B predicted at $r = 0.5$ by the
490 regression model. This corresponds to $B = 0.41$, so the B-value distribution was truncated
491 at $B > 0.41$. For consistency, this was done for both the human and fruit fly distributions,
492 even though human chromosomes had only 7% of regions where $B < 0.41$ versus 30% of
493 fruit fly autosomes. Simulations of BGS as genomically variable drift were run by sampling
494 B from these truncated distributions of B-values for each simulated genomic window (*i.e.*
495 simulation replicate).
496
497 ***Selective Sweeps***:
498 All simulations of positive selection were performed using the SLiM v. 3.5 software package
499 (Messer 2013, Haller et al. 2019). Unlike the coalescent-based models in msprime, SLiM
500 simulations run in forward time. The SLiM output consists of a tree structure (the
501 sequences and their genealogical history) of the site(s) under selection. To generate the
502 corresponding history of neutral sites that evolve in association with the sites evolving
503 under directional selection, we used msprime to simulate their coalescent history and link
504 them to the tree structure generated in SLiM, a process known as recapitation of the
505 genealogy, as described in Kelleher et al. (2018) and Haller et al. (2019).
506

507   Most of the SLiM simulations were initiated by introducing a beneficial mutation at a

508   randomly selected single site in the focal population in the first generation following its

509   split from its outgroup. The simulation runs were conditioned on either the fixation of the

510   beneficial allele or it reaching a specified terminal frequency $p_t < 1.0$. If the beneficial

511   mutation was lost from the population during the simulation, the simulation was

512   terminated and restarted. Thus, only the output tree structures of those replicates that

513   reached the desired terminal frequencies were retained.

514

515   Variations on this basic model of selective sweeps included conditioning on intermediate p_t

516   < 1.0 (partial sweeps), initial frequencies orders of magnitude greater than $1/2N$ (soft

517   sweeps), and simultaneous selection at the same locus in both the focal population and its

518   outgroups (parallel sweeps). All of the selective sweep scenarios described below were run

519   to generate 10,000 replicates for analysis.

520

521   A. *Complete local hard sweeps*: These simulations used selection coefficients $s = 0.025$ and

522   0.001 in the small and large populations, respectively (with the smaller s in the larger N_e

523   populations reflecting the expectation that weaker beneficial mutations can successfully be

524   favored by selection in larger populations). With the population size rescaling by 50x of the

525   large populations, the rescaled selection coefficient used in those simulations was $s' = 0.05$.

526

527   Because of the high probability that a single beneficial mutation will be lost or fail to be

528   fixed in a simulation with small N_e and $s << 1$, we used an initial frequency $p_0 = 0.001$

529   rather than $p_0 = 1/2N_e$ (set at the time of the focal population's split from its outgroup) in

530   the small N_e scenarios, to reduce the number of times that simulations had to be restarted

531   to achieve fixation. However, in the large N_e scenarios, we used $p_0 = 1/(2N_e)$ to avoid soft

532   sweep effects in the rescaled model populations. Previous studies, e.g. da Silva Ribeiro et al.

533   (2022) have shown that these initial frequencies are sufficiently low to generate hard

534   sweep outcomes (*i.e.* only one haplotype carrying the beneficial haplotype tends to

535   contribute to the sweep). Additionally, because many simulations didn't reach fixation by

536   the final generation, rather than restarting all simulations where the $p_t < 1.0$, we

537   implemented a strategy in which simulations with $p_t > 0.95$ were saved and only those

18

538   haplotypes with the beneficial mutation were sampled for F_{ST} estimation and other

539   analyses. These near-fixations provide a close approximation to a complete hard sweep,

540   and the approach was also used when modeling parallel and soft sweeps.

541

542   B. *Complete local soft sweeps*. Soft sweeps can occur when positive selection acts on alleles

543   that had already achieved a frequency significantly higher than $1/(2N_e)$ in the population

544   through genetic drift, and multiple initial haplotypes contribute to the sweep, contributing

545   to greater haplotype diversity after fixation than hard sweeps (Hermisson and Pennings

546   2005). We simulated soft sweeps with initial beneficial allele frequencies of $p_0 = 0.005$ and

547   0.01 in both the large and the small populations, using the same selection coefficients and

548   initiation times at the A,B split as in the complete sweep simulations.

549

550   C. *Partial hard sweeps*. In practice, many beneficial alleles may never become fixed, or they

551   may be sampled at time points prior to their fixation in a population. To simulate (hard)

552   partial sweeps, we applied positive selection to an allele in the focal population using the

553   same initial frequencies and selection coefficients as described for complete local sweeps.

554   To target sweeps reaching terminal frequency $p_t = 0.5$, we introduced the beneficial

555   mutations at $t = 1700$ and 3800 generations, i.e. 300 and 200 generations prior to

556   termination in the small and large populations, respectively; for $p_t = 0.8$, the beneficial

557   mutations were introduced at generation times 1500 and 3700. These initiation times were

558   chosen because they generated terminal frequencies close to the desired values.

559   Simulations were repeated until a terminal allele frequency between 0.45 and 0.55 was

560   observed in the last generation to simulate partial sweeps to 50% (or between 0.75 and

561   0.85 to approximate $p_t = 0.8$). Unlike for the complete sweeps, haplotypes were sampled

562   randomly from the entire population at the end of each saved simulation, as opposed to

563   only sampling haplotypes carrying the beneficial allele.

564

565   D. *Parallel sweeps*. We simulated parallel selective sweeps by introducing a distinct

566   beneficial mutation in all three populations at the same position. For hard complete parallel

567   sweeps, we used the same initial frequencies and selection coefficients as were used for

568   hard complete and partial local sweeps. Beneficial mutations were introduced at $t = 1000$

569 and $t = 2000$ at the same locus in all three populations (corresponding to the time of the
570 A,B population split in the small and large populations, respectively) in order to allow
571 maximum time to reach fixation. As with hard complete local sweeps, the simulations were
572 re-run until the beneficial mutations reached $p_t > 0.95$, and if not fixed, only haplotypes
573 carrying the beneficial mutation were sampled for analysis.

574

575 We also simulated soft parallel sweeps and partial parallel sweeps so that the resulting
576 allele frequencies and F_{ST}-based statistics could be compared to those generated by the
577 same positive selection models in local sweeps. The parallel soft and parallel partial sweeps
578 used the same initial frequencies, selection coefficients, initiation times for the introduction
579 of beneficial mutations, and targeted terminal frequencies as in the corresponding local
580 sweep simulations, but applied these conditions to all three populations rather than the
581 focal population alone.

582

583 *Simulating Migration*:
584 A subset of simulations included migration among all populations, in order to determine
585 the extent to which results on the relative efficacy of branch statistics were robust in the
586 presence of gene flow. We considered two demographic models with migration. One retains
587 the split times of the migration-free models (left tree in Figure 1B), with a symmetric
588 pairwise migration rate introduced so that $N_e m = 1$ per generation between all pairs of
589 populations. The migration rates were set so that each population received on average a
590 single migrant per generation independent of population size (such that m was smaller for
591 small population receiving migrants from a large population than the reverse) to avoid
592 source-sink effects. In the rescaling of the large populations, we set $N_e'm' = 1$ (where $N_e' =$
593 $N_e/50$, $m' = 50m$) so that there would be comparable numbers of migrants exchanged
594 among pairs of populations over the same number of coalescent time units.

595

596 Because migration is expected to reduce pairwise differentiation among populations, the
597 baseline expected F_{ST} at neutrally evolving sites will be lower in the presence of migration
598 than in its absence, so a high or low branch statistic in the absence of migration may not be
599 comparable to values observed in populations evolving with migration. Therefore, we also

600 implemented a longer branch demographic history for a three population model with

601 migration that generates very similar pairwise F_{ST} among all pairs as the migration-free

602 model. For the same population sizes, we introduced split times at $3x10^3$ and $5x10^3$

603 generations before the present for the small population models and $3x10^6$, $5x10^6$ for the

604 large population models (right tree in Figure 1B).

605 Simulations of genetic drift and selection with migration were executed in msprime and

606 SLiM with the same selection coefficients in the small and large populations as for the

607 models of selective sweeps as in the models without migration.

608

609 Because of migration, the selective sweeps are not expected to be complete, so we

610 conditioned the local sweeps on attaining a terminal frequency $p_t > 0.75$, and imposed the

611 same constraint on all three populations when simulating parallel sweeps in both the large

612 and small N_e scenarios. This threshold value was selected to eliminate cases where the

613 beneficial mutation was nearly lost while allowing for intermediate frequencies due to

614 migration (*e.g.* when simulating 1000 replicates of local hard sweeps, we found that 98% of

615 the replicates attained frequencies $p_t > 0.80$). As with the partial sweep simulations, all

616 haplotypes were included in the final sample rather than only those carrying the beneficial

617 mutation because we are not simulating fixation. Beneficial mutations were introduced at

618 the time of the split between populations A,B, as was done for the hard complete sweep

619 models without migration.

620

621 In simulations of migration with local sweeps, the beneficial mutation in the focal

622 population was neutral in the two non-focal populations. Similarly, when modeling parallel

623 selection with migration, the beneficial mutation for each population was neutral with

624 respect to the wildtype background in the other two populations (and thus relatively

625 deleterious as a migrant, compared to another beneficial mutation within its home

626 population).

627

628 *Model Genomes and Statistical Analysis:*

629 The msprime simulations, including the recapitated SLiM trees, generate a distribution of

630 multilocus genotypes from which we calculated allele frequencies in the three populations,

631   and F_{ST} at each locus for every population pair. We used the F_{ST} estimator in Reynolds et al

632   (1983) and Weir and Cockerham (1984) because it incorporates the contribution of

633   sampling variances and generalizes to a multilocus estimator for F_{ST} across multiple sites,

634   allowing F_{ST} to be calculated for genomic "windows". The branch statistics PBS, and PBSn1

635   for focal population A were calculated from the genomic window F_{ST} using Equations 1 and

636   2.

637

638   Here, each model genome consisted of a collection of simulated replicates that include both

639   a subset of local sweep replicates as well as replicates from one or more other evolutionary

640   models. We created four types of model genomes representing scenarios where the

641   majority of sites evolved either neutrally or under BGS, while a small fraction experienced

642   selective sweeps (Figure 2A). The genomes consisted of 25,000 loci (genomic windows

643   from independent simulation replicates) for both the small and large populations, roughly

644   aligning with the full genome size for human and *Drosophila*, respectively. In one scenario,

645   99% of the loci were sampled from the million replicates of genetic drift simulations, and

646   the remaining 1% from the 10,000 local sweep runs. A similar model had 99% of loci

647   generated under a BGS model rather than drift (we do not simulate cases that are mixtures

648   of BGS and drift, because our BGS models include a subset of sites with B close to 1 that

649   essentially evolve as under neutral drift). The other sets of model genomes combined

650   parallel and local sweeps, with 98% of loci sampled from the genetic drift (or BGS)

651   replicates, 1% from local sweeps, and 1% from parallel sweeps. Focal population PBE at

652   each locus was computed from the locus-specific PBS and T_{BC} and from the median PBS and

653   T_{BC} values across all loci (Eq. 3) within the model genome.

654

655   After the completion of each simulation, we sampled 25 diploid individuals from each

656   population and calculated pairwise F_{ST} and the population branch statistics. We quantified

657   the relative performance of F_{ST}, PBS, PBSn1, and PBE by calculating their probability of

658   identifying local sweep loci. In this context, true positives were defined as local sweep loci

659   in the upper 1% quantile of a statistic's distribution, while false positives were any non-

660   local sweep loci in this quantile (*i.e.* loci evolving under genetic drift, BGS, or parallel

661   sweeps). The ratio of true positives to all positives defined the statistical *precision* of each
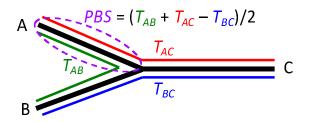
22

662   test. For the genomes with both local and parallel sweeps, we also calculated the fraction of

663   parallel sweep loci that contributed false negatives, and the fraction of false negatives that

664   were from parallel sweeps rather than drift or BGS.

665

666   Above, we describe the estimation of false positive rates in a simulated genome context. As

667   a complement to this approach, we also computed the statistical power of F_{ST} and the

668   branch statistics in a simpler and more conventional sense. Here, we separately compared

669   each simulated local sweep replicate against the full set of replicates simulated under a

670   given null model (either genetic drift or BGS). The statistical power was calculated as the

671   fraction of selective sweep replicates that were within a given statistic's upper 1% quantile

672   for the neutral or BGS replicates.

673

674

675   **Availability of Data and Materials:**

676   All of the msprime and SLiM scripts used to generate simulated data, along with scripts

677   used to analyze that data, are available at https://github.com/mshpak76/PBE.

678

679   **Acknowledgments:**

680   The authors thank the members of the Pool lab, Josep Comeron, and Aaron Ragsdale for

681   helpful discussion. We also thank Ben Haller, Peter Ralph, Jerome Kelleher and other

682   developers of SLiM and msprime in helping to resolve initial issues with the simulations.

683   This work was supported by NIH award R35 GM136306 and by NSF award DEB 1754745 to

684   JEP.

685

686   **Author Contributions:**

687   MS and JEP created the study design and wrote the manuscript, MS and KNL ran the

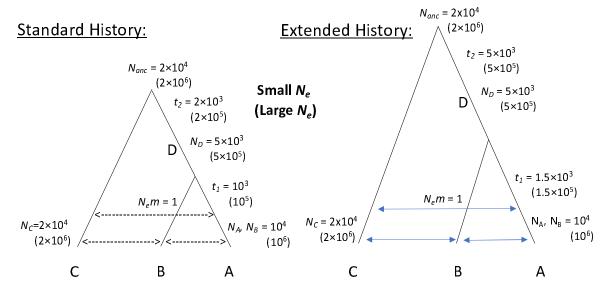688   simulations and analyzed the output. MS and JEP wrote the paper.

689 **Figures:**

(A) Unrooted Population Tree

$$PBS = (T_{AB} + T_{AC} - T_{BC})/2$$

(B) Simulated Demographic Histories

Standard History:

Extended History:

Small N_e (Large N_e)

(C) Genomic Parameters

Parameter	Small N_e	Large N_e
Mutation Rate	1.22×10^{-8}	5.21×10^{-9} [2.61×10^{-7}]
Recombination Rate	1.14×10^{-8}	1.09×10^{-8} [5.45×10^{-7}]
Gene Conversion Rate	5.9×10^{-8}	6.25×10^{-8} [3.13×10^{-6}]
Conversion Tract Length (bp)	100	518
Window Size (kilobases)	100	5

690

24

691   **Figure 1. Illustration of the demographic models and genomic parameters implemented in**

692   **population genetic simulations.** (A) Unrooted three population tree, with A representing the

693   the focal population. T values represent the genetic distances between each pair of

694   populations, based on a log transformation of F_{ST} (Eq. 1). PBS, as the estimated length of the

695   focal population branch, is then an intuitive function of these T values.

696   (B) The simulated three-population genealogies. N_i represents the size of the ith small

697   population, the value in parentheses is the size of the corresponding simulated large

698   population (the lineage D is the ancestral population for A,B before their split, N_{anc} is the

699   size of the population ancestral to all three sampled populations). t_1 and t_2 are the split

700   times for the inner and outer divergences, respectively. The left genealogy was used for the

701   migration-free simulations as well as the first set of migration simulations (a net

702   population migration rate of $N_e m = 1$ is represented by the dashed lines). The right

703   genealogy has branch lengths adjusted to generate the same pairwise F_{ST} as in the first

704   genealogy without migration under genetic drift alone. (C) The genomic parameters used

705   in simulations for the small and large populations (values in parentheses are the 50x

706   rescalings used in the simulations).

707

(A) Combining simulation replicates into simulated genomes

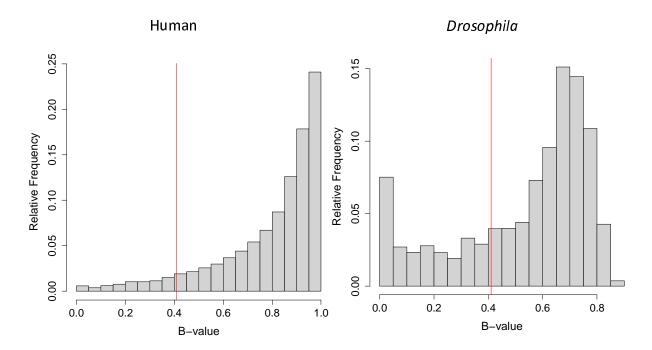## (B) Background selection B-value distributions

708

709 **Figure 2. Simulated genome models used in this study**. (A) Schematics representing the

710 model genomes each containing 25,000 simulated loci (representing genomic windows in

711 an empirical scan). The upper left genome has 99% of loci randomly sampled from the 10^6

712 genetic drift simulations and 1% from the 10^4 local selective sweep simulations. The upper

713 right genome has 98% neutral, 1% local selective sweep simulations, and 1% parallel

714 sweep simulations. The lower left genome has 99% BGS loci and 1% local sweeps. The

715 lower right genome has 98% BGS loci, 1% local sweeps, and 1% parallel sweep loci. (B) The

716 autosomal B-value distributions for human (left) and *D. melanogaster* (right) genomes are

717 shown, as used here for the simulations of small and large population simulations,

718 respectively. The vertical red line rep resents the truncation at $B = 0.41$ for each population

719 size scenario, in order to approximate genome scans in which low recombination regions

720 are excluded due to the difficulty in localizing targets.

721

722

	Other Loci Neutral				Other Loci Background Selection				Neutral + Parallel Sweeps				Backg. Selection + Parallel Sweeps			
SMALL POPULATION	F_{ST}	PBS	PBSn1	PBE	F_{ST}	PBS	PBSn1	PBE	F_{ST}	PBS	PBSn1	PBE	F_{ST}	PBS	PBSn1	PBE
Complete hard sweep	0.945	0.961	0.961	0.960	0.866	0.918	0.925	0.926	0.114	0.326	0.653	0.540	0.111	0.328	0.652	0.541
Partial sweep (80%)	0.877	0.921	0.919	0.918	0.718	0.831	0.842	0.845	0.152	0.417	0.628	0.625	0.146	0.411	0.616	0.612
Partial sweep (50%)	0.438	0.539	0.535	0.527	0.189	0.319	0.340	0.355	0.213	0.394	0.428	0.451	0.092	0.247	0.288	0.310
Soft sweep (0.005)	0.625	0.708	0.705	0.698	0.399	0.532	0.551	0.560	0.254	0.450	0.509	0.538	0.206	0.377	0.428	0.447
Soft sweep (0.01)	0.379	0.464	0.460	0.451	0.167	0.271	0.288	0.301	0.213	0.356	0.372	0.383	0.106	0.224	0.250	0.266
Migration (match F_{ST})	0.887	0.933	0.932	0.930	0.751	0.860	0.866	0.867	0.136	0.382	0.631	0.625	0.135	0.379	0.624	0.618
Migration (match tree)	0.939	0.961	0.960	0.959	0.847	0.915	0.919	0.922	0.128	0.381	0.651	0.659	0.124	0.376	0.647	0.654
LARGE POPULATION	F_{ST}	PBS	PBSn1	PBE	F_{ST}	PBS	PBSn1	PBE	F_{ST}	PBS	PBSn1	PBE	F_{ST}	PBS	PBSn1	PBE
Complete hard sweep	0.968	0.975	0.976	0.976	0.829	0.937	0.952	0.965	0.183	0.512	0.647	0.732	0.183	0.513	0.648	0.730
Partial sweep (80%)	0.864	0.922	0.925	0.925	0.551	0.744	0.781	0.842	0.061	0.271	0.403	0.540	0.059	0.269	0.402	0.537
Partial sweep (50%)	0.619	0.740	0.745	0.738	0.176	0.397	0.454	0.570	0.120	0.345	0.415	0.574	0.068	0.274	0.344	0.482
Soft sweep (0.005)	0.408	0.589	0.599	0.591	0.016	0.108	0.167	0.316	0.190	0.449	0.475	0.528	0.014	0.101	0.151	0.295
Soft sweep (0.01)	0.198	0.340	0.344	0.339	0.002	0.019	0.037	0.114	0.128	0.290	0.301	0.313	0.002	0.019	0.026	0.112
Migration (match F_{ST})	0.716	0.815	0.819	0.810	0.396	0.531	0.571	0.639	0.243	0.487	0.546	0.660	0.194	0.420	0.472	0.565
Migration (match tree)	0.851	0.921	0.922	0.919	0.533	0.723	0.750	0.813	0.242	0.487	0.546	0.678	0.239	0.481	0.539	0.660

Figure 3. Population branch statistics show greater precision to detect local selective sweeps than F_{ST} when other loci experience global positive or negative selection. Heat maps show the precision of F_{ST} and the population statistics PBS, PBSn1, PBE with respect to local sweeps, i.e. the fraction of local sweep loci among those contributing to the upper 1% quantile of each statistic. The table includes all demographic scenarios (large and small population, with and without migration), genomic backgrounds (genetic drift vs. BGS), and selection regimes (hard complete, partial, soft complete) considered in the study. Migration scenarios involved hard sweeps that might have fixed if not for gene flow.

	Other Loci Neutral				Other Loci Background Selection			
SMALL POPULATION	F_{ST}	PBS	PBSn1	PBE	F_{ST}	PBS	PBSn1	PBE
Complete hard sweep	0.888	0.675	0.349	0.459	0.891	0.675	0.347	0.461
Partial sweep (80%)	0.846	0.580	0.372	0.374	0.845	0.581	0.363	0.368
Partial sweep (50%)	0.705	0.451	0.363	0.284	0.544	0.321	0.246	0.201
Soft sweep (0.005)	0.741	0.530	0.452	0.406	0.683	0.466	0.383	0.352
Soft sweep (0.01)	0.660	0.418	0.355	0.398	0.481	0.289	0.245	0.207
Migration (match F_{ST})	0.866	0.621	0.370	0.378	0.864	0.618	0.366	0.373
Migration (match tree)	0.876	0.621	0.350	0.344	0.879	0.625	0.352	0.345
LARGE POPULATION								
Complete hard sweep	0.904	0.567	0.390	0.290	0.907	0.568	0.391	0.282
Partial sweep (80%)	0.940	0.731	0.598	0.459	0.938	0.729	0.595	0.460
Partial sweep (50%)	0.879	0.653	0.584	0.413	0.801	0.566	0.497	0.348
Soft sweep (0.005)	0.714	0.417	0.380	0.528	0.171	0.061	0.061	0.043
Soft sweep (0.01)	0.505	0.265	0.242	0.145	0.015	0.007	0.009	0.010
Migration (match F_{ST})	0.759	0.513	0.453	0.331	0.713	0.456	0.394	0.284
Migration (match tree)	0.759	0.516	0.458	0.327	0.759	0.514	0.454	0.324

Figure 4. PBE and PBSn1 are less likely than other statistics to register parallel sweep loci among their top outliers. For model genome scenarios in which 1% of loci are subject to parallel sweeps in all populations, the heat maps show the fraction of parallel sweep loci that contribute to the upper 1% quantile of F_{ST} and branch statistic distributions (i.e. the fraction of parallel sweep loci that are false positives) under different demographic and selection parameters.

References:

751 **References**:

752

753 Adrion JR, Hahn MW, Cooper BS. 2015. Revisiting classic clines in *Drosophila melanogaster*

754 in the age of genomics. Trends in Genetics 31:P434-444.

755

756 Akey JM. 2009. Constructing genomic maps of positive selection in humans: where do we go

757 from here? Genome Research 19:711-722.

758

759 Amato R, et al. 2009. Genome-wide scan for signatures of human population differentiation

760 and their relationship with natural selection, functional pathways, and diseases. PLoS

761 One:e7927.

762

763 Antonovics J, Bradshaw AD. 1970. Evolution in closely adjacent plant populations. VIII.

764 Clincal patterns in *Anthoxanthum odoratum* across a mine boundary. Heredity 25:249-362.

765

766 Baumdicker F, et al. 2022. Efficient ancestry and mutation simulation with msprime 1.0.

767 Genetics 220:iyab229.

768

769 Booker T, Yeaman S, Whitlock MC 2020. Variation in recombination rate affects detection of

770 outliers in genome scans under neutrality. Molecular Ecology: 29: 4274-4279.

771

772 Brodie ED Jr, Ridenhour BJ, Brodie ED III. 2002. The evolutionary response of predators to

773 dangerous prey: hotspots and coldspots in the geographic mosaic of coevolution between

774 garter snakes and newts. Evolution 56:2067-2082.

775

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

777 neutral genetic variation. Genetics 134:1289-1303.

778

779 Charlesworth D, Charlesworth B, Morgan MT. 1995. The pattern of neutral molecular

780 variation under the background selection model. Genetics 141:1619-1632.

781

782   Charlesworth B, Nordborg M, Charlesworth D. 1997. The effects of local selection, balanced

783   polymorphism, and background selection on equilibrium patterns of genetic diversity in

784   subdivided populations. Genetical Research 70:155-174.

785

786   Charlesworth B. 2012. The effects of deleterious mutations on evolution at linked sites.

787   Genetics 190:5-22.

788

789   Cheng JY, Stern AJ, Racimo F, Nielsen R. 2022. Detecting selection in multiple populations by

790   modeling ancestral admixture components. Molecular Biology and Evolution 39:msab294

791

792   Colosimo PF, et al. 2005. Widespread parallel evolution in sticklebacks by repeated fixation

793   of ectydysplasin alleles. Science 307:1928-1933.

794

795   Comeron JM, Ratnappan R, Bailin S. 2012. The many landscapes of recombination in

796   *Drosophila melanogaster*. PLoS Genetics 10:e1004434.

797

798   Comeron JM. 2014. Background selection as baseline for nucleotide variation across the

799   *Drosophila* genome. PLoS Genetics 10:e1004434

800

801   da Silva Ribeiro T, Galvan JA, Pool JE. 2022. SNP-level F_{ST} outperforms whole-window

802   statistics for detecting soft sweeps in local adaptation. Genome Biology and Evolution

803   14:evac143.

804

805   Ewing GB, Jensen JD. The consequences of not accounting for background selection in

806   demographic inference. Molecular Ecology 25:135-141.

807

808   Granka JM, et al. 2012. Limited evidence for classic selective sweeps in African populations.

809   Genetics 192:1049-64.

810

811   Gygax M, Rentsh AK, Rudman SM, Rennison DJ. 2018. Differential predation alters

812   pigmentation in threespine stickleback (*Gasterosteus aculeatus*). Journal of Evolutionary
813   Biology 31:1589-1598.

814

815   Haller BC, et al. 2019. Tree-sequence recording in SLiM opens new horizons for forward-
816   time simulations. Molecular Ecology Resources 19:552-566.

817

818   Hermisson J, Pennings PS. 2005. Soft sweeps: molecular population genetics of adaptation
819   from standing genetic variation. Genetics 169:2335-2352.

820

821   Hernandez RD et al. 2011. Classic selective sweeps are rare in recent human evolution.
822   Science 331:920-924.

823

824   Hoekstra HE, et al. 2006. A single amino acid mutation contributes to adaptive beach mouse
825   color pattern. Science 313:101-104.

826

827   Huber CD, DeGiorgio M, Hellman I, Nielsen R. 2015. Detecting recent selective sweeps while
828   controlling for mutation rate and background selection. Molecular Ecology 25:142-156.

829

830   Jiang X, Assis R. 2020. Population-specific genetic and expression differentiation in
831   Europeans. Genome Biology and Evolution 12:358-369.

832

833   Jonsson H, et al. 2017. Parental influence on human germline de novo mutations in 1,584
834   trios from Iceland. Nature 549:519-522.

835

836   Julian CG, Moore LG. 2019. Human genetic adaptation to high altitude: evidence from the
837   Andes. Genes 10:150

838

839   Kapun M, et al. 2020. Genomic analysis of European Drosophila melanogaster reveals
840   longitudinal structure, continent-wide selection, and previously unknown DNA viruses.
841   Molecular Biology and Evolution 37:2661-2678.

842

843   Kelleher J, Thornton KR, Ashander J, Ralph PL. 2018. Efficient pedigree recording for fast

844   population genetics simulation. PLoS Computational Biology 14:e1006581

845

846   Lack JB, et al. 2015. The *Drosophila* genome exus: a population genomic resource of 623

847   *Drosophila melanogaster* genomes, including 197 from a single ancestral range population.

848   Genetics 199:1229-1241.

849

850   Lack JB, et al. 2016. Life history evolution and cellular mechanisms associated with

851   increased size in high altitude *Drosophila*. Evolutionary Ecology 6:5893-5906.

852

853   Lange JD, Pool JE. 2018. Impacts of recurrent hitchhiking on divergence and demographic

854   inference in *Drosophila*. Genome Biology and Evolution 10:1882-1891.

855

856   Macnair MR. 1987. Heavy metal tolerance in plants: a model evolutionary system. Trends in

857   Ecology and Evolution 2:354-359.

858

859   Malaspinas A-S et al. 2016. A genomic history of Aboriginal Australia. Nature 538:207-

860   2014.

861

862   McVicker G, Gordon D, Davis C, Green P. 2009. Widespread genomic signatures of natural

863   selection in hominid evolution. PLoS Genetics 5:e1000471.

864

865   Messer PW. 2013. SLiM: Simulation evolution with selection and linkage. Genetics

866   194:1037-1039.

867

868   Miller SE, Metcalf D, Schluter D. 2015. Intraguild predation leads to genetically based

869   character shifts in the threespine stickleback. Evolution 69:3194-3203.

870

871   Nachman WM, Hoekstra HE, D'Agostino SL. 2003. The genetic basis of adaptive melanism in

872   pocket mice. Proceedings of the National Academy of Sciences USA 100:5268-5273.

873

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

875   *melanogaster*. Molecular Biology and Evolution 34:349-360.

876

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

878   a short-term genetic distance. Genetics 105:767-779.

879

880   Schiffels S, Durbin R. 2014. Inferring human population size and separation history from

881   multiple genome sequences. Nature Genetics 46:919-925.

882

883   Schlebusch CM, et al. 2020. Kohe-San genomes reveal unique variation and confirm the

884   deepest population divergence in *Homo sapiens*. Molecular Biology and Evolution.

885   37:2944-2954

886

887   Schmidt JM, et al. 2019. The impact of genetic adaptation on chimpanzee subspecies

888   differentiation. PLOS Genetics. 15:e1008485.

889

890   Siddiq MA, Thornton JW. 2019. Fitness effects but no temperature-mediated balancing

891   selection at the polymorphic *Adh* gene of *Drosophila melanogaster*. Proceedings of the

892   National Academy of Sciences USA 116:21634-21640.

893

894   Sprengelmeyer QD, Pool JE. 2021. Ethanol resistance in *Drosophila melanogaster* has

895   increased in parallel cold-adapted populations and shows variable genetic architecture

896   within and among populations. Ecology and Evolution 11:15364-15376.

897

898   Sprengelmeyer QD, et al. 2022. The evolution of larger size in high altitude *Drosophila*

899   *melanogaster* has a polymorphic genetic architecture. G3 12:jkab454.

900

901   Swetec N, Zhao L, Saelao P et al. 2015. Evidence that natural selection maintains genetic

902   variation for sleep in *Drosophila melanogaster*. BMC Evolutionary Biology 15:41

903

904   Vicuna L, et al. 2019. Adaptation to extreme environments in an admixed human population

905   from the Atacama Desert. Genome Biology and Evolution 11:2468-2479.

906

907   Voight BF, Kudaravalli S, Wen X, Pritchard JK. 2006. A map of recent positive selection in the

908   human genome. PLOS Biology 4:e72.

909

910   Weir BS, Cockerham CC 1984. Estimating F-statistics for the analysis of population

911   structure. Evolution 38: 1358-1370.

912

913   Williams AL, et al. 2015. Non-crossover gene conversions show strong CG bias and

914   unexpected clustering in humans. Elife 4:e04637.

915

916   Witt KE, Huerta-Sanchez E. 2019. Convergent evolution in human and domesticate

917   adaptation to high-altitude environments. Philosophical Transactions of the Royal Society

918   of London B Biological Sciences 374:20180235.

919

920   Yassin A, et al. 2016. Recurrent specialization on a toxic fruit in an island *Drosophila*

921   population. Proceedings of the National Academy of Sciences USA 113:4771-4776.