Impacts of Recurrent Hitchhiking on Divergence and Demographic Inference in *Drosophila*

Jeremy D. Lange and John E. Pool

Laboratory of Genetics, University of Wisconsin–Madison, Madison, WI, 53706

Running Title: Hitchhiking, Divergence, and Demographic Inference

Keywords: *Drosophila melanogaster*, recurrent hitchhiking, Hill-Robertson interference, demographic inference, forward simulation

Corresponding author:
John E. Pool
425-G Henry Mall
Madison, WI 53706
+1-608-265-1036
jpool@wisc.edu

Abstract

In species with large population sizes such as *Drosophila*, natural selection may have substantial effects on genetic diversity and divergence. However, the implications of this widespread non-neutrality for standard population genetic assumptions and practices remain poorly resolved. Here, we assess the consequences of recurrent hitchhiking (RHH), in which selective sweeps occur at a given rate randomly across the genome. We use forward simulations to examine two published RHH models for *D. melanogaster*, reflecting relatively common/weak and rare/strong selection, respectively. We find that unlike the rare/strong RHH model, the common/weak model entails a substantial degree of Hill-Robertson interference, which has implications for the rate of beneficial mutation and for the simulation of RHH models. We also find that the common/weak RHH model is more consistent with our genome-wide estimate of the proportion of substitutions fixed by natural selection between *D. melanogaster* and *D. simulans* (19%). Finally, we examine how these models of RHH might bias demographic inference. We find that RHH has relatively minor effects on the inference of recent between-population demographic parameters, while having stronger effects on inference of longer term ancestral population parameters. Thus, even for species with important genome-wide impacts of selective sweeps, neutralist demographic inference can have some utility in understanding the histories of recently-diverged populations.

Keywords: *Drosophila melanogaster*, recurrent hitchhiking, Hill-Robertson interference, demographic inference, forward simulations

41 **Introduction**

42 The advancement of DNA sequencing technology, along with computational capacity and
43 methodology, continues to revolutionize the field of population genetics. Harnessing the power
44 of whole genome datasets, researchers have begun to explore a wider variety of evolutionary
45 models. One such model that has received considerable attention recently is a model of recurrent
46 hitchhiking, where genetic diversity at neutral regions is reduced due to repeated selective
47 sweeps at nearby loci. This reduction in diversity has been explored theoretically (Kaplan *et al.*
48 1989; Stephan *et al.* 1992; Wiehe and Stephan 1993) showing that the expected reduction in
49 diversity can be approximated as a function of RHH model parameters: $\gamma = 2N_e s$ and λ, where
50 N_e is the effective population size, s is the selection coefficient, and λ is the rate of positively
51 selected substitutions. Subsequent studies have examined such RHH models using forward
52 simulation, focusing attention on how Hill-Robertson interference (HRI; Hill and Robertson
53 1966) between linked beneficial mutations on different haplotypes reduces the probability of
54 fixation (Gerrish and Lenski 1998, Chevin et al. 2008).

55 The impact of natural selection on genomic diversity may be particularly significant for
56 species with very large population sizes, such as *Drosophila melanogaster* (*e.g.* Sella *et al.* 2009;
57 Langley *et al.* 2012). In abundant taxa, the population adaptive mutation rate is elevated and the
58 weak influence of genetic drift may allow natural selection to favor alleles with modest selection
59 coefficients. By estimating RHH parameters, Jensen *et al.* (2008) suggested that selective sweeps
60 may reduce genomic diversity in *D. melanogaster* to half of neutral levels. While this study
61 implicated a model of relatively strong and infrequent sweeps, the study of Andolfatto (2007),
62 which focused on nonsynonymous sites, instead favored a model of substantially weaker but
63 more frequent adaptive substitutions. While these models should imply strongly different

64 proportions of substitutions driven by positive selection, their alignment with estimates of this
65 quantity from *Drosophila* data is unclear. And likewise, the predictions of each model for the
66 role of HRI and for the fixation of neutral variants have not been investigated. We therefore seek
67 to clarify the relationship between published *Drosophila* RHH models and divergence.
68 If linkage to natural selection substantially impacts *Drosophila* genetic diversity at
69 neutral sites, the accuracy of demographic inference methods that assume neutrality is not
70 assured. Most sites in the fly genome experience direct functional constraint (Halligan and
71 Keightley 2006), which may lead to an excess of rare alleles from deleterious polymorphisms.
72 Many sites that are not under as much direct selection pressure, such as synonymous sites and
73 middles of short introns, are by definition very close to nonsynonymous sites and other
74 functional sites that may experience natural selection. Selective sweeps could skew the genome-
75 wide allele frequency spectrum, in particular by generating a skew toward rare alleles
76 (Braverman *et al.* 1995) that may resemble the predictions of recent population growth. In line
77 with these concerns, Schrider *et al.* (2016) found that the presence of positive selection can bias
78 demographic parameter estimates for a single population's history, and can lead to
79 misidentification of demographic models. However, much interest centers on the inference of
80 demographic parameters between recently-diverged populations, and it remains unclear whether
81 *Drosophila*-like RHH on shorter time-scales is sufficient to bias parameters concerning
82 population divergence times, population-specific size changes, and migration rates. We therefore
83 use RHH simulations to investigate the impact of RHH on estimation of these parameters.
84

Materials and Methods

McDonald-Kreitman analysis

To estimate the proportion of substitutions in the *Drosophila melanogaster* genome that were fixed by natural selection, we applied a genome-wide McDonald-Kreitman analysis (McDonald and Kreitman1991). Here, we surveyed 197 genomes from a Zambian population of *Drosophila melanogaster* (Lack *et al.* 2015), which is believed to be within the ancestral range of the species (Pool *et al.* 2012). These genomes are masked for identity by descent, apparent heterozygosity, and recent cosmopolitan admixture. Thus, in our analysis, we required any given site to be called in at least 50% of the genomes. Further, for a site to be considered polymorphic, we required an allele to be segregating at least at 10% in our samples. Applying this filter to both putatively neutral and selected site classes should reduce bias from deleterious polymorphisms. To estimate the number of substitutions, we used a *Drosophila simulans* genome aligned to the *D. melanogaster* genome (Stanley and Kulathinal 2016).

In this analysis, we calculated the proportion of substitutions driven to fixation by natural selection as $\alpha = 1 - \frac{D_n P_s}{D_s P_n}$ (Smith and Eyre-Walker 2002). Here, D_n is the number of neutral substitutions, P_s is the number of non-neutral polymorphic sites, D_s is the number of non-neutral substitutions, and P_n is the number of neutral polymorphic sites. α was calculated for nine site classes (nonsynonymous, two-fold synonymous, three-fold synonymous, 5' untranslated regions, 3' untranslated regions, intron, intergenic, and RNA-coding) and individually for each major chromosome arm (2L, 2R, 3L, 3R, and X). Middles of short introns (Halligan and Keightley 2006) and four-fold synonymous sites were evaluated as proxies for neutral evolution. Site classes were taken from flybase.org for release 5.50 of the *D. melanogaster* genome.

Simulations

In this study, we are interested in the effects of recurrent hitchhiking on demographic inference. To examine this, we ran forward simulations using SLIM version 2.1 (Haller & Messer 2016) to model recurrent hitchhiking. Because full-forward simulations are memory intensive and slow when simulating large populations, it is necessary to rescale simulation simulation parameters. We started by running test simulations to get an idea of the largest population size that we could simulate in a reasonable amount of time. We concluded that diploid populations of 50,000 individuals was a sensible target. This results in 50X rescaling assuming an effective population size of roughly 2,500,000 (*e.g.* Duchen *et al.* 2013). Following the results of Uricchio & Hernandez (2014), we determined that under the RHH models of interest, a size reduction to 50,000 individuals should closely maintain the genetic variation of a non-rescaled population. Further, both algorithms provided in the cited paper yielded near identical scaled parameters when reducing the population size from 2,500,000 to 50,000 individuals. Because of this, we used the simpler algorithm 1 of Uricchio & Hernandez (2014). The main idea behind this rescaling method is that patterns of genetic diversity are maintained when population-scaled parameters $\theta = 4N_e\mu$, $\rho = 4N_e r$, and $\gamma = 2N_e s$ are fixed while N_e is varied. As such, if we decrease N_e=2,500,000 by 50X to N_e=50,000, then θ, ρ, and α must be increased by 50X. The algorithm is laid out in step form below.

- Let $s_0 = \alpha_0/2N_0$; $r_0 = \rho_0/4N_0$; $a = s_0/L_0 r_0$
- $\gamma_1 = \gamma_0$
- $s_1 = \gamma_1/2N_1$
- $r_1 = s_0/aL_1$
- $\lambda_1 = r_1\lambda_0/r_0$

Here, the subscripts refer to before rescaling (subscript 0) and after rescaling (subscript 1). s_1 is the selection strength, N_1 is the population size, r_1 is the per base pair per generation per chromosome recombination rate, and L_1 is the simulated locus length.

We ran simulations under two different models of RHH, both of which were estimated from *D. melanogaster* data. Since the rate of adaptive substitutions and the average selective advantage are highly confounded in terms of their impact on diversity levels, we wanted to examine complementary models. The first model we chose to study is from Jensen *et al.* (2008). Here, the rate of incoming adaptive substitutions (λ) is low (λ = 7.9E-12) and the average strength of selection (s = 0.011) is high. The second model, from Andolfatto (2007) consists of a high rate of adaptive substitutions (λ = 6.9E-10) with a very low average selection strength (s=1.2E-5). For both models, we used a mutation rate μ = 3.27E-9 (Schrider *et al.* 2013) and a recombination rate r = 2.5E-8. In forward simulations, it is not possible to directly specify the rate of adaptive substitutions. Instead, one must input the rate of beneficial mutations ν. In the absence of interference among selected mutations, this can be derived using λ and the probability of fixation (Kimura 1962):

$$\lambda = \nu P_{fix}$$

$$\nu = \frac{\lambda}{P_{fix}}$$

$$\nu = \lambda \Big/ \frac{1 - e^{-2s}}{1 - e^{-2\gamma}}$$

In our simulations, each beneficial mutation had its selection coefficient drawn randomly from an exponential distribution with a mean equal to the rescaled selection strength. This variation in selection coefficients helps to avoid the artificial scenario of interference between mutations with precisely identical fitness.

150 We wanted to simulate at least 10 kilobases (kb) for each simulation. Because a sweep
151 can affect regions far from the target of a sweep, we simulated extra flanking regions for each
152 side of the 10 kb that was used for the demographic inference, while analyzing only the middle
153 region. For the common weak sweep model, we simulated 480 base pairs on each side of the 10
154 kb for a total of 10,960 base pairs simulated. This follows a simple rule of thumb that any given
155 sweep will primarily affect diversity at regions within s/r base pairs from the sweep (where s is
156 the selection strength and r is the per base pair recombination rate). For the rare strong sweep
157 RHH model, such a rule of thumb would require 440,000 base pairs to be simulated. Such
158 simulations are computationally infeasible, so we simulated a total of 30 kb for this model.
159 10,000 simulations were ran for the common weak sweep model while 1,000 simulations were
160 ran for the rare strong sweep model. All simulation parameters are provided in Table 1.
161
162 *Demographies simulated*
163 Forward simulations require a burn-in period to generate appropriate genetic variation.
164 Thus, both recurrent hitchhiking models were run for 500,000 ($10N_e$) generations. These "trunk"
165 simulations were then used for the demographic simulations. There are two relevant
166 demographic models that we are interested in, two-population models where the populations split
167 after the burn-in period. These demographies include a bottleneck model and an isolation with
168 migration (IM) model. In our bottleneck model, the populations split and one population
169 experiences a bottleneck. The parameters of the bottleneck model were taken from Thornton and
170 Andolfatto (2006). In this model, a bottleneck occurs 0.0516 coalescent time units in the past and
171 lasts for 0.042 coalescent time units. During the bottleneck, the population decreases to 4.7% of
172 its original population size. The IM model consists of a population split $0.25N_e$ coalescent units

173 in the past and subsequent migration of $4N_em = 0.5$. We simulated two different kinds of IM
174 models, a "shared sweep" and a "private sweep" model. The "shared sweep" model allowed
175 selective mutations to have an equal selective advantage if they migrated into the other
176 population. The "private sweep" model multiplied *s* by -1 if it migrated into the other population,
177 making the allele deleterious.
178
179 *Demographic Inference*
180 To examine how recurrent hitchhiking affects demographic inference, we used δαδι
181 version 1.6.3 (Gutenkunst *et al.* 2009) to estimate demographic model parameters. We first
182 attempted to fit a two-epoch and three-epoch size change model to the trunk simulations (where
183 there were no size changes simulated) to examine whether recurrent hitchhiking can misidentify
184 demographic models. For the bottleneck simulations, we fit two and three parameter bottleneck
185 models. The three parameter bottleneck model consisted of a population size reduction, a length
186 of time as the reduced population size, and an instantaneous size change back to the original size
187 that occurs some time in the past. In the two population bottleneck model, the length of the
188 bottleneck was fixed and not optimized. We examined each bottleneck model both with and
189 without fitting the ancestral size change models as well. In this way, we could better parse
190 ancient parameters from more recent demographic parameters post population split. Finally, we
191 tested the IM models with both shared and private selective sweeps. In these cases, the timing of
192 the population split and the migration rate were estimated. As in the bottleneck cases, we tested
193 both IM models with and without ancestral size changes. In total, 14 demographic models were
194 investigated for both hitchhiking models and 11 demographic models were tested for the neutral
195 simulations (since there is no shared/private sweep distinction in the neutral case).

196

197 **Results**

198 *Effects of Hill-Robertson Interference*

199 In order to investigate the effects of recurrent hitchhiking on divergence and diversity in
200 *Drosophila*, we performed forward simulations reflecting two published models of RHH
201 representing relatively common/weak selection (Andolfatto 2007) and rare/strong selection
202 (Jensen *et al.* 2008), respectively. Before proceeding with further analysis, we checked to see if
203 our initial adaptive mutation rates based on these models were producing the prescribed rates of
204 adaptive substitution, or if instead an important impact of interference must be accounted for.
205 While the above studies assumed no interference between positively selected mutations,
206 Andolfatto (2007) suggested that an interesting next step would be to examine how the presence
207 of interference influences the observed versus expected adaptive substitution rate (λ) and
208 selection coefficient of fixed beneficial substitutions (s) in the simulations.

209 Under a model of no interference, we can use λ to set the adaptive mutation rate (ν) in our
210 simulations (Kimura 1962; Methods), and we can then expect approximately $2N_e\nu LgP_{fix}$
211 adaptive substitutions as a product of the total number of adaptive substitutions at L sites across g
212 generations and their probability of fixation. However, if multiple beneficial mutations are
213 sweeping simultaneously, competition among sweeping haplotypes will lead fewer mutations to
214 fix. Further, mutations that do fix will tend to have a higher selection strength than the input
215 distribution of selective effects. The dynamics of the common/weak and rare/strong models that
216 we tested are very different. In the rare/strong model, on average, we expect 8.88 beneficial
217 mutations to occur per simulation, with 3.42 of these mutations fixing. It is unlikely that any two
218 beneficial mutations would be sweeping at the same time across the 500,000 generations. In the

219 common/weak model, however, we expect 157,644 beneficial mutations per simulation. Under a
220 model of no interference, we would expect 189 of these beneficial mutations to fix. This makes it
221 very likely that more than one beneficial mutation would be sweeping at any given time. Thus,
222 we expect interference to be much more significant in the common/weak model relative to the
223 rare/strong model.

224 To test this expectation, we ran both the common/weak and the rare/strong models as
225 described in the methods section, using the published λ and s to generate input parameters. We
226 tracked every mutation that fixed across all simulations and recorded the average selection
227 coefficient. To accurately reflect the RHH models that we were simulating, our goal for our
228 simulations was to match the expected number of fixed beneficial mutations as described above.
229 Across the 1000 simulations of the the rare/strong model, an average of 3.21 beneficial mutations
230 fixed compared to an expectation of 3.42. For the common/weak model, however, we found that,
231 on average, 84.2 beneficial mutations fixed. This corresponds to 56% fewer fixed beneficial
232 mutations than we expected. Further, the average selection strength of the fixed beneficial
233 mutations was more than twice the input parameter of mean selection strength. We attribute this,
234 at least partially, to interference between positively selected mutations. We hypothesized that we
235 would have to increase the beneficial mutation rate and decrease the average selection strength in
236 order to emulate the properties of adaptive substitutions estimated for the common/weak model.
237 To test this, we ran batches of simulations with an increased beneficial mutation rate and
238 decreased average selection strength and calculated the average number of fixed beneficial
239 mutations and their associated average s (results not shown). Though increasing the beneficial
240 mutation rate will allow for more beneficial mutations to occur, it will also create more
241 interference among positively selected mutations. Thus, predictions may be nonlinear. For each

242 pair of beneficial mutation rate and *s*, we ran 150 simulations for 500,000 generations. We found
243 that by increasing the rate of beneficial mutation by 4X and decreasing the average selection
244 strength from 0.0006 to 0.000275 (Figure 1), we recover approximately the correct number of
245 fixed beneficial mutations and selection strength (on average, 146 beneficial mutations fixed
246 with an average selection coefficient of s= 0.00057). We used these retuned ν and *s* parameters
247 and ran 10,000 simulations as described in the methods section to more accurately reflect the
248 common weak sweep RHH model. The output of these simulations were used for the simulation
249 analyses detailed below. Note that for a non-rescaled *Drosophila* population, our adjusted RHH
250 mutational parameters for the Andolfatto (2007) model would correspond to an advantage of s =
251 5.50E-6 and a beneficial mutation rate of ν = 2.30E-11.
252
253 *Impacts of RHH on adaptive and neutral divergence*
254 A mutation can fix in a population in one of three ways: natural selection favors the allele
255 and it increases to fixation (our "driver" class), a neighboring mutation can be favored and the
256 focal mutation fixes via hitchhiking (a "passenger"), or the mutation can drift to 100% frequency
257 (a "drifter"). The simulations that we ran allowed us to track and classify each fixed mutation,
258 giving us insight into the dynamics of both RHH models. In order to have unambiguous criteria,
259 we define passengers narrowly as mutations that fix simultaneously with with beneficial
260 mutation. Fixation class proportions are given in Table 2, as expected showing much higher
261 proportions of adaptive substitutions in the common/weak RHH model relative to the rare/strong
262 model.
263 We can also obtain a prediction for the driver proportion (α), in the absence of HRI, using
264 the parameters of the model. The prediction, in this case, is simply $\frac{\lambda}{\lambda+\mu}$. Here, μ is the neutral

265 mutation rate, which is also equal to the rate of fixation of a neutral allele. In the rare/strong
266 RHH model, the expected driver proportion is 0.139%. Concordantly, simulations indicated that
267 0.137% of fixations were drivers. We expected a driver proportion of 17.4% in the common,
268 weak sweep RHH model. As expected, in the non-retuned case, we found a reduced proportion
269 (9.99%). This metric was mostly recovered after retuning parameters, with the driver proportion
270 of 16.7%. These driver proportions can be compared with the same quantity (α) estimated from
271 extended McDonald-Kreitman analyses of empirical data (Smith and Eyre-Walker 2002). Our
272 genome-wide analysis estimated that approximately 18.8% of substitutions in the *Drosophila*
273 *melanogaster* genome were driven to fixation by natural selection (Figure 2; Table S1). Our
274 estimates of α are largely concordant with other studies (Andolfatto 2005, Begun *et al.* 2007, but
275 see Mackay *et al.* 2012). This empirical analysis suggests that the common/weak RHH model
276 seems more compatible with adaptive divergence estimates in *D. melanogaster*. The specific
277 common/weak RHH model simulated in this study provides a convenient representative of such
278 a model.
279
280 *Hitchhiking effects on demographic inference*
281 The final goal of this study was to examine how models of recurrent hitchhiking affect
282 inferences on demography. We tested whether selective sweep models involving substantial
283 hitchhiking effects would violate assumptions of neutrality made by demographic inference tools
284 and bias parameter estimation. We ran our simulations for 500,000 generations as a single
285 population before we added a population split and distinct demographies. In Figure 3, we show
286 the single population site frequency spectrum for neutrality and both RHH models. There is a
287 slight increase in rare alleles under both RHH models compared to neutrality, which is consistent

288 with expectations (Braverman *et al.* 1995). $\pi, \theta_s, Tajima's\ D,$ and F_{ST} are provided in Table S2.
289 In every case, diversity is strongly reduced when selection is simulated.

290 For the "trunk" simulations, we first asked if δαδι would fit a model with population size
291 changes over the true model of constant size. Here, we tested both two and three epoch size
292 change models. The two epoch model consisted of two parameters to maximize: a single size
293 change (reduction or expansion) that occurs some time in the past. The three epoch model has
294 three parameters: a size change (reduction or expansion), a length of time at this new size, and
295 some time in the past when the population recovers. The parameters of these models were most
296 affected by the RHH models. Under neutrality, δαδι prefers a model of constant population size
297 (the true model) for both the two and the three epoch models. In the two epoch size change
298 model, δαδι infers a population expansion for both selection models, and it infers a population
299 bottleneck for the three epoch model (Figure 4). Qualititatively similar single-population results
300 have recently been reported (Schrider *et al.* 2016).

301 Our primary interest was to assess whether the demographic parameters that relate
302 recently-diverged populations are similarly biased by RHH. We therefore attempted to infer two
303 distinct bottleneck models occurring after the 500,000 generation burn-in: a two and a three
304 parameter model. Both models consisted of a population size contraction, a bottleneck length,
305 and a length in time since recovery back to original population size. All three parameters are
306 inferred in the three parameter model while bottleneck length is fixed and not inferred in the two
307 parameter model. Results showed that in the two parameter bottleneck model, bottleneck
308 strength was accurately inferred under neutrality and both selection models (Figure 5; Table S3).
309 In the three parameter bottleneck case, bottleneck strength and bottleneck duration were
310 accurately inferred in all cases. Time since recovery was the only parameter affected by the

311 presence of selection. Under neutrality, this parameter was accurately recapitulated for both the
312 two and three parameter bottleneck models. Under selection, however, this parameter was
313 overestimated nearly twofold. This result is in line with the fact that both post-bottleneck growth
314 and selective sweeps leave behind an excess of rare alleles. Thus, if both are occurring in our
315 simulation but δαδι assumes that only neutral events have occurred, it makes sense that this
316 inference method is biased towards overestimating bottleneck recovery times in the presence of
317 positive selection. Models were also run in which ancestral size change parameters and
318 bottleneck parameters were estimated simultaneously, with qualitatively similar results (Table
319 S4).

320 We simulated two distinct IM models: one with shared selective sweeps between
321 populations, and one simulating local adaptation. For both cases, we estimated the time since the
322 population split and the migration rate. Even under neutrality, δαδι moderately but consistently
323 underestimated the time since split (Figure 6; Table S3): the median value across 200 datasets
324 was 0.19 coalescent units, versus a true value of 0.25. Each selective regime showed a distinct
325 result for this split time parameter. For both RHH models, shared sweeps resulted in lower
326 divergence time estimates than under neutrality, while private sweeps resulted in higher
327 estimates. These results may reflect the greater/lesser genetic similarity between populations due
328 to selection in the shared/private sweep models decreasing/increasing divergence time estimates,
329 respectively. When inferring migration rate, there was little difference in estimates among
330 models. In each case, the correct order of magnitude was estimated (m = 1.25E-6 for $4N_em$ =
331 0.25). As with the bottleneck models, we also made demographic inferences by combining the
332 ancestral size change models with the recent divergence IM models, again with comparable
333 results (Table S4).

334

335 **Discussion**

336 Simulation provides us with flexible tools for studying the impact of selective sweeps on
337 genetic diversity and divergence. Here, we assess published *Drosophila* RHH models, we find
338 support for a model of frequent and weak positive selection, and we show that this model entails
339 substantial HRI. However, although we find agreement between the above model and an
340 empirical data analysis, we do not argue for a specific quantitative RHH model in light of some
341 important caveats. First, the model that gives α estimates in line with our genome-wide estimates
342 is based on an analysis of nonsynonynous sites specifically (Andolfatto 2007). While this model
343 may share traits in common with the true genomic RHH model for this species, it is best viewed
344 as a qualitative example of the type of model compatible with this aspect of data. Second, we are
345 assessing RHH models based on their general agreement with McDonald-Kreitman-based
346 estimates of α, but such estimates can be biased depending on the history of population size
347 change (Eyre-Walker 2002) and recombination rate change (Comeron 2014). Finally, estimates
348 of α derived from RHH models depend not only on adaptive parameters, but also upon the
349 *neutral* mutation rate that we use in simulations. The raw mutation rate may be somewhat higher
350 than our simulated μ (Schrider *et al.* 2013; Huang *et al.* 2016). However, more than half of these
351 mutations should be prevented from fixing by selective constraint (Halligan and Keightley
352 2006). If the μ that we use in our simulations is too high, the predictions for α may be too low,
353 and vice versa. In spite of these quantitative uncertainties, we argue that our analyses provide
354 general insight into the RHH models that are most plausible for this species, and particularly the
355 apparent importance of HRI in such models.

356 Our results also shed light on the impact of RHH on linked neutral variation and its
357 consequences for demographic estimation. It is clear that strong effects of natural selection
358 across the genome can violate the neutralist assumptions of typical demographic inference
359 methods (Ewing and Jensen 2016; Schrider *et al.* 2016). Our results support the biasing effect of
360 RHH on ancient parameter estimation and demographic model choice. This is in line with a
361 previous study using single population simulations (Schrider *et al.* 2016). However, we show
362 that there can be relatively less impact on recent parameters in two population models. Thus,
363 such methods may retain utility even for populous species like *D. melanogaster*, but the biases
364 that do exist should be borne in mind for this species and investigated for other taxa.
365 Our simulations emulate a population of *D. melanogaster*, a populous species with a
366 compact genome. It has been shown that selection may be prevalent in large fly populations
367 (Sella *et al.* 2009; Karasov *et al.* 2010; Langley *et al.* 2012) since there is more opportunity for
368 selection to occur. Further, due to the compactness of the genome, a selective sweep can affect
369 relatively large regions. It is less clear how much of an impact natural selection would have on
370 demographic inference in smaller N_e species, such as humans. It is presumed that much of the
371 genetic variation in the human genome is most affected by genetic drift instead of natural
372 selection. Recent studies, however, have argued that an appreciable rate of adaptive substitution
373 has shaped genetic variation in humans (Enard et al. 2014) and that soft sweeps play the
374 dominant role in adaption in human evolution (Schrider and Kern 2017). Thus, even though
375 theory suggests that demographic inference should be more accurate in a less populous species
376 where the effects of natural selection are lessened, further study is needed to delimit the
377 parameter space in which RHH biases demographic inference.

378 The simulations in this study entail specific, important caveats. First, we simulated data
379 reflecting the highly recombining portion of the *Drosophila* genome. The inclusion of low
380 recombination regions would presumably exacerbate the effects of selective sweeps, HRI, and
381 background selection on genetic variation and hence demographic estimates. However, one
382 factor mitigating these forces is gene conversion, which we could not include in our simulations.
383 Our simulations also did not model selective constraint. The biases we observed, therefore,
384 would presumably be worse if analyzed sites had an excess of rare alleles due to deleterious
385 variants.

386 Our simulations only modeled complete sweeps from new mutations with invariant
387 selection coefficients. It has been argued that this is not necessarily the dominant adaptive model
388 in nature (Pritchard *et al.* 2010; Hernandez *et al.* 2011; Kern and Schrider 2017), so it is worth
389 considering how other models of natural selection may alter conclusions drawn in this study. If
390 there are two simultaneous soft sweeps, for instance, it is more likely that there exists a
391 haplotype with both favorable variants prior to selection starting, reducing the impact of HRI.
392 Likewise, because soft sweeps have more limited impacts on genetic variation (Pennings and
393 Hermisson 2006), their impact on demographic estimation should be less severe as well.

394 It is also important to note that complete sweeps may only be a single component of
395 Darwinian selection in nature. Other models of selection may impact genetic variation in fly
396 populations with less effect on divergence, including fluctuating selection (Mustonen and Lässig
397 2010; Bergland *et al.* 2014) and diminishing selection (Vy *et al.* 2017). Thus, depending on the
398 modes of positive selection that are prevalent in nature, the total impact of hitchhiking on genetic
399 variation and demographic inference may be greater or lesser than simulated here – underscoring
400 the need for further investigation of this topic.

401

402 **Data Access**

403 This study produced no empirical data. All scripts necessary to recapitulate the analyses

404 presented can be found at http://github.com/jeremy-lange.

405

Literature Cited

Andolfatto, P., 2005 Adaptive evolution of non-coding DNA in *Drosophila*. Nature 437(7062):1149-52.

Andolfatto, P., 2007 Hitchhiking effects of recurrent beneficial amino acid substitutions in the Drosophila melanogaster genome. Genome Res. 17: 1755–1762.

Begun, D. J., A. K. Holloway, K. Stevens, L. W. Hillier, Y.-P. Poh et al., 2007 Population genomics: whole-genome analysis of Polytene Chromosome Maps in Drosophila 1651 polymorphism and divergence in Drosophila simulans. PLoS Biol. 5: 2534–2559.

Bergland, A. O., E. L. Behrman, K. R. O'Brien, P. S. Schmidt, and D. A. Petrov, 2014 Genomic evidence of rapid and stable adaptive oscillations over seasonal time scales in Drosophila. PLoS Genet. 10: e1004775.

Braverman, J. M., R. R. Hudson, N. L. Kaplan, C. H. Langley and W. Stephan, 1995 The hitchhiking effect on the site frequency spectrum of DNA polymorphisms. Genetics 140: 783–796.

Chevin, L. M., and F. Hospital, 2008 Selective sweep at a quantitative trait locus in the presence of background genetic variation. Genetics 180: 1645–1660.

Comeron, J. P., 2014 Background selection as baseline for nucleotide variation across the Drosophila genome. PLoS Genet. 10: e1004434.

Duchen, P., D. Zivkovic, S. Hutter, W. Stephan, and S. Laurent, 2013 Demographic Inference Reveals African and European Admixture in the North American *Drosophila melanogaster* Population. Genetics 193(1): 291–301.

Enard, D., P. W. Messer, and D. A. Petrov, 2014 Genome-wide signals of positive selection in human evolution. Genome Res. 24: 885–895.

Ewing, G. B., and J. D. Jensen, 2016 The consequences of not accounting for background selection in demographic inference. Mol. Ecol. 25: 135–141.

Eyre-Walker, A., 2002 Changing effective population size and the McDonald-Kreitman test. Genetics 162: 2017–2024.

Gerrish, P., and R. Lenski, 1998 The fate of competing beneficial mutations in an asexual population. Genetica 102/103: 127–144.

Gutenkunst, R. N., R. D. Hernandez, S. H. Williamson, and C. D. Bustamante, 2009 Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. PLoS Genet. 5: e1000695.

Haller, B. C. and P. W. Messer, 2016 Slim 2: Flexible, interactive forward genetic simulations. Molecular Biology and Evolution 34: 230–240.

Halligan, D. L., and P. D. Keightley, 2006 Ubiquitous selective constraints in the Drosophila genome revealed by a genome-wide interspecies comparison. Genome Res. 16: 875–884.

Hernandez, R. D., J. L. Kelley, E. Elyashiv, S. C. Melton, A. Auton et al., 2011 Classic selective sweeps were rare in recent human evolution. Science 331: 920–924.

Hill, W. G., and A. Robertson, 1966 The effect of linkage on limits to artificial selection. Genet. Res. 8: 269.

Huang, W., R.F. Lyman, R.A. Lyman, M.A. Carbone, S.T. Harbison et al., 2016 Spontaneous mutations and the origin and maintenance of quantitative genetic variation. eLife 5:e14625.

Jensen, J. D., K. R. Thornton and P. Andolfatto, 2008 An approximate Bayesian estimator suggests strong, recurrent selective sweeps in Drosophila. PLoS Genet. 4: e1000198.

Kaplan, N. L., R. Hudson, and C. Langley, 1989 The "hitchhiking effect" revisited. Genetics 123: 887–899.

Karasov, T., S. W. Messer, and D. A. Petrov, 2010 Evidence that adaptation in Drosophila is not limited by mutation at single sites. PLoS Genet. 6: e1000924.

Kimura, M., 1962 On the probability of fixation of mutant genes in a population. Genetics 47(6): 713–719.

Lack, J. B., C. M. Cardeno, M. W. Crepeau, W. Taylor, R. B. Corbett-Detig 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: 1229–1241.

Langley, C. H., K. Stevens, C. Cardeno, Y. C. Lee, D. R. Schrider et al., 2012 Genomic variation in natural populations of Drosophila melanogaster. Genetics 192: 533–598.

Mackay, T. F. C., S. Richards, E. A. Stone, A. Barbadilla, J. F. Ayroles et al., 2012 The Drosophila melanogaster Genetic Reference Panel. Nature 482: 173–178.

McDonald, J. H., and M. Kreitman, 1991 Adaptive protein evolution at the Adh locus in Drosophila. Nature 351: 652–654.

Mustonen, V., and M. Lässig, 2010 Fitness flux and ubiquity of adaptive evolution. Proc. Natl. Acad. Sci. USA 107: 4248–4253.

Pennings, P. S., and J. Hermisson, 2006 Soft sweeps III: the signature of positive selection from recurrent mutation. PLoS Genet. 2: e186.

Pool, J. E., R. B. Corbett-Detig, R. P. Sugino, K. A. Stevens, C. M. Cardeno et al., 2012 Population genomics of Sub-Saharan Drosophila melanogaster: African diversity and non-African admixture. PLoS Genet. 8: e1003080.

Pritchard, J. K., J. K. Pickrell and G. Coop, 2010 The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. Curr. Biol. 20: R208–R215.

Schrider, D. R., D. Houle, M. Lynch, and M. W. Hahn, 2013 Rates and genomic consequences of spontaneous mutational events in *Drosophila melanogaster*. Genetics 194: 937–954.

Schrider, D. R., A. G. Shanku, and A. D. Kern, 2016 Effects of linked selective sweeps on demographic inference and model selection. Genetics 204: 1207–1223.

Schrider, D.R. and A.D. Kern, 2017 Soft Sweeps are the dominant mode of adaptation in the human genome. Molecular Biology and Evolution 34: 1863–1877.

Sella, G., D. A. Petrov, M. Przeworski, and P. Andolfatto, 2009 Pervasive natural selection in the

Drosophila genome? PLoS Genet. 5: e1000495.

Smith, N. G. C., and A. Eyre-Walker, 2002 Adaptive protein evolution in Drosophila. Nature 415: 1022–1024.

Stanley, C.E. and R.J. Kulathinal, 2016 Genomic signatures of domestication on neurogenetic genes in *Drosophila melanogaster*. BMC Evolutionary Biology 16:6.

Stephan, W., T. H. Wiehe, and M. W. Lenz, 1992 The effect of strongly selected substitutions on neutral polymorphism: analytical results based on diffusion theory. Theor. Popul. Biol. 41: 237–254.

Thornton, K., P. Andolfatto, 2006 Approximate Bayesian inference reveals evidence for a recent, severe bottleneck in a Netherlands population of Drosophila melanogaster. Genetics 172(3):1607-19.

Uricchio, L.H. and R.D. Hernandez, 2014 Robust forward simulations of recurrent hitchhiking. Genetics 197(1):221-36.

Vy, H.M.T., Y.J. Won, and Y. Kim, 2017 Multiple modes of positive selection shaping the patterns of incomplete selective sweeps over African populations of *Drosophila melanogaster*. Molecular Biology and Evolution.

Wiehe, T., and W. Stephan, 1993 Analysis of a genetic hitchhiking model, and its application to DNA polymorphism data from Drosophila melanogaster. Mol. Biol. Evol. 10: 842–854.

406 **Table 1.** Simulation parameters for the rare/strong and common/weak recurrent hitchhiking
407 models are provided. "Initial" values refer to parameters relevant to *Drosophila* populations,
408 while "rescaled" values were used directly in forward simulations. Parameter values for the
409 "retuned" common/weak model are also provided, in which simulated selection coefficients and
410 adaptive mutation rates were adjusted to match model expectations in the absence of Hill-
411 Robertson interference.

	Initial	**Rescaled**	**Initial**	**Rescaled**	**Initial**	**Rescaled**
	Rare strong sweeps		**Common weak sweeps**		**Retuned common weak sweeps**	
Population Size	2.50E+06	5.00E+04	2.50E+06	5.00E+04	2.50E+06	5.00E+04
Recombination Rate	2.50E-08	1.25E-06	2.50E-08	1.25E-06	2.50E-08	1.25E-06
Neutral Mutation Rate	3.27E-09	1.64E-07	3.27E-09	1.64E-07	3.27E-09	1.64E-07
Adaptive Mutation Rate	7.26E-17	5.92E-15	5.75E-12	2.88E-10	2.30E-11	1.15E-09
Selection Coefficient	1.10E-02	5.50E-01	1.20E-05	6.00E-04	5.50E-06	2.75E-04
Predicted Adaptive Substitution Rate	7.90E-12	3.95E-10	6.90E-10	3.45E-08	1.27E-09	6.32E-08

Table 2. The proportion of substitutions driven by positive selection (drivers, α), hitchhiking to fixation along with them (passengers), and reaching fixation by genetic drift (drifters) are shown. Here, the passenger class is defined narrowly as substitutions that fix simultaneously with beneficial mutations.

RHH Model	Driver	Passenger	Drifter
Rare/strong	0.00137	0.00151	0.99711
Common/weak	0.09993	0.03155	0.86862
Common/weak (retuned)	0.16683	0.04425	0.78892

Figure Legends

Figure 1: The expected and observed rate of adaptive substitution is shown for two hitchhiking models, with expected values reflecting theoretical predictions (Kimura 1962) based on the adaptive mutation rates and selection coefficients simulated. Left columns show our initial version of the common/weak RHH model (Andolfatto 2007). A strong deficit of observed versus expected driver substitutions for this model implied substantial HRI. To obtain the desired adaptive substitution rate, we therefore increased the adaptive mutation rate (right columns). The observed adaptive substitution rate was again much less than predicted in the absence of HRI, but it now aligned with the published model's rate.

Figure 2: The estimated proportion of substitutions driven by positive selection (α), as calculated from *Drosophila* genomic data, is shown for each chromosome arm and site functional class, as well as the genome-wide average across all arms.

Figure 3: Falsely-inferred population size changes based on RHH simulations are illustrated. Population size change estimates are shown for the trunk simulations (no true size changes). In the two-epoch scenario, we allow δαδι to infer a single population size change. In the three-epoch scenario, we allow δαδι to infer a size change and require it to return to its original effective population size.

Figure 4: Demographic parameter estimates are shown for bottleneck simulations based on the true model shown on the left. Simulations with RHH showed relatively little bias for the duration and time since the bottleneck, but moderate bias for the time since recovery.

Figure 5: Demographic parameter estimates for simulations of the depicted isolation- migration model are shown. Some bias was observed for this relatively ancient population split time even under neutrality. However, differences from neutral estimates for RHH cases with shared sweeps or local sweeps were consistent with the effects of decreased or increased genetic differentiation, respectively. Migration rate estimates were relatively unaffected by RHH.

412

Figure 1: The expected and observed rate of adaptive substitution is shown for two hitchhiking models, with expected values reflecting theoretical predictions (Kimura 1962) based on the adaptive mutation rates and selection coefficients simulated. Left columns show our initial version of the common/weak RHH model (Andolfatto 2007). A strong deficit of observed versus expected driver substitutions for this model implied substantial HRI. To obtain the desired adaptive substitution rate, we therefore increased the adaptive mutation rate (right columns). The observed adaptive substitution rate was again much less than predicted in the absence of HRI, but it now aligned with the published model's rate.

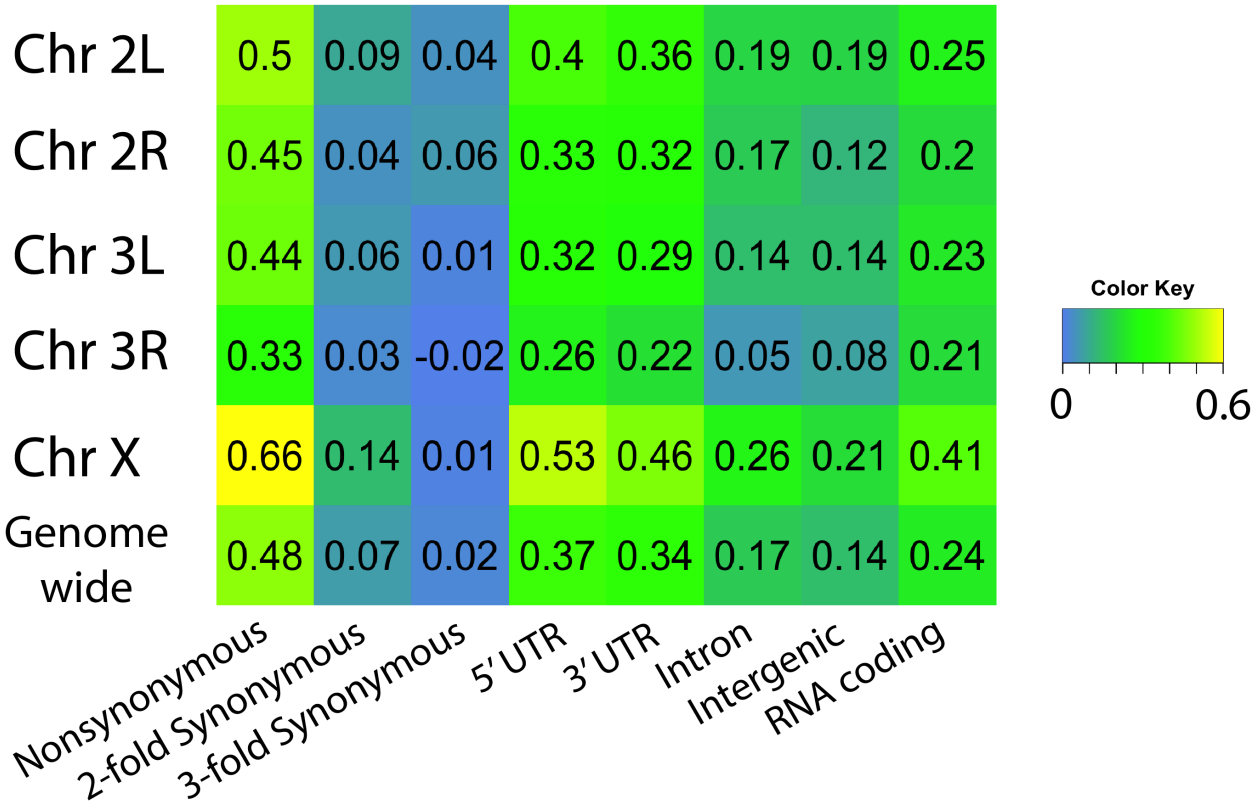

Figure 2: The estimated proportion of substitutions driven by positive selection (α), as calculated from *Drosophila* genomic data, is shown for each chromosome arm and site functional class, as well as the genome-wide average across all arms.

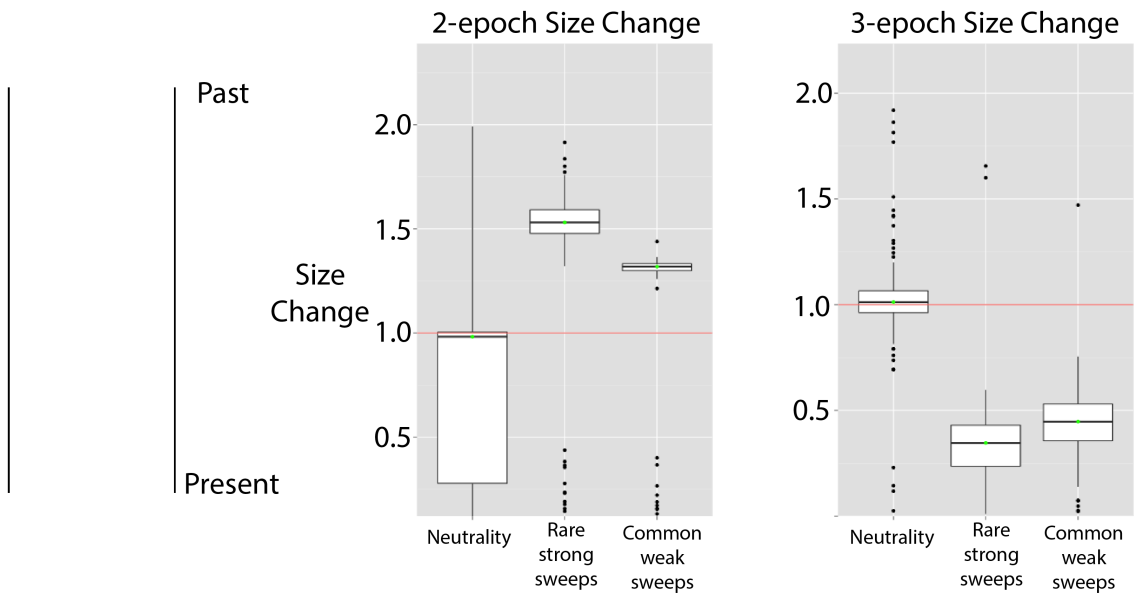

Figure 3: Falsely-inferred population size changes based on RHH simulations are illustrated. Population size change estimates are shown for the trunk simulations (no true size changes). In the two-epoch scenario, we allow δαδι to infer a single population size change. In the three-epoch scenario, we allow δαδι to infer a size change and require it to return to its original effective population size.

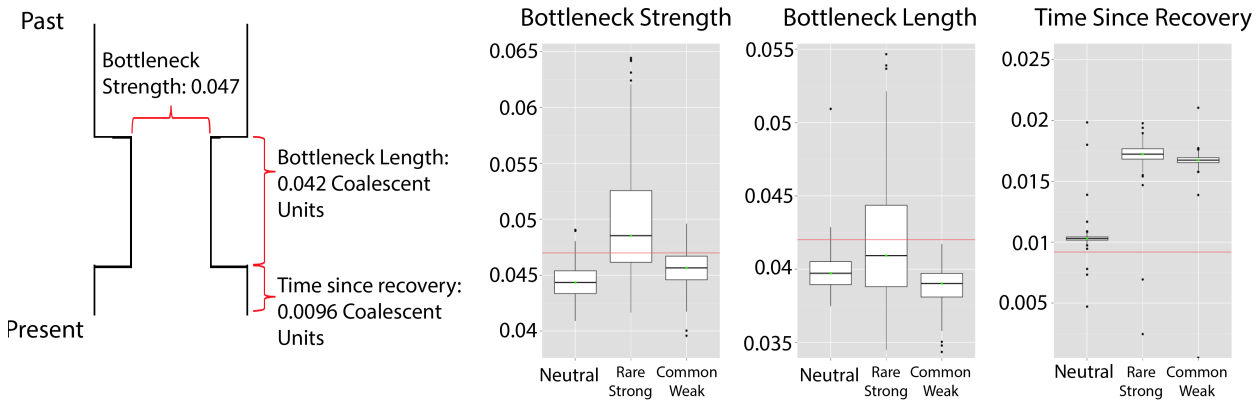

Figure 4: Demographic parameter estimates are shown for bottleneck simulations based on the true model shown on the left. Simulations with RHH showed relatively little bias for the duration and time since the bottleneck, but moderate bias for the time since recovery.

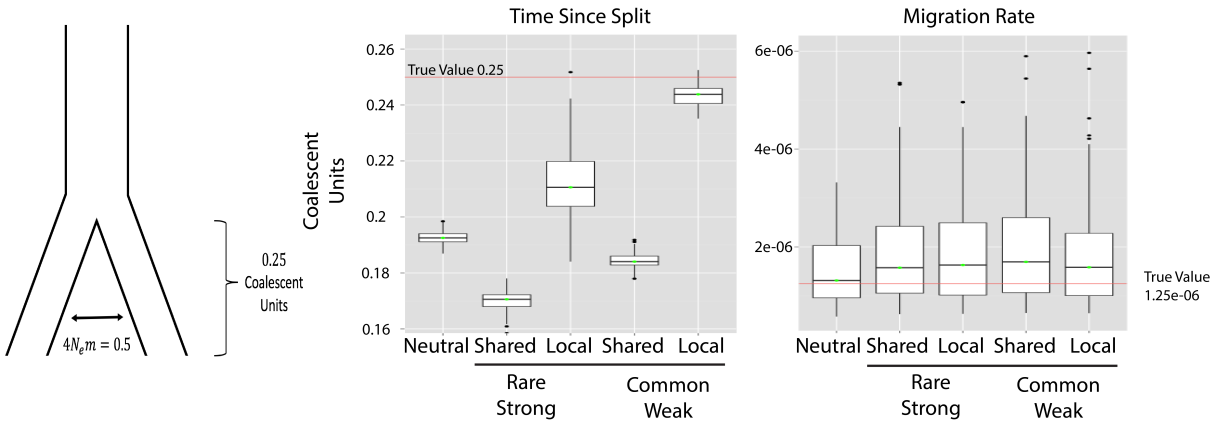

Figure 5: Demographic parameter estimates for simulations of the depicted isolation- migration model are shown. Some bias was observed for this relatively ancient population split time even under neutrality. However, differences from neutral estimates for RHH cases with shared sweeps or local sweeps were consistent with the effects of decreased or increased genetic differentiation, respectively. Migration rate estimates were relatively unaffected by RHH.