

MOLECULAR ECOLOGY

Background selection and F_{ST} : consequences for detecting local adaptation

Journal:	<i>Molecular Ecology</i>
Manuscript ID	MEC-19-0061
Manuscript Type:	Original Article
Date Submitted by the Author:	17-Jan-2019
Complete List of Authors:	Matthey-Doret, Remi; University Of British Columbia, Department of Zoology Whitlock, Michael; University of British Columbia, Department of Zoology
Keywords:	Adaptation, Evolutionary Theory, Population Genetics - Theoretical, Population Structure, Local Adaptation

SCHOLARONE™
Manuscripts

1

2 **Title:** Background selection and F_{ST} : consequences for detecting local adaptation

3

4 **Authors:** Remi Matthey-Doret¹ and Michael C. Whitlock

5

6 **Affiliation:** Department of Zoology and Biodiversity Research Centre, University of
7 British Columbia, Vancouver, British Columbia V6T 1Z4, Canada

8

9 ¹**Corresponding author:** matthey@zoology.ubc.ca

10

11

12 Abstract

13 Background selection is a process whereby recurrent deleterious mutations cause a
14 decrease in the effective population size and genetic diversity at linked loci. Several
15 authors have suggested that variation in the intensity of background selection could
16 cause variation in F_{ST} across the genome, which could confound signals of local
17 adaptation in genome scans. We performed realistic simulations of DNA sequences,
18 using recombination maps from humans and sticklebacks, to investigate how
19 variation in the intensity of background selection affects F_{ST} and other statistics of
20 population differentiation in sexual, outcrossing species. We show that, in
21 populations connected by gene flow, Weir & Cockerham's (1984) estimator of F_{ST} is
22 largely insensitive to locus-to-locus variation in the intensity of background
23 selection. Unlike F_{ST} , however, d_{XY} is negatively correlated with background
24 selection. Background selection does not greatly affect the false positive rate in F_{ST}
25 outlier studies. Overall, our study indicates that background selection will not
26 greatly interfere with finding the variants responsible for local adaptation.

27

28 **Introduction**

29 Maynard Smith & Haigh (1974) recognized the influence of selection on linked
30 neutral sites, proposing that strong positive selection could reduce genetic diversity
31 at nearby sites. This process is now referred to as a 'selective sweep'. Much later,
32 Charlesworth et al. (1993) proposed that deleterious mutations could also affect
33 genetic diversity at nearby sites, because some haplotypes would be removed from
34 the population as selection acts against linked deleterious alleles. They named this
35 process background selection (BGS). Both selective sweeps and background
36 selection affect genetic diversity; they both reduce the effective population size and
37 distort the site frequency spectrum (SFS) of linked loci. Empirical evidence of a
38 positive correlation between genetic diversity and recombination rate has been
39 reported in several species (Cutter and Payseur, 2013), including *Drosophila*
40 *melanogaster* (Begin & Aquadro, 1992; Elyashiv et al., 2016), humans (Spencer et
41 al., 2006), collared flycatchers, hooded crows and Darwin's finches (Dutoit et al.,
42 2017; see also Vijay et al., 2017).

43 BGS is also expected to affect F_{ST} (Charlesworth et al., 1997; Cutter & Payseur, 2013;
44 Cruickshank & Hahn, 2014; Hoban et al., 2016). The negative relationship between
45 effective population size N_e and F_{ST} is captured in Wright's classical infinite island
46 result; $F_{ST} = \frac{1}{1 + 4Ne(m + \mu)}$ (Wright, 1943), where m is the migration rate and μ is the
47 mutation rate. One might therefore expect that loci under stronger BGS would show
48 higher F_{ST} .

Charlesworth (1998), i.e. the G_{ST} equivalent suggested by Nii (1973)

I do not see why this is done, as if this is simpler formula on l. 46.

Many authors have also argued that, because BGS reduces the within-population diversity, it should lead to high F_{ST} (Cutter & Payseur, 2013; Cruickshank & Hahn, 2014; Hoban et al., 2016). Expressed in terms of heterozygosities, $F_{ST} = \frac{H_T - H_S}{H_T}$ = $1 - \frac{H_S}{H_T}$, where H_T is the expected heterozygosity in the entire population and H_S is the average expected heterozygosity within subpopulations (H_S and H_T are also sometimes called π_S and π_T ; e.g. Charlesworth, 1998). All else being equal, a decrease of H_S would indeed lead to an increase of F_{ST} . However, all else is not equal; H_T is also affected by BGS (Charlesworth et al., 1997). Therefore in order to understand the effects of BGS on F_{ST} , we must understand the relative impact of BGS on both H_S and H_T .

$$H_T = \frac{1}{2} H_S + \frac{1}{2} H_B \\ = \frac{1}{2} H_S + \frac{1}{2} H_S + \frac{1}{4m} \mu = H_S + \frac{1}{4m} \mu$$

Performing numerical simulations, Charlesworth et al. (1997) report that BGS reduces the within population heterozygosity H_S slightly more than it reduces the total heterozygosity H_T , causing a net increase in F_{ST} . The effect on F_{ST} reported is quite substantial, but, importantly, their simulations were not meant to be realistic. The authors highlighted their goal in the methods:

"The simulations were intended to show the qualitative effects of the various forces studied [...], so we did not choose biologically plausible values [...]. Rather, we used values that would produce clear-cut effects".

For example, talking about their choice for the deleterious mutation rate of 8×10^{-4} per site:

Slater (1993) : $F_{ST} = \frac{\bar{\tau}_1 - \bar{\tau}_0}{\bar{\tau}_0 + \bar{\tau}_1} = \frac{\pi_D}{\pi_S + \pi_B} = \frac{\pi_D}{\frac{2\pi_S + \pi_D}{4}}$

$$\pi_B = \pi_S + \pi_D$$

69 "This unrealistically high value was used in order for background selection to
70 produce large effects [...]"

71 Much of the literature on the effect of BGS on F_{ST} is based on the results in

72 Charlesworth et al. (1997), even though they only intended to show proof of concept

73 (see also Zeng & Charlesworth, 2011 and Zeng & Corcoran, 2015). They did not

74 attempt to estimate how strong of an effect BGS has on F_{ST} in real genomes.

| provide a

} show to

} resolve

this co-

nundrum

75 The intensity of BGS varies throughout the genome as a consequence of variation in

76 recombination rate, selection pressures and mutation rates. Therefore, if BGS

77 significantly affects F_{ST} , we should expect that baseline F_{ST} to vary throughout the

78 genome. It is important to distinguish two separate questions when discussing the

79 effect of BGS on F_{ST} ; 1) How does BGS affect the average genome-wide F_{ST} ? and 2)

80 **How does locus-to-locus variation in the intensity of BGS affect locus-to-locus**

| only

} partially

agree.

| There is

an inter-

mediate

scale

| that may be

just as

important for

detecting local

selection or

selective

| sweeps.

| directly

81 variation in F_{ST} ? The second question is of particular interest to those trying to

82 identify loci under positive selection (local selection or selective sweep). Locus-to-

83 locus variation in F_{ST} due to BGS potentially could be confounded with the F_{ST} peaks

84 created by positive selection. **In this paper, we focus on this second question.**

85 The identification of loci involved in local adaptation is often performed via F_{ST}

86 outlier tests (Lotterhos & Whitlock, 2014; Hoban et al., 2016). Other tests exist to

87 identify highly divergent loci such as cross-population extended haplotype

88 homozygosity (XP-EHH; Sabeti et al., 2007), comparative haplotype identity (Lange

89 & Pool, 2016) and cross-population composite likelihood ratio (XP-CLR; Chen et al.,

*speaking, it is not clear what exactly is meant by
locus-to-locus variation in F_{ST} .*

90 2010). F_{ST} outlier tests, such as FDist2 (Beaumont & Nichols, 1996), BayeScan (Foll
91 & Gaggiotti, 2008) or FLK (Bonhomme et al., 2010), look for genomic regions
92 showing particularly high F_{ST} values to find candidates for local adaptation. If BGS
93 can affect F_{ST} unevenly across the genome, then regions with a high intensity of BGS
94 could potentially have high F_{ST} values that could be confounded with the pattern
95 caused by local selection (Charlesworth et al., 1997; Cruickshank & Hahn, 2014).

96 BGS could therefore inflate the false positive rate when trying to detect loci under
97 local selection.

98 The potential confounding effect of BGS on signals of local adaptation has led to an
99 intense effort trying to find solutions to this problem (Bank et al., 2014; Huber et al.,
100 2016; Aeschenbacher et al., 2017). Many authors have understood from
101 Cruickshank and Hahn (2014) that d_{XY} should be used instead F_{ST} in outlier tests
102 (e.g. McGee et al., 2015; Yeaman, 2015; Whitlock & Lotterhos, 2015; Brousseau et al.,
103 2016; Picq et al., 2016; Payseur & Rieseberg, 2016; Hoban et al., 2016; Vijay et al.,
104 2017; see also Nachman & Payseur, 2012). F_{ST} is a measure of population divergence
105 relative to the total genetic diversity, while d_{XY} is an absolute measure of population
106 divergence defined as the probability of non-identity by descent of two alleles
107 drawn in the two different populations averaged over all loci (Nei, 1987; Nei, 1987
108 originally called it D_{XY} but, here, we follow Cruickshank and Hahn's, 2014
109 terminology by calling it d_{XY}). The argument is that because F_{ST} is a measure of
110 divergence relative to the genetic diversity and d_{XY} an absolute measure of
111 divergence and because BGS reduces genetic diversity (Cruickshank & Hahn, 2014;

112 Hoban et al., 2016), then BGS must affect F_{ST} but not d_{XY} , a claim that we will
113 investigate in this paper.

114 Whether BGS can affect genome-wide F_{ST} under some conditions is not in doubt
115 (Charlesworth et al., 1997), but whether locus-to-locus variation in the intensity of
116 BGS present in natural populations substantially affects variation in F_{ST} throughout
117 the genome is very much unknown. Empirically speaking, it has been very difficult
118 to measure how much of the genome-wide variation in genetic diversity is caused by
119 BGS, as opposed to selective sweeps or variation in mutation rates (Cutter &
120 Payseur, 2013; see also attempts in humans by Cai et al., 2009, McVicker et al. 2009
121 and Elyashiv et al., 2016). We are therefore in need of realistic simulations that can
122 give us more insight into how BGS affects genetic diversity among populations and
123 how it affects the statistics of population divergence.

124 In this article, we investigate the effect of BGS in structured populations with
125 realistic numerical simulations. Our two main goals are 1) to quantify the impact of
126 locus-to-locus variation in the intensity of BGS on F_{ST} (Weir & Cockerham, 1984) and
127 d_{XY} (Nei, 1987) and 2) to determine whether BGS inflates the false positive rate of
128 F_{ST} outlier tests.

129 Methods

130 Our goal is to perform biologically plausible simulations of the local genomic effects
131 of background selection. BGS is expected to vary with strength of selection (itself
*This should
be more
precise*
7

132 affected by gene density), mutation rate and recombination rate across the genome.

133 We used data from real genomes to simulate realistic covariation in recombination

134 rates and gene densities. We chose to base our simulated genomes on two eukaryote

135 recombination maps, sticklebacks and humans, because these two species have

136 attracted a lot of attention in studies of local adaptation and because sticklebacks have

137 a variance in recombination rate which is almost 15 times higher than humans (data

138 not shown), allowing us to test vastly different types of eukaryotic genomes. The

139 recombination rate variation in humans is extremely fine scale, but it presents the

140 potential issue that it is estimated from linkage disequilibrium data. As selection

141 causes linkage disequilibrium to increase, estimates of recombination rate at

142 regions under strong selection may be under-estimated, which might bias the

143 simulated variance in the intensity of BGS. Although the recombination map for

144 stickleback is much less fine-scaled, the estimates are less likely to be biased as they

145 are computed from pedigrees. (2 But they may not reflect long-term rates)

146 Our simulations are forward in time and were performed using the simulation

147 platform SimBit version 3.69. We simulated non-overlapping generations,

148 hermaphroditic, diploid individuals and random mating within patches. Selection

149 occurred before dispersal. The code and user manual are available at

150 <https://github.com/RemiMattheyDoret/SimBit>. The rational for using a new

151 simulation platform is because all existing simulation platforms today were too slow

152 for our needs. (See Appendix A.) To double check our results, we also ran some

153 simulations with SLiM (Haller & Messer, 2017) and Nemo (Guillaume & Rougemont,

| of

{ use the
non-LD
based one

(which corne-
lates well
with the
ID-based
one)

one)

*
Appendix A

154 2006) (see Appendix A), confirming that we get consistent distributions of genetic | obtain
155 diversity and of F_{ST} among simulations. For simulations with SLiM and Nemo,
156 independent simulations were parallelized with GNU parallel (Tange, 2011).

157 *Genetics*

158 For each simulation, we randomly sampled a sequence of about 10 cM from either | Why
159 the stickleback (*Gasterosteus aculeatus*) genome or the human genome (see "about"
160 treatments below) and used this genomic location to determine the recombination Better be
161 map and exon locations for a simulation replicate. For the stickleback genome, we more
162 used the gene map and recombination map from Roesti et al. (2013). Ensembl- precise?
163 retrieved gene annotations were obtained from Marius Roesti. For the human
164 genome, we used the recombination map from The International HapMap
165 Consortium (2007) and the gene positions from NCBI and positions of regulatory
166 sequences on Ensembl (Zerbino et al., 2017). We excluded sex chromosomes to
167 avoid complications with haploid parts of the genomes. As estimates of mutation
168 rate variation throughout the genome are very limited, we assumed that the haploid
169 mutation rate varies from site to site following an exponential distribution with
170 mean of 2.5×10^{-8} per generation (mean estimate from Nachman & Crowell, 2000). } Was there
| a control
| with a
| uniform
| mutation
| rate?

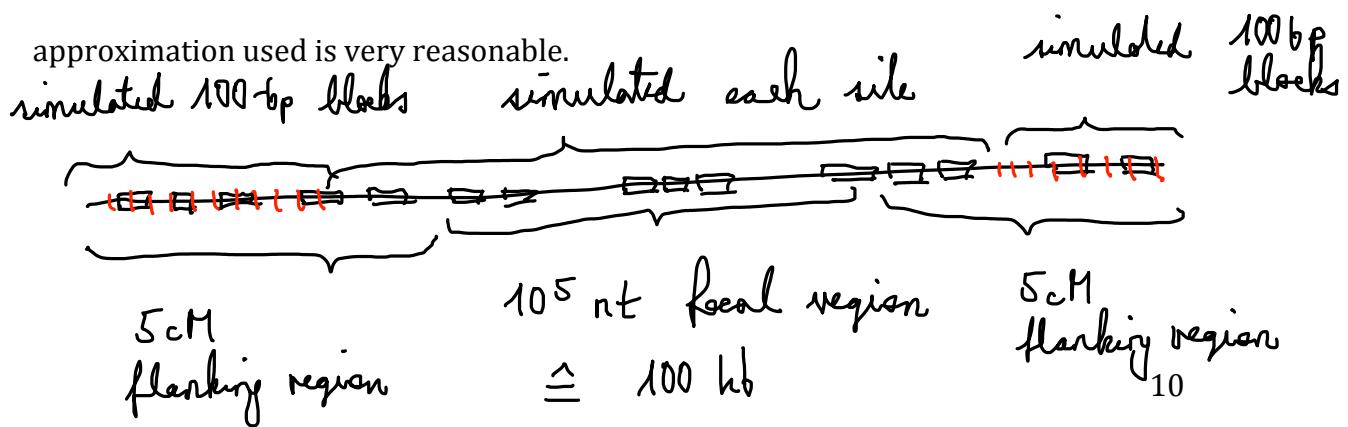
171 More specifically, we first randomly sampled a sequence of 10^5 nucleotides, which
172 we will refer to as the focal region. All of the statistics (defined under the section
173 Statistics below) are calculated only on the focal region of each simulation. Yes.

174 Nucleotides that occur in locations determined to be exons in the sampled genomic

density of selection targets ~ exon density in focal region

175 map are subject to selection (see *Selection*), while all other nucleotides are assumed
 176 to be neutral. The focal region itself contained on average ~0.44 genes for the
 177 human genome and ~3.15 genes for the stickleback genome.

178 We simulated a 5 cM region on each side of the focal region (resulting in a window
 179 of 10 cM plus the map distance covered by the specific focal region of 10^5 sites) in
 180 order to capture the local effects of background selection. In these 10 cM flanking
 181 regions, we only tracked exons. In the nearest 1 cM on each side of the focal region,
 182 as well within the focal region, we individually simulated each nucleotide as a bi-
 183 allelic locus. On the remaining outer 4 cM, to improve the speed and RAM usage of
 184 the process, we tracked the number of mutations in blocks of up to 100 nucleotides.
 185 For these blocks, we tracked only the number of mutations but not their location
 186 within the block. Ignoring recombination within a block likely had little effect on the
 187 results because the average recombination distance between the first and last site of
 188 a block is of the order of 10^{-6} cM. The expected number of segregating sites within a
 189 block is $4N\mu\sum_{i=1}^{2N-1}(1/i)$, which for a mutation rate per block of 10^{-6} and a
 190 population size of $N = 10,000$ is ~0.42. The probabilities of having more than one
 191 mutation and more than two mutations (based on a Poisson approximation) are
 192 therefore only approximately 6.7% and 0.9%, respectively. Overall, the level of
 193 approximation used is very reasonable.



194 Selection

195 As we are interested in the effect of BGS, we modelled the effects of purifying
 196 selection against novel deleterious mutations. Each nucleotide in the exons (and
 197 regulatory sequences for the human genetic map) is subject to purifying selection
 198 with a selection coefficient against mutant alleles determined by a gamma
 199 distribution described below. For focal regions that include exons, statistics are
 200 computed over a sequence that is at least partially under direct purifying selection.

| A

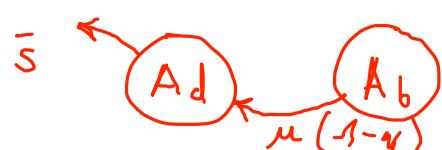
201 To create variance in selection pressures throughout the genome, each exon (and
 202 regulatory sequence for the human genetic map) has its own gamma distribution of
 203 heterozygous selection coefficients s . The mean and variance of these gamma
 204 distributions are drawn from a bivariate uniform distribution with correlation
 205 coefficient of 0.5 (so that when the mean is high, so is the variance) bounded
 206 between 10^{-8} and 0.2 for both the mean and the variance. These bounds were
 207 inspired by the methodology used in Gilbert et al. (2017). The gamma distributions
 208 are bounded to one. Figure S1 shows the overall distribution of selection coefficient
 209 s , with 2% of mutations being lethal and an average deleterious selection coefficient
 210 for the non-lethal mutations of 0.07. In the treatment *Low selection pressure* (see
 211 treatments below), the upper bounds for the mean and variance of the gamma
 212 distributions were set to 0.1 instead of 0.2. To improve the performance of our
 213 simulations, we assumed multiplicative fitness interactions among alleles, where the

Introducing
variance in
selection
pressure along
the
genome.

* Fig.S1

multiplicative
fitness
interactions

$$\bar{s}_d = 0.07 \quad \bar{\mu} = 2.5 \times 10^{-8}$$



$$\hat{q} = -(\bar{s} - \mu) / \mu = 1 - \frac{\bar{s}}{\mu}$$

214 fitness of heterozygotes at locus i is $1 - s_i$ and the fitness of the double mutant is $(1 - s_i)^2$. Any mutation changes the state of the locus to the other possible allele.

|
| 2 diminishing
| neutral
| epistasis

216 As a consequence of our parameter choices, our genome-wide deleterious mutation
217 rate was about 1.6 in sticklebacks and about 3 in humans. 9.8% of the stickleback

218 genome and 2.7% of the human genome was under purifying selection. For
219 comparison, the genome-wide deleterious mutation rate is estimated at 2.2 in
220 humans (Keightley, 2012) and 0.44 in rodents (Keightley & Gaffney, 2003). To our
221 knowledge, there is currently no such estimation for sticklebacks. Note however
222 that the above estimates cannot reliably detect mutations that are quasi-neutral (s
223 $<< 1/2N$). By our distribution of selection coefficients, 49% of all deleterious
224 mutations have a heterozygote selection coefficient lower than $1/2N_e$ when $N_e =$
225 1,000 (42% when $N_e = 10,000$). The fraction of selection coefficients that are of
226 intermediate effect (between $1/2 N_e$ and $10/2 N_e$) is 10% when $N_e = 1,000$ (7%
227 when $N_e = 10,000$).

} surround
2N by
brackets
for
clarity

228 It is worth noting that, in rodents, about half of the deleterious mutations occur in
229 non-coding sequences (Keightley & Gaffney, 2003). Our simulations using human
230 genetic map had all exons and all regulatory sequences under purifying selection.
231 With our simulations based on the stickleback genome, however, only exons were
232 under purifying selection. It is therefore possible that we would have over-
233 estimated the deleterious mutation rate in gene-rich regions and under-estimated
234 the deleterious mutation rate in other regions, especially in stickleback. This would

... and
we have
shown
that up
to 85%
of the
human

genome might be
affected by BGS

235 artificially increase the locus-to-locus variation in the intensity of BGS in our
 236 simulations, which is conservative to our conclusions.

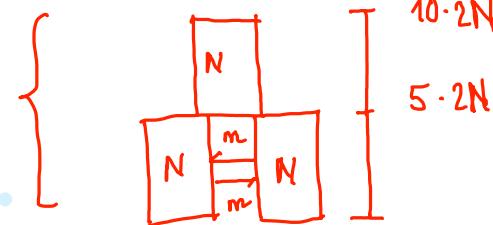
Constraining deleterious mutations does not only affect the locus-to-locus variation, but also the correlation between N and μ_{bp} !

237 *Demography*

238 In all simulations, we started with a burn-in phase with a single population of N
 239 diploid individuals, lasting $5 \times 2N$ generations. The population was then split into
 240 two populations of N individuals each with a migration rate between them equal to
 241 m . After the burn-in phase, each simulation was run for $5 \times 2N$ more generations for
 242 a total of $10 \times 2N$ generations.

243 *Treatments*

BGS operating in both periods



244 We explored the presence and absence of deleterious mutations over two patch sizes, three migration rates, two genomes, and three selection scenarios. We considered a basic design and explored variations from this design. The basic design had a population size per patch of $N = 1000$, a migration rate of $m = 0.005$ and used the stickleback genome for its recombination map and gene positions. As deviations from this basic design, we explored modification of every variable, one variable at a time. The Large N treatment has $N = 10000$. The Human Genetic Map treatment uses the human genome for gene positions, regulatory sequences and recombination map. The treatments No Migration and High Migration have migration rates of $m = 0$ and $m = 0.05$, respectively. The Constant μ treatment assumes that all sites have a

$$Nm = 5$$

Note: The design might not conform to the assumptions of the BGS model assumed by CN97 ("shaky BGS")

$$s = 0.02 : 1 - 0.02 = 0.98 ; \frac{(1-0.02)^2}{1-2 \cdot 0.02 + 0.0004} > 0.96$$

The CNC97

scenario deviates from the multiplicative scheme!

254 mutation rate of 2.5×10^{-8} per generation. The **Low selection pressure treatment**
 255 simulates lower selection coefficients (see section *Selection* above).

256 To test the robustness of our results and because it may be relevant for inversions,
 257 we also performed simulations where the recombination rate for the entire 10cM
 258 region was set to zero. As a check against previous work, we qualitatively replicated
 259 the results of Charlesworth et al. (1997) by performing simulations with similar
 260 assumptions as they used. We named this **treatment CNC97**. In our *CNC97*
 261 simulations, $N=2000$, $m=0.001$, and 1000 loci were all equally spaced at 0.1 cM apart
 262 from each other with constant selection pressure with heterozygotes having fitness
 263 of 0.98 and double homozygotes fitness of 0.9 and constant mutation rate $\mu =$
 264 0.0004. We performed further checks against previous works that are presented in
 265 Appendix A. A full list of all treatments can be found in table 1.

266 In all treatments (except *Large N*), we performed 4000 simulations; 2000
 267 simulations with BGS and 2000 simulations without selection (where all mutations
 268 were neutral). For *Large N*, simulations took more memory and more CPU time. We
 269 therefore could only perform 2000 simulations for *Large N*; 1000 simulations with
 270 background selection and 1000 simulations without selection.

271 We set the generation 0 at the time of the split. The state of each population was
 272 recorded at the end of the burn-in period (generation -1) and at generations $0.001 \times$
 273 $2N$, $0.05 \times 2N$, $0.158 \times 2N$, $1.581 \times 2N$ and $5 \times 2N$ after the split. For $N=1000$, the
 274 sampled generations are therefore -1, 2, 100, 316, 3162 and 10000.

275 Predicted intensity of Background Selection

276 In order to investigate the locus-to-locus correlation between the predicted
 277 intensity of BGS and various statistics, we computed B , a statistic that approximates
 278 the expected ratio of the coalescent time with background selection over the
 279 coalescent time without background selection ($B = \frac{T_{BGS}}{T_{neutral}}$). B quantifies how strong
 280 BGS is expected to be for a given simulation (Nordborg et al. 1996). A B value of 0.8
 281 means that BGS has caused a drop of genetic diversity of 20% compared to a
 282 theoretical absence of BGS. Lower B values indicate stronger BGS.

283 Both Hudson & Kaplan (1995) and Nordborg et al. (1996) have derived the
 284 following theoretical expectation for B .

$$285 B = \exp\left(-\sum_i \frac{u_i s_i}{(s_i + r_i)^2}\right) = \frac{N_e^{(BGS)}}{N_e}$$

$F_{ST} = \frac{1}{1 + 8N_e^{(BGS)} m}$
 L 16 in the case of Nei's (1973) G_{ST} analog

What are the underlying assumptions?

286 where r_i is the recombination rate between the focal site and the i^{th} site under
 287 selection, and s_i is the heterozygous selection coefficient at that site, and u_i is the
 288 (haploid) mutation rate at the i^{th} site. By this formula, B is bounded between 0 and 1,
 289 where 1 means no BGS at all and low values of B mean strong BGS. We computed B

290 for all sites in the focal region and report the average B for the region.

Clarify: Were all mutations taken into account? See l. 180.

291 For the stickleback genome, B values ranged from 0.03 to 0.99 with a mean of 0.84
 292 (Figure S2). For the human genome, B values ranged from 0.20 to 1.0 with a mean at

#1 - B is computed over the focal region only.

L Does the flanking region have no effect?

- alternatively, B might vary over a smaller scale.

293 0.91. In the *No Recombination* treatment, B values range from 10^{-10} to 0.71 with a
294 mean of 0.07.

295 Excluding the treatments *No recombination* and *CNC97*, we observed that
296 simulations with BGS have a genetic diversity (whether H_T or H_S ; H_T data not shown)
297 6% to 25% lower than simulations without BGS. Messer & Petrov (2013) simulated
298 a panmictic population, looking at a sequence of similar length inspired from a gene-
299 rich region of the human genome, and reported a similar decrease in genetic
300 diversity. Under the *No Recombination* treatment, this average reduction of genetic
301 diversity due to BGS is 53%. In empirical studies, linked selection are estimated to
302 reduce genetic diversity by up to at 6% according to Cai et al. (2009) or 19-26% | ok
303 according to McVicker et al. (2009) in humans. In *Drosophila melanogaster*, where
304 gene density is higher, the reduction in genetic diversity due to linked selection is
305 estimated at 36% when using Kim & Stephan (2000)'s methodology and is
306 estimated at 71% reduction using a composite likelihood approach (Elyashiv et al.,
307 2016). In mice, BGS alone cannot explain fully the reduction in genetic diversity at
308 low recombination sites, and selective sweeps due to positive selection are
309 responsible for the majority of the reduction in diversity due to linked selection
310 (Booker & Keightley, 2018). It is worth noting that, because we were interested in
311 the locus-to-locus variation of various statistics in response to varying intensity of
312 BGS, we did not simulate a whole genome worth of BGS and hence the overall
313 reduction in genetic diversity that we observe should not be understood as a
314 genome-wide effect of BGS.

$$\text{From Charlesworth (1998) Eq. (4c)} : F_{ST} = \frac{1}{4Nm \frac{n^2}{(n-1)^2} + 1} \quad \left. \right\} N_e (1973)$$

$$\Rightarrow (4Nm \frac{n^2}{(n-1)^2} + 1) F_{ST} = 1$$

315 F_{ST} outlier tests
 $\Rightarrow 4Nm \frac{n^2}{(n-1)^2} F_{ST} = 1 - F_{ST} \Rightarrow m = (1 - F_{ST}) / [4 \frac{n^2}{(n-1)^2} N F_{ST}] \quad \checkmark$

316 In order to know the effect of BGS on outlier tests of local adaptation, we used a

317 variant of FDist2 (Beaumont & Nichols, 1996). We chose FDist2 because it is a
 318 simple and fast method for which the assumptions of the test match well to the
 319 demographic scenario simulated here. Because the program FDist2 is not available
 320 through the command line, we rewrote the FDist2 algorithm in R and C++. Source
 321 code can be found at <https://github.com/RemiMattheyDoret/Fdist2>.

322 Our FDist2 procedure is as follows; first, we estimated the migration rate from the

323 average F_{ST} of the specific set of simulations considered ($m = \frac{1 - F_{ST}}{4 \left(\frac{d}{d-1}\right)^2 N F_{ST}}$);

324 Charlesworth, 1998) and then running 50000 simulations each lasting for 50 times

325 the half-life to reach equilibrium F_{ST} given the estimated migration rate (Whitlock,

326 1992). For each SNP, we then selected the subset of FDist2 simulations for which

327 allelic diversity was less than 0.02 away from the allelic diversity of the SNP of

328 interest. The P -value is computed as the fraction of FDist2 simulations within this

329 subset having a higher F_{ST} than the one we observed. The false positive rate is then

330 defined as the fraction of neutral SNPs for which the P -value is lower than a given α

331 value, using $\alpha = 0.05$. We confirmed that the results were similar for other α values.

332 For the outlier tests, to avoid issues of pseudo-replication, we considered only a

333 single SNP (randomly sampled from the focal region) per simulation whose minor

334 allele frequency is greater than 0.05. Then, we randomly assembled SNPs from a

\bullet This equation assumes drift-migration equil. !

We've only F_{ST} values

after $5 \times 2N$ gen.

and?

Did you

use N or

N_e ?

One should

use N_e ;

if N is

used, m

is likely

deflated,

and the

P -value

inflated.

This inflation

applies particularly to the

BGS

simulations

17

335 given treatment into groups of 500 SNPs to create the data file for FDist2. We have
336 4000 simulations (2000 with BGS and 2000 without BGS) per treatment (*Large N* is
337 an exception with only 2000 simulations total), which allowed 8 independent false
338 positive rate estimates per treatment (4 estimates with BGS and 4 without BGS). In
339 each treatment, we tested for different false positive rate with and without BGS with
340 both a Welch's *t*-test and a Wilcoxon test.

341 *Statistics*

342 F_{ST} and d_{XY} are both measures of population divergence. In the literature there are
343 several definitions of F_{ST} , and we also found potential misunderstanding about how
344 d_{XY} is computed. We want to clarify here these definitions and what we mean when
345 we use the terms F_{ST} and d_{XY} .

check if this
is correct.

346 There are two main estimators of F_{ST} in the literature; G_{ST} (Nei, 1973) and θ (Weir &
347 Cockerham, 1984). In this article, we focus on θ as an estimator of F_{ST} (Weir &
348 Cockerham, 1984). There are also two methods of averaging F_{ST} over several loci.
349 The first method is to simply take an arithmetic mean over all loci. The second
350 method consists at calculating the sum of the numerator of θ over all loci and
351 dividing it by the sum of the denominator of θ over all loci. Weir and Cockerham
352 (1984) showed that this second averaging approach has lower bias than the simple
353 arithmetic mean. We will refer to the first method as the "average of ratios" and to
354 the second method as "ratio of the averages" (Reynolds et al. 1983; Weir &
355 Cockerham, 1984). In this article, we use F_{ST} as calculated by "ratio of the averages",

The
formulae
given above
for F_{ST}
seem to
be those
pertaining
to the
 G_{ST} -type
 F_{ST} !

I wonder if the focal region is so large
that variation in F_{ST} due to BGS is smoothed out.

356 as advised by Weir and Cockerham (1984). To illustrate the effects of BGS on the
 357 biased estimator of F_{ST} , we also computed F_{ST} as a simple arithmetic mean ("average
 358 of the ratio"), and we will designate this statistic with a subscript $F_{ST, \text{average of ratios}}$.
 359 d_{XY} is a measure of genetic divergence between two populations X and Y. Nei (1987)
 360 defined d_{XY} as

$$361 \quad d_{XY} = \frac{\sum_{l=1}^L \left(1 - \sum_{k=1}^{A_l} x_{l,k} y_{l,k} \right)}{L}$$

362 where L is the total number of sites, A_l is the number of alleles at the l^{th} site and $x_{l,k}$
 363 and $y_{l,k}$ are the frequency of the k^{th} allele at the l^{th} locus in the population X and Y
 364 respectively.

365 Some population genetics software packages (e.g., EggLib; De Mita and Siol, 2012)
 366 average d_{XY} over polymorphic sites only, instead of averaging over all sites, as in
 367 Nei's (1987) original definition of d_{XY} . This measure averaged over polymorphic
 368 sites only will be called d_{XY-SNP} ; otherwise, we use the original definition of d_{XY} by Nei
 369 (1987).

370 We report the average F_{ST} , d_{XY} , and within population genetic diversity

$$371 \quad H_S = \sum_{l=1}^L \left(1 - \sum_{k=1}^{A_l} x_{l,k}^2 \right) / L. \text{ Our main results lie in the comparison between simulations}$$

372 with BGS and simulations without BGS within each treatment. Because theoretical

373 expectations exist for the strength of BGS on genetic diversity within populations,
374 we also investigated the relationships between this theoretical expectation, B , and
375 F_{ST} , F_{ST} (*average of ratios*), d_{XY} , d_{XY-SNP} and H_S in five independent tests for each treatment
376 and at each generation. For this, we used Pearson correlation test, Spearman
377 correlation tests, ordinary least squares regressions, robust regressions (using M-
378 estimators; Huber, 1964), and permutation tests. The results were systematically
379 consistent. Permutations tests of Pearson's correlation coefficients were performed
380 with 50,000 iterations. Because all tests were congruent, we only report the Pearson
381 correlation coefficient and the P -values from permutation tests.



382 Results

383 The distributions of F_{ST} values from simulations with BGS are extremely similar to
384 the distribution of F_{ST} values of simulations where all mutations were neutral. This
385 remains true even in the most extreme treatment with no recombination. This
386 general result is exemplified in Figure 1 by comparing the *Default* treatment to the
387 *No Recombination* treatment. As we have a large number of simulations, the means
388 of the distributions of F_{ST} are significantly different between simulations with BGS
389 and simulations without BGS for both the *Default* (Wilcoxon tests: $W=47875000$;
390 $P=0.00002$) and the *No Recombination* (Wilcoxon tests: $W=47804000$; $P=0.002$)
391 treatments but the increase in mean F_{ST} due to BGS is only of 4.3% for the *Default*
392 treatment and of 2.6% in the *No Recombination* treatment.



393 Figure 2 shows the means and standard errors for F_{ST} , d_{XY} and H_S for the treatments
394 *Default, High Migration, Large N, Human Genetic Map, Low selection pressure,*
395 *Constant μ and No Migration.* Similar graphs for the treatments *No Recombination*
396 and *CNC97* can be found in Figure S2.

*



The correlation may have been smoothed out by the size of the focal region

397 Figure 3 shows the correlation between B and the statistics F_{ST} , d_{XY} and H_S for
398 *Default* at the last generation. F_{ST} is not correlated with B ($P = 0.24$, $r = -0.02$). The
399 strongest correlations with B are observed for the statistics d_{XY} ($P = 4 \times 10^{-5}$, $r =$
400 0.06) and H_S ($P = 4 \times 10^{-5}$, $r = 0.06$). In fact, the two statistics d_{XY} and H_S are very
401 highly correlated ($P < 2.2 \times 10^{-16}$, $r = 0.99$). This high correlation explains the
402 resemblance between the central and right graphs of figure 3. All correlations
403 between the statistics H_S , F_{ST} , F_{ST} (average of ratios), d_{XY} , and d_{XY-SNP} and B , are summarized
404 in tables S1, S2, S3, S4 and S5, respectively.

405 When looking at correlations between B and the statistics of population divergence,
406 the *No Migration* treatment is an exception to the other treatments. For the *No*
407 *Migration* treatment, F_{ST} is not significantly correlated with B at early generations
408 but become slightly correlated as divergence rises to an F_{ST} of 0.6 and higher. d_{XY}
409 shows an opposite pattern. d_{XY} is very significantly correlated with B at early
410 generations and seemingly independent of B at the last generation. Note that for F_{ST}
411 all correlation coefficients are always very small. The largest r^2 observed for F_{ST} is
412 $r^2=0.01$ (found for *F_{ST} No Migration*).

413 As expected, in the CNC97 simulations, there is a strong difference between
414 simulations with BGS and simulations without BGS for all three statistics (F_{ST} , d_{XY} ,
415 and H_S) at all generations (Welch's t -tests; all $P < 2.2 \times 10^{-16}$; Figure S3). For more
416 simulations and discussion with similarly unrealistic genetic parameters, please see
417 Appendix A.



418 F_{ST} averaged over loci as advised by Weir and Cockerham (1984) was generally less
419 sensitive to BGS than F_{ST} calculated as an average of ratios (compare tables S2 and
420 S3). This effect is again partially visible in the correlations with B . Figure S3
421 illustrates the sensitivity of F_{ST} (average of ratios) in the worst case, the *No Recombination*
422 treatment. This sensitivity is driven largely by rare alleles and goes away when
423 minor alleles below a frequency of 0.05 are excluded.

424 The observed false positive rate for the F_{ST} outlier test is relatively close to the α
425 values except for *No Migration* (with and without BGS) and *CNC97* (with BGS).
426 Excluding the unrealistic treatment *CNC97*, we do not see more significant
427 differences between the FPR with and without BGS than we would expect by chance.
428 (Figure 4 and figure S4). There are other statistics of interest that one can consider
429 to investigate whether BGS causes F_{ST} outliers. Among all treatments (excluding
430 *CNC97*), the fraction of SNPs that are associated with a F_{ST} that is greater than 10
431 times the average F_{ST} in its particular treatment is 0.075% with BGS and 0.085%
432 without BGS. These numbers go up to 1.7% and 1.8% for SNPs that have a F_{ST} 5
433 times greater than the average F_{ST} , for treatments with BGS and without BGS,

Please discuss why!

FPR not previously defined.

434 respectively. We have also computed, in each treatment, the ratio of the largest F_{ST}
 435 to the average F_{ST} . Among treatments without BGS, the largest F_{ST} was on average
 436 12.2 times the average F_{ST} , while among treatments with BGS, the largest F_{ST} was on
 437 average 12.1 times the average F_{ST} . With an alpha threshold of 0.001, we observe
 438 that 0.080% and 0.083% of the SNPs turn out false positives among treatments
 439 without BGS and among treatment with BGS, respectively. For the conditions
 440 considered in this paper, BGS does not increase the rate of F_{ST} outliers.

Be more careful !

Yes, but
how sensitive
are the rates
to the size
of the focal
region?

441 Discussion

with the genetic architectures a.
dmg. recessions

442 In agreement with previous works (e.g. Charlesworth, 2012; Elyashiv et al., 2016;
 443 Messer & Petrov, 2013; Nordborg et al., 1996; Vijay et al., 2017; Zeng &
 444 Charlesworth, 2011), we show that background selection reduces genetic diversity,
 445 both within and among populations. The magnitude of this effect is very similar to
 446 previous realistic simulations (Messer & Petrov, 2013). The effect of BGS on F_{ST} is
 447 however rather small and does not seem to impact the overall distribution of F_{ST} in
 448 the sexual outcrossing species that we study here (Figure 1). The relative
 449 robustness of F_{ST} to variation in the intensity of BGS is contrary to what has been
 450 found in less realistic simulations (Charlesworth et al., 1997; Zeng & Corcoran,
 451 2015). *Please avoid, say simulations w.
much stronger BGS than the one simulated here. The admission*
 452 F_{ST} was also generally not significantly correlated with B . The only exception is for
 453 the *No Migration* treatment, where, after many generations, as the average F_{ST}
 454 becomes very high ($F_{ST} > 0.5$), we observe a slight, yet significant, negative



Inconsistent

w. l. 387
to l. 391,
where the
author
mention signif.
different

mean F_{ST} .

Also not in
line with
the admission
of a genome-
wide effect of
BGS.

455 correlation between the expected effects of BGS, B , and F_{ST} (intense BGS lead to high
456 values of F_{ST}). This highlights that F_{ST} is not completely insensitive to BGS, but F_{ST} is
457 largely robust to BGS. The observed correlation coefficients are always very small
458 with not a single r^2 value greater than 1%. It is important to highlight that B has not
459 been defined in order to estimate the effect of BGS onto F_{ST} but only for the effect of
460 BGS on H_S in a panmictic population. Here, we consider these correlations to
461 consolidate the evidence brought by comparing simulations with and without BGS
462 and to clarify how BGS affect the different statistics of population divergence.

Could this just be a transient correlation, as F_{ST} → 1 anyway?

463 In this study, we investigated the locus-to-locus variation in the intensity of BGS and
464 how it affects F_{ST} . Future research is needed to attempt a theoretical estimate of the
465 genome-wide effect of BGS on F_{ST} (but see Charlesworth et al., 1997; Zeng &
466 Corcoran, 2015). Our work has been restricted to the stickleback and human
467 recombination maps. While these two genomes are good representatives of many
468 cases of eukaryotic genomes, they are not good representatives of more compact
469 genomes such as bacterial genomes or yeasts. Our simulations used randomly
470 mating diploid populations. Non-random mating, selfing, and asexual reproduction
471 could also affect our general conclusion, and potentially strongly increase the effects
472 of BGS on F_{ST} (Charlesworth et al. 1997). Also, we did not explore the effect of
473 haploid selection as we only considered autosomes (see Charlesworth; 2012). We
474 have explored two population sizes, but we could not explore population sizes of the
475 order of a million individuals (like *Drosophila melanogaster*) and still realistically
476 simulate such long stretch of DNA. It is not impossible that a much greater

Internal: Could this be overcome using theory?

477 population size or a more complex demography could result in BGS having a greater
478 effect on F_{ST} than what we observed here (Torres et al. 2017). In the *No*
479 *Recombination* treatment, we have explored cases of complete suppression of
480 recombination over stretches of DNA of, on average, 0.8% of the stickleback genome
481 and our results were still consistent. However, we have not explored the effect of
482 suppression of recombination over greater regions, such as a whole chromosome
483 that does not recombine. We have also not explored such suppression of
484 recombination in perfectly isolated populations as we mainly focused on
485 interconnected populations. It is not impossible that in such cases we might observe
486 a greater impact of BGS on F_{ST} such as those observed in Zeng & Corcoran (2015)
487 and reproduced in appendix A.

Does
appendix A
explore
isolated
popul. ?

488 Some have argued that, because BGS reduces the within population diversity, it
489 should lead to high F_{ST} (Cutter & Payseur, 2013; Cruickshank & Hahn, 2014; Hoban
490 et al., 2016). All else being equal, this statement is correct. However, BGS reduces H_T
491 almost as much as H_S (Figure S6). It is therefore insufficient to consider only one
492 component, and we must consider the ratio of these two quantities captured by the
493 definition of F_{ST} , $F_{ST} = 1 - \frac{H_S}{H_T}$. This ratio, as we have shown, appears to be relatively
494 robust to BGS. While genome-wide BGS might eventually be strong enough to cause
495 departures with F_{ST} values, it appears that locus-to-locus variation in the intensity of
496 BGS is not strong enough to have much impact on F_{ST} in outcrossing sexual
497 organisms, at least as long as populations are not too highly diverged. In theoretical
498 studies, it is also possible to consider the coalescent time definition of F_{ST} (Slatkin,

This is
not so much
a question
of coal.
line def.
vs. for by site.

499 1991). In appendix A, we show that, for sites that are directly under selection, the
 500 coalescent definition yield to different estimates than the definition based on allelic
 501 state.

502 We also investigated the consequences of BGS on the widely-used but imperfect
 503 estimator, F_{ST} (*average of ratios*), for which F_{ST} measures for each locus are averaged to
 504 create a genomic average. It is well known that F_{ST} (*average of ratios*) is a biased way to
 505 average F_{ST} over several loci (Weir & Cockerham, 1984); however, its usage is
 506 relatively common today. In our simulations, F_{ST} (*average of ratios*) is more affected by
 507 BGS than F_{ST} . Interestingly, F_{ST} (*average of ratios*) is most often higher with weaker BGS.
 508 The directionality of this correlation may seem unintuitive at first. To understand
 509 this discrepancy, remember that BGS affects the site frequency spectrum; we
 510 observed that BGS leads to an excess of loci with low H_T (results not shown but see
 511 Charlesworth et al., 1995; see also contrary expectation in Stephan, 2010). Loci
 512 associated with very low H_T also have low F_{ST} (figure S5), a well-known result
 513 described by Beaumont and Nichols (1996). As BGS creates an excess of loci with
 514 low H_T and loci with low H_T tend to have low F_{ST} , BGS can actually reduce F_{ST} (*average of*
 515 *ratios*). After filtering out SNPs with a minor allele frequency lower than 5%, most of
 516 the correlation between F_{ST} (*average of ratios*) and B is eliminated (Figure S3).

517 The absolute measure of divergence d_{XY} is more sensitive to BGS than F_{ST} (Figure 2;
 518 Figure 3; Tables S2 and S4). Regions of stronger BGS are associated with low d_{XY} .

519 This is in agreement with correlation tests between B and d_{XY} . The effect, although

- 1) Distorts from the main focus
- 2) Logic does not make sense (H_T red. $\rightarrow F_{ST}$ red.)
- 3) Misleading of SFS does not make sense.

$$\text{For } N \gg 0 \text{ and } t \rightarrow \infty: d_{XY} \rightarrow 4N\mu + \frac{\mu}{m}$$

This expectation only applies to the complete isolation model.

The generic expression is more complicated (see Wilkinson-herbst 2008)

Better:

Directly consult Wakeley (1986)

J. genetics eq. (11)

This only holds for the split model.

Time for the split model, but I am not sure about the generic model before the equilibrium is reached
→ check in Wilkinson-herbst

520 significant, is of relatively small size. The expected d_{XY} for neutral loci is $d_{XY} = 4N\mu +$
 521 $2t\mu$ (Nei, 1987), where t is the time in generations since the populations started to
 522 diverge. $4N\mu$ is the expected heterozygosity in the ancestral population (before
 523 splitting) and $2t\mu$ is the expected number of mutations fixed over time in either
 524 population since the population split. BGS does not affect the rate of fixation of
 525 neutral mutations arising after the populations diverged, but BGS affects the
 526 expected heterozygosity. Therefore, BGS should affect d_{XY} by its effect on the
 527 expected heterozygosity, and this effect should be greater early in divergence when
 528 the $4N\mu$ term is large relative to the fixation term. This is consistent with the results
 529 of our simulations. This result is in agreement with Vijay et al. (2017) who reported
 530 a strong correlation between H_S and d_{XY} when F_{ST} is low ($F_{ST} \approx 0.02$), but this
 531 correlation breaks down when studying more distantly related populations ($F_{ST} \approx$
 532 0.3). Previous works have even suggested a potential negative correlation between
 533 F_{ST} and d_{XY} (Nachman & Payseur, 2012; Irwin et al. 2016). Of course, any potential
 534 effect of BGS on F_{ST} should disappear as F_{ST} approaches one as the statistic is
 535 saturated.

536 As BGS also leads to a reduction of the number of polymorphic sites, BGS has an
 537 even stronger effect on d_{XY-SNP} than on d_{XY} (Figure S3). (The measure that we call d_{XY-}
 538 SNP is d_{xy} improperly calculated based only on polymorphic sites, as is done in some
 539 software packages.) This result highlights the importance of not blindly trusting the
 540 output of a given software package.



541 Outside the effect of BGS on N_e , there are at least two other possible factors that can
 542 potentially affect the correlation between B and μ : the effect of deleterious
 543 mutations on the effective migration rate and the auto-correlation of μ . Because
 544 most deleterious mutations are recessive (García-Dorado and Caballero, 2000;
 545 Peters et al., 2003; Shaw & Chang, 2006), the offspring of migrants, who enjoy an
 546 increased heterozygosity compared to local individuals, will be at a selective
 547 advantage. The presence of deleterious mutations therefore lead to an increase in
 548 the effective migration rate (Ingvarsson & Whitlock, 2000). This increases the
 549 effective migration rate and hence, leads to a decrease in F_{ST} . In our simulations
 550 however, mutational allelic effects are close to additive and hence, we should not
 551 expect to see much effect of BGS on the effective migration rate.

$$N_e \approx N e \exp\left(-\frac{\mu_d}{\tau}\right) \approx N\left(1 - \frac{\mu_d}{\tau}\right)$$

552 As mutation rate is auto-correlated throughout the genome, neutral sequences
 553 closely linked to sequences that frequently receive deleterious mutation are also
 554 likely to experience frequent neutral mutations. As a high mutation rate leads to low
 555 F_{ST} values ($F_{ST} \approx \frac{1}{1 + 4Ne(m + \mu)}$, Wright 1943), autocorrelation in mutation rate may
 556 also act to reduce the effect of BGS on F_{ST} . This effect is however likely to be
 557 negligible as long as $m \gg \mu$.

558 Recently, evidence of a correlation between recombination rate and F_{ST} has been
 559 interpreted as likely being caused by deleterious mutations rather than positive
 560 selection, whether the divergence between populations is very high (e.g.
 561 Cruickshank & Hahn, 2014), moderately high (Vijay et al., 2017) or moderately low

even though they 1) earlier state that they
 are not interested in genome-wide
 patterns; 2) do not show any results on F_{ST} vs.
 r (and gene density)

Read
this
paper!
↓
What are the
remaining
open questions?

The authors
talk about
corr. between
 F_{ST} and
recomb. rate,

28

562 (Torres et al., 2017). Here we showed the BGS is unlikely to explain all of these
563 correlations between F_{ST} and recombination rate. As positive selection has been
564 shown to also have important effect on genetic diversity (Eyre-Walker & Keightley,
565 2009; Hernandez, Kelley, Elyashiv, Melton, & Auton, 2011; Macpherson, Sella, Davis,
566 & Petrov, 2007; Sattath, Elyashiv, Kolodny, Rinott, & Sella, 2011; Wildman, Uddin,
567 Liu, Grossman, & Goodman, 2003), it would be important to investigate whether
568 positive selection (selective sweeps and local adaptation) could be an important
569 driver of the correlations between F_{ST} and recombination rate. More research would
570 be needed to investigate whether this is true.

✓ again the authors talk about genome-wide correlation

571 McVicker et al. (2009) attempted an estimation of B values in the human genome
572 (see also Elyashiv et al., 2016). They did so using equations from Nordborg et al.
573 (1996). As there is little knowledge about the strength of selection throughout the
574 genome, to our understanding, this estimation of B values should be highly
575 influenced by the effects of beneficial mutations as well as deleterious mutations.
576 Torres et al. (2017) reused this dataset and found a slight association between B and
577 F_{ST} among human lineages. It is plausible that this correlation between B and F_{ST}
578 could be driven by positive selection rather than by deleterious mutations.

✓ Check if McVicker et al. (2009) estimated B and S or if they assumed values.

579 Our FDist2 analysis shows that the false positive rate does not differ in simulations
580 with BGS or without BGS (figure 4). The only exceptions concern the unrealistic
581 CNC97 treatment (figure S4). The average F_{ST} at the last generation of the No
582 Migration treatment is greater than 0.8. With such high F_{ST} , the F_{ST} outlier method,

583 Fdist2, does not seem to perform well and both the simulations without BGS and
584 with BGS lead to very high false positive rates (0.472 without BGS and 0.467 with
585 BGS; Figure S4). While other issues may intervene in F_{ST} outlier methods (Lotterhos
586 & Whitlock, 2014), our results show that BGS should not represent any significant
587 issue for outcrossing sexual species with moderately low mean F_{ST} .

*Include
restriction to loci-
to-
locus
variation*

588 We have shown that BGS affects H_S and d_{XY} but has only a very minor effect on F_{ST}
589 among sexual outcrossing populations connected by gene flow. Many authors
590 (Cutter & Payseur, 2013; Hoban et al., 2016) have raised concerns that BGS can
591 strongly reduce our ability to detect the genomic signature of local adaptation. Our
592 analysis shows that BGS is not a strong confounding factor to F_{ST} outlier tests.

593 Acknowledgments

594 Many thanks to Loren H. Rieseberg, Sarah P. Otto and Amy L. Angert for their help in
595 discussing the design of the project and for feedback. Thanks to Sarah P. Otto,
596 Darren E. Irwin, and Bret A. Payseur as well as two anonymous reviewers for helpful
597 comments on the manuscript. We also thank Yaniv Brandvain and Tom Booker for
598 their feedback and Marius Roesti for help with the Ensembl-retrieved gene
599 annotations. We also acknowledge ComputeCanada that provided the hardware
600 resources for running our simulations.

601 Funding

602 The work was funded by NSERC Discovery Grant RGPIN-2016-03779 to MCW and
603 by the Swiss National Science Foundation via the fellowship Doc.Mobility
604 P1SKP3_168393 to RMD.

605 **References**

- 606 Aeschbacher, S., Selby, J. P., Willis, J. H., & Coop, G. (2017). Population-genomic
607 inference of the strength and timing of selection against gene flow. *Proceedings
608 of the National Academy of Sciences*, 114(27), 7061–7066.
609 <https://doi.org/10.1073/pnas.1616755114>
- 610 Bank, C., Ewing, G. B., Ferrer-Admetla, A., Foll, M., & Jensen, J. D. (2014). Thinking
611 too positive? Revisiting current methods of population genetic selection
612 inference. *Trends in Genetics*, 30(12), 540–546.
613 <https://doi.org/10.1016/j.tig.2014.09.010>
- 614 Beaumont, M. A., & Nichols, R. A. (1996). Evaluating loci for use in the genetics
615 analysis of population structure. *Proceedings of the Royal Society B: Biological
616 Sciences*, 263, 1619–1626.
- 617 Begun, D. J., & Aquadro, C. F. (1992). Levels of naturally occurring DNA
618 polymorphism correlate with recombination rate in *D. melanogaster*. *Nature*,
619 356.
- 620 Bonhomme, M., Chevalet, C., Servin, B., Boitard, S., Abdallah, J., Blott, S., &

- 621 SanCristobal, M. (2010). Detecting Selection in Population Trees: The Lewontin
622 and Krakauer Test Extended. *Genetics*, 186(1), 241–262.
- 623 <https://doi.org/10.1534/genetics.110.117275>
- 624 Booker, T. R., & Keightley, P. D. (2018). Understanding the factors that shape
625 patterns of nucleotide diversity in the house mouse genome. *BioRxiv*, 35(12),
626 2971–2988. <https://doi.org/10.1101/275610>
- 627 Brousseau, L., Postolache, D., Lascoux, M., Drouzas, A. D., Källman, T., Leonarduzzi, C.,
628 ... Vendramin, G. G. (2016). Local adaptation in European firs assessed through
629 extensive sampling across altitudinal gradients in southern Europe. *PLoS ONE*,
630 11(7), e0158216. <https://doi.org/10.1371/journal.pone.0158216>
- 631 Cai, J. J., Macpherson, J. M., Sella, G., & Petrov, D. A. (2009). Pervasive hitchhiking at
632 coding and regulatory sites in humans. *PLoS Genetics*, 5(1), e1000336.
633 <https://doi.org/10.1371/journal.pgen.1000336>
- 634 Charlesworth, B. (1998). Measures of divergence between populations and the effect
635 of forces that reduce variability. *Molecular Biology and Evolution*, 15(5), 538–
636 543. <https://doi.org/10.1093/oxfordjournals.molbev.a025953>
- 637 Charlesworth, B. (2012). The role of background selection in shaping patterns of
638 molecular evolution and variation: Evidence from variability on the Drosophila
639 X chromosome. *Genetics*, 191(1), 233–246.

- 640 <https://doi.org/10.1534/genetics.111.138073>
- 641 Charlesworth, B., Morgan, M. T., & Charlesworth, D. (1993). The effect of deleterious
642 mutations on neutral molecular variation. *Genetics*, 134(4), 1289–1303.
- 643 Charlesworth, B., Nordborg, M., & Charlesworth, D. (1997). The effects of local
644 selection, balanced polymorphism and background selection on equilibrium
645 patterns of genetic diversity in subdivided populations. *Genetical Research*,
646 70(2), 155–174. <https://doi.org/10.1017/S0016672397002954>
- 647 Charlesworth, D., Charlesworth, B., & Morgan, M. T. (1995). The pattern of neutral
648 molecular variation under the background selection model. *Genetics*, 141(4),
649 1619–1632.
- 650 Chen, H., Patterson, N., & Reich, D. E. (2010). Population differentiation as a test for
651 selective sweeps. *Genome Research*, 20(3), 393–402.
652 <https://doi.org/10.1101/gr.100545.109>
- 653 Cruickshank, T. E., & Hahn, M. W. (2014). Reanalysis suggests that genomic islands
654 of speciation are due to reduced diversity, not reduced gene flow. *Molecular
655 Ecology*, 23(13), 3133–3157. <https://doi.org/10.1111/mec.12796>
- 656 Cutter, A. D., & Payseur, B. A. (2013). Genomic signatures of selection at linked sites :
657 unifying the disparity among species. *Nature Publishing Group*, 14(4), 262–274.
658 <https://doi.org/10.1038/nrg3425>

- 659 De Mita, M. S. (2012). EggLib: processing, analysis and simulation tools for
660 population genetics and genomics. *BMC Genetics*, 13(1), 27.
661 <https://doi.org/10.1186/1471-2156-13-27>
- 662 Dutoit, L., Vijay, N., Mugal, C. F., Bossu, C. M., Burri, R., Wolf, J., & Ellegren, H. (2017).
663 Covariation in levels of nucleotide diversity in homologous regions of the avian
664 genome long after completion of lineage sorting. *Proceedings of the Royal
665 Society B: Biological Sciences*, 284(1849), 20162756.
666 <https://doi.org/10.1098/rspb.2016.2756>
- 667 Elyashiv, E., Sattath, S., Hu, T. T., Strutsovsky, A., McVicker, G., Andolfatto, P., ... Sella,
668 G. (2016). A Genomic map of the effects of linked selection in *Drosophila*. *PLoS
669 Genetics*, 12(8), e1006130. <https://doi.org/10.1371/journal.pgen.1006130>
- 670 Eyre-Walker, A., & Keightley, P. D. (2009). Estimating the rate of adaptive molecular
671 evolution in the presence of slightly deleterious mutations and population size
672 change. *Molecular Biology and Evolution*, 26(9), 2097–2108.
673 <https://doi.org/10.1093/molbev/msp119>
- 674 Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci
675 appropriate for both dominant and codominant markers: a Bayesian
676 perspective. *Genetics*, 180(2), 977–993.
677 <https://doi.org/10.1534/genetics.108.092221>

- 678 García-Dorado, A., Caballero, A., Bateman, A. J., Bregliano, J.-C., Laurencon, A.,
679 Degroote, F., ... Crow, J. F. (2000). On the average coefficient of dominance of
680 deleterious spontaneous mutations. *Genetics*, 155(4), 1991–2001.
681 <https://doi.org/10.1080/09553005914550241>
- 682 Gilbert, K. J., Sharp, N. P., Angert, A. L., Conte, G. L., Draghi, J. A., Guillaume, F., ...
683 Whitlock, M. C. (2017). Local adaptation interacts with expansion load during
684 range expansion: Maladaptation reduces expansion load. *The American
685 Naturalist*, 189(4), 368–380. <https://doi.org/10.1086/690673>
- 686 Guillaume, F., & Rougemont, J. (2006). Nemo: An evolutionary and population
687 genetics programming framework. *Bioinformatics*, 22(20), 2556–2557.
688 <https://doi.org/10.1093/bioinformatics/btl415>
- 689 Haller, B. C., & Messer, P. W. (2017). SLiM 2: Flexible, interactive forward genetic
690 simulations. *Molecular Biology and Evolution*, 34(1), 230–240.
691 <https://doi.org/10.1093/molbev/msw211>
- 692 Hernandez, R. D., Kelley, J. L., Elyashiv, E., Melton, S. C., & Auton, A. (2011). Classic
693 Selective Sweeps were rare in Recent Human Evolution. *Science*, 331(6019),
694 920–924.
- 695 Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., ...
696 Whitlock, M. C. (2016). Finding the genomic basis of local adaptation: pitfalls,

- 697 practical solutions, and future directions. *The American Naturalist*, 188(4), 379–
698 397. <https://doi.org/10.1086/688018>
- 699 Huber, C. D., Degiorgio, M., Hellmann, I., & Nielsen, R. (2016). Detecting recent
700 selective sweeps while controlling for mutation rate and background selection.
701 *Molecular Ecology*, 25, 142–156. <https://doi.org/10.1111/mec.13351>
- 702 Huber, P. J. (1964). Robust Estimation of a Location Parameter. *The Annals of
703 Mathematical Statistics*, 35(1), 73–101.
704 <https://doi.org/10.1214/aoms/1177703732>
- 705 Hudson, R. R., & Kaplan, N. L. (1995). Deleterious background selection with
706 recombination. *Genetics*, 141(4), 1605–1617.
- 707 Ingvarsson, P. K., & Whitlock, M. C. (2000). Heterosis increases the effective
708 migration rate. *Proceedings of the Royal Society B: Biological Sciences*,
709 267(1450), 1321–1326. <https://doi.org/10.1098/rspb.2000.1145>
- 710 Irwin, D. E., Alcaide, M., Delmore, K. E., Irwin, J. H., & Owens, G. L. (2016). Recurrent
711 selection explains parallel evolution of genomic regions of high relative but low
712 absolute differentiation in a ring species. *Molecular Ecology*, 25(18), 4488–
713 4507. <https://doi.org/10.1111/mec.13792>
- 714 Keightley, P. D. (2012). Rates and fitness consequences of new mutations in humans.
715 *Genetics*, 190(2), 295–304. <https://doi.org/10.1534/genetics.111.134668>

- 716 Keightley, P. D., & Gaffney, D. J. (2003). Functional constraints and frequency of
717 deleterious mutations in noncoding DNA of rodents. *Proceedings of the National
718 Academy of Sciences*, 100(23), 13402–13406.
719 <https://doi.org/10.1073/pnas.2233252100>
- 720 Kim, Y., & Stephan, W. (2000). Joint effects of genetic hitchhiking and background
721 selection on neutral variation. *Genetics*, 155(3), 1415–1427.
- 722 Lange, J. D., & Pool, J. E. (2016). A haplotype method detects diverse scenarios of
723 local adaptation from genomic sequence variation. *Molecular Ecology*, 25(13),
724 3081–3100. <https://doi.org/10.1111/mec.13671>
- 725 Lotterhos, K. E., & Whitlock, M. C. (2014). Evaluation of demographic history and
726 neutral parameterization on the performance of FST outlier tests. *Molecular
727 Ecology*, 23(9), 2178–2192. <https://doi.org/10.1111/mec.12725>
- 728 Macpherson, J. M., Sella, G., Davis, J. C., & Petrov, D. A. (2007). Genomewide spatial
729 correspondence between nonsynonymous divergence and neutral
730 polymorphism reveals extensive adaptation in *Drosophila*. *Genetics*, 177(4),
731 2083–2099. <https://doi.org/10.1534/genetics.107.080226>
- 732 Maynard-Smith, J., & Haig, J. (1974). The hitch-hiking effect of a favourable gene.
733 *Genetical Research*, 23(1), 23–35.
734 <https://doi.org/10.1017/S0016672300014634>

- 735 McGee, M. D., Neches, R. Y., & Seehausen, O. (2015). Evaluating genomic divergence
736 and parallelism in replicate ecomorphs from young and old cichlid adaptive
737 radiations. *Molecular Ecology*, 25(1), 260–268.
- 738 <https://doi.org/10.1111/mec.13463>
- 739 McVicker, G., Gordon, D., Davis, C., & Green, P. (2009). Widespread genomic
740 signatures of natural selection in hominid evolution. *PLoS Genetics*, 5(5),
741 e1000471. <https://doi.org/10.1371/journal.pgen.1000471>
- 742 Messer, P. W., & Petrov, D. a. (2013). Frequent adaptation and the McDonald-
743 Kreitman test. *Proceedings of the National Academy of Sciences of the United
744 States of America*, 110(21), 8615–8620.
- 745 <https://doi.org/10.1073/pnas.1220835110>
- 746 Nachman, M. W., & Crowell, S. L. (2000). Estimate of the mutation rate per
747 nucleotide in humans. *Genetics*, 156(1), 297–304.
- 748 Nachman, M. W., & Payseur, B. A. (2012). Recombination rate variation and
749 speciation: theoretical predictions and empirical results from rabbits and mice.
750 *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1587),
751 409–421. <https://doi.org/10.1098/rstb.2011.0249>
- 752 Nei, M. (1973). Analysis of Gene Diversity in Subdivided Populations. *Proceedings of
753 the National Academy of Sciences*, 70(12), 3321–3323.

- 754 https://doi.org/10.1073/pnas.70.12.3321
- 755 Nei, M. (1987). *Molecular Evolutionary Genetics*. New York: Columbia University
756 Press.
- 757 Nordborg, M., Charlesworth, B., & Charlesworth, D. (1996). The effect of
758 recombination on background selection. *Genetical Research*, 67(02), 159–174.
759 https://doi.org/10.1017/S0016672300033619
- 760 Payseur, B. A., & Rieseberg, L. H. (2016). A genomic perspective on hybridization
761 and speciation. *Molecular Ecology*, 25(11), 2337–2360.
762 https://doi.org/10.1111/mec.13557
- 763 Peters, A. D., Halligan, D. L., Whitlock, M. C., & Keightley, P. D. (2003). Dominance and
764 Overdominance of Mildly Deleterious Induced Mutations for Fitness Traits in
765 *Caenorhabditis elegans*. *Genetics*, 165(2), 589–599.
- 766 Picq, S., Mcmillan, W. O., & Puebla, O. (2016). Population genomics of local
767 adaptation versus speciation in coral reef fishes (*Hypoplectrus* spp,
768 Serranidae). *Ecology and Evolution*, 6(7), 2109–2124.
769 https://doi.org/10.1002/ece3.2028
- 770 Reynolds, J., Weir, B. S., & Cockerham, C. C. (1983). Estimation of the coancestry
771 coefficient: Basis for a short-term genetic distance. *Genetics*, 105(3), 767–779.

- 772 Roesti, M., Moser, D., & Berner, D. (2013). Recombination in the threespine
773 stickleback genome - Patterns and consequences. *Molecular Ecology*, 22(11),
774 3014–3027. <https://doi.org/10.1111/mec.12322>
- 775 Sabeti, P. C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., ... Stewart, J.
776 (2007). Genome-wide detection and characterization of positive selection in
777 human populations. *Nature*, 449(7164), 913–918.
778 <https://doi.org/10.1038/nature06250>
- 779 Sattath, S., Elyashiv, E., Kolodny, O., Rinott, Y., & Sella, G. (2011). Pervasive adaptive
780 protein evolution apparent in diversity patterns around amino acid
781 substitutions in drosophila simulans. *PLoS Genetics*, 7(2), e1001302.
782 <https://doi.org/10.1371/journal.pgen.1001302>
- 783 Shaw, R. G., & Chang, S. M. (2006). Gene action of new mutations in Arabidopsis
784 thaliana. *Genetics*, 172(3), 1855–1865.
785 <https://doi.org/10.1534/genetics.105.050971>
- 786 Slatkin, M. (1991). Inbreeding coefficients and coalescence times. *Genetics Research*,
787 58(2), 167–175. <https://doi.org/10.1017/S0016672300029827>
- 788 Spencer, C. C. A., Deloukas, P., Hunt, S., Mullikin, J., Myers, S., Silverman, B., ...
789 McVean, G. (2006). The influence of recombination on human genetic diversity.
790 *PLoS Genetics*, 2(9), e148. <https://doi.org/10.1371/journal.pgen.0020148>

- 791 Stephan, W. (2010). Genetic hitchhiking versus background selection : the
792 controversy and its implications, 365(1544), 1245–1253.
793 <https://doi.org/10.1098/rstb.2009.0278>
- 794 Tange, O. (2011). GNU Parallel: The Command-Line Power Tool. ;*Login*;, 36(1), 42–
795 47. <https://doi.org/10.5281/zenodo.16303>
- 796 The International HapMap, C. (2007). A second generation human haplotype map of
797 over 3.1 million SNPs. *Nature*, 449(7164), 851–861.
798 <https://doi.org/10.1038/nature06258.A>
- 799 Torres, R., Szpiech, Z. A., & Hernandez, R. D. (2018). Human demographic history has
800 amplified the effects background selection across the genome. *PLoS Genetics*,
801 14(6), e1007387. <https://doi.org/10.1371/journal.pgen.1007387>
- 802 Vijay, N., Weissensteiner, M., Burri, R., Kawakami, T., Ellegren, H., & Wolf, J. B. W.
803 (2017). Genomewide patterns of variation in genetic diversity are shared
804 among populations, species and higher-order taxa. *Molecular Ecology*, 26(16),
805 4284–4295. <https://doi.org/10.1111/mec.14195>
- 806 Weir, B. S., & Cockerham, C. C. (1984). Estimating F-Statistics for the Analysis of
807 Population Structure. *Evolution*, 38(6), 1358–1370.
- 808 Whitlock, M. C. (1992). Temporal Fluctuations in Demographic Parameters and the
809 Genetic Variance among Populations. *Evolution*, 46(3), 608–615.

- 810 <https://doi.org/10.2307/2409631>
- 811 Whitlock, M. C., & Lotterhos, K. E. (2015). Reliable Detection of Loci Responsible for
812 Local Adaptation: Inference of a Null Model through Trimming the Distribution
813 of F_{ST} . *The American Naturalist*, 186(S1), S24–S36.
- 814 <https://doi.org/10.1086/682949>
- 815 Wildman, D. E., Uddin, M., Liu, G., Grossman, L. I., & Goodman, M. (2003).
816 Implications of natural selection in shaping 99.4% nonsynonymous DNA
817 identity between humans and chimpanzees: enlarging genus Homo.
818 *Proceedings of the National Academy of Sciences of the United States of America*,
819 100(12), 7181–7188. <https://doi.org/10.1073/pnas.1232172100>
- 820 Wright, S. (1943). Isolation by distance. *Genetics*, 28(2), 114–138.
- 821 Yeaman, S. (2015). Local Adaptation by Alleles of Small Effect. *The American
822 Naturalist*, 186(S1), S74–S89. <https://doi.org/10.1086/682405>
- 823 Zeng, K., & Charlesworth, B. (2011). The joint effects of background selection and
824 genetic recombination on local gene genealogies. *Genetics*, 189(1), 251–266.
825 <https://doi.org/10.1534/genetics.111.130575>
- 826 Zeng, K., & Corcoran, P. (2015). The effects of background and interference selection
827 on patterns of genetic variation in subdivided populations. *Genetics*, 201(4),
828 1539–1554. <https://doi.org/10.1534/genetics.115.178558>

- 829 Zerbino, D. R., Achuthan, P., Akanni, W., Amode, M. R., Barrell, D., Bhai, J., ... Flicek, P.
- 830 (2018). Ensembl 2018. *Nucleic Acids Research*, 46(D1), D754–D761.
- 831 <https://doi.org/10.1093/nar/gkx1098>

832

For Review Only

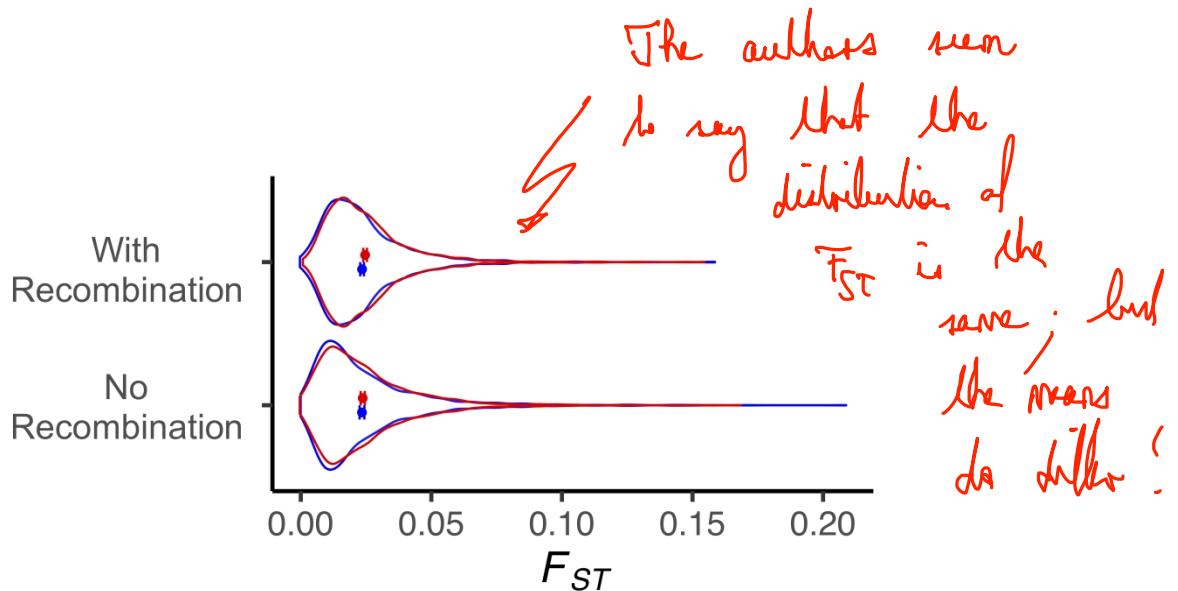


Figure 1: Violin plots showing the distribution of F_{ST} values for the *Default* treatment (labelled "With Recombination") and for the *No Recombination* treatment. Simulations with BGS are shown in red, and simulations without BGS are in blue. The means and standard errors are displayed with dots and error bars (although the error bars are barely visible).

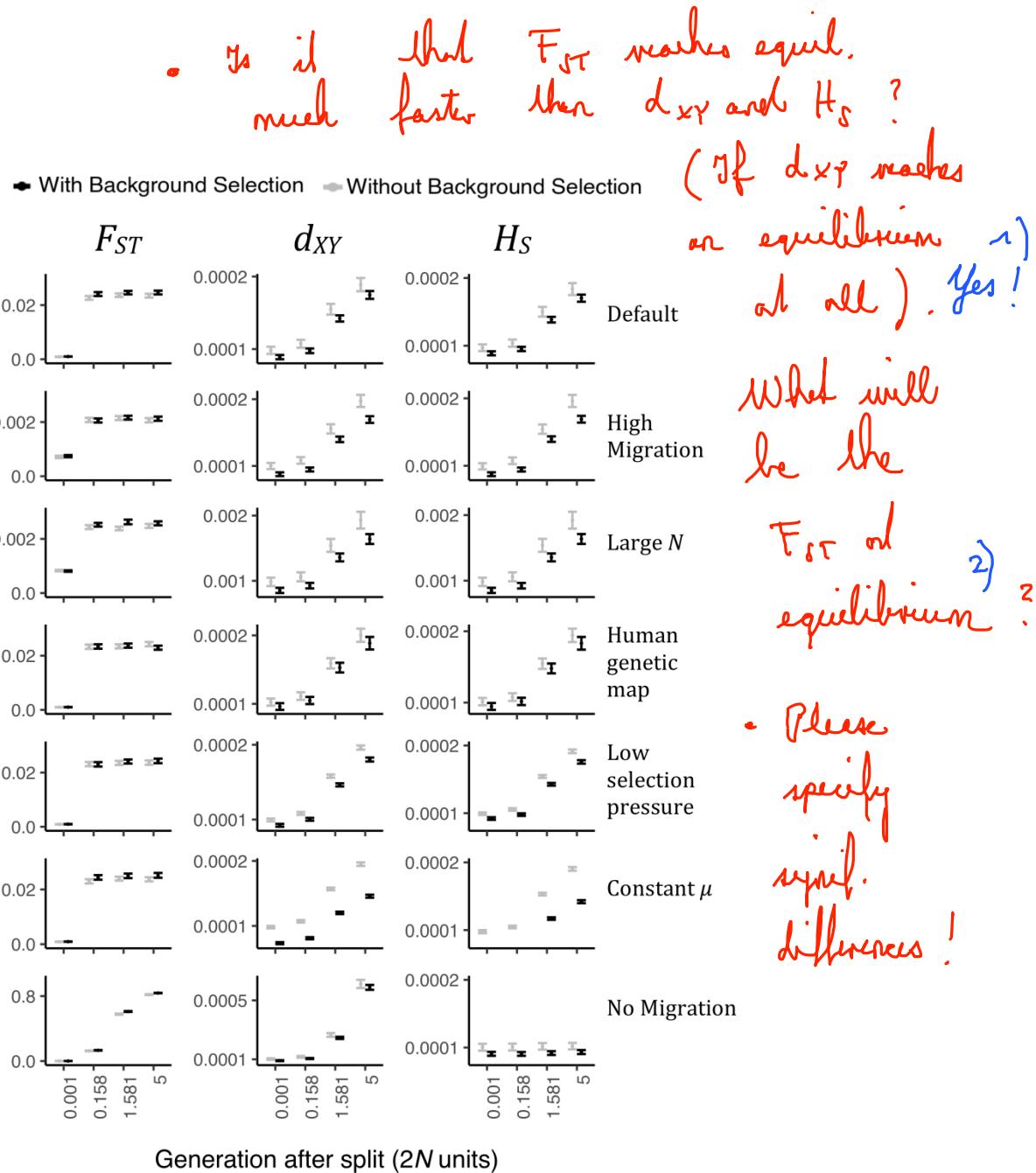


Figure 2: Comparisons of mean F_{ST} (left column), d_{XY} (central column), and H_S (right column) between simulations with (black) and without (grey) BGS. Similar graphs for the treatments *No Recombination* and *CNC97* are in figure S3. Error bars are 95% CI.

1) E.g. Plathkin (1991)

2) $\hat{F}_{ST} = \frac{1}{1 + 8Nm}$

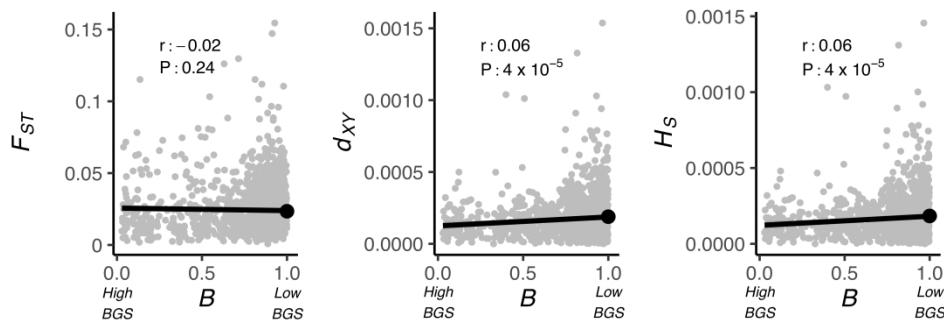


Figure 3: Correlation between B and F_{ST} , H_s , and d_{XY} for the last generation ($5 \times 2N$ generations after the split) of the *Default* treatment. Each grey dot is a single simulation where there is BGS. The large black dot is the mean of the simulations with no BGS. The P -values are computed from a permutation test and r is the Pearson's correlation coefficient. P -values and r are computed on both simulations with and without BGS. Results are congruent when computing the correlation coefficients and P -value on the subset of simulations that have BGS.

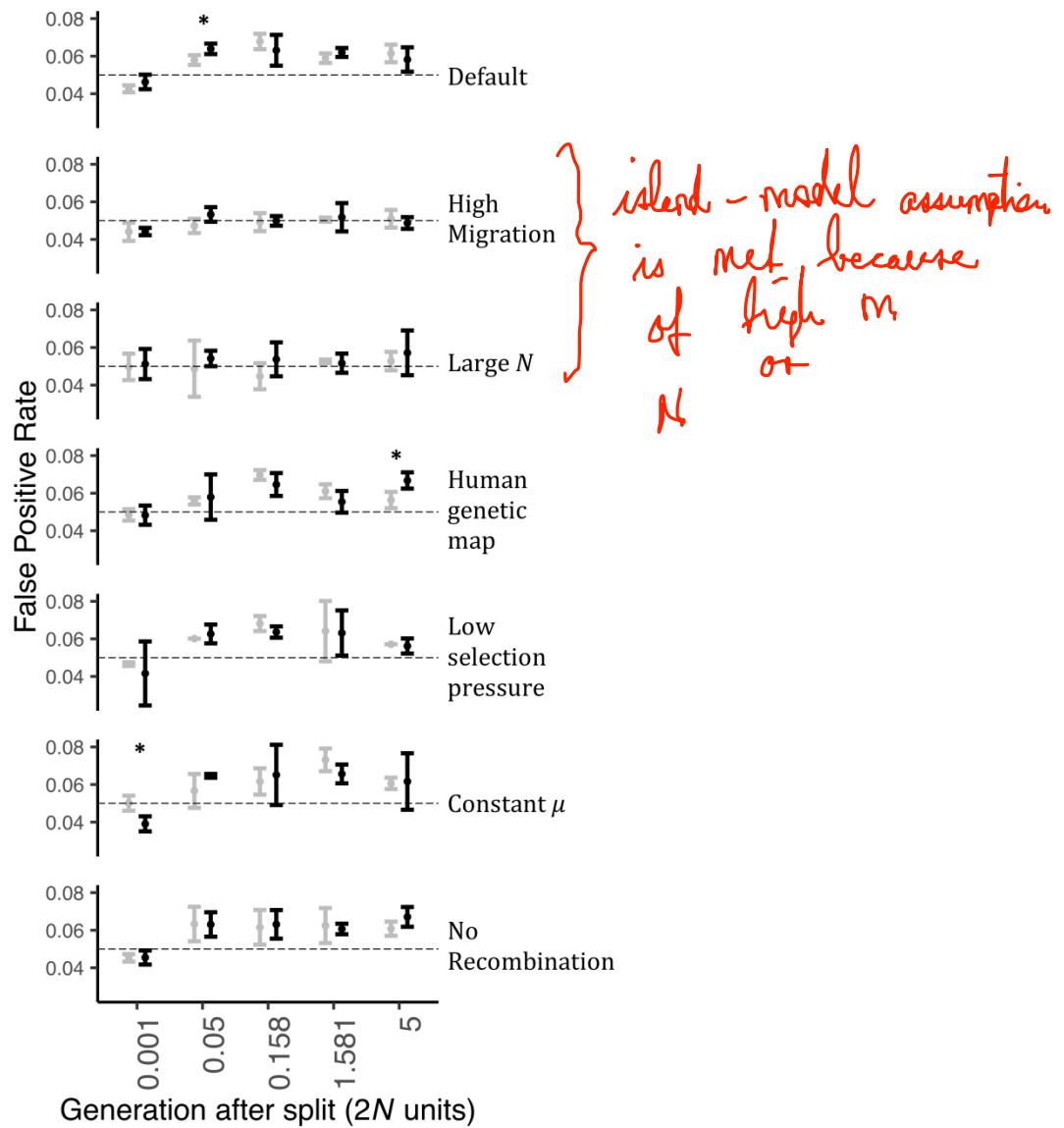


Figure 4: Comparison of false positive rate (FPR) returned by FDist2 between simulations with BGS (black) and without BGS (grey) for all treatments by generation. The significance level is 0.05 and is represented by the horizontal dashed line. Significance based on a Welch's t -test is indicated with stars (** $P < 0.001$; ** $P < 0.01$; * $P < 0.05$). With Wilcoxon tests, none of the treatments displayed here comes out as significant. Treatments No Migration and CNC97 are presented in Figure S4.

<i>Treatment</i>	<i>N</i>	<i>m</i>	<i>Genome</i>	<i>Other</i>	<i>BGS</i>
<i>Default</i>	1000	0.005	<i>Stickleback</i>	“Normal”	Yes
					No
<i>High Migration</i>	1000	0.05	<i>Stickleback</i>	“Normal”	Yes
					No
<i>Large N</i>	10000	0.005	<i>Stickleback</i>	“Normal”	Yes
					No
<i>Human genetic map</i>	1000	0.005	<i>Human</i>	“Normal”	Yes
					No
<i>Low selection pressure</i>	1000	0.005	<i>Stickleback</i>	<i>Low selection pressure</i>	Yes
					No
<i>Constant μ</i>	1000	0.005	<i>Stickleback</i>	<i>Constant mutation rate</i>	Yes
					No
<i>No Migration</i>	1000	0	<i>Stickleback</i>	“Normal”	Yes
					No
<i>No Recombination</i>	1000	0.005	<i>Stickleback</i>	<i>Absence of crossover</i>	Yes
					No
<i>CNC97</i>	2000	0.001	NA	<i>See methods section</i>	Yes
					No

Table 1: Summary of treatments. For all treatments but *CNC97*, the average mutation rate was set to 2.5×10^{-8} per site, per generation and the mean heterozygous selection coefficient to 0.1.