

MOLECULAR ECOLOGY

Background selection and the statistics of population differentiation: consequences for detecting local adaptation

Journal:	<i>Molecular Ecology</i>
Manuscript ID	MEC-18-0562
Manuscript Type:	Original Article
Date Submitted by the Author:	20-May-2018
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

Only

1

2 **Title:** Background selection and the statistics of population differentiation:
3 consequences for detecting local adaptation

4

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

6

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

9

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

11

12

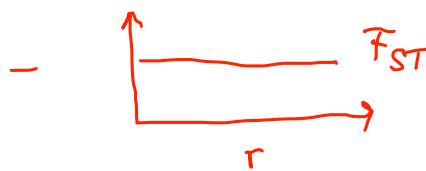
13 Abstract

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

*also
disrupts
the SFS*

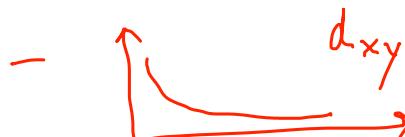
28

Claim :



$$F_{ST} \approx \frac{1}{1 + 8Nm}$$

$$T_w = 2N$$



$$T_B \approx 2N + \frac{1}{2m}$$

$$F_{ST} = \frac{H_T - \bar{H}_S}{H_T} = \frac{1 - f_T - (1 - \bar{f}_S)}{1 - f_T}$$

$$\bar{T} \approx \frac{1}{2} T_w + \frac{1}{2} T_B$$

$$= \frac{\bar{f}_S - f_T}{1 - f_T} = \frac{2N - (2N + \frac{1}{4m})}{1 - 2N - \frac{1}{4m}} = \frac{-1}{4m - 8Nm - 1} \approx \frac{1}{1 + 8Nm}$$

$$= N + N + \frac{1}{4m} = 2N + \frac{1}{4m}$$

$$4m \ll 8Nm$$

29 Introduction

30 Natural selection affects patterns of genetic diversity throughout the genome. How
31 selection affects genetic diversity on a single isolated locus is relatively easy to
32 model; however, when a large number of linked loci are considered, interactions
33 between evolutionary pressures at different sites render the task of modelling much
34 more difficult.

35 Maynard Smith & Haigh (1974) recognized the influence of selection on linked
36 neutral sites, proposing that strong positive selection could reduce genetic diversity
37 at nearby sites. This process is now referred to as a 'selective sweep'. Much later,
38 Charlesworth et al. (1993) proposed that deleterious mutations could also affect
39 genetic diversity at nearby sites, because some haplotypes would be removed from
40 the population as selection acts against linked deleterious alleles. They named this
41 process background selection (BGS). Both selective sweeps and background
42 selection affect genetic diversity; they both reduce the effective population size of
43 linked loci. Empirical evidence of a positive correlation between genetic diversity
44 and recombination rate has been reported in several species (Cutter and Payseur,
45 2013), including *Drosophila melanogaster* (Begin & Aquadro, 1992; Elyashiv et al.,
46 2016), humans (Spencer et al., 2006), collared flycatchers, hooded crows and
47 Darwin's finches (Dutoit et al., 2017; see also Vijay et al., 2017).

and distort
the SFS,
at least

49 BGS is also expected to affect F_{ST} (Charlesworth et al., 1997; Cutter & Payseur, 2013;
 50 Cruickshank & Hahn, 2014; Hoban et al., 2016). At low effective population size,
 51 different populations may randomly fix different alleles, increasing F_{ST} , while at high
 52 population size, allele frequency changes less through time and different
 53 populations are more likely to have comparable allele frequencies, keeping F_{ST} low.
 54 This negative relationship between effective population size N_e and F_{ST} is captured
 55 in Wright's classical infinite island result; $F_{ST} = \frac{1}{1+4N_e(m+\mu)}$ (Wright, 1943). One
 56 might therefore expect that loci under stronger BGS would show higher F_{ST} .

57 Many authors have also argued that, because BGS reduces the within-population
 58 diversity, it should lead to high F_{ST} (Cutter & Payseur, 2013; Cruickshank & Hahn,
 59 2014; Hoban et al., 2016). Expressed in terms of heterozygosities, $F_{ST} = \frac{H_T - H_S}{H_T} =$
 60 $1 - \frac{H_S}{H_T}$, where H_T is the expected heterozygosity in the entire population and H_S is
 61 the average expected heterozygosity within subpopulations (H_S and H_T are also
 62 sometimes called π_S and π_T ; e.g. Charlesworth, 1998). All else being equal, a
 63 decrease of H_S would indeed lead to an increase of F_{ST} . However, all else is not equal;
 64 H_T is also affected by BGS (Charlesworth et al., 1997). Therefore in order to
 65 understand the effects of BGS on F_{ST} , we must understand the relative impact of BGS
 66 on both H_S and H_T .

67 Performing numerical simulations, Charlesworth et al. (1997) report that BGS
 68 reduces the within population heterozygosity H_S slightly more than it reduces the
 69 total heterozygosity H_T , causing a net increase in F_{ST} . The effect on F_{ST} reported is

70 quite substantial, but, importantly, their simulations were not meant to be realistic.

71 The authors highlighted their goal in the methods:

check whl

F_{bp} was used

72 "The simulations were intended to show the qualitative effects of the various

73 forces studied [...], so we did not choose biologically plausible values [...].

74 Rather, we used values that would produce clear-cut effects".

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

76 per site:

77 "This unrealistically high value was used in order for background selection to

78 produce large effects [...]"

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

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

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

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

83 It is important to distinguish two separate questions when discussing the effect of

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

85 locus-to-locus variation in the intensity of BGS affect locus-to-locus variation in F_{ST} ?

86 The second question is of particular interest to those trying to identify loci under

87 positive selection (local selection or selective sweep). Locus-to-locus variation in F_{ST}

V | Due
to

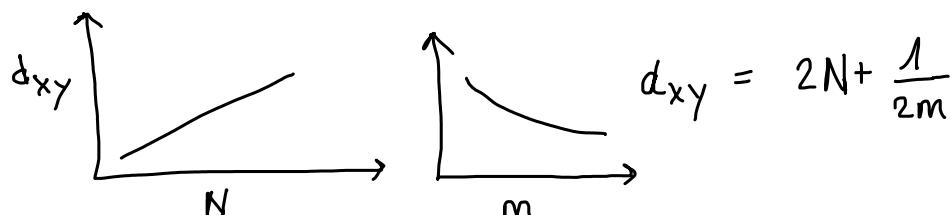
BGS

[reformulate ?]

88 potentially could be confounded with the F_{ST} peaks created by positive selection. In
 89 this paper, we focus on this second question.

90 The identification of loci involved in local adaptation is often performed via F_{ST}
 91 outlier tests (Lotterhos & Whitlock, 2014; Hoban et al., 2016). Other tests exist to
 92 identify highly divergent loci such as cross-population extended haplotype
 93 homozygosity (XP-EHH; Sabeti et al., 2007), comparative haplotype identity (Lange
 94 & Pool, 2016), cross-population composite likelihood ratio (XP-CLR; Chen et al.,
 95 2010). F_{ST} outlier tests, such as FDist2 (Beaumont & Nichols, 1996), BayeScan (Foll
 96 & Gaggiotti, 2008) or FLK (Bonhomme et al., 2010), look for genomic regions
 97 showing particularly high F_{ST} values to find candidates for local adaptation. If BGS
 98 can affect F_{ST} unevenly across the genome, then regions with a high intensity of BGS
 99 could potentially have high F_{ST} values that could be confounded with the pattern
 100 caused by local selection (Charlesworth et al., 1997; Cruickshank & Hahn, 2014).
 101 BGS could therefore inflate the false positive rate when trying to detect loci under
 102 local selection.

103 The potential confounding effect of BGS on signals of local adaptation has led to an
 104 intense effort trying to find solutions to this problem (Bank et al., 2014; Huber et al.,
 105 2016). Many authors have understood from Cruickshank and Hahn (2014) that d_{xy}
 106 should be used instead F_{ST} in outlier tests (e.g. McGee et al., 2015; Yeaman, 2015;
 107 Whitlock & Lotterhos, 2015; Brousseau et al., 2016; Picq et al., 2016; Payseur &
 108 Rieseberg, 2016; Hoban et al., 2016; Vijay et al., 2017; see also Nachman & Payseur,



109 2012). F_{ST} is a measure of population divergence relative to the total genetic
 110 diversity, while d_{XY} is an absolute measure of population divergence defined as the
 111 probability of non-identity by descent of two alleles drawn in the two different
 112 populations averaged over all loci (Nei, 1987; Nei, 1987 originally called it D_{XY} but,
 113 here, we follow Cruickshank and Hahn's, 2014 terminology by calling it d_{XY}). The
 114 argument is that because F_{ST} is a measure of divergence relative to the genetic
 115 diversity and d_{XY} an absolute measure of divergence and because BGS reduces
 116 genetic diversity, then BGS must affect F_{ST} but not d_{XY} , a claim that we will
 117 investigate in this paper.

d_{XY} : not affected vs. reduced

118 Whether BGS can affect genome-wide F_{ST} under some conditions is not in doubt
 119 (Charlesworth et al., 1997), but whether locus-to-locus variation in the intensity of
 120 BGS present in natural populations substantially affects variation in F_{ST} throughout
 121 the genome is very much unknown. Empirically speaking, it has been very difficult
 122 to measure how much of the genome-wide variation in genetic diversity is caused by
 123 BGS, as opposed to selective sweeps or variation in mutation rates (Cutter &
 124 Payseur, 2013; see also attempts in humans by Cai et al., 2009 and McVicker et al.
 125 2009). We are therefore in need of realistic simulations that can give us more
 126 insight into how BGS affects genetic diversity among populations and how it affects
 127 the statistics of population divergence.

128 In this article, we investigate the effect of BGS in structured populations with
 129 realistic numerical simulations using parameter estimates from humans and

Do we really need simulations
 for this?

depends
on N,
h!
Is this
really
what has
been
argued?
I do not
think so,
and if
would
be wrong

see / cite
more
literature
that
has looked

into this
(e.g. Elyashiv et al. a.
vs. Moreen)

130 stickleback. Our two main goals are 1) to quantify the impact of locus-to-locus
131 variation in the intensity of BGS on d_{XY} (Nei, 1987) and F_{ST} (Weir & Cockerham,
132 1984) and 2) to determine whether BGS inflates the false positive rate of F_{ST} outlier
133 tests.



How is the intensity of
BGS defined?

134 Methods

135 Our goal is to perform biologically plausible simulations of the local genomic effects
136 of background selection. BGS is expected to vary with gene density, mutation rate
137 and recombination rate across the genome. We used data from real genomes to
138 simulate realistic covariation in recombination rates and gene densities. We chose
139 to base our simulated genomes on two eukaryote genomes, sticklebacks and
140 humans, because these two species have attracted a lot attention in studies of local
141 adaptation and because sticklebacks have a variance in recombination rate which is
142 almost 15 times higher than humans (data not shown), allowing us to test vastly
143 different types of eukaryotic genomes. The recombination rate variation in humans
144 is extremely fine scale, but it presents the potential issue that it is estimated from
145 linkage disequilibrium data. As selection causes linkage disequilibrium to increase,
146 estimates of recombination rate at regions under strong selection may be under-
147 estimated, hence artificially increasing the simulated variance in the intensity of
148 BGS. Although the recombination map for stickleback is much less fine-scaled, the
149 estimates are less likely to be biased as they are computed from pedigrees.

and sd!

BGS
at start
& sentence

BT:
There
are ran-
dom based
maps!

which
might
bias our
inference...

150 Our simulations are forward in time and were performed using the simulation
151 platform SimBit version 3.69. The code and user manual are available at
152 <https://github.com/RemiMattheyDoret/SimBit>. To double check our results, we
153 also ran some simulations with SFS_code (Hernandez, 2008), confirming that we get
154 consistent distributions of genetic diversity and of FST among simulations (results
155 not shown). Generations are non-overlapping, individuals are hermaphrodites,
156 mating is random within patches and selection occurs before dispersal.

show,
or
do not
mention.

157 Genetics

158 For each simulation, we randomly sampled a sequence of about 10 cM coming either
159 from the stickleback (*Gasterosteus aculeatus*) genome or from the human genome
160 (see treatments below) and used this genomic location to determine the
161 recombination map and exon locations for a simulation replicate. For the stickleback
162 genome, we used the gene map and recombination map from Roesti et al. (2013).
163 Ensembl-retrieved gene annotations were obtained from Marius Roesti. For the
164 human 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 (Nachman & Crowell, 2000).

random
w. r. t.
what?
(w. r. t.
gen. map?)
o

justification
for exponential
distrib.?
Why
not
constant?

$$r_{bp} = 10^{-8} \text{ M/bp} \hat{=} 1 \text{ cM/Mb}$$

$$5 \text{ cM} \hat{=} 5 \text{ Mb} \rightarrow 0.05 \text{ Mb exons} \rightarrow$$

$$\text{guess } \mu \sim 5 \times 10^4 \cdot 2.5 \times 10^{-8} = 1.25 \cdot 10^3$$

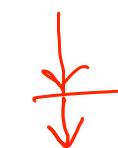
Ruminate on
this sampling regime!

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

6/11/2019

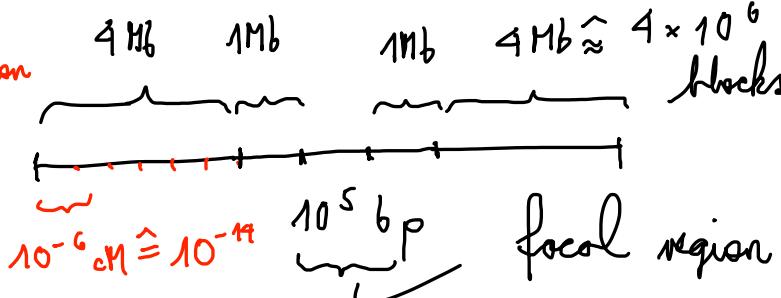
{ 2 nuclear;

see prev.
page 8



178 We simulated a 5 cM region on each side of the focal region (resulting in a window
179 of 10 cM plus the recombination rate present in the specific focal region of 10^5 sites)
180 in 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 with 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

target of
purifying selection



{ 2
What's the
Re pr

10

stickleback?

show distrib. of nr. of exonic sites
per flanking 5 Mb-region

specify what appears
you refer to (biallelic
sites or no recomb.?)

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 genome) is subject to purifying selection with a
198 selection coefficient against mutant alleles determined by a gamma distribution
199 described below. For focal regions that include exons, statistics are computed over a
200 sequence that is at least partially under direct purifying selection.

201 To create variance in selection pressures throughout the genome, each exon (and
202 regulatory sequence for the human genome) 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. To improve the performance of our
211 simulations, we used multiplicative dominance, where the fitness of heterozygotes
212 is at locus i is $1 - s_i$ and the fitness of the double mutant is $(1 - s_i)^2$.

Why such
a complicated
model?

cf.
of this?

How
does this
improve
performance?

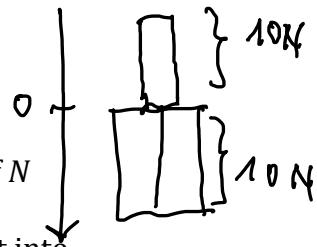
grossly overestimated?

e.g. McVicker et al. (2008): $\bar{s} \approx 0.0025$ term used elsewhere
for exonic regions; 10^{-5} for non-exonic ... ?

213 As a consequence of our parameter choices, our genome-wide deleterious mutation
214 rate was about 1.6 in sticklebacks and about 3 in humans. 9.4% of the stickleback
215 genome and 2.6% of the human genome was under purifying selection. For
216 comparison, the genome-wide deleterious mutation rate is estimated at 2.2 in
217 humans (Keightley, 2012) and 0.44 in rodents (Keightley & Gaffney, 2003). To our
218 knowledge, there is currently no such estimation for sticklebacks. Note however
219 that the above estimates cannot reliably detect mutations that are quasi lethal ($s \ll$ | 'neutral'?
220 $1/2N$). By our distribution of selective coefficients, 49% of all deleterious mutations
221 have a heterozygote selective coefficient lower than $1/2N_e$ when $N_e = 1,000$ (42% |
222 when $N_e = 10,000$). 6/21/2018

223 It is worth noting however that, in rodents, about half of the deleterious mutations s |
224 rate occurs in non-coding sequences (Keightley & Gaffney, 2003). Our simulations |
225 using human genome had all exons and all regulatory sequences under purifying |
226 selection. With our simulations based on the stickleback genome, however, only |
227 exons were under purifying selection. It is therefore possible that we would have |
228 over-estimated the deleterious mutation rate in gene-rich regions and under- |
229 estimated the deleterious mutation rate in other regions, especially in stickleback. |
230 This would artificially increase the locus-to-locus variation in the intensity of BGS in |
231 our simulations.

232 *Demography*



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

238 *Treatments*

239 We explored ~~the~~ presence and absence of deleterious mutations over two patch sizes,
 240 three migration rates, and two genomes.

241 We do not have a full factorial design. We considered a basic design and explored
 242 variations from this design. The basic design had a population size per patch of $N =$
 243 1000, a migration rate of $m = 0.005$ and used the stickleback genome for its
 244 recombination map and gene positions. As deviations from this basic design, we

245 explored modification of every variable, one variable at a time. The *Large N*
 246 treatment has $N = 10000$. The *Human* treatment uses the human genome for gene

247 positions, regulatory sequences and recombination map. The treatments No
 248 *Migration* and *High Migration* have migration rates of $m = 0$ and $m = 0.05$,
 249 respectively.

250 To test the robustness of our results and because it may be relevant for inversions,
 251 we also had unrealistic simulations where recombination rate for the entire genome
 252 was set at zero. As a check against previous work, we qualitatively replicated the

- basic
design
(default)

- large N
- human

$Nm \in$

{ 0, 5, 50 }
- no migration
- high migration

|
|

Note: CNC97 focus on equilibrium situation!

1 of 2

253 results ✓ Charlesworth et al. (1997) by performing simulations with similar
 254 assumptions as they used. We named this treatment *CNC97*. In our *CNC97*
 255 simulations, $N=2000$, $m=0.001$, and 1000 loci were all equally spaced at 0.1 cM apart
 256 from each other with constant selection pressure with heterozygotes having fitness
 257 of 0.98 and double homozygotes fitness of 0.9 and constant mutation rate $\mu =$
 258 0.0004.

$$Nm = 2$$

$$r = 10^{-9}$$

$$S = 0.02$$

s correct? h = ?

$$(1-s)^2 = \\ 1 - 2st + s^2$$

$$= 1 - 0.04$$

$$+ (0.04)^2$$

$$\neq 0.9$$

259 In all treatments (except *Large N*), we performed 4000 simulations; 2000
 260 simulations with BGS and 2000 simulations without selection (where all mutations
 261 were neutral). For *Large N*, simulations took more memory and more CPU time. We
 262 therefore could only perform 2000 simulations for *Large N*; 1000 simulations with
 263 background selection and 1000 simulations without selection. That represents a
 264 total of 26,000 simulations for 7 treatments. A full list of all treatments can be found
 265 in table 1.

266 We set the generation 0 at the time of the split. The state of each population was
 267 recorded at the end of the burn-in period (generation -1) and at generations
 268 $0.001 \times 2N$, $0.05 \times 2N$, $0.158 \times 2N$, $1.581 \times 2N$ and $5 \times 2N$ after the split. For
 269 $N=1000$, the sampled generations are therefore -1, 2, 100, 316, 3162 and 10000.

270 *Predicted intensity of Background Selection*

271 In order to investigate the locus-to-locus correlation between the predicted
 272 intensity of BGS and various statistics, we computed *B*, a statistic that approximates

273 the expected ratio of the coalescent time with background selection over the
 274 coalescent time without background selection ($B = \frac{T_{BGS}}{T_{neutral}}$). B quantifies how
 275 strong BGS is expected to be for a given simulation (Nordborg et al. 1996). A B value
 276 of 0.8 means that BGS has caused a drop of genetic diversity of 20% compared to a
 277 theoretical absence of BGS. Lower B values indicate stronger BGS.

278 Both Hudson & Kaplan (1995) and Nordborg et al. (1996) have derived theoretical
 279 expectations for B . We applied both methods and found that the predictions of the
 280 two formulas are highly correlated (Figure S2). Because the Hudson & Kaplan
 281 (1995) approach has been more popular in the literature (cited almost twice as
 282 often), we show only the B values computed following Hudson & Kaplan (1995):

$$283 B = \exp\left(-\sum_i \frac{u_i s_i}{(s_i + r_i(1-s_i))^2}\right)$$

284 where r_i is the recombination rate between the focal site and the i^{th} site under
 285 selection, and s_i is the heterozygous selection coefficient at that site. u_i is the
 286 mutation rate at the i^{th} site. By this formula, B is bounded between 0 and 1, where 1
 287 means no BGS at all and low values of B mean strong BGS. We computed B for all
 288 sites in the focal region and report the average B for the region.

289 For the stickleback genome, B values ranged from 2×10^{-6} to 1.0 with a mean of
 290 0.937 (Figure S2). For the human genome, B values ranged from 0.45 to 1.0 with a
 291 mean at 0.975. There is indeed less variability and much fewer extremely low B

↳ appears too high ; McVicker: mean 0.7 ± 0.81

292 values in the human genome. In the unrealistic *No Recombination* treatment, B
 293 values range from 0.00003 to 0.84 with a mean of 0.17.

294 F_{ST} outlier tests

295 In order to know the effect of BGS on outlier tests of local adaptation, we used a
 296 variant of FDist2 (Beaumont & Nichols, 1996). We chose FDist2 because it is a
 297 simple and fast method for which the assumptions of the test match well to the
 298 demographic scenario simulated here. Because the program FDist2 is not available
 299 through the command line, we rewrote the FDist2 algorithm in R and C++. Source
 300 code can be found at <https://github.com/RemiMattheyDoret/Fdist2>.

Explain!

301 Our FDist2 procedure is as follows; first, we estimated the migration rate from the
 302 average F_{ST} ($m = \frac{F_{ST}-1}{8 \cdot F_{ST} \cdot N}$; Charlesworth, 1998) and then running 50000 simulations
 303 each lasting for 50 times the half-life to reach equilibrium F_{ST} given the estimated
 304 migration rate (Whitlock, 1992). For each SNP, we then selected the subset of FDist2
 305 simulations for which allelic diversity was less than 0.02 away from the allelic
 306 diversity of the SNP of interest. The P -value is computed as the fraction of FDist2
 307 simulations within this subset having a higher F_{ST} than the one we observed. The
 308 false positive rate is then defined as the fraction of neutral SNPs for which the P -
 309 value is lower than a given α value. The α values explored are 0.1, 0.05, 0.01, 0.001,
 310 0.0001 and 0.00001.

$$\bar{F}_{ST} = \frac{1}{1+8M}$$

$$1+8M = \frac{1}{\bar{F}_{ST}}$$

$$8M = \frac{1}{\bar{F}_{ST}} - 1$$

$$= \frac{1 - \bar{F}_{ST}}{\bar{F}_{ST}}$$

~~•~~ $F_{ST} = \frac{1}{1+8Nm}$; $N \downarrow$ due to BGS $\Rightarrow \bar{F}_{ST}$ overestimated
 the scaled effective migration rate can be estimated ($N_{BGS}m$) but it contains the effect of BGS and is hence underestimated¹⁶

311 For the outlier tests, to avoid issues of pseudo-replication, we considered only a
 312 single SNP per simulation whose minor allele frequency is greater than 0.05. Then,
 313 we randomly assembled SNPs from a given treatment into groups of 500 SNPs to
 314 create the data file for FDist2. We have 4000 simulations (2000 with BGS and 2000
 315 without BGS) per treatment (*Large N* is an exception with only 2000 simulations
 316 total), which allowed 8 independent false positive rate estimates per treatment (4
 317 estimates with BGS and 4 without BGS). In each treatment, we tested for different
 318 false positive rate with and without BGS with both a Welch's *t*-test and a Wilcoxon
 319 test.

320 *Statistics*

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

325 There are two main estimators of F_{ST} in the literature; G_{ST} (Nei, 1973) and θ (Weir &
 326 Cockerham, 1984). In this article, we focus on θ as an estimate of F_{ST} (Weir &
 327 Cockerham, 1984). There are also two methods of averaging F_{ST} over several loci.
 328 The first method is to simply take an arithmetic mean over all loci. The second
 329 method consists at calculating the sum of the numerator of θ over all loci and
 330 dividing it by the sum of the denominator of θ over all loci. Weir and Cockerham
 331 (1984) showed that this second averaging approach has lower bias than the simple

WC : lower variance, higher bias } 17
 alt : higher variance, lower bias } 2)

But this
biases
towards
lower
 F_{ST}
seems
like
small
sample
sizes for
a *t* test

θ is a
param; $\hat{\theta}$
is the estimator

This is
the result
of what I
thought;
check!

332 arithmetic mean. We will refer to the first method as the “average of ratios” and to
333 the second method as “ratio of the averages” (Reynolds et al. 1983; Weir &
334 Cockerham, 1984). In this article, we use F_{ST} as calculated by “ratio of the averages”,
335 as advised by Weir and Cockerham (1984). To illustrate the effects of BGS on the
336 biased estimator of F_{ST} , we also computed F_{ST} as a simple arithmetic mean (“average
337 of the ratio”), and we will designate this statistic with a subscript $F_{ST(average\ of\ ratios)}$.

338 d_{XY} is a measure of genetic divergence between two populations X and Y. Nei (1987)
339 defined d_{XY} as

$$340 \quad d_{XY} = \frac{\sum_{l=1}^L \left(1 - \sum_{k=1}^{A_l} x_{l,k} y_{l,k} \right)}{L}$$

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

344 Some population genetics software packages (e.g., EggLib; De Mita and Siol, 2012)
345 average d_{XY} over polymorphic sites only, instead of averaging over all sites, as in
346 Nei's (1987) original definition of d_{XY} . This measure averaged over polymorphic
347 sites only will be called d_{XY-SNP} ; otherwise, we use the original definition of d_{XY} by Nei
348 (1987).

349 The statistics reported are the average F_{ST} , d_{XY} , and within population genetic
 350 diversity $H_S = \sum_{l=1}^L \left(1 - \sum_{k=1}^{A_l} x_{l,k}^2\right) / L$. For each treatment and at each generation, we
 351 computed five independent Pearson correlation tests between B and F_{ST} , F_{ST} (average of
 352 ratios), d_{XY} , d_{XY-SNP} and H_S . We compared our correlation tests with ordinary least
 353 squares regressions and robust regressions (using M-estimators; Huber, 1964), and
 354 the results were consistent.

Better to
use
{ Spearman
linear
regression
{ does not
use
line

6/22/2018

355 Results

356 Correlations between the statistics H_S , F_{ST} , F_{ST} (average of ratios), d_{XY} , and d_{XY-SNP} and B , are
 357 summarized in tables S1, S2, S3, S4 and S5, respectively. Figure 1 shows the means
 358 and standard errors for the treatments *Default*, *High Migration*, *Large N*, *Human* and
 359 *No Migration*. The same graphs for the treatments *No Recombination* and *CNC97* can
 360 be found in Figure S3.

361 Genetic Diversity

362 Genetic diversity within populations (H_S) is very similar among the treatments
 363 *Default*, *High Migration* and *Human* (around $H_S = 1.9 \times 10^{-4}$) but is about 1.9 times
 364 lower in the *No Migration* treatment ($H_S = 1.0 \times 10^{-4}$) and about 10 times higher in
 365 the *Large N* treatment ($H_S = 1.9 \times 10^{-3}$; Figure 1).

366 B is significantly correlated with genetic diversity within populations (H_S) for all
367 treatments using the stickleback genome (and at almost all generations) but not
368 with the *Human* treatment (table S1). Excluding the unrealistic treatments (*No*
369 *Recombination* and *CNC97*), simulations with BGS have a genetic diversity 4% to
370 20% lower than simulations without BGS (Figure 1, right column and Figure S3,
371 right column). In the *Human* treatment, there is no significant correlation of B and
372 H_S . Note that Pearson's correlation coefficients between B and H_S are always very
373 small even when the effect is highly significant. The largest R^2 observed in realistic
374 simulations (excluding in the *No Recombination* and *CNC97* treatments) between B
375 and H_S is $R^2 \approx 0.0121$.

6/29/2018
50 min.



376 *Statistics of population divergence*

377 Figure 2 shows the correlation between B and the statistics F_{ST} , d_{XY} and H_S for
378 *Default* at the last generation. These graphs highlight the general tendencies of the
379 treatments *Default*, *High Migration*, *Large N* and *No Recombination*. The strongest
380 correlation with B is observed for the statistics d_{XY} ($P = 3.28 \times 10^{-5}$, $R = 0.093$) and H_S
381 ($P = 3.1 \times 10^{-5}$, $R = 0.093$). In fact, the two statistics d_{XY} and H_S are very highly
382 correlated ($P < 2.2 \times 10^{-16}$, $R = 0.99$). This high correlation explains the
383 resemblance between the central and right graphs of figure 2. F_{ST} is not correlated
384 with B ($P = 0.99$, $R = 10^{-4}$). All correlation tests between B and the statistics F_{ST} , F_{ST}
385 (*average of ratios*), d_{XY} , d_{XY-SNP} can be found in Tables S2, S3, S4 and S5, respectively.

We do not need simulation to know this?

\$\swarrow\$
linear correlation
not appropriate!

$$F_{ST} = \frac{\pi_T - B\pi_0}{\pi_T} ; d_{XY} = \left(\frac{1}{2m} + 2B\pi_0 \right) 2m$$

386 The *No Migration* treatment is an exception to the other treatments. F_{ST} is not
 387 significantly correlated with B at early generations but become slightly correlated as
 388 divergence rises to 0.6 and higher. d_{XY} shows an opposite pattern. d_{XY} is very
 389 significantly correlated with B at early generations and seemingly independent of B
 390 at the last generation. Note that for both d_{XY} and F_{ST} , all correlation coefficients are
 391 always very small. The largest R^2 observed is $R^2=0.0121$ in realistic simulations
 392 (found for F_{ST} *No Migration* and for d_{XY} Large N; Tables S2, S4) and $R^2=0.0256$ for
 393 the *No Recombination* treatment (found for d_{XY} ; Table S4).

394 As expected, in the CNC97 simulations, there is a strong difference between
 395 simulations with BGS and simulations without BGS for all three statistics (F_{ST} , d_{XY} ,
 396 and H_S) at all generations (Welch's t -tests; all $P < 2.2 \times 10^{-16}$; Figure S3).

397 F_{ST} calculated as advised by Weir and Cockerham (1984) was generally less sensitive
 398 to BGS than F_{ST} calculated as an average of ratios (compare tables S2 and S3). Figure
 399 S4 illustrates the sensitivity of F_{ST} (*average of ratios*) in the worst case, the *No*
 400 *Recombination* treatment. This sensitivity is driven largely by rare alleles and goes
 401 away when minor alleles below a frequency of 0.05 are excluded.

402 *F_{ST} outlier tests*

403 The observed false positive rate is relatively close to the α values except for *No*
 404 *Migration* (with and without BGS) and *CNC97* (with BGS). With the exception of
 405 treatments *No Migration* and *CNC97*, there is no significant difference in false

This is
the
carrying
problem

because
Hc does
not increase
anymore

appropriate
to filter?

406 positive rates between simulations with BGS and those without BGS for α of 0.05

407 (Figure 3). In the treatment, No Migration, the false positive rates between ||

408 simulations with and without BGS are significantly different for the latter

409 generations. In this treatment, the false positive rate for both simulations with and

410 without BGS are much higher than the α value of 0.05. Results remain very

411 congruent for other α values (results not shown).

412 **Discussion**

m might be too high

*explanation
for high
false
positive
rate w/o
BGS?*

413 Background selection reduces genetic diversity, both within and among populations,

414 but the effect on F_{ST} is rather small. In simulations of interconnected populations

415 with realistic parameters, F_{ST} is insensitive to BGS while the absolute measure of

416 divergence d_{XY} is affected by BGS. On the other hand, in highly diverged populations

417 unconnected by migration, BGS can have a greater impact on F_{ST} and a lower impact

418 on d_{XY} . The effects of BGS, when observed, are always very small with R^2 never over

419 1.1%.

*This is
because
m is
too high
in all
but the
'No migr.'
scenario.
In above?*

420 BGS impacts both total and within population genetic diversity. Excluding the

421 treatments *No recombination* and *CNC97*, we observe that simulations with BGS

422 have a genetic diversity (whether H_T or H_S ; H_T data not shown) 6% to 16% lower

423 than simulations without BGS. Messer & Petrov (2013) simulated a panmictic

424 population, looking at a sequence of similar length inspired from a gene-rich region

425 of the human genome, and reported a similar decrease in genetic diversity. Under

426 the *No Recombination* treatment, this average reduction of genetic diversity due to

427 BGS is 53%. Although empirical estimates are very complex and can hardly
428 disentangle BGS from selective sweeps, our results are also comparable with
429 empirical estimates. Reduction in genetic diversity in humans between regions
430 under high BGS are estimated at 6% according to Cai et al. (2009) or 19-26%
431 according to McVicker et al. (2009). In *Drosophila melanogaster*, where gene density
432 is higher, the reduction in genetic diversity due to BGS is estimated at 36% when
433 using Kim & Stephan (2000)'s methodology and is estimated at 71% reduction using
434 a composite likelihood approach (Elyashiv et al., 2016) and is hence closer to our **No**
435 **Recombination treatment** than to the other **treatments**. It is worth noting that,
436 because we were interested in simulating variance among sites in effects of BGS, we
437 only simulated local effects and therefore underestimate the expected genome-wide
438 effect of BGS on H_T .

This
does
not
make
sense

439 In contrast to measures of heterozygosity, F_{ST} was generally not significantly
440 correlated with B . The only exception is for the *No Migration* treatment, where, after
441 many generations, as the average F_{ST} becomes very high ($F_{ST} > 0.5$), we observe a
442 slight, yet significant, negative correlation between the expected effects of BGS, B ,
443 and F_{ST} (intense BGS lead to high values of F_{ST}). This highlights that F_{ST} is not
444 completely insensitive to BGS, but F_{ST} is largely robust to BGS.

not
surprising
given
the
high m

2 the scope

of the

study

does not

justify

such a

23

generic
statement

study the literature! → Corbett - Delig

448 genomes, they are not good representatives of more compact genomes such as
449 bacterial genomes or yeasts. Our simulations used randomly mating diploid
450 populations. Non-random mating, selfing, and asexual reproduction could also affect
451 our general conclusion, and potentially strongly increase the effects of BGS on F_{ST}
452 (Charlesworth *et al.* 1997). We have explored two population sizes, but we could not
453 explore population sizes of the order of a million individuals (like *Drosophila*
454 *melanogaster*) and still realistically simulate such long stretch of DNA. It is not
455 impossible that a much greater population size or a more complex demography
456 could yield to BGS having a greater effect on F_{ST} than what we observed here (Torres
457 et al. 2017).

- But
| can you
| use a
| scaling
| argument?
- What
| about
| Elyashiv?

458 Some have argued that, because BGS reduces the within population diversity, it
459 should lead to high F_{ST} (Cutter & Payseur, 2013; Cruickshank & Hahn, 2014; Hoban
460 et al., 2016). All else being equal, this statement is correct. However, BGS reduces H_T
461 almost as much as H_S (Figure 4). It is therefore insufficient to consider only one
462 component, and we must consider the ratio of these two quantities captured by the
463 definition of F_{ST} , $F_{ST} = 1 - \frac{H_S}{H_T}$. This ratio, as we have shown, appears to be relatively
464 robust to BGS. While genome-wide BGS might eventually be strong enough to cause
465 departures with F_{ST} values, it appears that locus-to-locus variation in the intensity of
466 BGS is not strong enough to have much impact on F_{ST} as long as populations are not
467 too highly diverged.

| Only with
| high migra-
| tion!

| if there
| are no
| local
| effects,

468 We also investigated the consequences of BGS on the widely-used but imperfect
 469 estimator, F_{ST} (*average of ratios*), for which F_{ST} measures for each locus are averaged to
 470 create a genomic average. It is well known that F_{ST} (*average of ratios*) is a biased way to
 471 average F_{ST} over several loci (Weir & Cockerham, 1984); however, its usage is
 472 relatively common today. In our simulations, F_{ST} (*average of ratios*) is more affected by
 473 BGS than F_{ST} . Interestingly, F_{ST} (*average of ratios*) is most often higher with weaker BGS.
 474 The directionality of this correlation may seem unintuitive at first. To understand
 475 this discrepancy, remember that BGS affects the site frequency spectrum; we
 476 observed that BGS leads to an excess of loci with low H_T (results not shown but see
 477 Charlesworth et al., 1995; see also contrary expectation in Stephan, 2010). Loci
 478 associated with very low H_T also have low F_{ST} (figure S5), a well-known result
 479 described by Beaumont and Nichols (1996). As BGS creates an excess of loci with
 480 low H_T and loci with low H_T tend to have low F_{ST} , BGS can actually reduce F_{ST} (*average of*
 481 *ratios*). After filtering out SNPs with a minor allele frequency lower than 5%, most of
 482 the correlation between F_{ST} (*average of ratios*) and B is eliminated (Figure S4).

So far,
my perception
was that
Weir &
Cockerham
is more
often used!
And : is
this really
the
average
of ratios ?

483 The absolute measure of divergence d_{XY} is sensitive to BGS. Regions of stronger BGS
 484 are associated with low d_{XY} . The effect, although significant, is of relatively small
 485 size. The expected d_{XY} for neutral loci is $d_{XY} = 4N\mu + 2t\mu$ (Nei, 1987), where t is the
 486 time in generation since the populations started to diverge. $4N\mu$ is the expected
 487 heterozygosity in the ancestral population (before splitting) and $2t\mu$ is the expected
 488 number of mutations fixed over time in either population since the population split.
 489 BGS does not affect the rate of fixation of mutations arising after the populations

BUT:
BGS
also,
and mostly,
reduces
 H_S !

↓
explanation
not satisfactory

not true;
BGS does not
affect $t \propto \mu$

490 diverged, but BGS affects the expected heterozygosity. Therefore, BGS should affect
 491 d_{XY} by its effect on the expected heterozygosity, and this effect should be greater
 492 early in divergence when the $4N\mu$ term is large relative to the fixation term. This is
 493 consistent with the results of our simulations. This result is in agreement with Vijay
 494 et al. (2017) who reported a strong correlation between H_S and d_{XY} when F_{ST} is low
 495 ($F_{ST} \approx 0.02$), but this correlation breaks down when studying more distantly related
 496 populations ($F_{ST} \approx 0.3$).

BUT:
 $4N\mu > \frac{1}{m}$
 if m
 is large
 at least /
 for $t \rightarrow \infty$
 where t
 is the
 split time

497 Interestingly, d_{XY} becomes less sensitive to BGS when F_{ST} becomes more sensitive.
 498 While our methodology does not allow us to test the efficiency of d_{XY} in outlier tests,
 499 it is possible that d_{XY} could be used for highly divergent lineages, but not in cases
 500 when divergence is relatively low. Cruickshank and Hahn (2014) suggested relying
 501 more on d_{XY} than F_{ST} for finding highly divergent loci. Their conclusion was based on
 502 analysis of a dataset involving highly divergent populations only (F_{ST} values range
 503 from about 0.38 to about 0.8). Based on our simulations, their conclusion is not valid
 504 when populations are not highly diverged.

C But this is not well supported!
 be measured
 not by
 F_{ST} !

505 As BGS also leads to a reduction of the number of polymorphic sites, BGS has an
 506 even stronger effect on d_{XY-SNP} than on d_{XY} (Figure S4). (The measure that we call d_{XY-SNP}
 507 is d_{XY} improperly calculated based only on polymorphic sites, as is done in some
 508 software packages.) This result highlights the importance of not blindly trusting the
 509 output of a given software package.

510 Outside the effect of BGS on N_e , there are at least two other possible factors that can

$$\begin{array}{cccccc} \text{AA} & \text{Aa} & \text{aa} \\ 1 & 1-s & (1-s)^2 \end{array} \quad \text{e.g. } s=0.2 : \quad \begin{array}{ccc} 1 & 0.8 & 0.64 \end{array}$$

- 511 potentially affect the correlation between B and μ : the effect of deleterious mutations on the effective migration rate and the auto-correlation of μ . Because most deleterious mutations are recessive (García-Dorado and Caballero, 2000; Peters et al., 2003; Shaw & Chang, 2006), the offspring of migrants, who enjoy an increased heterozygosity compared to local individuals, will be at a selective advantage. The presence of deleterious mutations therefore lead to an increase in the effective migration rate (Ingvarsson & Whitlock, 2000). This increases the effective migration rate and hence, leads to a decrease in F_{ST} .
- opposite might be happening! also, see Harris et al. (2016) a. Turic et al. (2016)*
- 519 As mutation rate is auto-correlated throughout the genome, neutral sequences closely linked to sequences that frequently receive deleterious mutation are also likely to experience frequent neutral mutations. As a high mutation rate leads to low F_{ST} values ($F_{ST} \approx \frac{1}{1+4Ne(m+\mu)}$, Wright 1943), autocorrelation in mutation rate may also impact the correlation between B and F_{ST} . This effect is likely to be negligible as long as $m \gg \mu$.
- 525 Recently, evidence of a correlation between recombination rate and F_{ST} has been interpreted as likely being caused by deleterious mutations rather than positive selection, whether the divergence between populations is very high (e.g. Cruickshank & Hahn, 2014), moderately high (Vijay et al., 2017) or moderately low (Torres et al., 2017). Here we showed the BGS is unlikely to explain all of these correlations. It can be hypothesized that positive selection (selective sweeps and local adaptation) could be the main cause of this correlation.

BUT:
In your simulations, heterozygous individuals are still suffering a disadvantage!

BUT:
also a higher u_d → $F_{ST} \uparrow$

Where is the

27 evidence for the sweeps?

532 McVicker et al. (2009) attempted an estimation of B values in the human genome
533 (see also Elyashiv et al., 2016). They did so using equations from Nordborg et al.
534 (1996). As there is little knowledge about the strength of selection throughout the
535 genome, to our understanding, this estimation of B values should be highly
536 influenced by the effects of beneficial mutations as well as deleterious mutations.

537 Torres et al. (2017) reused this dataset and found a slight association between B and
538 F_{ST} among human lineages. It is plausible that this correlation between B and F_{ST}
539 could be driven by positive selection rather than by deleterious mutations.

540 Our FDist2 analysis shows that the false positive rate does not differ in simulations
541 with BGS or without BGS. The only exceptions concern the unrealistic CNC97
542 treatment and the *No Migration* treatment after many generations (Figure 3). The
543 average F_{ST} at the last generation of the *No Migration* treatment is greater than 0.8.
544 With such high F_{ST} both the simulation without BGS and with BGS lead to very high
545 false positive rates (0.45 without BGS and 0.5 with BGS). This difference in false
546 positive rates is significant, but the observed high false positive rate should make it
547 clear that for such highly diverged populations, F_{ST} outlier tests are not
548 recommended in general.

549 Many authors (Cutter & Payseur, 2013; Hoban et al., 2016) have raised concerns
550 that BGS can strongly reduce our ability to detect the genomic signature of local
551 adaptation. Our analysis shows that BGS is not a strong confounding factor to F_{ST}
552 outlier tests of populations that are not too highly diverged.

| BUT:
| you should
| provide
| evidence!

| Also:
| Elyashiv
| et al. 2016
| nicely
| show
| that BGS
| is needed
| to explain
| the Div.
| data!

| validate
| using
| some
| theory

553 **Acknowledgments**

554 Many thanks to Loren H. Rieseberg, Sarah P. Otto and Amy L. Angert for their help in
555 discussing the design of the project and for feedback. Thanks to Sarah P. Otto,
556 Darren E. Irwin, and Bret A. Payseur for helpful comments on the manuscript. We
557 also thank Yaniv Brandvain and Tom Booker for their feedback and Marius Roesti
558 for help with the Ensembl-retrieved gene annotations.

559 **Funding**

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

563 **References**

564 Bank C, Ewing GB, Ferrer-Admetla A, Foll M, Jensen JD (2014) Thinking too
565 positive? Revisiting current methods of population genetic selection
566 inference. *Trends in Genetics*, **30**, 540–546.

567

568 Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of
569 population structure. *Proceedings of the Royal Society B: Biological Sciences*, **263**,
570 1619–1626.

571

572 Begun DJ, Aquadro CF (1992) Levels of naturally occurring DNA polymorphism
573 correlate with recombination rates in *D. melanogaster*. *Nature*, **356**, 519–520.

| d.

574

575 Bonhomme M, Chevalet C, Servin B *et al.* (2010) Detecting selection in population
576 trees: the Lewontin and Krakauer test extended. *Genetics*, **186**, 241–262.

577

578 Brousseau L, Postolache D, Lascoux M *et al.* (2016) Local adaptation in European
579 Firs assessed through extensive sampling across altitudinal gradients in southern
580 Europe. *Plos One*, **11**.

page range ?

|

581

582 Cai JJ, Macpherson JM, Sella G, Petrov DA (2009) Pervasive hitchhiking at coding and
583 regulatory sites in humans. *PLoS Genetics*, **5**.

page range ?

(

584

585 Charlesworth B, Morgan MT, Charlesworth D (1993) The effect of deleterious
586 mutations on neutral molecular variation. *Genetics*, **134**, 1289–1303.

587

588 Charlesworth B (1998) Measures of divergence between populations and the effect
589 of forces that reduce variability. *Molecular Biology and Evolution*, **15**, 538–543.

590

591 Charlesworth B, Nordborg M, Charlesworth D (1997) The effects of local selection,
592 balanced polymorphism and background selection on equilibrium patterns of
593 genetic diversity in subdivided populations. *Genetical Research*, **70**, 155–174.

594

595 Charlesworth D, Charlesworth B, Morgan MT (1995) The pattern of neutral
596 molecular variation under the background selection model. *Genetics*, **141**, 1619–
597 1632.

598

599 Chen H, Patterson N, Reich D (2010) Population differentiation as a test for selective
600 sweeps. *Genome Research*, **20**, 393–402.

601

602 Cruickshank TE, Hahn MW (2014) Reanalysis suggests that genomic islands of
603 speciation are due to reduced diversity, not reduced gene flow. *Molecular
604 Ecology*, **23**, 3133–3157.

605

606 Cutter AD, Payseur BA (2013) Genomic signatures of selection at linked sites:
607 Unifying the disparity among species. *Nature Reviews Genetics*, **14**, 262–274.

608

609 Dutoit L, Vijay N, Mugal CF *et al.* (2017) Covariation in levels of nucleotide diversity
610 in homologous regions of the avian genome long after completion of lineage
611 sorting. *Proceedings of the Royal Society B: Biological Sciences*, **284**, 20162756. (?)

612

613 Elyashiv E, Sattath S, Hu TT *et al.* (2016) A genomic map of the effects of linked
614 selection in *Drosophila*. *PLOS Genetics*, **12**. (Page range ?)

615

616 Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci
617 appropriate for both dominant and codominant markers: A Bayesian
618 perspective. *Genetics*, **180**, 977–993.

619

620 García-Dorado A, Caballero A (2000) On the average coefficient of dominance of
621 deleterious spontaneous mutations. *Genetics*, **155**, 1991–2001.

622

623 Gilbert KJ, Sharp NP, Angert AL *et al.* (2017) Local adaptation interacts with
624 expansion load during range expansion: Maladaptation reduces expansion load. *The*
625 *American Naturalist*, **189**, 368–380.

626

627 Hernandez RD (2008) A flexible forward simulator for populations subject to
628 selection and demography. *Bioinformatics*, **24**, 2786–2787.

629

630 Hoban S, Kelley JL, Lotterhos KE *et al.* (2016) Finding the genomic basis of local
631 adaptation: Pitfalls, practical solutions, and future directions. *The American*
632 *Naturalist*, **188**, 379–397.

633

634 Huber CD, Degiorgio M, Hellmann I, Nielsen R (2016) Detecting recent selective
635 sweeps while controlling for mutation rate and background selection. *Molecular*
636 *Ecology*, **25**, 142–156.

637

638 Huber PJ (1964) Robust estimation of a location parameter. *The Annals of
639 Mathematical Statistics*, **35**, 73–101.

640

641 Hudson RR, Kaplan NL (1995) Deleterious background selection with
642 recombination. *Genetics*, **141**, 1605–1617.

643

644 Ingvarsson PK, Whitlock MC (2000) Heterosis increases the effective migration
645 rate. *Proceedings of the Royal Society B: Biological Sciences*, **267**, 1321–1326.

646

647 International HapMap Consortium (2007) A second generation human haplotype
648 map of over 3.1 million SNPs. *Nature*, **449**, 851.

649

650 Keightley PD, Gaffney DJ (2003) Functional constraints and frequency of deleterious
651 mutations in noncoding DNA of rodents. *Proceedings of the National Academy of
652 Sciences*, **100**, 13402–13406.

653

654 Keightley PD (2012) Rates and Fitness Consequences of new mutations in
655 humans. *Genetics*, **190**, 295–304.

656

657 Kim Y, Stephan W (2000) Joint effects of genetic hitchhiking and background
658 selection on neutral variation. *Genetics*, **155**, 1415–1427.

659

660 Lange JD, Pool JE (2016) A haplotype method detects diverse scenarios of local
661 adaptation from genomic sequence variation. *Molecular Ecology*, **25**, 3081–3100.

662

663 Lotterhos KE, Whitlock MC (2014) Evaluation of demographic history and neutral
664 parameterization on the performance of F_{ST} outlier tests. *Molecular Ecology*, **23**,
665 2178–2192.

666

667 Maynard Smith J, Haigh J (1974) The hitch-hiking effect of a favourable
668 gene. *Genetical Research*, **23**, 23.

1 page range ?

669

670 McGee MD, Neches RY, Seehausen O (2015) Evaluating genomic divergence and
671 parallelism in replicate ecomorphs from young and old cichlid adaptive
672 radiations. *Molecular Ecology*, **25**, 260–268.

673

674 McVicker G, Gordon D, Davis C, Green P (2009) Widespread genomic signatures of
675 natural selection in hominid evolution. *PLoS Genetics*, **5**.

676

677 Messer PW, Petrov DA (2013) Frequent adaptation and the McDonald-Kreitman
678 test. *Proceedings of the National Academy of Sciences*, **110**, 8615–8620.

679

680 Mita SD, Siol M (2012) EggLib: processing, analysis and simulation tools for
681 population genetics and genomics. *BMC Genetics*, **13**, 27.

682

683 Nachman MW, Crowell SL (2000) Estimate of the mutation rate per nucleotide in
684 humans. *Genetics*, **156**, 297–304.

685

686 Nachman MW, Payseur BA (2012) Recombination rate variation and speciation:
687 Theoretical predictions and empirical results from rabbits and mice. *Philosophical
688 Transactions of the Royal Society B: Biological Sciences*, **367**, 409–421.

689

690 Nei M (1973) Analysis of gene diversity in subdivided populations. *Proceedings of
691 the National Academy of Sciences*, **70**, 3321–3323.

692

693 Nei M (1987) *Molecular evolutionary genetics*. Columbia University Press, New York
694 (N.Y.).

695

696 Nordborg M, Charlesworth B, Charlesworth D (1996) The effect of recombination on
697 background selection. *Genetical Research*, **67**, 159.

698

699 Payseur BA, Rieseberg LH (2016) A genomic perspective on hybridization and
700 speciation. *Molecular Ecology*, **25**, 2337–2360.

701

702 Peters AD, Halligan DL, Whitlock MC, Keightley PD (2003) Dominance and
703 overdominance of mildly deleterious induced mutations for fitness traits in
704 *Caenorhabditis elegans*. *Genetics*, **165**, 589–599.

705

706 Picq S, Mcmillan WO, Puebla O (2016) Population genomics of local adaptation
707 versus speciation in coral reef fishes (*Hypoplectrus spp*, *Serranidae*). *Ecology and*
708 *Evolution*, **6**, 2109–2124.

709

710 Reynolds J, Weir BS, Cockerham CC (1983) Estimation of the coancestry coefficient:
711 Basis for a short-term genetic distance. *Genetics*, **105**, 767–779.

712

713 Roesti M, Moser D, Berner D (2013) Recombination in the threespine stickleback
714 genome-patterns and consequences. *Molecular Ecology*, **22**, 3014–3027.

715

716 Sabeti PC, Varilly P, Fry B *et al.* (2007) Genome-wide detection and characterization
717 of positive selection in human populations. *Nature*, **449**, 913.

718

719 Shaw RG, Chang S-M (2006) Gene action of new mutations in *Arabidopsis*
720 *thaliana*. *Genetics*, **172**, 1855–1865.

721

722 Spencer CCA, Deloukas P, Hunt S *et al.* (2006) The influence of recombination on
723 human genetic diversity. *PLoS Genetics*, **2**.

724

725 Stephan W (2010) Genetic hitchhiking versus background selection: the controversy
726 and its implications. *Philosophical Transactions of the Royal Society B: Biological
727 Sciences*, **365**, 1245–1253.

728

729 Torres R, Szpiech ZA, Hernandez RD (2017) Human demographic history has
730 amplified the effects of background selection across the genome.

| complete
now null
it is out |

731

732 Vijay N, Weissensteiner M, Burri R *et al.* (2017) Genomewide patterns of variation in
733 genetic diversity are shared among populations, species and higher-order
734 taxa. *Molecular Ecology*, **26**, 4284–4295.

735

736 Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population

/

737 structure. *Evolution*, **38**, 1358.

page range ?

738

739 Whitlock MC (1992) Temporal fluctuations in demographic parameters and the

/

740 genetic variance among populations. *Evolution*, **46**, 608.

page range ?

741

742 Whitlock MC, Lotterhos KE (2015) Reliable detection of loci responsible for local

/

743 adaptation: Inference of a null model through trimming the distribution of F_{ST} . *The*

744 *American Naturalist*, **186**.

page range ?

745

746 Wright S (1943) Isolation by distance. *Genetics*, **28**, 114–138.

747

748 Yeaman S (2015) Local adaptation by alleles of small effect. *The American*

/

749 *Naturalist*, **186**.

page range ?

750

751 Zeng K, Corcoran P (2015) The effects of background and interference selection on

/

752 patterns of genetic variation in subdivided populations. *Genetics*, **201**, 1539–1554.

753

754 Zeng K, Charlesworth B (2011) The joint effects of background selection and genetic
755 recombination on local gene genealogies. *Genetics*, **189**, 251–266.

756

757 Zerbino DR, Achuthan P, Akanni W *et al.* (2017) Ensembl 2018. *Nucleic acids*
758 *research*, **46**, D754–D761.

759

760

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.

Treatment	N	m	Genome	Recombination	BGS
<i>Default</i>	1000	0.005	<i>Stickleback</i>	Yes	Yes
					No
<i>No Migration</i>	1000	0	<i>Stickleback</i>	Yes	Yes
					No
<i>High Migration</i>	1000	0.05	<i>Stickleback</i>	Yes	Yes
					No
<i>Large N</i>	10000	0.005	<i>Stickleback</i>	Yes	Yes
					No
<i>Human</i>	1000	0.005	<i>Human</i>	Yes	Yes
					No
<i>No Recombination</i>	1000	0.005	<i>Stickleback</i>	No	Yes
					No
<i>CNC97</i>	2000	0.001	NA	Yes	Yes
					No

$$Nm \in \{0, 0.00\delta N, 0.05N\}$$

$$N=1000 = \{0, 5, 50\}$$

unrealistically
high migration
for humans!

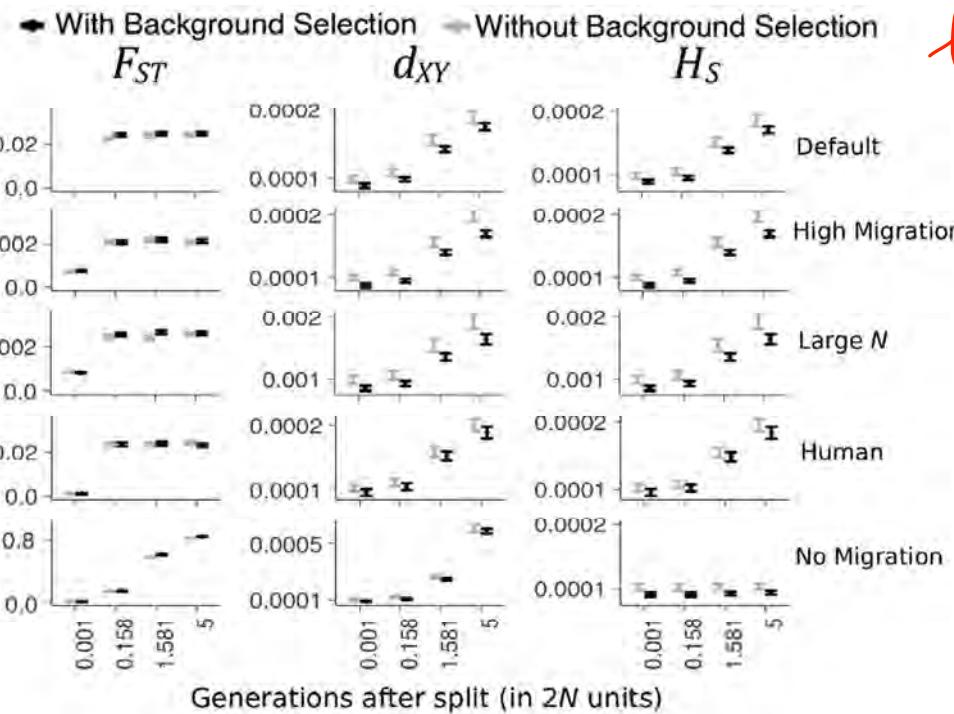


Figure 1: Comparisons of means F_{ST} (left column), d_{XY} (central column), and H_S (right column) between simulations with (black) and without (grey) BGS. Similar graphs for the treatments *No Recombination* and *CNC97* are in figure S3. Error bars are 95% CI.

$$d_{XY} = \left(\frac{1}{2m} + 2N \right) \cdot 2u$$

✓ hard to believe

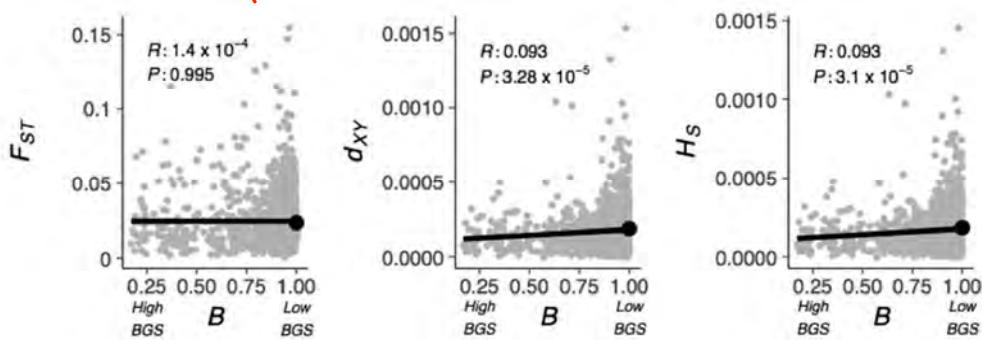


Figure 2: Correlation between B and F_{ST} , H_S , and d_{XY} for the last generation ($5 \times 2N$ generations after the split) of the *Default* treatment. Each grey dot is a single simulation where there is BGS. The large black dot is the mean of the simulations where BGS was artificially turned off. The P -values are computed from a Pearson's correlation test. P -values and R are computed on the simulations with BGS (grey dots) only.

$$F_{ST} \approx \frac{\bar{\pi}_T - \bar{\pi}_0 \cdot B}{\bar{\pi}_T} = 1 - B \cdot \frac{\bar{\pi}_0}{\bar{\pi}_T}$$

$\bar{\pi}_T \geq \bar{\pi}_0$

↳ changes!

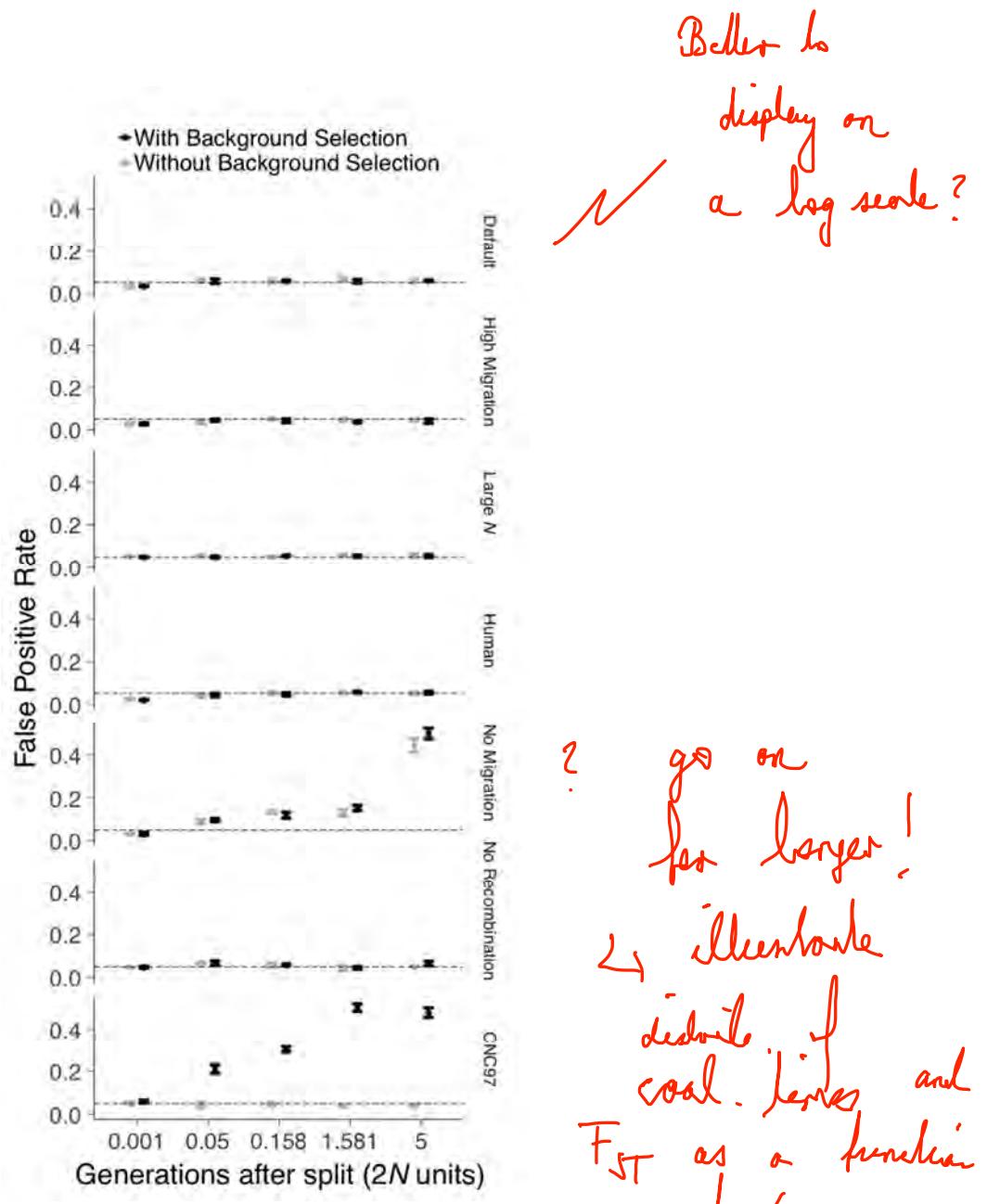


Figure 3: Comparison of false positive rate (FPR) returned by FDist2 between simulations with BGS (black) and without BGS (grey) for all treatments by generation. The significance level is 0.05 and is represented by the horizontal dashed line. Significance based on a Welch's *t*-test is indicated with stars ($2 \times 10^{-16}^{***}$, 0.001 '**', 0.01 **', 0.05 *', 0.1 ' ', 1). Significance levels are the same with Wilcoxon tests.

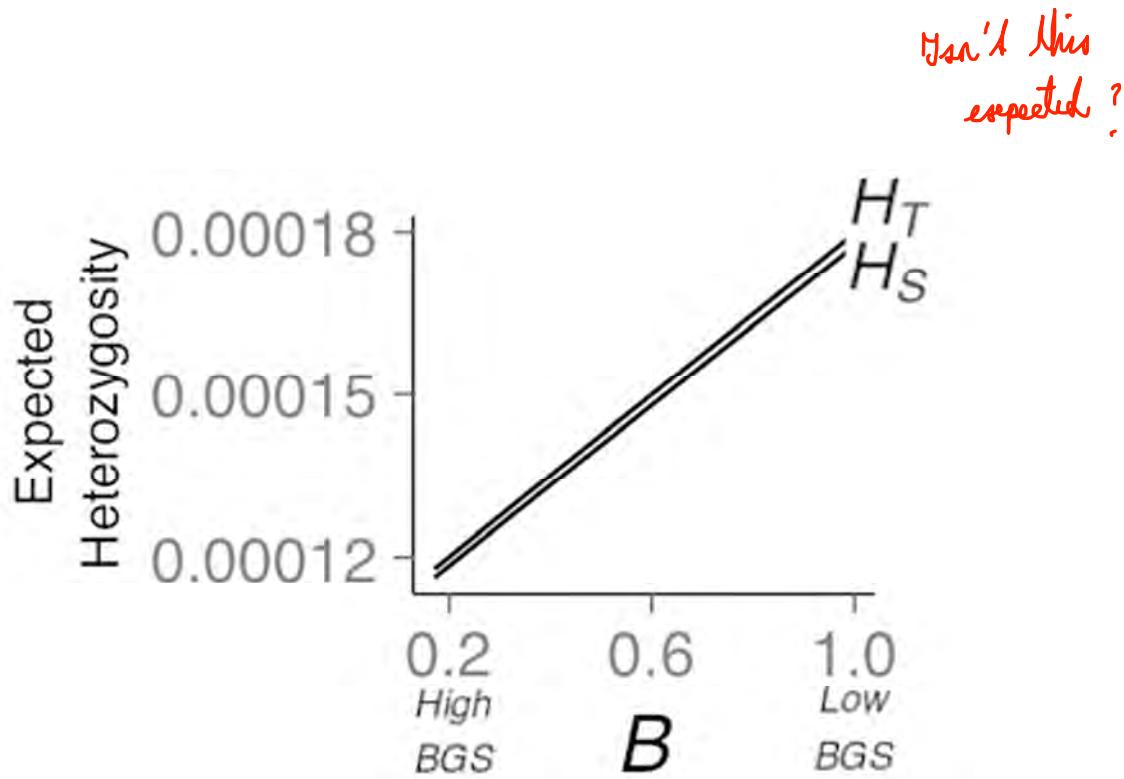


Figure 4: Regressions of total (H_T ; upper line) and within (H_S ; lower line) population expected heterozygosity on the coefficient of BGS (B) for the last generation of the **Default** treatment. The two regression lines are not exactly parallel with H_S tending to H_T as B goes to low values (more intense BGS).

$$B = \frac{H_S}{H_0} = a \cdot H_S$$

$$a = \frac{1}{H_0}$$

$$\Rightarrow H_S = \underbrace{\frac{1}{a}}_b \cdot B$$