

Review of manuscript MEC-19-0061 by Matthey-Doret & Whitlock

Summary

Old

In their manuscript “Background selection and the statistics of population differentiation: consequences for detecting local adaptation”, Remi & Whitlock study the effect of variation in the intensity of background selection (BGS) along the genome, and address whether this variation affects the detectability of local adaptation. Accepting that BGS can affect genome-wide average patterns of genetic diversity, the authors focus on small-scale (locus-to-locus) variation in the intensity of BGS as it supposedly arises in real genomes of natural populations. This question is of great interest to the field, as sensible interpretation of patterns of genomic variation requires a detailed understanding of the joint signals of various evolutionary forces, and of the confounding effects individual forces have on the genomic signature of other forces. Remi & Whitlock use their own software to perform forward simulations of an ancestral population that splits into two populations t generations ago. These two daughter populations then may or may not exchange gene flow at a rate m . As such, this is the classic and well-studied isolation-with-migration (IM) model. To make these simulations ‘realistic’, the authors simulate subsamples of the human and stickleback genome. Remi & Whitlock track commonly used summary statistics of genetic diversity within and between populations, including the pairwise nucleotide difference d_{xy} and the relative population differentiation F_{ST} . The authors compare two estimators of F_{ST} with different statistical properties, and find that the one introduced by Weir & Cockerham (1984) is largely insensitive to variation in BGS intensity if populations are connected by gene flow. Using their simulations under BGS and neutrality, Remi & Whitlock also find that the false positive rate of an F_{ST} outlier test aimed at detecting loci under divergent positive selection is not much affected by BGS.

While Remi & Whitlock address a very important issue in population genomics, the manuscript in its current stage falls short of providing the necessary evidence to justify some of the stronger claims that are made. I feel that more justification is required, as otherwise the work presented here is at risk of being misinterpreted. Along these lines, I would request that the authors embed their hypotheses, findings, and interpretations more carefully in the recent literature about the effects of various forms of selection at linked sites, and use analytical results for simple marginal cases of their scenario to support their simulation results. Unfortunately, the lack of evidence is paired with a number of potential technical flaws (at least, I found it hard to convince myself of the contrary based on the material presented) and some apparent misinterpretations of previous work. Overall, I feel that this amounts to a major revision at least, and requires substantial efforts in any case.

Major issues

Old

1. I was dissatisfied by the fact that this study solely uses simulations, while substantial existing theory on BGS could be paired with the structured coalescent under the IM model to explore a much wider parameter space. I am not against simulations, as these are necessary to model complex genomic architectures and demographies. However, without a connection to well-understood marginal cases, I find it difficult to assess the simulated outcomes. The motivation behind the fairly complex simulations done here is to simulate ‘realistic’ scenarios. However, I am afraid that this ‘realism’ is compromised by crucial assumptions that should be better supported before they can be used to justify the complexity of the simulations. This was a general impression. I refer to specific issues below.

2. To what extent can the findings obtained here for humans and stickleback be extrapolated to other species? How generally do the findings apply to sticklebacks, even? I presume that e.g. marine vs freshwater ecotypes, and to some extent lake vs stream populations, have quite different demographic histories. Since demographic history may affect the efficacy of selection (including BGS, see e.g. Torres

et al. 2017), I am concerned about the generality of the conclusions. As these conclusions are currently stated quite generally, I suggest the authors make it more clear that they are based on two exemplary genomes and a single, simple demographic scenario.

3. The authors have written an apparently new software (SimBit) for their simulations. While they state that results were double-checked against SFS_code (Hernandez 2008), they do not show these checks (l. 152–155). I suggest the authors show these comparisons, as it would make the verification of their results more transparent.

4. Sampling of genomic architectures uniformly w.r.t. the physical map (l. 171 ff) likely results in a non-uniform representation of the recombination landscape. If a small proportion of the genome exhibits low recombination, low-recombination regions – which are expected to show stronger signals of BGS – will be underrepresented. Therefore, the sampling regime currently implemented by the authors may fail to detect stronger effects of BGS simply because the weighting is uniform w.r.t. to the physical map, but not the linkage map. From a practical point of view, if BGS were to create a false signal of local adaptation only in low-recombination regions, this might be enough to produce false positives in an empirical study at a rate that leads to wrong conclusions. To address this concern, I suggest the authors complement their existing analyses with separate analyses for, say, the 5% of the genome with the highest and lowest recombination rates, respectively.

5. I did not see any obvious justification for why mutation rate should vary according to an exponential distribution along the genome (l. 167–170). In fact, I think that this assumption might introduce far too much variation and drive some of the observations. Specifically, because of the strong leptocurtic shape of the exponential, the current simulation scheme might assign too much weight to lower-than-average mutation rates, and so underestimate the effects of BGS. To address this concern, I suggest the authors repeat some of their simulations with a constant mutation rate and compare the results to the ones obtained with an exponential distribution.

6. The authors assumed a distribution of selection coefficients that amounts to an average selection coefficient per non-lethal deleterious mutation of 0.07 (l. 208–212). This seems very high and is more than a magnitude higher than what McVicker et al. (2009) inferred for *conserved* exonic sites (0.0025) in humans. I therefore suspect that deleterious mutations are so rapidly removed from the simulated genomic regions that the effect of BGS on (at least the human) genome is strongly underestimated (except for the scenario without recombination). The reduction in effective population size (or B) due to BGS scales inversely with the selection coefficient s : in other words, the higher s the higher B . Indeed, Remi & Whitlock find that the mean B across their simulated focal regions is much higher (0.975) than the one inferred by McVicker et al. (2009) if selection is restricted to conserved exonic sites (90% CI ranging from 0.74 to 0.81).

If I assume a genomic region of length 10 Mb spanning a total map length of 1 cM, which roughly corresponds to a genomic region simulated by the authors, and that 1.75% of the sites in this region are exonic with a deleterious mutation rate of 2.5×10^{-8} per exonic site, then the B statistic calculated using Eq. 8 of Hudson & Kaplan (1995) [HK95] roughly connects the value pairs (s, B) pertaining to Remi & Whitlock and McVicker (2009), respectively (see Fig. 1 [MEC_Remi&Whitlock_fig1.png]). In other words, by simply changing the distribution of selection coefficient such that the mean is ~ 0.0025 rather than 0.07, but keeping the other parameters constant, the authors might obtain more plausible BGS simulations.

If, instead of Eq. 8 of HK95, I use the equation on l. 283 given by the authors, then the mean selection coefficient would have to be reduced from 0.07 to at least about 0.015 in order to reproduce the mean B found by McVicker et al. (2009) (Fig. 2 [MEC_Remi&Whitlock_fig2.png]). Overall, since the mean selection coefficient used in the simulations is about an order of magnitude too high, and the mean B predicted by these simulations is about 20% below previous estimates of the mean B in humans, I am concerned that simulations might not be appropriately representing the pressure of purifying selection in the human genome.

7. The authors assumed migration rates of $m \in \{0, 0.005, 0.05\}$ and population sizes of $N \in \{1000, 10\,000\}$. In the “Default” and “Human” scenarios studied by the authors, the scaled migration rate is $M = 4Nm = 4 \times 1000 \times 0.005 = 20$, which seems very high for humans, at least on an inter-continental scale. Akey et al. (2002) and others (Elhaik 2012 PLoS One) reported average F_{ST} estimates on the order of 10%, which is an order of magnitude higher than the F_{ST} for the “Default” and “Human” scenarios reported by the authors in Fig. 1 of their manuscript (left column). In the limit of high migration, BGS has a decreasing effect on F_{ST} , and so if m is chosen too high, then there is little scope for detecting BGS at all.

To illustrate this, I have plotted the equilibrium $F_{ST} = 1/(1 + 8BNm)$ (note that this applies only to the later stages of the simulations, i.e. a large enough time since the population split, and so likely represents an overestimate) as a function of $M = 4Nm$ in Fig. 3 for three values of B . The vertical grid line indicates $M = 20$, and the dashed and dot-dashed horizontal grid lines represent an F_{ST} of 0.02 and 0.1, respectively. My impression is that the combination of an inflated selection coefficient (and, hence, an inflated mean B ; see point **6.** above) with too high a migration rate m explains why the authors find no local signal of BGS. In contrast, with a value for $m \approx 0.0012$, F_{ST} would be on the order of 0.1. Then, combined with the average B inferred by McVicker et al. (2009) of about 0.775, the effect of BGS on F_{ST} would be expected to be on the order of a 20% reduction compared to $B = 1$ (compare black to blue curves in Fig. 3 [MEC_Remi&Whitlock_fig3.png]). These considerations apply to the *average*; regions of low recombination may well exhibit a much stronger increase in F_{ST} . I am not an expert in stickleback demographic history, but presume that similar arguments could be made there.

8. I was uncertain about how the migration rate m was determined for the FDist2 simulations. Depending on the answer of the authors to my question related to l. 301–310 (see below), the estimated false positive rates might not be appropriate, and this could explain why scenarios with and without BGS showed similar false positive rates except for a few extreme scenarios.

Minor comments

C: Comment Q: Question S: Suggested change R: Requested change

Introduction

Methods

Results

Discussion

References

Figures and tables

Appendix A

- [Paragraph 1] S: ‘...get the same...’ → ‘...obtain the same...’
- [Paragraph 2] R: Please fix the grammar in the first sentence (‘simulations which results are’). C: The second sentence is confusing, as Figures A1 and A2 are arranged vertically, but referred to as “left” and “right” apparently.

Supporting material