Deleterious variation mimics signatures of genomic incompatibility and adaptive introgression

Sent
21-Dec-2017
From
EiC.MBE@gmail.com, EAssist.MBE@gmail.com
To
klohmueller@ucla.edu, klohmueller@g.ucla.edu
CC
EiC.MBE@gmail.com, EAssist.MBE@gmail.com
Subject
Editorial Decision to Reject MBE-17-1030
Body
21-Dec-2017

MS: MBE-17-1030
Title: Deleterious variation mimics signatures of genomic incompatibility and adaptive introgression

Dear Dr. Lohmueller:

Thank you for submitting your manuscript to Molecular Biology and Evolution (MBE). We regret to inform you that it did not receive high enough priority for publication after an in-depth review by the editors and the peer reviewers. Specific comments from the editors and/or external reviewers are included below.

In general, MBE seeks to publish research, methods, and resources of broad significance in molecular evolutionary biology. Even when the external reviewers find a manuscript to be scientifically and technically sound, the ultimate priority for publication is determined based on the novelty and impact of the work presented. MBE does not publish manuscripts judged by the reviewers to contain largely descriptive work, confirmatory results, and discoveries with a limited gene and taxonomic scope. All of these factors were considered in deciding the publication priority for your manuscript.

Thank you for considering Molecular Biology and Evolution, and please continue to consider MBE as a venue for the publication of your best work.

Sincerely,

Board of Editors
Molecular Biology and Evolution

Comments for the author:
Associate Editor
Editors' comments to the author:
The effect of deleterious mutations on gene flow is a very important issue in interpreting the signature of introgression or admixture. However, since the significance of this problem was already recognized and

discussed in the previous studies, one would expect a subsequent study to 1) provide broader theoretical insight for general cases of admixture or 2) analysis of particular data to confirm or reject the role of deleterious mutations in gene flow. Regarding the first direction, both reviewers comment that scenarios and parameters were not fully explored, partly because it depends a lot on simulations but not on any analysis. The second kind of study is difficult because not so much information is available for the nature of deleterious mutations in actual populations. Because of these limitations, both reviewers judged the broader scientific impact of this work to be only medium. I think a substantial addition to the current manuscript is needed in revision to make it suitable for publication in MBE.
Reviewer(s)' comments to author:
Reviewer: 1

Comments to the Author
This manuscript by Kim, et al. presents a simulation study of how selection on weakly deleterious genetic variation can shape the landscape of introgressed fragments in the wake of gene flow between diverged populations. This is an important topic; over the past decade, a lot of attention has been paid to introgression that is adaptive, neutral, or shaped by reproductive incompatibilities, but the conversation has only recently started to include more discussion of how ordinary deleterious alleles affect the persistence of gene flow.

There are a lot of interesting insights in here, but because the introgression of deleterious alleles is a broad topic with potential relevance to many different biological systems, I think the authors have been tempted to shallowly traverse too much ground. As a result, the analyses seem a bit scattered and thin and fall short of the ambitious promises that are made in the introduction and discussion. The results section reads too much like a laundry list of simulations that were done, not like a comprehensive answer to a few memorable questions. Each simulation is clearly described, but not always clearly motivated. I often found myself wondering why various choices were made, e.g. why do the authors choose to simulate 5 Mb sequences? Is there some reason that 5 Mb should suffice to capture all the linkage dynamics that exist? On page 11, what motivates the choice of a split time of 50 generations and a split/bottleneck time of 20,000 generations? It seems like particular model systems must be motivating these choices, unless they are instead motivated by features of the coalescent (e.g. the amount of time it takes for x% of segregating sites to become fixed differences). It would be helpful to give readers intuitive benchmarks of these numbers, e.g. if a certain split time is motivated by the human/Neanderthal split, saying that will give readers intuition about what level of divergence the results apply to.

In setting up their work, the authors say that “few [previous] studies have investigated these questions outside of demographic models specific to a species or system,” but their simulation framework doesn’t actually seem much more general than what has been done before. All simulations are done with effective population sizes ranging from 10,000 to 1,000, which makes sense given the computational constraints of SLiM but means that they are effectively simulating humans and Neanderthals, not fruit flies. The divergence times, chromosome maps, and distribution of mutational fitness effects used are all either directly inferred from human data or closely resemble divergence events that are relevant to human history. I don’t necessarily see anything wrong with focusing on human-inspired parameters, since there’s only so much room in one paper, but it also doesn’t seem right to claim that the results in this paper are much more general than what has been done before. There is certainly material in here that is novel and different from previous work, including examinations of the effects of recombination rate variation, but it

would be more accurate to claim that they are examining certain specific questions for the first time, not that this study does much better than previous work at generalizing beyond particular systems/species.

There are several specific points in the discussion that don't seem to line up with what is actually done in the paper. For example, the discussion explicitly cites fruit fly systems as examples of settings where the results of this paper are relevant. If they are going to discuss fly papers, the authors need to note that they are not simulating fly-appropriate population size regimes and comment on whether their results might differ if population sizes were 1-2 orders of magnitude larger. Similarly, although maize is mentioned, there are no simulations of maize-like genome architecture where exons are interspersed among huge quantities of TEs. Lines 515-517 of the discussion say that "our study also highlights the fundamental importance of understanding the distribution of selection coefficients and the relationship to dominance coefficients in natural populations (the h-s relationship)," but I don't believe that the paper ever explores the dependence of the results on the DFE. The same DFE is used for every analysis, and there is no exploration of how h-s relationships other than a constant h=0.5 or constant h=0.

Some of the most interesting results, in my opinion, are the dependence of introgression dynamics on recombination rate and divergence time. A focus on these two issues would be less sprawling for the reader to keep track of and would be plenty of material for one paper. For the analyses on pages 7-8 saying that "short-term bottlenecks" have a weaker effect than "long-term population contractions," a key missing detail is what defines "short-term" versus "long-term," ideally in terms of a quantity like FST that would give the result some true generality. Currently, it looks like a single demographic scenario is used to represent a short-term bottleneck, and a single scenario is used to represent a long-term bottleneck, with no justification about why these parameters were chosen and how the results might interpolate to other sets of parameters. There should really be multiple divergence time datapoints showing how the short-term-bottleneck regime transitions into the long-term-contraction regime. Also, since the length of the bottleneck is coupled to the divergence time in these scenarios, it is not clear how much of the difference between the two scenarios is a function of the difference in divergence versus the difference in diversity between the populations. Similarly, in the section "increasing population split times enhances the effect of heterosis," not enough values of t_s are sampled for us to see when the effect of heterosis starts to plateau (it must stop increasing and plateau eventually as the diversity in the bottlenecked population approaches equilibrium for its new population size).

The point about heterosis mimicking adaptive introgression is important to think about, but something is lost here in considering just a single 5 Mb locus. If this kind of simulation were done at a genome-wide scale, how likely is it that a small number of genes would pop out as apparent candidates for adaptive introgression? Instead, wouldn't the admixture fraction be inflated across the entire genome, making it look more like the initial introgression fraction was high than like adaptive introgression was going on?

Reviewer: 2

Comments to the Author
Summary
=======

Here, Kim and co-authors use population genetic simulations to study the impact of admixture and

dominance on deleterious variation under alternative demographic scenarios and as a function of recombination. They show that in a model with two demes of equal size that split some time ago and then experience a single pulse of unidirectional gene flow, admixture has essentially no impact if deleterious mutations are 'additive' (h = 0.5). However, with recessive deleterious mutations (h = 0), admixture causes an immediate increase in mean fitness in the recipient population relative to the source population. This increase in relative fitness (or, equivalently, reduction in relative load) is due to heterosis resulting from a protective effect that wild-type alleles convey in hybrid offspring. This heterosis effect is in line with both classical theory (Crow 1948), as well as more recent empirical work on Neanderthal introgression in humans (Harris & Nielsen 2016). The protective effect diffuses over time as recombination breaks up the association between introgressing and local alleles, though. A short bottleneck in the recipient population causes no qualitative and little quantitative difference from this dynamics. With recessive mutations, in both demographic scenarios, the proportion of introgression-derived ancestry in the source population is found to be highest if recombination rate is lowest.

In contrast to a short bottleneck, long-term differences in population size prior to admixture cause a strong difference in the mean population fitness (and hence the mutation load) between populations before admixture. The smaller population (either the source or recipient) accumulates more deleterious mutations, both with 'additive' and recessive mutations. With 'additive' mutations, admixture has essentially no impact on the purging of relative load in the recipient population if the source population is the smaller of the two populations. However, the proportion of introgression-derived ancestry now increases with recombination rate. This is in line recent work on selection against putatively deleterious variants introgressed from Neanderthals into humans (Harris & Nielsen 2016; Juric et al. 2016). With recessive mutations, mean fitness in the recipient population increases more strongly if recombination is low, but admixture as such has little impact on the increase in relative fitness. The proportion of introgression-derived ancestry is found to be negatively correlated with recombination rate. If, in contrast, the source population is larger than the recipient, then admixture leads to an immediate increase of mean population fitness in the recipient relative to the source population. This increase is due to deleterious mutations in the recipient population being masked by wild-type alleles on introgressing haplotypes. The change in relative fitness strongest for low recombination, but with low recombination the mean fitness also decreases rapidly because the recipient population is so small.

In a fifth model, the authors study the impact of population-size recovery after admixture in a recipient population that went through a strong and prolonged bottleneck. The recovery to a large population size in the recipient population prevents the accumulation of further load. Mean fitness in the recipient population recovers and approaches the mean fitness in the source population. This holds for 'additive' and recessive mutations, although the effects are much more pronounced for recessive ones. The proportion of introgression-derived ancestry is found to be highest if recombination rate is lowest.

Overall, the authors show that the correlation between the proportion of introgression-derived ancestry and recombination rate can be positive or negative depending on the combination of dominance and demography. Admixture and purifying selection can thus mimic both the patterns expected under divergent selection against locally maladapted alleles (i.e. a positive correlation between recombination and the proportion of introgressed locally deleterious ancestry), as well as under adaptive introgression (i.e. a negative correlation between recombination and the proportion of introgressed adaptive ancestry).

The question of how and to what extent unconditionally deleterious mutation can mimic the signal of other modes of selection in and among populations that are connected through gene flow is of interest. The paper is very well written, with a clear structure. However, I had a number of concerns that I think should be addressed (see below). Among these are a somewhat problematic definition of 'additivity' (i.e. no dominance), a potential lack of generality due to the separation of dominance and the distribution of fitness effects, and some missing clarity about the difference between genetic incompatibilities vs. selection against locally maladaptive introgressing alleles.

Major comments
==========

1) The admixture proportion has been fixed, although it is not clear if the effect of this parameter on the observed quantities (population mean fitness, proxies of mutational burden) is monotonic. A detailed exploration of the effect of the admixture proportion might be beyond the scope of this paper, but it should be mentioned as a direction for future work. Similarly, only recessivity and additivity, but no intermediate dominance, have been studied. In Model 2, the opposite effects of admixture on the frequency of introgression-derived ancestry (p_I) under recessivity vs. additivity suggests that there is an intermediate degree of dominance at which selection should have no effect on p_I. This particular degree of dominance should depend on the recombination rate. The fact that this study is entirely based on simulations under a restricted range of biological scenarios therefore limits the insight about the role of important parameters. There is no immediate problem with that, but the limitations should be more clearly addressed in the Discussion.

2) In light of the on-going debate about the usefulness of proxies measuring the burden of deleterious mutations, it is noteworthy that the number of deleterious variants per individual seems to be much less sensitive to changes in mean fitness than the number of homozygous derived genotypes. This applies in particular to the case of recessive deleterious mutations. Conversely, the number of derived variants does seem to track -- at least qualitatively -- the reduction in fitness relatively well in the case of additive deleterious mutations. It might strengthen the Discussion if the observed patterns were related more extensively to the ones found in previous recent literature that ignored admixture (e.g. Lohmueller et al. 2008; Lohmueller 2014; Do et al. 2015, Simons et al. 2014; Simons and Sella 2016).

3) Introgression-derived ancestry was tracked here based on neutral marker mutations introduced at the very beginning into the source population only. There were no mutations at these marker sites, and these were used to track the dynamics of the proportion of introgression-derived ancestry (p_I). Besides these marker mutations, other neutral mutations were apparently also simulated (e.g. l.139), but it is unclear whether these were used, and if so for what. The dynamics of the frequency of introgression-derived ancestry (p_I) was related to the findings of studies that modelled the effect of selection at linked sites on diversity at neutral sites without restrictions on neutral mutation rate (e.g. l.375-376; but see Juric et al. 2016). So, to me, it is not clear how the authors draw the connection between non-mutating ancestry markers in their simulations, and predictions/results from other studies about the diversity at neutral sites linked to sites under selection. Specifically, the authors caution that purifying selection could mimic patterns expected under divergent selection against gene flow (e.g. a negative correlation between diversity and recombination). It would be great if the authors could address the distinction between the

frequency of foreign ancestry vs. linked neutral diversity.

4) The distribution of fitness effects (DFE) used here was inferred by Kim et al. (2016) under the assumption of additivity. Yet, the same DFE is used in this study here for both additive and recessive deleterious mutations. Since theory and empirical studies suggest that more deleterious mutations tend to be more recessive, I am concerned that the combinations of selection coefficients and dominance coefficients simulated here may not be representative of real populations. It is therefore difficult to assess the relevance of the results. If the authors could convincingly show/argue how a negative correlation between dominance and selection coefficients would affect p_I, fitness, and proxies of mutation load, this would strengthen the manuscript a lot. The problem is touched on l.514-523 and l.562-656 in the Discussion, but the main issue remains not fully addressed. I have reservations about the usefulness (biological relevance) of considering dominance and the DFE separately.

5) What Fig. 5 and the related analyses seem to ignore is that recombination rate and gene density *jointly* determine the impact of selection on the proportion of introgression-derived ancestry, p_I. The results shown only reflect the specific recombination map and gene locations and, in particular, the correlation between recombination and gene density, typical of the chosen subset of chromosome 1. It is not clear whether one can draw general conclusions. And it is therefore hard to judge how generic the predictions in Table 2 are, as long as qualitatively different correlations between gene density and recombination are not explicitly investigated. I actually do believe that Table 2 holds, but I do not think that simulating just one randomly chosen segment of a human chromosome provides sufficient evidence. If I am not mistaken, gene density and recombination rate are slightly positively correlated along the human genome. Hence, in scenarios where increasing recombination is predicted to decrease p_I under uniform gene density (e.g. in Model 4), an (strong) underlying positive correlation between recombination and gene density could reverse -- or at least attenuate -- the expected pattern of a negative correlation between p_I and recombination in reality. My suspicion is that the correlation between recombination and gene density is not strong enough in humans to lead to such an inversion. But this potential issue seems worth being discussed. See also Slotte (2014, Brief Funct Genomics) or Aeschbacher et al. (2017).

6) I wonder if in the title 'genomic incompatibilities' should be replaced by 'divergent selection in the face of gene flow' or similar. Genomic incompatibilities (in the classical sense of DMIs) per se are not a very effective barrier to gene flow unless the underlying alleles are also locally adapted (see e.g. Bank et al. 2012, Genetics).

7) The way fitness is computed in the 'additive' model (l.629) is inconsistent with the use of h as 'dominance coefficient' and with the use of 'additive mutations' for mutations with h = 0.5 (see e.g. l.144). With the definition on l.629, h = 0.5 does not result in additivity. In a similar way, the justification given on l.625-628 for the definition of w_j on l.629 is in itself inconsistent: "Because our intent with an additive fitness model was to make the fitness effect of each variant independent of its genotype, we computed additive fitness as purely multiplicative across all deleterious variants,...". Additivity and multiplicativity exclude each other; dominance is a property of the genotypic state (i.e. of the alleles involved), and trying to make dominance independent of genotypic state therefore sounds strange. In fact, what the currently used fitness regime implies is that what the authors call the 'additive' case actually involves slight underdominance. The extent of underdominance depends on the magnitude of s. To see this, compare (1 - s)(1 - s) = 1 - 2s + s^2 to 1 - 2s. The former homozygote fitness is larger than the latter by s^2. The

fitness of the heterozygote is 1 - s, which is smaller than the average between 1 and 1 - 2s + s^2.

Minor issues
==========

l.132-135: It was not fully clear to me how you estimated the proportion of introgression-derived ancestry (p_I). Is it a multi-locus average over all maker sites of the frequency of the allele indicating introgressed ancestry? In the simulations with a realistic map of recombination and gene density (a subset of the human chromosome 1), is p_I computed as the average over marker sites within each window?

l.200: Please check if the references to Fig. 2 and S1 are needed.

l.339: "has increases to" -> "increases to"

l.355: I think it should say "are analogous to Models 0, 1, and 4" instead of "are analogous to Models 0 and 4". However, reading the paragraph starting on l.358, it was not clear to me why Model 1 was invoked here at all.

l.359: "for these three models" -> "for Models 0 and 4". Or if indeed three models were simulated, then l.355 must be adjusted, as only two models are listed there.

l.372-377: I suggest to separately list the literature concerned with either genomic incompatibility or maladaptive alleles. For instance, neither Brandvain et al. (2014) nor Aeschbacher et al. (2017) were concerned with classical genetic incompatibilities (i.e. epistasis at the level of genotypes), but with maladaptive alleles.

l.377-378: Aeschbacher et al. (2017) explicitly accounted for the confounding effect of selection on deleterious variation (i.e. background selection in their case). This comment also applies to l.476-481.

l.505-513: I was confused by this paragraph. Isn't a positive correlation between recombination and the frequency of introgressed ancestry predicted by Model 2 under additivity? Model 2 therefore would not only seem to explain the situation in swordtail fish, but also observed patterns of Neanderthal introgression into modern humans.

l.532-547: For Model 3 (introgression from a large source into a small recipient population), the statement that the fitness in the admixed population increases the most for the lowest recombination rate (l.537-538) seems to be appropriate only immediately after the admixture (Fig. 2) if deleterious mutations are recessive. Interestingly, on l.540 this scenario is used to explain why recombination is apparently found to be more effective in removing deleterious mutations in regions of low recombination (as opposed to the prediction under Hill-Robertson interference). From Fig. 2, Model 3, left-hand panel, it seems as if fairly soon after admixture, the prediction of the Hill-Robertson interference does apply. Please clarify.

l.562: Aeschbacher et al. (2017) include background selection in their null model.

Fig. 2: The caption mentions both mean and median fitness without clarifying why different terms are used.

Fig. 5: It would be great to see the recombination rate along the genome, superimposing the existing plot on an appropriate scale, or in a separate panel on top.

Fig. 5: I was a bit puzzled by an apparent lack of parallelism in the loess curves for the recessive mutations under Models 0, 2, and 4. Specifically, I would have expected to see the red curves peak more consistently across all Models in regions of high exonic density. Instead, local peaks are not shared very consistently among Models. Is this just a consequence of a moderate number of replications, insufficient to smooth out variation induced by the sampling of selection coefficients from the DFE? If so, this might be worth a comment.

Fig. S1: The y-axis label says "median fitness" is shown; the caption mentions both "mean fitness" and "median fitness". Please clarify.

Figs. S1, S2, S3, and Fig. 3: Please consider providing additional versions of these figures in which the range of the y axis is identical between different rows within the same figure. Currently, the scales are different, which emphasises the qualitative patterns, but distracts from the quantitative differences among Models.