The authors present an inference model that expands on their previously published iSMC model, that infers how nucleotide diversity in the genome is jointly influenced by genealogical history, recombination, and large-scale regional spatial variation in mutation rates. Applying this model to Drosophila, they argue that variation in mutation rates is the primary driver of nucleotide diversity at varying spatial resolutions throughout the genome. This has important implications for various inference tasks in population genomics.

Overall, I found the approach innovative and interesting and so the work presents a valuable contribution to the field. At the same time, I remain skeptical about the biological conclusions about Drosophila, in light of the stark difference from the long history of work arguing for the key importance of linked selection. While the importance of mutation rate variation on different scales seems intuitive to me, I remain unsure whether the variance decomposition framework provides adequate support for the authors' claims.

Major comment
-=-=-=-=-=-=-=-

The authors model pi (pairwise nucleotide diversity at SNPs) as a response variable, linearly related to their estimates of several parameters: population scaled recombination rate, mutation rate and coalescent time and the interaction of coalescent time with mutation rate (a term for which the justification is not entirely clear). The parameters themselves are estimated in an iSMC framework on a small part of the genome. From the population genetic perspective, this seems like a step back compared to work in the last decade and a half to explain variation in nucleotide diversity along the genome. Of course, if prediction is king, then any approach that does the job well is merited. However:

(a) It is hard to interpret results, even qualitative ones. For example, my sense is that the transition probabilities of 1 for recombination rate and mutation rate (lines 170-185)---estimates that are at the limit of the possible range for the estimand—suggest that the analysis window size needs to be explored in more depth, and the model is ill-specified for 10kb windows. I also do not see how this provides support for the claims on the relative importance of smaller-scale factors (e.g. nucleotide context) that are not modelled at all, to the best of my understanding, versus larger-scale factors (e.g. replication timing).

(b) I think the bar has to be set high here, given the careful modeling done by Sattath et al. PLoS Genetics 2011 to consider effects of sweeps and demography while controlling for mutation rate variation (albeit imperfectly, by "normalizing" by divergence) and then Elyashiv et al., who extended the results to include (again, coalescent-based) effects of background selection and incorporated information about functional importance of individual sites and genomic regions. Here, in a variance decomposition framework, I can't help but wonder whether the reason the mutation rate factor explains the most variation

is just that it is the most free to vary across adject windows, such that variance unaccounted for by the model can be easily absorbed in this term.

(Here, simulations with little spatial mutation rate variation, and with or without local (e.g. nucleotide-specific, or just smaller than the analysis window) would be useful.)

(c) Some comparisons with previous results of the landscape of mutation, recombination and selection in Drosophila is needed. It's one thing for the model to do well on simulations where the generative process is picked by the researchers; it is another to show that the variation plausibly reflects real biology. e.g. Assaf et al. 2017, recombination maps, and estimates of the intensity of selection with Elyashiv et al. 2016.

Other comments

-=-=-=-=-=-=-=-=-

Does increasing this granularity improve model fit? (Conversely, does using a model with fewer parameters improve performance in smaller window sizes?) See for example Michaelson et al., 2012,) using an HMM to model effects of regional mutation rate on de-novo mutation counts in human trios,

The authors mention in the discussion (line 523-525) that it would be interesting to assess the extent to which large-scale variation in mutation rates is explained by the composition of sequence motifs (with different mutation rates) in each genomic window. This is more or less already a consensus in the human mutation rate literature (most recently, Fang et al. showed that motif-specific mutation rate estimates explain a significant portion of large-scale variation in mutation rates in humans). It would be straightforward and worth including some simple analyses to demonstrate the relevance of motif-specific mutation rates in the regional mutation trends. One idea might be to identify particular 3-mers that are overrepresented in the windows with the highest and lowest genetic diversity. (the Drosophila mutation spectrum characterized by Assaf et al., 2017 would be useful here).

Relatedly, I'm curious how the authors expect this model might generalize to mammalian species whose genomes are significantly influenced by the extreme hypermutability of methylated CpG sites. Since Drosophila is not affected by mCpG hypermutability, I'd expect the overall mutation landscape to look more "flat" than in mammals, so if applied to a mammalian species, is it possible that this model would simply recapitulate regional variation in mCpG content? Stratifying some of the analyses by mutation (sub)type might yield some interesting insights into how theta, rho, & tau are interacting to shape genome diversity.

Line 172-174: Does fitting on a single chromosome arm sufficiently model genome-wide mutation rate heterogeneity and/or accommodate large genomic regions that may harbor unusually high or low mutation rates in drosophila? An interesting counterexample might be chr8p in humans, which is known to have a much higher mutation rate than elsewhere in the genome.

Minor comments
-=-=-=-=-=-=-=-=-=

Line 50-51: “new stab” is a bit vague—what exactly are these recent papers attempting to clarify/augment RE: Lewontin’s Paradox?

Line 490: “according to their position in the genome” is a bit too reductive here, as it implies that mutation patterns in functional loci are purely a function of the surrounding genomic features—mechanisms like transcription-associated mutagenesis and transcription-coupled repair uniquely impact the mutation landscape of transcribed loci, independently of those loci’s spatial positioning in the genome.

Line 540-545: Venkat et al. is another good ref here, as they provide a compelling example of how a more detailed mutation model (in this case, one that accounts for multi-nucleotide mutations) can confound selection scans.

Fig 2. Please add labels on the left for the 50kb/200kb/1Mb scales

Fig 5. Legend: typo in sentence “in A and B, results are displayed according [to] simulated parameters”

Fig 6. Please make font sizes of panel labels larger

Fig 7. As in Fig 2, please add labels above each column of panels for the scale

Line 227: typo, simulatios -> simulations