We thank the reviewers for providing thoughtful comments that have helped us to improve this manuscript. We have incorporated changes that address the reviewer’s concerns, and we believe have significantly improved the clarity of the manuscript. We hope that you find the revised manuscript suitable for publication. Below we have included each of the reviewer's comments in blue and our responses in black.

**Reviewer1**

In absence of source code/data and assuming these will be in good shape, I can only raise minor concerns on the discussion scope.

All code and supporting data are now available as a github repository: https://github.com/coleoguy/microsat

First, brief literature search suggested that several recent studies on large-scale insect genome evolution exist focusing on gene content (https://genomebiology.biomedcentral.com/articles/10.1186/s13059-019-1925-7), transposable elements (https://bmcevolbiol.biomedcentral.com/articles/10.1186/s12862-018-1324-9), and probably on other aspects as well. It would be interesting to compare the order level evolution rate trends observed here with the results for other genome element types.

We agree that encouraging the integration of these different types of studies could be illuminating. However, we feel that it is beyond the scope of this manuscript. We do include a sentence at the end of the manuscript suggesting that with new and better assemblies performing integrated analyses that look at a variety of these genomic characteristics is now possible:

Approaches like those we have used that allow for an evolving trait to impact the rate of evolution of a second trait could leverage these genomes to reveal the impact of a broad range of characters (e.g. TEs, structural elements, codon bias, recombination rates) on the evolution of microsatellite content.

Second, authors note that microsatellite instability is linked to slippage during DNA replication, however, the discussion only deals with meiotic recombination. My quick literature search failed, but maybe there are some studies giving insights into the evolutionary differences in the replication control for insects? At least this can be an interesting field to follow up on.

We agree that differences in DNA replication might lead to differences in the rates of microsatellite evolution. We have added a mention of this in the discussion:

“Another potential cause of differences in monocentric and holocentric species could be differences in DNA replication mechanics. Some evidence suggests that the distribution of translation initiation sites may vary based on centromere type (Heckmann *et al.* 2013).”

Lines 388-391

Apart from that, I would appreciate some clarifications in the Methods and/or Results sections:

- What is the "most complete assembly" for a species? Did you access BUSCO scores at this point? If you account for genome size only, this estimate might be skewed many factors, including contamination or duplicate/repeat resolution strategy.

Our wording here was unnecessarily vague with regard to genome assemblies we have changed this sentence to read:

“The most recent assembly..”

Line 80

The question regarding BUSCO scores is clarified by our clarification described above

How do you infer assembly category in case of mixed data available?

We have clarified this in the text:

“If mixed data was available for a genome assembly (e.g. long read sequencing with short read polishing) this was classified as long read for our categories.”

Lines 88-90

The evolutionary rate estimation section in methods looks a bit vague. Do you estimate evolutionary rates directly from the microsatellite content and the phylogeny? (from the ace function docs it seems like that).

We have added clarification to our description of the rate estimations:

“All rate estimates described were generated using the restricted maximum likelihood approach using the ace function in the R package APE (Paradis *et al.* 2004). This function takes observed microsatellite content and the phylogeny, and returns an ancestral state estimate for every node in the tree and the maximum likelihood estimate for the rate of evolution. For comparison, we fit the same model using the fitContinuous function in the R package Geiger v2.0.6.4 (Harmon *et al.* 2008). Rate estimates between these two approaches were qualitatively identical.”

Lines 169-175

In tip rate estimates, was the ancestral state inferred from order-level analysis results? Also, it would be good to indicate rate units at the sigma-square estimates.

We have clarified the description of our tip rate analysis. With regard to units for sigma-squared, this model parameter is customarily not reported with units. In fact, a review of the literature developing this comparative test we found no mention of units that would be appropriate (O'Meara *et al.* 2006). We would argue that if desired units of per MY might be appropriate since this is the units of the phylogeny on which the data is analyzed. However, to maintain consistency within the field we have chosen to report it simply as a model parameter without units.

“Using the ancestral state estimates from our combined analysis of all data, tip rates were estimated by taking the difference in microsatellite content of a species and the ancestral state estimate for the node from which it descends.”

Lines 179-180

- Suggest adding some information on the reconstructed taxonomy to the results section rather than keeping it all in methods.

We have moved a paragraph from the methods to the results that described the results of the phylogenetic analysis.

Lines 252-259

I also hope that the data files released with the paper will include the information collected for the wide range of insect genome assemblies: assembly category, BUSCO scores, microsatellite content, centromere type, chromosome number, genome size

All code and supporting data are now available as a github repository: https://github.com/coleoguy/microsat

Minor points:

l. 49 - suggest substituting "diseases" with something like "thus leading to diseases" to avoid listing disease alongside the molecular processes.

Corrected

Lines 51

l. 50 - term "slippage repair errors" seems to be non-existent

corrected

Lines 53

l. 184 - more chromosome -> more chromosomes

corrected

Lines 204-205

- Location of Fig3C&D versus Fig4A&B is a bit confusing.

We appreciate that we do present tip rate first followed by content in figure 3 while we present content followed by tip rate in figure 4. However, the flow of the discussion of results worked best with this order of presentation and so if possible, we would like to retain the current ordering of panels within these figures.

- BUSCO score cutoff of 90% was seemingly chosen arbitrarily. Some reasoning is needed to justify its recommendation for further studies.

We have added a bit more to our reasoning behind picking the cutoff of 90%

Lines 244-247

**Reviewer 2**

It is striking that such a large range of rates for microsatellite evolution is observed (line 235). The authors may wish to comment in the discussion about the magnitude of this range.

We have expanded our discussion of this variation in the discussion.

Lines 416-418

The finding related to centromere type is limited by the number of clades studied and by the fact that the trend is not totally consistent across clades. The author’s appropriately qualify the limitations of this analysis in the discussion, but may wish to reword the final sentence of the abstract to reflect this limitation. Perhaps species with monocentric chromosomes ‘tend to’ evolve faster.

Corrected

Lines 21

Line 45: plan an important ROLE

Corrected

Lines 46

Line 45: and centromeres and telomeres (seems like a word is missing – centromere/telomere function?)

Corrected

Lines 48

**Reviewer 3**

From my point of view, the selection of sequenced genomes that could be used in this study is important part of results and it would be better to place it in the ‘Results’..

In the methods we state that we downloaded all genomes assemblies available at the time the study began (lines 76-78). In the results section we describe the process that we used to reduce this down to genomes that appeared of sufficient quality to be included in our study (lines 237-250).

Lines 20-21: ‘those species with monocentric chromosomes evolve faster than species with holocentric chromosomes’. *Probably, microsatellite sets in the species with monocentric chromosomes evolve faster. Not the species themselves.*

Corrected

Line 21

Lines 99-100: ‘This yielded a dataset of 221 operational taxonomic units (OTUs) representing members of 12 of the 24 insect orders’*. It would be better to show the number of OTUs in any of 12 insect orders.*

The OTUs in each of the insect orders that were utilized for the study are now present in figure 2B along with the microsatellite content results.

Figure 2

*Selection of 221 OTUs for the investigation would be better to include in Results.*

We agree that the OTU numbers of 231 and 221 are technically a result of our interpretation of the BUSCO analysis and overlap with genbank. However, we retain the mention of it on line 106 to describe the sequence data that was selected for use in the phylogenetic inference. We do more fully dive into the filtering process that we undertook using BUSCO in the results section on lines 237-250.

Lines 216-217: ‘we feared that some insect genomes may be more poorly assembled than the vertebrate genomes we examined’. *Is it about all 303 genomes or about 221 OTUs? However, below authors wrote*‘we discarded 83 genomes and all downstream analyses are performed on the remaining 221 genomes. As mentioned above this was further reduced to 201 for all analyses involving our phylogeny due to elevated taxonomic instability scores during tree inference’. *It should be written more clearly*.

We have made numerous changes that we believe clarify our inclusion filtering

Lines 238-239, 242, 244-247

Lines 223-225: ‘Using this threshold, we discarded 83 genomes and all downstream analyses are performed on the remaining 221 genomes’. 83+221=304. However, only 303 genomes were involved in the study.

This was an error on our part the original downloaded number was 304 insect genomes (supplemental file Table 1).

Lines 237-239: ‘for most species were normally distributed around zero. However, two hemipterans, *Pseudococcus longispinus*and *Paracoccus marginatus,*both exhibited strikingly negative tip rate values (-3.3x10-6 and -3.7x10-6 respectively)’. *From my point of view, -3.3x10-6 and -3.7x10-6 are very close to the zero.*

We agree that the magnitude of these numbers is small however the importance is the comparison of these rates relative to all other tip rates as calculated. We have clarified this by describing the magnitude of difference relative to the mean tip rate observed in our dataset.

Lines 280-281

Lines 260-264: ’Figure 2. Comparing microsatellite content and rates of evolution among orders. Both y-axes are measured in a log scale. The centromere type present in an order is indicated with an H or M at the top of the plot for holocentric and monocentric respectively. Orders are indicated on the horizontal axis. **(a)**The rate of microsatellite evolution for all orders with at least ten representatives. **(b)**Microsatellite content for all species included in comparative analyses’*. For many orders the circles are fused into smear and they are difficult to analyze. It would be better to show the number of species for all included insect orders.*

Points in figure 2A have been spread to allow for better visualization. Additionally, in 2B we included the number of values for each order that is represented in the figure.

Figure 2

Lines 372-373: ‘Centromeric and telomeric regions are both normally heterochromatic…’. *It is not quite clear what ‘centromeric and telomeric regions’ mean. Is* *centromeric region* only centromere? Or does *centromeric region include also pericentromeric heterochromatin? Usually pericentromeric heterochromatic regions are white spots in the draft of genome assembling. Istelomeric region* only telomere? Or does *telomeric region include also region enriched for telomere associated repeats?*

We have clarified our writing to make it clear that we are not speaking strictly of the centromere and telomere but rather the regions approaching these areas of the genome which are often characterized by increased levels of repetitive sequences.

Lines 422-423

Lines 381-384: ‘we note that centromeric and telomeric regions are difficult to assemble regions of the genome and may become more difficult to assemble as the number of chromosomes increases. As such, the use of whole genome assemblies rather than raw reads may reduce our ability to detect a concentration of microsatellites in these regions’. *Probably, in most genomes involved in the study only their euchromatic part was assembled and authors should write about microsatellite content and evolution only in euchromatic part of insect genomes.*

This is a good point. We have added a sentence to our discussion to point out that our results are likely most applicable to describing the mode and tempo of microsatellite evolution in euchromatic portions of the genome.

Lines 434-436

*Conclusions in the manuscript look like ‘Discussion’. Furthermore, I do not agree that the main conclusion of this study is* ‘As more insect genome assemblies become available and the quality of assemblies increases, future studies could use approaches similar to ours to understand the impact of a variety of genomic characters (i.e., codon bias, transposable elements, recombination rates) on microsatellite landscapes’. *Conclusions should be rewritten.*

We agree that the conclusion needed work. We have rewritten this paragraph. We now present clear conclusions from our study followed by what we hope to see as the quality and completeness of available genomes improves.

Lines 477-492

Harmon, L. J., J. T. Weir, C. D. Brock, R. E. Glor and W. Challenger, 2008 GEIGER: investigating evolutionary radiations. Bioinformatics 24**:** 129-131.

Heckmann, S., J. Macas, K. Kumke, J. Fuchs, V. Schubert *et al.*, 2013 The holocentric species L uzula elegans shows interplay between centromere and large‐scale genome organization. The Plant Journal 73**:** 555-565.

O'Meara, B. C., C. Ane, M. J. Sanderson and P. C. Wainwright, 2006 Testing for different rates of continuous trait evolution using likelihood. Evolution 60**:** 922-933.

Paradis, E., J. Claude and K. Strimmer, 2004 APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20**:** 289-290.