Warren et al. provide a novel look into the role of protein identity and macro- and microsynteny in speciation processes. In this updated manuscript, they improve their work in which they utilize congeneric species pairs to identify genomic processes that lead to divergence of species and seek to determine whether or not there are universal patterns that drive speciation. In the updated version, little has been changed with the exception of minor text edits. In some cases, reviewers expressed concerns with the methodology (not just minor text details), and the authors dismiss the commentary rather than making substantive edits. A few that clearly still need addressing:

- 1) Several places were suggested for expansion in the introduction and clarification of the methods and discussion. Only the minor, "low-hanging fruit" edits were changed.
- 2) Two reviewers highlighted that the title was vague and doesn't accurately capture the extent of the work.
- 3) Two reviewers expressed concerns about the criteria for species selection and taxonomic resolution. In many cases, the selected species seem somewhat arbitrary (e.g., why is the "mammals" represented predominantly by primates, rather than the diverse array of mammals that are known to have huge chromosomal rearrangements such as Equidae or Cervidae?).
- 4) Reviewer 2 raised a valid point with the correlations between protein identity, macrosynteny, and microsynteny. Without the "insect" comparisons, the already loose correlation would disappear. Rather than adding justifications in text, the author responds to "check back in 5 years" when additional references are available. The authors' response that there is an insufficient number of references available is also categorically false. As an example, when the authors initially submitted the manuscript to bioRxiv, there were four *Daphnia* species with NCBI annotated, chromosome-level assemblies, but choose only to look at two species. It is entirely valid to decide that additional analyses are not possible at this time due to time constraints, etc. It is also entirely valid to not use all the available resources if a research group is simply seeking to explore a phenomenon at a high level. But typically if authors have concerns with portions of the analyses, it requires changes and justifications within the text, rather than snarky responses.
- 5) The authors did not address reviewer concerns about the divergence time of the taxa in the analyses. The authors respond saying that it is unnecessary due to issues with dating and that divergence times were not available for all species pairs in TimeTree (TimeTree is not the only source of estimates and often does not aligned with peer-reviewed literature). The issue is that the divergence times *may* span orders of magnitude different time intervals which is concerning regardless of molecular clock and fossil calibrations. Pairs that may be diverged by hundreds of millions of years

(*Daphnia*) are **not** apples-to-apples comparisons with pairs diverged by a few million years (*Drosophila*) even though we as a research community have arbitrarily binned them as congeners. This does *not* need additional analyses, only additional text.

Taken together, the work presented here has intellectual merit and is novel, but needs more work. I also have little new to contribute as the manuscript is nearly identical to the first round, but below are additional comments that have occurred to me since the first time reading the manuscript:

Additional comments:

Minor comments:

Page 4, Line 2-5: Authors should use full binomial Latin names instead of a mix of genus and common names.

General:

- 1) Assemblies are not perfect and the criteria for selection needs to be described in the main text. The assemblies simply being "chromosome-level" is not enough. Although these are all chromosome-level assemblies, quality within this classification can still vary considerably. Many of the assemblies here are from a similar methodology from the same institution (HiFi assembly with Hi-C scaffolding by Wellcome Sanger), but many are using totally different methods or are missing meaningful assembly information on the NCBI assembly pages and their associated bioprojects do not link to a publication. In some cases it's unclear whether some of these references are even suitable for the questions at hand. A short read assembly that was scaffolded to chromosome-level using linkage information (at least one of the assemblies, Daphnia magna) would have drastically higher local misassembly rates compared with a long-read assembly scaffolded with Hi-C. This would inherently bias the microsynteny estimates in pairwise comparisons where a short-read derived assembly is compared with a long-read assembly. I did not look through all the assemblies listed in the GitHub, but assemblies scaffolded using synteny against congeners can also meet the criteria to be considered "chromosome-level NCBI RefSeq", which would be problematic for the questions in this manuscript. In all cases local misassemblies are more likely than large-scale misassemblies, particularly with scaffolded short read assemblies and could inflate the number of microsyntenic changes.
- 2) Several of the assembly hyperlinks in the GitHub direct to "species" hubs on NCBI with multiple assembly versions with different assembly methods. The authors

should make sure to include the GenBank or RefSeq assembly accessions ("GCF_XXXXX", "GCA_XXXX") such that readers are directed to a static assembly rather than multiple.

- 3) Several assemblies are missing from the list of species in the GitHub. For example, the manuscript Table 1 mentions that two species of Perca were used. In the GitHub "input_datasets" section, only Perca flavescens is listed in the markdown which links to the NCBI species pages. Perca fluviatilis is mentioned in another supplemental document, but as far as I can see there is no link nor text that describes the accession used. All species and accessions need to be listed in one spot not scattered throughout.
- 4) The species in the *Crassostrea* comparisons have been reclassified into two genera, *Magallana* and *Crassostrea*. The *C. virginica* v. *C. angulata* comparison in the main text is no longer a congeneric comparison. *M. angulata* v. *M. gigas* would have to take its place to remain a congeneric comparison.