**Author comments are in bold and start with \*\*\*.**

Referee #1 (Comments to the Author):

I have carefully re-read the revised manuscript, and I can see and appreciate all the hard work that went into the revision. Unfortunately, many of my original concerns remain. However, before I get into it, I want to emphasize that I think the authors are exploring a very interesting and important topic. It is just that the topic is also very difficult to tackle, and the data/patterns need careful consideration before jumping to conclusions.

This study aims to understand the processes underlying recruitment patterns into host-associated microbial communities. The central finding is that the order of arrival (i.e., recruitment/detection) is non-random. Specifically, that species which are more closely related to community members recruit sooner than species which are less closely related to community members. While I am pleased to see many improvements to the manuscript, I still have significant lingering concerns.

My foremost concern, as I stated in my initial review, is that, “I don’t think it is acceptable nowadays to use patterns of phylogenetic relatedness in communities to infer community assembly processes in a dynamic system.” There is now strong theoretical evidence that species in a community may be underdispersed (clustered) due to either competition or environmental filtering. The link between community patterns and process depends critically on the traits and nature of competition/environmental filtering occurring in the community (e.g., Mayfield and Levine 2010).

**\*\*\* We understand the reviewer’s concern that if underdispersion can arise from either environmental filtering or from competition, it is not possible to separate the two using community phylogenetics.** We believe in a Popperian sense that we can never prove things in science – in particular \*prove\* the process behind the pattern - but only disprove or, in a Bayesian sense, inform a posterior with evidence. Here, we disprove the hypothesis that there will be competitive exclusion of close relatives in the assembly of microbiome communities, insofar as a direct consequence of “competitive exclusion of close relatives” is overdispersed arrival. For Bayesian brains, we provide evidence modifying a prior on the relative importance of close relatives competitively excluding one-another, or, for other models, our evidence modifies the prior distribution of the intensity of competition as a function of phylogenetic distance. We have changed the abstract and introduction to make this important epistemological point more clear – that our phylogenetic analysis disproves some models of community assembly, leaving us with a subset of possible processes behind the pattern - and appreciate your careful attention to this point.

Moreover, in this paper, we are two steps further removed from process than in many other studies on community patterns: (1) we are ignoring the importance of species abundances to community and immigration dynamics,

**\*\*\*We have addressed this issue by adding the null model the reviewer requested in their second comment and found that the microbiome datasets we analyzed are all underdispersed regardless whether or not we used an abundance-weighted null model. We’ve included these results in the SI for those readers similarly curious about species- vs. individual-neutral models of arrival (in the particular assumption in which the abundances we observe reflect metacommunity abundances or propagule pressures).**

and (2) we are ignoring the fact that phylogenetic relatedness does not always correlate with functional similarity (Cavendar-Bares et al 2014).

As Robert MacArthur said,

*Ecological patterns, about which we construct theories, are only interesting if they are repeated. They may be repeated in space or in time, and they may be repeated from species to species. A pattern which has all of these kinds of repetition is of special interest because of its generality, and yet these very general events are only seen by ecologists with rather blurred vision. The very sharp-sighted always find discrepancies and are able to say that there is no generality, only a spectrum of special cases. This diversity of outlook has proved useful in every science, but it is nowhere more marked than in ecology.*

Here, we show the repetition of underdispersion across microbiome datasets which statistically overcomes scenarios in which phylogenetically distant species are unusually similar. It’s both true that 16S phylogeny correlates with function (Langille et al. 2013, Zanevelt et al. 2010) and a sharp-sighted observation that scenarios, from horizontal gene transfer to convergent evolution, do exist to disrupt the correlation between phylogenetic relatedness and functional similarity. However, the repetition of this pattern, even absent our ability to pin down precisely the process, is the foundation for theory regarding the subset of possible processes at play and predictive power of 16s phylogeny for assembly-related functions. We don’t claim that 16S phylogeny perfectly predicts function any more than a weather forecast predicts the time and place a raindrop will fall, but the existence of statistical patterns discovered here may be useful for making bets about the processes behind microbial community assembly.

In other words, I am not convinced that there is anything we can conclude about process here (or at least I haven’t been sufficiently convinced of it by the Introduction). As such, I do not agree with the overall framing of the study in the Abstract and Introduction; in particular, I do not think it is worthwhile to bring up overdispersion / underdispersion / naturalization / preadaptation hypotheses if they are not readily useful for connecting pattern and process. Now, it may be that reporting the pattern of phylogenetically underdispersed recruitment is useful from an applied standpoint (and the authors raise this point in the Discussion). However, in my opinion, the Abstract/Introduction should focus on that applied angle, and not on trying to provide support for conceptual frameworks that are ultimately unable to unambiguously connect pattern and process.

**\*\*\* As stated above, we understand where the reviewer is coming from. We have edited the abstract to focus on the empirical results, i.e. that species are more likely to be recruited into human gut and skin microbiomes if a close relative is already present. We have also edited the first part of our Introduction similarly. We have changed our Introduction with the reviewer’s comment in mind, especially in regard to making sure that our overdispersion / underdispersion / neutral hypotheses are clearly in regard to empirical patterns in data. We humbly add, in private, that it is the reviewer’s respected opinion, and not scientific fact, whether or not it is “worthwhile” to present such a clearly repeated pattern which reduces the set of possible processes driving microbiome community assembly to an informative subset. We hope reviewer 1 and the editor can appreciate the value of our work, and take reviewer 2 as additional evidence that people outside our list of authors find this to be a worthwhile pursuit.**

A closely related concern involves the assumptions that go into the neutral/null model, which interfere with my ability to fully believe the overall conclusions. First: the neutral/null model in this study assumes that all species in the regional species pool are equally likely to arrive. The authors then treat any deviations from this null model as a non-random pattern of arrival. However, in reality, a true neutral/null model should assume that all \*individuals\* are equally likely to arrive, not all species; in other words, the relative abundances of species (zOTUs) in the regional species pool should significantly impact colonization patterns, and thus should affect null expectations about arrival times.

**\*\*\* As stated above, we now include a null model that samples individual observations instead of sampling species. This null model is very similar to what is used in Raup-Crick or Beta-MNTD analyses, although subsequent analysis of simulated communities is obviously different in our implementation. Per the reviewer’s comment, this “individual” null model uses relative abundance information from a rarefied OTU table. For the Moving Pictures dataset, we now include a supplementary figure that shows the results of this null model. In this case, the results of the “individual” null model strongly support the results of our original analysis, in that palm and fecal communities exhibited significant underdispersion, but tongue communities did not. For the other two datasets, all communities exhibited significant underdispersion under the “individual” null model, which mostly aligns with what we found earlier. We thank the reviewer for suggesting this alternative approach, because as the reviewer noted, readers may worry that underdispersion may be a conclusion determined by the relative frequencies of species within a data set, and these results address that concern.**

Indeed, in some cases the composition of the community itself will partly arise from the relative abundances of taxa in the regional species pool, e.g., if some populations persist or are common simply due to propagule pressure.

**\*\*\* We agree with the reviewer that community composition can partially arise from the relative abundances of taxa in the real regional species pool, and that this species pool isn’t truly reflected in the species observed in a data set.**

A second assumption of the null/neutral model is that the regional species pool is static over time, even though it likely is not; this assumption isn’t as problematic to me, but it should probably be mentioned.

**\*\*\* We agree with the reviewer that this is worth mentioning. We have included this in our discussion section, next to the aforementioned discussion of regional species pool being pre-underdispersed.**

Overall, I feel the authors need to explore these assumptions, and how they might impact their results. Are we sure that these assumptions do not fundamentally undermine the authors’ main conclusion of underdispersion?

**\*\*\* We found that the individual-neutral model corroborates our previous findings. We believe these changes address your revisions and will improve the robustness and impact of our paper.**

A few other minor things:

- I like the added section about detection and recruitment in the Discussion (L330-L343). Indeed, these details seem critical to understand the study. Because of their importance for comprehension, I think this clear description should be present in the Introduction, not the Discussion.

**\*\*\* We agree with the reviewer that this is an important point. We have expanded our treatment of detection vs. recruitment in our introduction.**

- Related: The shift to using “detection” is an important one, but it has been overdone in my opinion. I think it is important to state early and often that there are methodological limits to our ability to detect recruitment. However, I think this only needs to be mentioned a few times, and then “recruitment” should be used thereafter. If the paper is ultimately trying to understand recruitment patterns, then it is confusing to not use the word recruitment when discussing hypotheses and results. As it stands, the reader has to mentally translate “detection” to “recruitment” over and over. It leads the reader to repeatedly ask the question: If detection ability is such a problematic experimental limitation to inferring recruitment, then how can we justify using it to understand recruitment patterns?

**\*\*\*We agree with the reviewer that the use of “detection” everywhere is perhaps overdone. We have made the distinction between the two clearer in the introduction and discussion sections as advised by the reviewer, and using “recruitment” in those sections following said distinction except where the word is used to describe empirical data.**

- It should be made clear in the abstract that overdispersion/underdispersion results in this study refer to the \*order\* of arrival.

**\*\*\*We agree with the reviewer and have made this change.**

Referee #2 (Comments to the Author):

I thank the authors for carefully and thoughtfully addressing all of my major concerns. The authors very clearly modified the manuscript to identify the major limitations and assumptions of the study and to avoid unsupported generalisations. I believe the revised manuscript presents a very strong study and will be of broad interest to those in microbial ecology.

**\*\*\*We hope the current revisions to our manuscript also meet with the reviewer’s approval.**