Comments to the authors.

I appreciate the efforts from the authors to address my previous comments.

I think the manuscript has improved but it still needs some changes and has a lot of potential for additional data mining.

Venn diagrams.

Unfortunately, Figure 3 is still very much incorrect, which I find a bit worrying.

First of all, I would not include the Jaccard results in the Venn Diagram. The authors should use the number of BINs.

Secondly, and most important, the this is not how a Venn diagram is presented: each intersecting space should have its own value (in this case a number of BINs), but for each site, the sum of values should be 100% of the diversity (total number of BINs for that site).

If the intersection of the three sites shows 63% of BINs in common, that means that each of the site should have a remaining % of BINs of 37% (100-63). So, the values currently included in the diagram that show a % of similarity between 73-75 (which are supposed to be the shared BINs between two sites) should be reported as 75-63=12% and 73-63=10%. Because these are the BINs in common between the two sites minus the BINs that are in common to all sites. Similarly, if site SHI (for example) has 63% of BINs share with the other sites, 10% shared only with GAN and 10% shared only with QIU, then it should have a total of 17% BINs present only in SHI (where you have included the H value).

The authors should use the number of BINs and, in bracket, the %.

For example, at the centre of the diagram, they should put the number of BINs shared by all sites: (this number needs to be calculated correctly, I am using here an approximation) ~1386 BINs (63%), at the intersection of GAN and SHI they should put ~220 BINs (10%), and so on. Remove from the Venn diagram both Jaccard and Shannon indexes, those can be reported/discussed in the results/discussion.

Alpha diversity measures.

I appreciate the use of Shannon and Jaccard diversity measures (for alpha-diversity), but:

- The authors should specify in the materials and methods what software they used to calculate the alpha diversity measures.

Beta diversity measures and testing.

Since the authors have calculated the alpha diversity, additional tests could explore the beta diversity across the different sites, including metrics calculating abundance, and phylogenetic distance. This would strongly benefit their work, allowing to determine if the different sites had a different beta diversity (were some species more present in some of the sites?). Adding phylogenetic metrics (such as Philr) would also allow the authors to determine if the insect diversity changed across sites based on the different insect groups (e.g., are there more dipteran insects in one of the sites?). This would add much more value to the work, since it would provide information on the composition of the different traps and the abundance of each BIN.

HTS/NGS.

High throughput sequencing (HTS) and Next Generation Sequencing (NGS) are substantially synonyms. However, HTS has been recently replacing NGS (which is considered outdated). This is due to the fact that NGS was originally used to highlight the novelty of the technology, which has now been used for more than 15 years.. and it is not that novel anymore. The author should use consistently HTS and remove mention to NGS.

Discussion.

As pointed out by other reviewers, it seems that the discussion is not really discussing in depth the finding of the work. This is a pity, since the finding are very valuable put are not put in any context. They are just listed, with no comments. A few examples:

- **Lines 246-248:** “Interestingly, the number of BINs revealed at each of the three sites were similar, but 25% of the BINs were unique to each site.” Why/how is this interesting? How does this result compare with other works? Would the author have expected a different result?
- **Lines 248-249 (just following the comment above):** “This result supports the utility of BINs as an effective approach for counting species (Hebert et al. 2016) and assessing insect diversity (Telfer et al. 2015).” How is this result supporting the utility of BINs? Discuss! If the authors had not used BINs, the results would have changed? I am not saying that BINs proved useful.. but it is not “this result” that supported the choice of using BINs.
- **Lines 249-255:** This is good! The authors have presented their results and discussed the findings in light of known/referenced works. This is how the whole discussion should be!
- **Lines 258-262:** If more than half of the BINs have been encountered only once, this is **NOT** seasonality. These are “rare species” or “low abundance” species. If it was seasonality, you would have larger populations at some time points of the sampling. For example, seasonality of some species is linked to the flowering of the plant, in Spring, and you would have recorded high number of individuals of these species during Spring vs a low number in Winter. But recording a single specimen across a two-years experiment, means that is not linked to the seasonality. So, why did the records show so many insects recorded only once? Where these species not associated with Citrus, windblown on the orchard?
- **Lines 263-270:** The authors went through all the effort of classifying their BINs as pests or beneficials but it seems they didn’t really use this data in their analysis. Does the percentage of pests/beneficial changes across sites? If yes, why do you think that is?