In their manuscript, the authors explore the complex relationship between rates of protein adaptation and population size. The authors provide evidence that a positive correlation between the rate of adaptation and population size is visible only within the lowest range of population sizes where evolution is mutation-limited. This is indeed shown very clearly by figure 2. The authors also suggest that there might be a negative correlation between the rate of adaptation and group-level, overall population size that is compatible with a Fisher's geometric model of protein adaptation where proteins in small populations tend to be further away from their optimum, thus leaving more space for adaptive steps.

Despite some limitations that are well acknowledged by the authors themselves, the manuscript represents an important milestone for the understanding of how adaptation is influenced by population size. The main limitation of the manuscript is the "small" number of data points in figure 1 for the group-level analysis. Despite the impressive sampling and sequencing effort, the ten data points do not make it possible to conclude firmly that there is a negative correlation between adaptation rate and population size at this scale. But the sampling effort that would be required is unrealistically large and cannot be asked of the authors. This preliminary evidence is still extremely valuable and will certainly pave the way for future studies. I could see other reviewers pointing out that the small number of data points does not provide enough power, but this would be missing the point of the message of the paper. It is a first foray with an impressive, yet still inconclusive sample size, but that shows the way to the field for making progress.

This is what I like the most about this study. It really puts some order to the former literature mess, and really shows a clear path toward understanding the problem at hand. The introduction is also an excellent recap of the recent progress made.

In addition to the reasonable explanation of Fisher's Geometric model, the authors could also discuss the possibility that adaptation itself could have decreased π_s, which could contribute to the group-level negative correlation. The authors should also discuss the possibility that the effect of adaptation itself on diversity might hide a group-level positive correlation, and that at the very least, future simulations will be needed to see the selective/population size regime where this could, or could not happen.

Always about figure 1, the authors mention that for figure 1B, omega-a-a is unbiased estimate of the adaptive rate. However, the authors do not specify if for figure 1A, omega-a-p is also an unbiased estimate. This should be specified because as it is, the reader is left wondering.

My main remaining concern has to do with weakly advantageous mutations that do not fix fast enough that they can be neglected in the estimation of the number of nonsynonymous polymorphism. If there are more weakly advantageous mutations with a smaller intensity of selection (2Ns) in smaller populations, and if the still-segregating adaptive variants bias the estimate of wa downwards, then this could also explain the within-group positive correlation particularly visible in groups with small π_s. This possibility really depends on the ability of Grapes to deal properly with weakly advantageous mutations. From Galtier Plos Genetics 2016

S1 text, it looks like weakly advantageous mutations that still segregate are well taken into account, with a simulated intensity of 2Ns=20. However, it would be great to see the performance of Grapes across a wider range of selection intensities, and also when coding sequences experience a mix of weak and strong selection intensities. This would lift a small remaining doubt I have about the robustness of Grapes relative to selection intensity, and how this could influence the results presented in the manuscript. Maybe the authors just need to provide more information about Grapes in their manuscript to address this.

Finally, in order to make the manuscript even more thorough than it already is, the authors could add a paragraph of discussion about how interference between nearby advantageous mutations could potentially decrease the rate of adaptation when π_s is high.

Overall, this manuscript represents a very solid contribution. Most limitations are well acknowledged already, and the few things left unanswered are easy to address. It should also be pointed out that this manuscript adds to a growing body of work that highlights the relevance of Fisher's geometric model regarding protein evolution (for example, recent papers from the Lohmueller lab). It is reinsuring to see that different labs and approaches are converging to a similar conclusion that Fisher's geometric model may explain differences in both adaptive and deleterious rates across species with distinct complexity and population sizes.