

Reviewer #1 (General assessment and major comments (Required)):

The authors use dN/dS-based statistics to show that the S1 domain has been adapting in coronaviruses OC43 and 229E, but not in NL63. For at least OC43 and 229E, this agrees qualitatively with earlier studies. The authors should make a detailed, quantitative comparison with these earlier results.

Comments/questions:

1. I had a hard time understanding exactly what was going on in the simulations, and whether, e.g., they produced phylogenies that looked like the real ones. But perhaps the simulations should be cut or de-emphasized. From my understanding of the Bhatt et al papers and the authors' code, their selection-detection method is entirely site-based and doesn't use the tree at all, and so should be insensitive to recombination a priori. If this is correct, the simulations are largely unnecessary (but see the next point).
2. It looks like there is substantial geographic population structure, as well as time-varying geographic biases in sampling. This seems like it could generate false signals of substitutions, where standing spatial variation appears to be temporal variation. Can the authors bound how large this contribution could be? If not, perhaps they could still say that they've found signatures of either global positive selection or local adaptation, both of which are interesting.
3. Why is synonymous divergence also much higher in OC43 S1 than it is in RDRP or S2?
4. As I mention in the general assessment above, the authors should compare their results to previous studies. This includes both overall substitution rates as well as specific sites found to have repeated substitutions. For NL63, how does the null result compare to what's known? I believe Kiyuka et al, J of Infectious Diseases 2018 found widespread reinfection by similar NL63 genotypes, which may be relevant.

Reviewer #1 (Minor Comments):

1. Fig 2 caption: Maybe clarify that 192 and 262 are just shown as illustrative examples---when I first read this, I started looking through the paper to see what was special about them.
2. Fig 3: Maybe make the axis ranges uniform across subplots?
3. Fig 4 seems like it should be a supplement to Fig 5, along with the corresponding plots for OC43 lineage B and 229E.
4. It would be helpful to have a more detailed description of the method for the TMRCA analysis. Not repeating what's in the PACT paper, but just explaining how it's applied in this dataset and how it's affected by recombination.
5. I love that the authors made an accompanying nextstrain page and comprehensive github repo. I had some trouble navigating the nextstrain page (trying to change strains between OC43 and 229E using the dropdown menu under Dataset resulted in an error page). The github repo should probably be cleaned up a bit at some point (remove system files, fix error messages, more function descriptions in, e.g., bhatt\_bootstrapping.ipynb).

Reviewer #2 (General assessment and major comments (Required)):

The paper by Kistler and Bedford explores whether adaptive evolution has led to diversification of coronaviruses responsible for the common cold in the human population. The paper is very well written and presents evidence for adaptive evolution in some strains and lack of evidence in others. I enjoyed reading the paper, and all my comments are relatively minor.

Reviewer #2 (Minor Comments):

1. Line 141-2 states that more repeated mutations are expected if they confer immune avoidance. This is indeed correct, but it is also worth mentioning that an alternative hypothesis for repeat mutations is high mutation rate and genetic drift. This has in fact been shown for many recurrent mutations observed in SARS-CoV-2, and is also purportedly associated with higher mutation rate at certain types of sites (see Van Dorp et al. Nature Communications 2020).
2. Figure 3 shows higher synonymous rates at OC43 lineages for Spike/S1 and. Could the authors discuss reasons? How does this affect their results and interpretations on non-synonymous divergence?
3. The authors employed an alternative to the canonical dN/dS metric for adaptive evolution, which is convincing. It would be nevertheless nice to see results on dN/dS using codeml or some similar model-based phylogenetic approach, which takes into account tree topology and many other factors. Fig. 3 seems to support higher dN/dS in line with the authors hypotheses but it would be useful to see the results.

4. The authors detect recombination in the data, how does this affect the results shown in Figure 3?
5. All in all, figure 3 is a key figure, but it is very small (much smaller than similar supplementary figures). Why the lack of data for RdRp in OC43 lineage B?
6. Line 192 specifies a linear accumulation of adaptive substitutions. What is the interpretation ?
7. It would be useful to keep the colors consistent across different figures, for example, Fig. 1A and 2A with regards to lineages A and B.
8. I am a little concerned that divergence was calculated relative to the first sequence ("founder"), since this does not necessarily represent the true ancestral sequence. Due to the ladder-like tree structure it probably is a good approximation, but could this be validated by ancestral sequence reconstruction?
9. Table 1 show lineage A results only, what about lineage B?
10. Line 262: differences in TMRCA indicate selective pressure and/or recombination, but not necessarily both.
11. It is intriguing that NL63 does not display evidence for adaptive evolution and this merits more discussion why the differences among different strains that all infect in a similar manner and presumably should be under similar selective pressures.
12. The title reads as if all seasonal coronaviruses experience adaptive evolution; whereas the authors do not find this. I would thus suggest changing to "Evidence for adaptive evolution ... of some seasonal coronaviruses".

Reviewer #3 (Evaluation Summary (Required)):

This paper analyzing potential adaptive evolution in seasonal coronaviruses ("common colds") is highly relevant with well-supported conclusions. Its results will be widely applicable to myriad fields, including evolutionary biology, epidemiology and public health, vaccinology, immunology, and virology.

Reviewer #3 (Public Review (required)):

Kistler and Bedford present a timely and highly relevant analysis of adaptive evolution in seasonal "common cold" coronaviruses. Overall, I find the research compelling, well-performed, and mostly well-presented. The research contains sufficient statistical rigor with commendable computational reproducibility to be considered highly reliable. The authors conclude that at least two of the four known common cold coronaviruses have been undergoing adaptive evolution in their human hosts. These results may shed light on the emerging long-term evolutionary dynamics of SARS-CoV-2, the causative agent of COVID-19, as it continues circulating in humans, as well as informing ongoing vaccine design.

Reviewer #3 (Recommendations for the authors (required)):

Kistler and Bedford present a timely and highly relevant analysis of adaptive evolution in seasonal "common cold" coronaviruses. Overall, I find the research compelling, well-performed, and mostly well-presented (though some minor organizational changes are needed). As always, the commitment to open code and data from the Bedford lab is admirable and successfully performed/communicated. In my comments below, I offer advice to improve the clarity and presentation of the paper, with a few small requested analyses or need for further explanations.

1. On Page 3 line 105: "We have arbitrarily labeled these lineages 'A' and 'B' (Fig. 1)." The figure then shows A/B panels for different hCoV strains, which is a little confusing at first until you orient to the figure presentation. I recommend labeling the phylogeny in panel A with "A" and "B" rather than just using colors, and including that lineage information in the caption.

2. The caption for Figure 1 needs \_units\_ for the clock rate. Presumably this is the codon sub rate per year, which is used elsewhere? Or is this nucleotide? Similarly, units are also needed...

-->Page 7 line 204 for the parenthetical "(or 0.45 adaptive substitutions each year)". Are the authors converting to nucleotide?

--> Page 14 lines 388-389 page 14

3. In general, I find the references to supplementary figures in the text confusing. For example, I first read the phrase "Fig 1 Supplement 1A" to mean both Fig 1 and Supplement 1A. Writing this as, "Supplement 1A, Fig 1" will make it more clear that Fig 1 of the main text is not being referenced.

4. Page 3 lines 119-122 sentences are not well ordered. Should be in order: 1) Though..., 2) Because..., 3) This....

5. The last two sentences in the paragraph spanning page 3-4 seem tacked on and not immediately relevant to recombination. Please move these sentences or include a paragraph break.

6. Related to the previous comment, the Results section as a whole will benefit from improved organization, specifically by creating subsections. I highly recommend adding these to improve readability.

7. Comments for Figure 2:

--> Unless it becomes too busy, small indicators (or at least in the caption to avoid figure noise) might be added to emphasize the RBD domain within S1 specifically.

--> I recommend changing the color of the asterisks in panel A to match the lineage A (red) color.

8. Page 5 line 160 authors write, "from that lineage's common ancestor." It would be more precise to say the "from that lineage's MOST RECENT common ancestor."

9. Figure 3, specifically in panel A for RdRp, leaves some ambiguous interpretations: Is the line missing for OC43 lineage B because there is no RdRp data after the early 1990s (seems unlikely?), or because there were no adaptive substitutions, in which case the orange lines should remain steady at 0? This aspect of the figure should be clarified or fixed.

--> There is a similar situation for panel C, HKU1 lineage B, in fig s3. In addition, the "C" for that panel is cut off at the bottom, so this figure needs to be slightly reformatted.

10. Page 6 line 185, the authors introduce the H3N2 analysis, but the actual analysis is not really presented or described for another 2 pages. The H3N2 comparison is definitely not part of figure 4, which this paragraph is introducing. I think this sentence is likely misplaced? Again, this comment shows that adding subsections to the Results section will be helpful for overall organization.

11. Page 6 line 179, I would like to see more details about what constitutes an \_adaptive substitution\_ in the Bhatt method, which is not as widely used as dN/dS, within the main text itself. A couple additional sentences briefly and "birds eye view" explaining what constitutes adaptive will help orient readers. The easiest way to this - just move lines 418-422 from page 14 ("Briefly,.....each of these timepoints.") to the results section.

12. Jumping off of the last comment - why wasn't dN/dS done? Given that dN/dS is more commonly applied, I think a comparison of these results to standard dN/dS analysis is merited. In fact, including a dN/dS analysis may bolster the authors' overall conclusions and/or contribute to justifying using the Bhatt method, especially if dN/dS is not sufficient sensitive for this data.

13. In general, the authors should clarify their use of the terms "positive selection" and "adaptive substitutions." The former is traditionally associated with interpreting dN/dS, which isn't calculated in the manuscript, and the latter term is more mechanistically-oriented regarding effects of mutations. Therefore, what is meant by "positive selection" in the simulations, and how does this definition/implementation compare to the authors' measurements of "adaptive substitutions"?

14. Figure 7 and its associated analysis raised some minor concerns for me. It seems like only simulation 5 replicates were performed for each condition. Is there a reason so few simulations were performed (eg too computationally expensive?). Further, mean and CI bars for only 5 replicates in Figure 7 gives the impression that there are more than 5 replicates. A strip plot would be more forthcoming about the analyses conducted here, and some additional explanation about why only 5 replicates were performed per condition would help.

15. Table 1 and its associated analysis:

--> CI's or some measure of statistical bounds should be included in Table 1.

--> Where is OC43B in the table? Was the analysis not performed on this lineage, and if so why not?

--> The authors motivate this analysis by explaining how TMRCA is meaningful for H3N2. Can the authors perform this analysis for H3N2 proteins as well to provide further context for the HCoV values, just as they did for these analyses associated with Fig 5?

16. Page 11 paragraph beginning on line 283. To motivate this analysis, the authors may also wish to cite this paper, co-authored by Trevor, that uses a downsampling strategy from empirical to study 2009 H1N1 dynamics, and shows time dependency in evolutionary metrics. <https://bedford.io/papers/meyer-time-dependence/>

--> In addition, why did the authors use simulated data here? If we have HCoV sequence data since at least the 1990s, it seems possible to have used real data here. Further explanation/justification is therefore needed.

--> All that said, looking at the CI's (assuming these are CI's - the caption needs to add this info) in Fig s7, the bounds across time points are often overlapping. One might expect that CIs would be wider as sample size decreases, which is not always the case. To my eye, the ONLY panel in this figure that truly shows the authors' conclusion is the "no recombination/high positive selection" panel.

17. Please add a reference for the sentence page 12 line 333-334 about transmissibility/pathology correlates.

18. Line 384 of page 14: IG-Tree typo should be IQ-TREE. In addition, The authors may also wish to confirm on their own end which IQ-TREE version was used - a major version 2 was released in 2020 and has a different citation. Either way, please indicate the IQ-TREE version used and make sure the citation is right for whichever was used.

19. Figure 2 supplement 1 panel C - please choose different colors. Since some transparency is used for points, it's very hard to distinguish precisely light from dark purple.

20. Grammar and spelling:

-->Page 1 line 27, there should be a comma after 229E (as in, "...two species of HCoV, OC43 and 229E, were identified...")

-->Page 2 sentence beginning on 55 "Some human respiratory illness..." is a runon sentence. Please add a comma before ",while others,".

-->Caption of figure 1: Sentence in the middle of line 134 is a runon. Please add a comma at "...for each viral gene, and those..."

-->Page 5 line 163 typo, "spikenand" → "spike and"

-->Page 7 sentence beginning on line 203 is a run-on sentence. Please add a comma before ",while the rate of adaptation..."

-->Page 12 line 335-336, again a comma is needed before "while."