Review 2: "Host genome analysis of structural variations by Optical Genome Mapping provides clinically valuable insights into genes implicated in critical immune, viral infection, and viral replication pathways in patients with severe COVID-19"

Eva Chan¹

¹NSW Health Pathology: New South Wales Health Pathology, Genomics, Australia

Published on: Feb 23, 2021

DOI: 10.1162/2e3983f5.a6e4e994

License: [Creative Commons Attribution 4.0 International License \(CC-BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

RR:C19 Evidence Scale rating by reviewer:

- **Potentially informative.** The main claims made are not strongly justified by the methods and data, but may yield some insight. The results and conclusions of the study may resemble those from the hypothetical ideal study, but there is substantial room for doubt. Decision-makers should consider this evidence only with a thorough understanding of its weaknesses, alongside other evidence and theory. Decision-makers should not consider this actionable unless the weaknesses are clearly understood and there are other theory and evidence to further support it.

Review:

This is an interesting article as it is possibly the first to delve into host structural variants as a genetic risk to covid-19 severity. Using optical genome mapping, a genomic technology that can capture large SVs, the authors identified candidate genes with plausible associations to covid-19 severity predisposition. The research presented highlights the potential complementarity of examining SVs in addition to small variants when evaluating genetic risk.

However, I have reservations regarding some of the analytical methods used.

1. Comorbidities confound association. It is unclear whether identified SVs (particularly rare SVs) relate to pre-existing (non-covid-19 related) conditions or are true predispositions to covid-19 severity.
2. Confusion around the variant frequency
 - E.g. The duplication of CCL4L2 is more common in BNGO AFR controls (45%) than in severely ill covid-19 patients (10/37=27%), suggesting the duplication is associated with African-American ancestry and not necessarily with genetic risk to covid-19. An assessment of the frequency of this duplication in asymptomatic/non-severe covid-19 patients of African-American ancestry would be important to ascertain genetic risk.
 - E.g. The heterozygous deletion disrupting APOEBC3A/B was observed in 3 of 37 severely ill patients but is present in high frequency in controls (15% in BNGO AFR, 20% in BNGO controls, 4.5% DGV, 11% gnomAD), potentially implying this deletion is LESS common in severely ill covid-19 patients. Again, it would be important to evaluate the relative frequency of this variant in asymptomatic patients.
3. Method 2: common SV analysis

- Of the 4 common SVs proposed to confer genetic risk to covid-19 severity, only two were observed in $\geq 20\%$ of the patient cohort. Furthermore, all four SVs were detected at patient frequency below the BNGO AFR controls. This implies these SVs are less common in the patient cohort thus contrary to the speculation that the SVs may confer risk for covid-19 severity.
- Interestingly, two of the four SVs failed qPCR verification across all 5 (MUC4) & 11 (MX1) patients, respectively. Explanation on whether this is likely a technical challenge due to the nature of the SV is warranted. If verification of these inversion & insertion variants is difficult, an assessment of transcript or protein expression levels may provide clues on their impact.

4. Method 3: GWAS

- Calculation of odds ratio at the gene-level irrespective of SV type does not seem appropriate because SVs can result in loss or gain of function.
- It's not immediately obvious if multiple testing was accounted for in this analysis. This is particularly important as the study sample size is quite modest compared to the number of genes tested.

5. Expression analysis (STK26):

- While I appreciate the challenge in patient recruitment, the demographic and clinical discrepancies between the control (asymptomatic) and case (severely ill) groups is nonetheless a concern. In particular, the asymptomatic group is ~20 years younger and dominated by females compared to the case cohort – variables that could significantly impact gene expression levels. Similarly, treatments for covid-19 in the severely ill patient cohort, as well as symptoms and treatments for other comorbidities, could also contribute to differences in gene expression levels. This is not to suggest STK26 has no association with covid-19 severity, but the arguments would be much strengthened if these factors can be appropriately accounted for.
- Intriguingly, the expression level of STK26 is much higher in the severely ill cohort even though only one of the 11 patients harbored a duplication. It would be interesting, and strengthen the argument for this gene's involvement as a risk factor for severe disease, if other variant types (e.g. SNV/indel) at STK26 could be identified in the other 10 patients.