Dear Dr Gonché,

Thank you for the revised version of your manuscript. Only few minor points remained to be done before I can recommend your article. Please consider carefullly the propositions of the reviewer.

Sincerely

Emmanuelle Lerat

Reviews

Reviewed by Gabriel Wallau, 08 Jul 2023 14:51

Danesh and collaborators reviewed the manuscript adding and adjusting parameters of the COVFflow pipeline, clarifying some sections, highlighting some of the limitations of the pipeline configurations and analysis along with more robust analyses. Therefore, I recomment its acceptance after minor edits as below.

More specifically, the authors improved the results section regarding the comparison with nextstrain. I recommend the authors to include the information present in one of their answers: "For example, it can select data if a column contains a certain word, allowing the user to filter data that may contain spelling mistakes or to select data from a group of laboratories that contain a common word (in our case CERBA) but don't have the same names". Only including "COVflow allows a more flexible filtering stage using the JSON file" (page 8 - line 169) don't make it clear.

=> We made the change in the manuscript to make it more clear.

The authors also created a test dataset and updated the workflow documentation accordingly. There is a divergence in documentation and the test data. The test files zip on repository are covflow_test_dataset.zip that englobes covflow_test_metadata.tsv and covflow_test_sequences.fasta and in the documentation is informed: "In the data directory, the compressed archive data_test.zip contains a fasta file (sequences.fasta) and tsv file (metadata.tsv)." So I recommend the authors to correct the name of test files, or the documentation.

=> Thank you. We modified the name of the files as it is informed in the documentation in Gitlab.

Page 8 line 166 - change Nextrain to Nextclade

Page 30 line 261 - "from raw sequence data to phylodynamics analyses." its looks like that covflow perform raw sequence reads analysis and genome assembly which is not the case. Please correct it.

=> Thank you. These are now corrected in the manuscript.