# Suggestions for SARS-CoV-2 variant analysis of Sanger sequencing traces using SeqScape Software

Detecting and confirming mutations in newly arising SARS-CoV-2 strains is an important strategy for controlling the spread of new viral lineages. Recently we published four protocols that use Sanger sequencing to analyze SARS-CoV-2 isolates. Although the data contained in the sequence files can be compared against known sequences using a BLAST™ search, in some cases a simpler and more definitive workflow is needed. In this protocol, we present options for SARS-CoV-2 variant analysis using Sanger sequencing files and Applied Biosystems™ SeqScape™ Software for secondary analysis. Although this workflow is not comprehensive, it is presented to help users begin to better understand their results. To take advantage of all the complete features of the software, please see the user guide.

1. Import files and templates
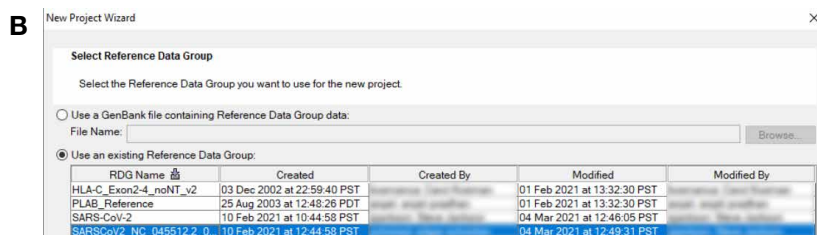
A project template, DemoProject, can be downloaded from **thermofisher.com/us/en/home/life-science/sequencing/sanger-sequencing/applications/sars-cov-2-research.html**. This template contains the SARS-CoV-2 reference sequence and known mutations in variants of concern. **(A)** Once the template is downloaded and SeqScape Software has been started, go to the Tools menu and select the SeqScape Software 4 Manager. Go to the Project Templates tab and click Import at the bottom of the window. Select the DemoProject.pt.ctf file and click Import. To load a project template in the SeqScape Software, select File –> New Project. **(B)** When the New Project window opens, select DemoProject from the list and enter a name for the project, and click New. The named project using the selected template will open.

2. Import specimens, samples, and reference sequences to a project
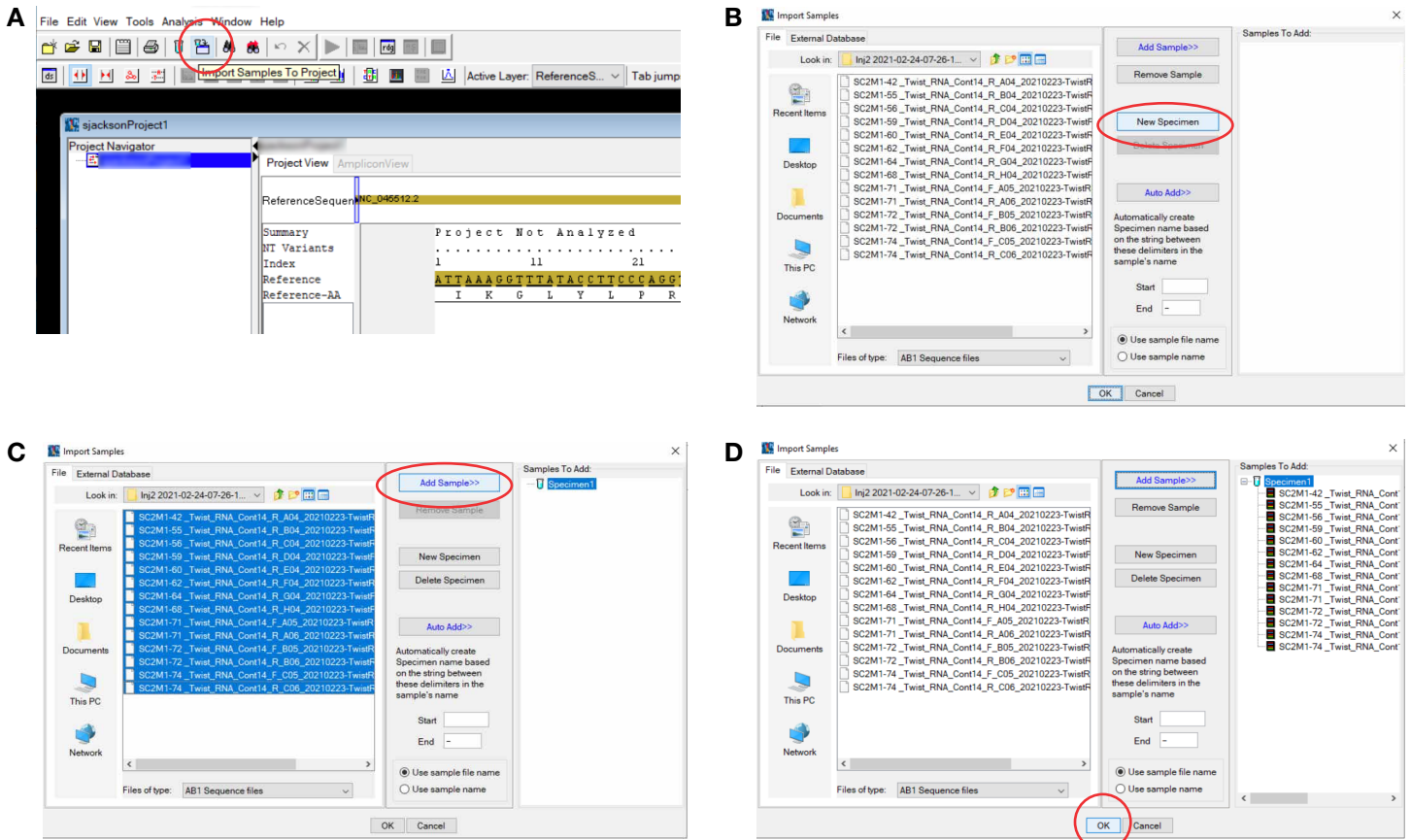
In SeqScape Software, a specimen is a biological sample that could have many sequencing traces for different genomic regions assigned to it. A sample is a trace (.ab1 file) that was derived from a specimen. **(A)** To add specimens and samples, click the Import Samples to Project (red oval) button. **(B)** In the Import Samples window, click the New Specimen button (red oval). **(C)** A specimen (test tube) should appear in the Samples To Add window. The specimen can be renamed by double-clicking on the name. Once the specimen appears in the window, navigate to the location containing the appropriate set of .ab1 files. Highlight all the files for that specimen and choose Add Sample (red oval). Note that the software can create a specimen based on file names; for more information see the SeqScape Software 4.0 User Guide. **(D)** Once all files have been added, they should appear under the specimen name. Repeat for other specimens and samples as needed. Once all specimens and samples have been imported, click OK (red oval).

3. Analyze samples
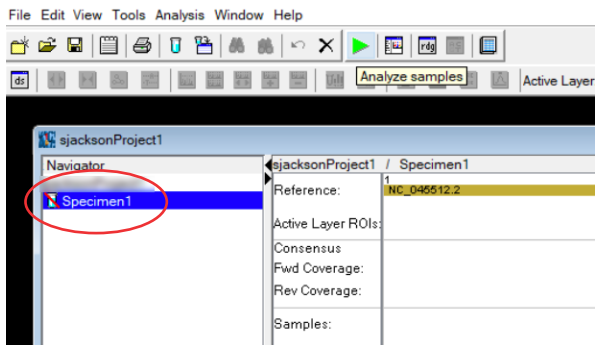
The specimen is ready to be analyzed if the green arrow icon for Analyze samples is active. To analyze your specimen, click the green arrow. SeqScape Software will align all of the imported traces to the reference genome. The specimen icon in the Navigator window will have a red line across it (red oval) until the analysis is complete.

4. View results

When the analysis is complete, SeqScape Software will display all the aligned and unaligned (unassembled) traces in the leftmost Project Navigator panel. In the Project View panel, the reference sequence is shown (top), with red and blue vertical bars above it. The red bars highlight the positions of variants defined in the Reference Data Group (RDG) template; the blue bars highlight variants all variants, including those not defined in the RDG template.

5. View assembled sequences
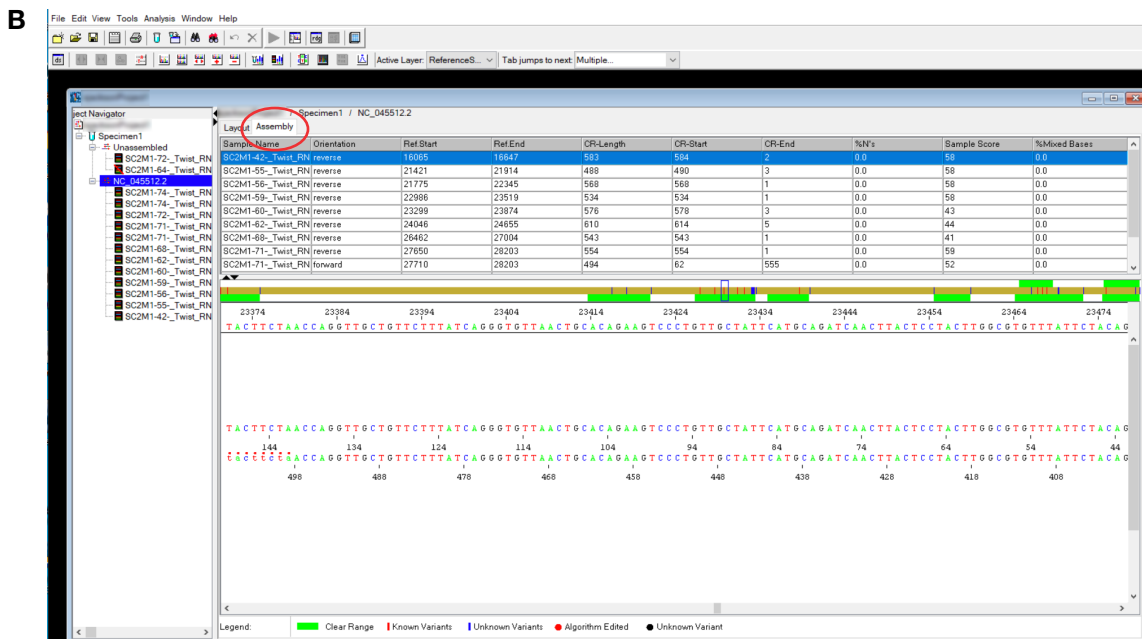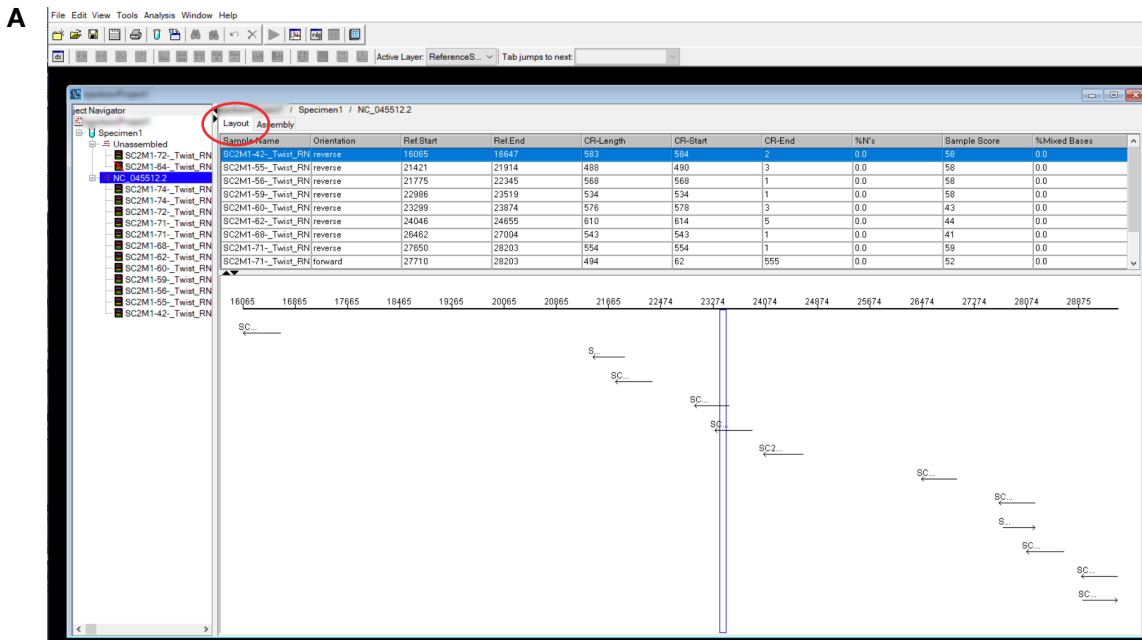
(A) Highlighting the area with the aligned sequences gives details about how the traces align to the reference. The top of the window shows sequencing quality metrics for each of the traces. In the Layout tab (red oval), the position of each of the amplicons relative to the reference can be displayed. (B) In the Assembly tab (red oval), the sequences are shown. The highlighted position can be changed by sliding the box or moving the scroll bar at the bottom of the window.

4

6. View individual electropherograms

The quality of a sequencing run can be visualized by viewing the electropherograms. To see an analyzed electropherogram, highlight the sample listed on the left, under the specimen name. Choose the Electropherogram tab (red oval) at the top of the Results window.

7. Generate report

A

B

A report listing the nature of variants can be generated. **(A)** From any window, click on the Report Manager button (red oval). **(B)** Various reports, including quality control metrics and variants identified, can be seen in the resulting screen. To view variants, click on Mutations Report. A list of all variants identified in each specimen is generated. Note that variants that are defined in the RDG file are shown as Yes in the Known column; all other variants are shown as No in that column. The blue text in the Base Change column indicates that it contains a link to the variant's electropherogram and alignment to the reference sequence. Click on the base change to view the sequence trace. To print the Mutations Report, choose File –> Print. Alternatively, choose File –> Export –> Report and export the Mutations Report as a PDF file.

8. Examine variants

A

B
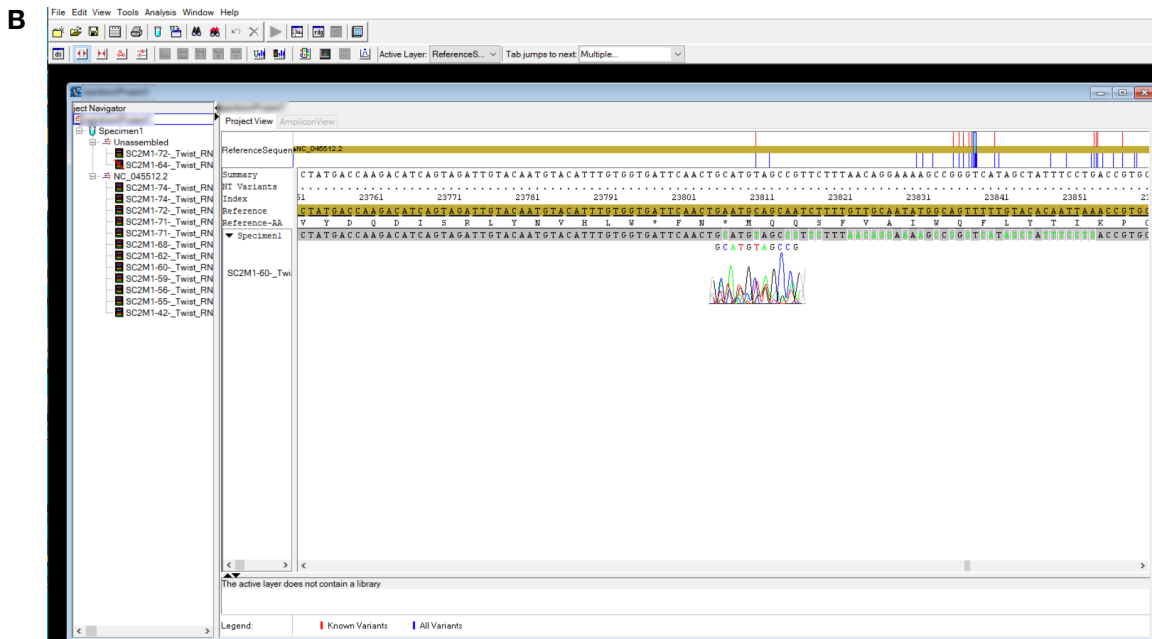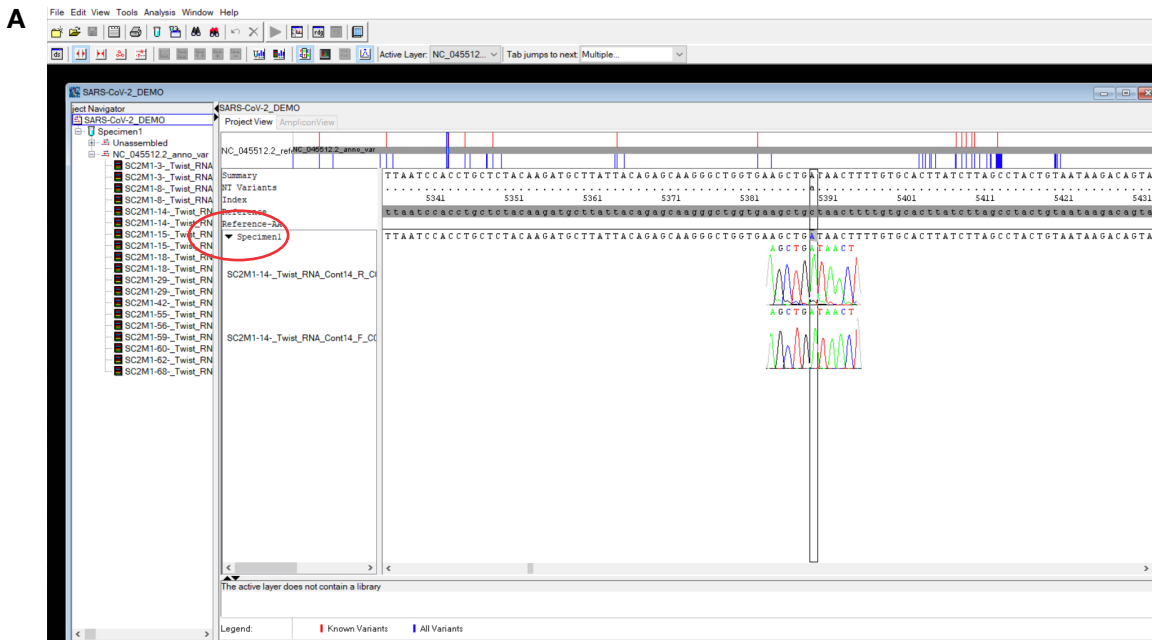
Clicking on blue text in the Base Change column of the report will display the relevant portion of the electropherogram aligned to the reference sequence. **(A)** Note that the Specimens triangle needs to be expanded (red oval) to view the electropherograms. This view can be used to quickly examine the quality of the variant definition. An example of an acceptable variant (5388C>A) is shown. **(B)** A variant that is inconclusive (and should be rejected) is displayed.

9. Mapping sequences to known strain lineages

A

B

C
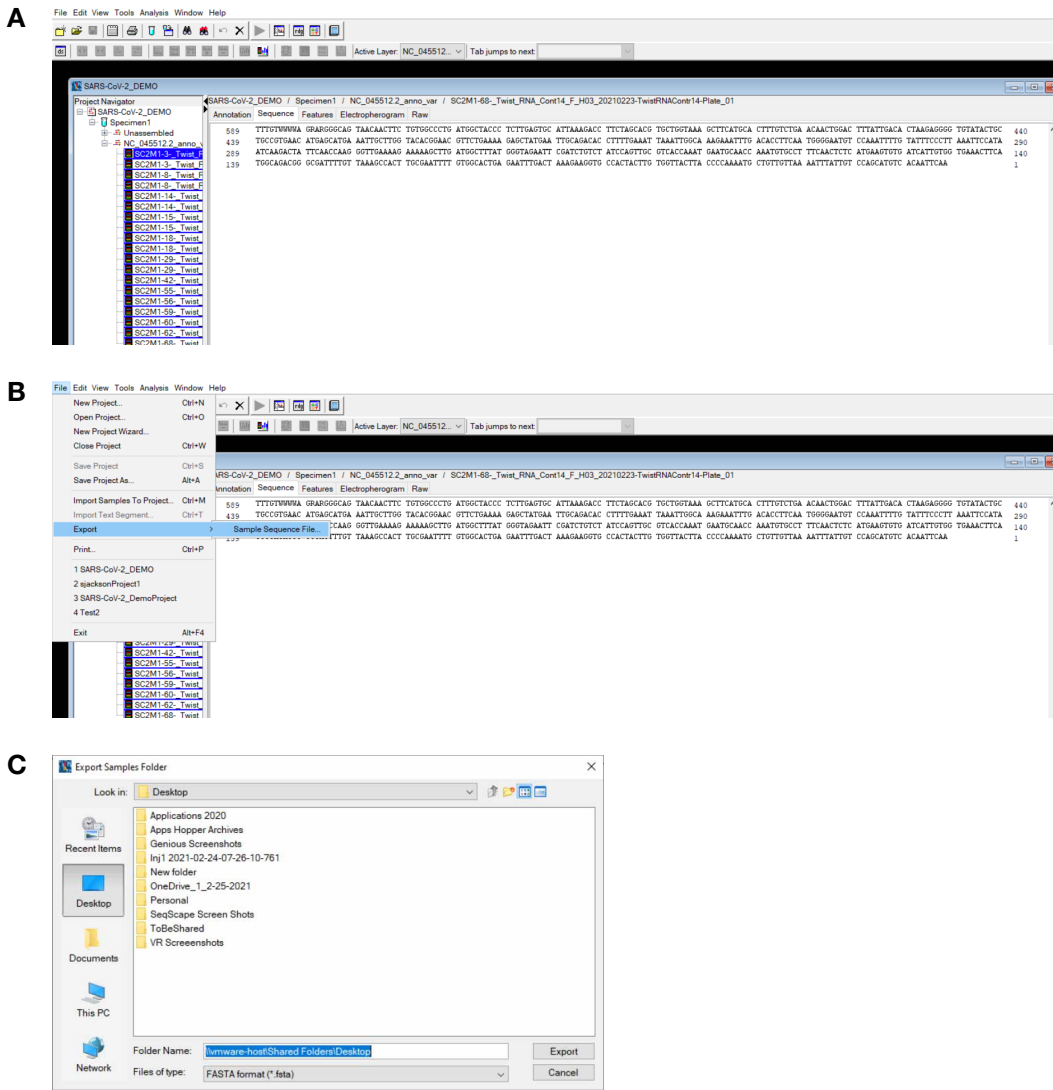
Results can be exported into FASTA files for querying public databases. **(A)** To generate a FASTA file, select all the files to export in the Project Navigator window. **(B)** From the main menu, select File –> Export –> Sample Sequence File. **(C)** Specify the location for the files. A separate FASTA text file will be generated for each of the files selected. Sequences in these files can then be used to query strain lineage databases. One of these, the GISAID database (**epicov.org/epi3/frontend#637476**), provides BLAST search results against all known SARS-CoV-2 strains. Up to 10 FASTA sequences can be queried at once. Registration is required, but it is free to use. The Coronavirus Typing Tool (**genomedetective.com/app/typingtool/cov**) also accepts FASTA files. Single FASTA files can be queried at no charge, but batch uploading requires premium access.

Find out more at **thermofisher.com/sangercoronavirus**