Feb 10, 2021

Sanger sequencing of a part of the SARS-CoV-2 spike protein

Tue Sparholt Jørgensen[1]

[1]Technical University of Denmark

1 *Works for me* dx.doi.org/10.17504/protocols.io.bsbdnai6

tuspjo

SUBMIT TO PLOS ONE

DOI

dx.doi.org/10.17504/protocols.io.bsbdnai6

DOCUMENT CITATION

Tue Sparholt Jørgensen 2021. Sanger sequencing of a part of the SARS-CoV-2 spike protein. **protocols.io** https://dx.doi.org/10.17504/protocols.io.bsbdnai6

CREATED

Feb 10, 2021

LAST MODIFIED

Feb 10, 2021

DOCUMENT INTEGER ID

47173

Protocol for Sanger sequencing of a part of the SARS-CoV-2 spike protein

1. RT-PCR on purified RNA from SARS-CoV-2 positive samples
2. Preparing PCR product for Sanger sequencing
3. Ship samples
4. Analyze results

This protocol is designed to be as simple, fast, cheap, and flexible as possible. The protocol has been running since 2021-01-16 at The Technical University of Denmark (DTU), and we continue to streamline the process to maximize the sensitivity and minimize the

time to result.
RT-PCR on purified RNA from SARS-CoV-2 positive samples
Materials
1. PCR machine
2. 96 well PCR plate with lids or film cover
3. Filter pipette tips: 1-10µL+ 10-100µL
4. 1-10µL pipettor+ 10-100µL
5. RT-PR mastermix
a. Tested with One Step PrimeScript III RT-PCR (Takara Bio)
b. This workflow should work with the same mastermix (no primers) as the normal RT-q-PCR tests. If you get amplification, you can use this protocol!
6. Primers
7. Nuclease free water
8. Extracted RNA from SARS-CoV-2 positive samples
Primers
https://github.com/artic-network/artic-ncov2019/blob/master/primer_schemes/nCoV-2019/V3/nCoV-2019.tsv
These ARTIC Network protocol primers give a detection window that includes N439K to T716I of the spike protein and all positions between them, allowing identification of all current VOCs (B.1.1.7, B.1.351, and P.1)
Tested with RPC purified primers from TAGCopenhagen (http://tagc.dk):
nCoV-2019_76_LEFT_alt3 GGGCAAACTGGAAAGATTGCTGA
nCoV-2019_78_RIGHT TGTGTACAAAAACTGCCATATTGCA
Primer stock concentration: 100uM from TAGCopenhagen

Controls
We recommend that each 96 well PCR plate be set up with the following controls:
1. 4 No Template Controls (NTC)
2. 1 positive control
a. A high titer patient sample, which has been diluted and aliquoted and is stored at -80 °C is recommended. Note that a failed Sanger reaction will NOT give a wrong mutation call, why the positive control is not as important as in e.g. the RT-qPCR SARS-CoV-2 detection.

RT-PCR setup
Setup reactions in 96 well PCR plates. Make mastermix of Primescript, water, and primers for all samples to be run and distribute to the wells before adding the samples.
Reagent Volume Final concentration
One Step PrimeScript III 10µL 1x
Diluted left primer xx µL 0.4 µM
Diluted right primer xx µL 0.4 µM
RNA sample 5µL
Water to 20µL
Notes: the amount of RNA input can be varied; 5µL is what DTU uses in the qPCR SARS-CoV-2 test.

RT-PCR program

52 °C 5 min Reverse Transcription
95 °C 10s Hotstart
95 °C 5 s Denaturation |
58 °C 30 s Annealing | repeat for a total of 45 cycles(!)
72 °C 1 min Elongation |
72 °C 5 min
4 °C ∞
Notes: the annealing temperature can be varied. 60 °C is the ARTIC protocol standard, from which the primers originated.
The PCR is supposed to run to completion, to ensure approximate equimolarity of the samples. This may require more than 45 cycles. Remember that Ct occurs many cycles before the maximum product amount is achieved in the PCR reaction, and that the 1kb product of this RT-PCR amplifies less efficiently than a shorter fragment would, so a sample with Ct35 might require 45 cycles or more before the reaction run to the end.
Optional, but highly recommended: after the RT-PCR, measure the DNA concentration of one of the NTC and one of the positive

Citation: Tue Sparholt JÃÂ,rgensen (02/10/2021). Sanger sequencing of a part of the SARS-CoV-2 spike protein. https://dx.doi.org/10.17504/protocols.io.bsbdnai6

controls as well as 5-10 samples on a fluorometric system (e.g. Qubit) or run an AGE gel of 1.5µL PCR product to confirm amplification. Nanodrop is strongly discouraged for DNA concentration measurements. The concentration of the RT-PCR product should be above 50 ng/µL for reactions, which have run to completion. IF ANY OF THE NTCS HAVE AMPLIFICATION THE RT-PCR NEEDS TO BE RERUN.

Preparing PCR product for Sanger sequencing

This protocol has been tested on purified and unpurified RT-PCR product, and both work fine. Because it saves time and resources to use unpurified RT-PCR product, DTU recommends this.

For Sanger sequencing, DTU has used Eurofins Genomics who deliver overnight results. Any Sanger sequencing provider should be fine.

The setup below assumes 75-85ng/µL RT-PCR product, it is not necessary to measure each sample.

The setup below approximates the concentrations suggested in the Mix2Seq Overnight service at Eurofins

The Mix2Seq tubes have non-sequencial barcodes, please make sure to record the position of each Mix2Seq tube as well as the sample ID in an excel sheet or similar, this is necessary to connect the sample ID and the Mix2Seq barcodes.

Please discuss sample pickup with Eurofins when ordering Mix2Seq tubes. At DTU, we have a daily pickup by UPS at 15:00, and we get the results before noon the following day.

THE POST PCR WORK SHOULD NOT BE PERFORMED IN THE SAME ROOM OR WITH THE SAME EQUIPMENT AS THE PCR SETUP or by people going back to the pre PCR area.

Primers

For the Sanger reaction, we recommend using the left RT-PCR primer (nCoV-2019_76_LEFT_alt3). Alternatively, both the left and right RT-PCR primer can be used in two seperate reactions.

Materials

a. Pipettors 1-10µL and 10-100µL
b. 1-10µL pipette tips
c. 10-100µL pipette tips
d. Mix2Seq OverNight sample tubes or PlateSeq OverNight plates
e. Nucease free water
f. RT-PCR product
g. Left primer from the RT-PCR reaction

Setup per tube

15µL nuclease free water
2µL 10µM Left primer (nCoV-2019_76_LEFT_alt3)
1. Alternatively, we recommend diluting primer stocks to 1.2 µM and adding 17ul of this directly to the Mix2Seq tube to minimize pipetting and pipetting error.
1.5µL unpurified RT-PCR product
Cap samples and arrange for DHL pickup

Analysis of the Sanger sequencing results

The accompanying analysis program basecalls the ab1-format Sanger traces to fastq files, maps the read to the reference sequence, and search for common mutations. We will update the list of mutations regularly, after which you need to update the program on your computer, unless you use the webapp. The program is designed to be lightweight enough to run locally on any laptop. Full analysis of 96 samples should take 20s-5min depending on the system.

We recommend that you first do the analysis, and afterwards connect the Mix2Seq barcode to the sample ID.

Web app

Statens Serum Institut has set up a self contained web app (https://ssi.biolib.com/app/covid-spike-classification/run to ease the data processing by removing the need for command line interaction. the trace data is not uploaded from this webapp, it runs entirely locally for data privacy reasons

Command line interface

Requirements for command line program: Unix system, for example from the Ubuntu terminal app on windows 10 machines (press Windows button and open the Windows store. Search for 'Ubuntu', download and open, after activating the subsystems for windows (search "turn windows features on or off" in the windows search bar, scroll to Windows Subsystems for Linus, and tick the box)).

Download the program Covid-spike-classification from https://github.com/kblin/covid-spike-classification. Follow the instructions

Citation: Tue Sparholt JÃÂ,rgensen (02/10/2021). Sanger sequencing of a part of the SARS-CoV-2 spike protein. https://dx.doi.org/10.17504/protocols.io.bsbdnai6

for installing and running the program on the Github page. You can copy files into your Ubuntu path with “cp /mnt/c/User/your_username/data_location/data.zip . “

Output:
The program will output reads in fastq format, and mutation calls in csv format
The csv file looks similar to the box below, where 0=WT/no mutation, 1=mutation found, and NA=Sanger reaction failed to produce a product of sufficient quality.

```
sample,N439K,Y453F,E484K,N501Y,P681H,D614G,H655Y
EF32201820_EF32201830,0,NA,0,0,NA,NA,0
EF32201875_EF32201831,0,0,0,0,0,1,0
EF32201832_EF56456422,0,0,0,0,0,1,0
EF32201873_EF32343422,0,0,0,1,1,1,0
E343222234_EF33432243,0,0,0,0,0,1,0
EF32201222_EF32201222,NA,NA,NA,NA,NA,NA,NA
EF32201222_EF32201134,0,0,0,0,0,1,0
EF32201222_EF32201465,1,0,0,0,0,1,0
EF32201222_EF32201355,0,0,0,0,0,1,0
EF32201222_EF32201123,0,0,1,1,0,1,0
```

Sensitivity of the assay
At DTU, we have an efficiency of the Sanger sequencing assay of ca 85% overall., whereas ca 15% of samples do not produce a mappable Sanger read, mostly samples with very high Ct values from the diagnostic qPCR assay.

Citation: Tue Sparholt JÃÂ,rgensen (02/10/2021). Sanger sequencing of a part of the SARS-CoV-2 spike protein. https://dx.doi.org/10.17504/protocols.io.bsbdnai6