Customer Collaboration - SARS-CoV-2 Virus Sample Preparation Suggestions for PacBio® SMRT® Sequencing

Contents

Overview and Background

The SARS-CoV-2 viral ssRNA genome is positive strand, 5' adenylated cap and 3' poly-A tailed. While it is capped and polyadenylated like a eukaryotic mRNA, the overall length (~29 kb) makes it too long to get contiguous sequence in one pass by the Iso-Seq® method or possibly full-length directed RT-PCR.

An overlapping tiled approach will be the best path for retrieving full-length genomes, both for RT and PCR. We recommend targeting 5 kbp sections, though up to 10 kbps could be achievable. The limit is RT and PCR, not Single Molecule, Real-Time (SMRT) Sequencing.

Since many viral sequences from this outbreak have already been deposited in Genbank and in GISAID, customers can design their own RT-PCR primers based on conserved sequences. At this time, many groups have also published RT-PCR primers and amplicons, any of which are suitable for SMRT Sequencing.

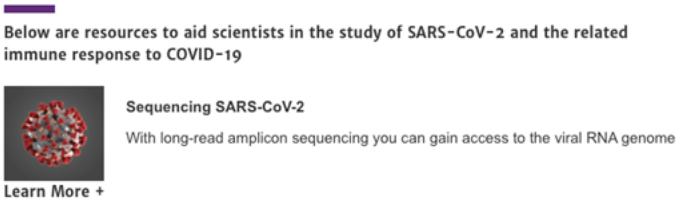
Amplicons that are already barcoded can be pooled and run on our Sequel® Systems as-is. Amplicons that are not yet barcoded can be barcoded during SMRTbell® library preparation with the following procedure:

- Procedure & Checklist – Preparing SMRTbell Libraries using PacBio Barcoded Overhang Adapters for Multiplexing Amplicons

A final option is to modify the PCR primers in a posted procedure by adding PacBio M13 sequences, using one of the two following procedures:

- Procedure & Checklist – Preparing SMRTbell Libraries using PacBio Barcoded M13 Primers for Multiplex SMRT Sequencing for multiplexing up to 1,024 samples
- Procedure & Checklist – Preparing SMRTbell Libraries using PacBio Barcoded Universal Primers for Multiplexing Amplicons for multiplexing up to 96 samples with the PacBio-supplied 96 Barcoded F/R Universal Primers Plate

We have posted links to several customer-generated PCR methods with primers adapted for PacBio barcoding on our COVID-19 webpage (Expand the 'Learn More+' link). If you have any questions or need more information, please contact support@pacb.com.



General Recommendations for Developing Your Own PCR Protocols

1. Purify RNA from cultured or uncultured sample (e.g. Trizol extraction or this kit because it has an integrated poly-A selection step). We cannot advise on the exact lab safety protocols, but certainly all virus handling would need to be done in a controlled environment to make sure the technicians are safe from infection risk. Viral or infectious disease labs should be well apprised of the necessary safety precautions.

2. Reverse transcribe the RNA* with one of the many RT enzymes available. We recommend using Superscript III or IV or Maxima RT (all available from Thermo) as these can be incubated at elevated temperatures (e.g. 55°C) to help get through RNA secondary structures.

   We are also testing MarathonRT, available in limited quantities from Kerafast. Many RT kits are available here that contain dNTPs, random hexamers and oligo-dT primer. A viral specific primer that binds to the 3' end of the RNA would work well also. Many of these kits will also have a specific "how to make long cDNA" adaptation of the protocol. A template switch oligo (TSO) like we use in standard Iso-Seq method is not necessary.

3. At the end of the RT reaction, the enzyme is heat denatured. Additionally, we suggest destroying the template RNA at this point by RNase H or RNase A. Those could be added to the heat denatured RT reaction and then incubated for an extra 10 min at 37°C.

4. Often you can do PCR directly from an aliquot of cDNA, but with especially long amplicons, adding cDNA purification may make PCR more efficient. Purify the cDNA by 1X AMPure purification or a silica column kit like this.

5. PCR amplify from the purified sample with the various primer sets against the viral cDNA. SARS-CoV-2 sequences can be found in Genbank or at GISAID.

6. Barcoding options
   - One option is to use the barcoding scheme you already have supplies for. PacBio HiFi reads are highly accurate and we can accommodate barcodes commonly used with other sequencing technologies.
   - Procedure & Checklist – Preparing SMRTbell Libraries using PacBio Barcoded M13 Primers for Multiplex SMRT Sequencing for multiplexing up to 1,024 samples
   - Procedure & Checklist – Preparing SMRTbell Libraries using PacBio Barcoded Universal Primers for Multiplexing Amplicons for multiplexing up to 96 samples with the PacBio-supplied 96 Barcoded F/R Universal Primers Plate
   - For multiplexing lower numbers of samples, or if you have not added PacBio M13 sequences to your PCR primers, you can use this strategy: https://www.pacb.com/wp-content/uploads/Procedure-Checklist-Preparing-SMRTbell-Libraries-using-PacBio-Barcoded-Overhang-Adapters-for-Multiplexing-Amplicons.pdf

*RNA input is often quite flexible with these types of cDNA synthesis protocols. It's not clear what mass can be isolated from available samples or what % of that RNA is viral mRNA, but a good general guideline is to try to go into the cDNA reaction with >100 ng of total RNA.

Internal Work on SARS-CoV-2 Sequencing

We are worked on comparing the customer generated PCR methods currently posted on our website, with a view to understanding evenness and completeness of viral genome coverage against control samples, and suitability for multiplexing with our M13 BUP approach. As we have results, we will share them on our COVID-19 webpage.

Sequencing Resources

The protocol for using a barcoded universal primer approach can be found here.

Customers can follow recommendations and link to protocols for Targeted Sequencing / Amplicon Sequencing after performing RT and cDNA amplification steps.

- Overview - Sequel Systems Application Options and Sequencing Recommendations

Customers should follow the recommendations for amplicons >3 kb, which calls for using Binding Kit 2.0 (not 2.1), loading at higher concentration, and switching to Pro-Nex bead purification.

- Quick Reference Card - Loading and Pre-Extension Recommendations for the Sequel II System
- Quick Reference Card - Loading and Pre-Extension Time Recommendations for the Sequel System

ATCC – BSL-1 standards

- Quantitative Synthetic SARS-CoV-2 RNA: ORF, E, N* – *Compatible with the CDC assay*
- Quantitative Synthetic SARS-CoV-2 RNA: 5' Spike
- Quantitative Synthetic SARS-CoV-2 RNA: 3' Spike
- Genomic RNA from 2019 Novel coronavirus strain 2019-nCoV/USA-WA1/2020
- Heat-inactivated 2019 Novel coronavirus strain 2019-nCoV/USA-WA1/2020

Twist

- Twist Synthetic SARS-CoV-2 RNA Control 1 (MT007544.1) - SKU: 102019
- Twist Synthetic SARS-CoV-2 RNA Control 2 (MN908947.3) - SKU: 102024
- Twist SARS-CoV-2 Research Panel, 2 Reactions, Kit - SKU: 102016
- Twist SARS-CoV-2 Research Panel, 12 Reactions, Kit - SKU: 102017
- Twist SARS-CoV-2 Research Panel, 96 Reactions, Kit - SKU: 102018

All Twist resources can be found on their coronavirus research tools page.

The Twist synthetic controls cover the 30 kb full viral genome. The full sequences are available here:
https://www.ncbi.nlm.nih.gov/nuccore/MT007544
https://www.ncbi.nlm.nih.gov/nuccore/MN908947

The full-length sequence is in 6 pieces, 5 kb each. There are no overlaps between the pieces (so under no circumstances can the sequences re-combine and form an intact genome). The 5 kb dsDNA pieces were synthesized, transcribed into ssRNA, and pooled. Each tube has 100 µL (plus some overage), @ 1 million copies per µL. Twist has verified that RNA target enrichment followed by sequencing recovers >99.9% of the bases.

Arbor Biosciences

Arbor Biosciences is providing complete SARS-CoV-2 hybridization-based capture kits free of charge to researchers worldwide who are studying the epidemiological and phylogenetic properties of this novel virus. The design of the probes is compatible with enriching both long and short NGS library molecules, even from degraded targets. We are actively accepting requests for this research use only panel, which is available for immediate shipment. Request a panel here.