An Improved Multiplex Targeted Amplicon Sequencing of SARS-CoV-2 Using Oxford Nanopore Technology

Luo Sun, Kaylinnette Pinet, Lynne Apone, Fiona J. Stewart, Christine Sumner, Eileen T. Dimalanta, and Theodore B. Davis

New England Biolabs, Inc.

Introduction

In December 2019, an unknown emerging pathogen that causes severe acute respiratory syndrome started to spread in Wuhan China. By mid-January 2020, there were at least 41 confirmed cases and one death. On January 12th, the genome sequence of this pathogen was published, and it was identified as a novel type of coronavirus (SARS-CoV-2) that is related to SARS-CoV and MERS-CoV. Since December, this RNA virus has spread to the rest of the world. Currently, in the U.S. alone, there are nearly 25 million confirmed cases and over400,000 deaths. Thus, developing fast, reliable, and accurate methods for SARS-CoV-2 sequencing has become a worldwide necessity.

The current and most widely used protocol for SARS-CoV-2 library preparation is based on the ARTIC Network's nanopore sequencing protocols for real-time detection of viral outbreaks, such as Ebola and influenza. As SARS-CoV-2 infections spread, the ARTIC Network nCoV-2019 (i.e. SARS-CoV-2) sequencing protocol was quickly developed and updated. NextGen sequencing technologies have been invaluable in helping to identify, confirm, monitor, and trace SARS-CoV-2 mutations. Among these NextGen sequencing tools, Oxford Nanopore Technology (ONT) has been one of the most widely accepted platforms. This is due to ONT's remarkable portability, ease of operation, and fast turnaround time. Here we report that we have optimized the ARTIC Network SARS-CoV-2 sequencing protocol for ONT platforms by simplifying steps, reducing costs, and improving amplicon coverage.

With this approach, we have assembled an all-in-one kit (NEBNext® ARTIC SARS-CoV-2 Companion Kit) that can be used in conjunction with the ONT Native Barcoding Kit/Ligation Sequencing Kit to generate targeted amplicon libraries for MinION, GridION and other ONT sequencing platforms. This all-in-one kit demonstrates improved SARS-CoV2 genome coverage and variant calling. The kit can also be easily adopted for use with current and previous ARTIC Network nCov-2019 sequencing protocol versions, as well as the PCR tiling of COVID-19 virus protocol from ONT. Furthermore, this kit can be modified by swapping out our primer pools with any other amplicon specific primer sets to sequence any other known viral genomes.

Methods

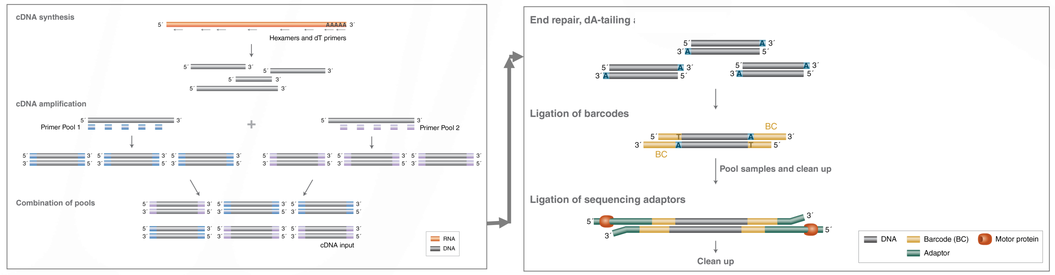


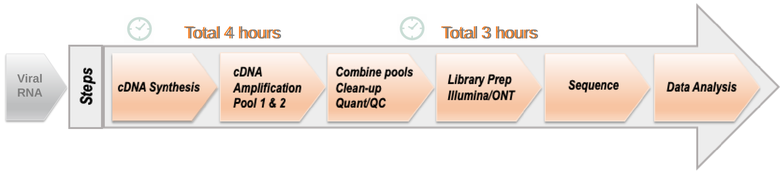
Figure 1. Workflow based on ARTIC network nCoV-2019 sequencing protocol for preparing cDNA sequencing library on ONT platforms. cDNA from SARS-CoV-2 RNA is prepared by reverse transcription and PCR amplification using balanced v3 nCoV-2019 PrimalSeq sequencing Primers. Pools of overlapping tiled amplicons are generated by two PCR reactions using two different set of primers. 50 ng of the amplified amplicons are end prepped and dA tailed followed by barcode ligation. Then, barcoded samples are pooled and cleaned by beads followed by ligation to the Nanopore sequencing adapter using 5ng from each barcoded sample. 6-24 samples can be processed and pooled for one sequencing run on MinION or GridION.

Figure 2. Components in the kit

PRODUCT

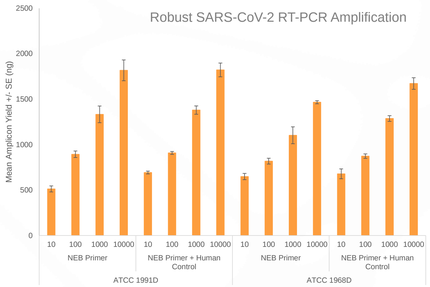
LunaScript RT SuperMix (5X)
Q5 Hot Start High-Fidelity 2X Master Mix
NEBNext Ultra II End Prep Enzyme Mix
NEBNext Ultra II End Prep Reaction Buffer
Blunt/TA Ligase Master Mix
NEBNext Quick Ligase
NEBNext Quick Ligation Reaction Buffer
NEBNext ARTIC SARS-CoV-2 Primer Mix 1
NEBNext ARTIC SARS-CoV-2 Primer Mix 2
NEBNext ARTIC Human Primer Mix 1
NEBNext ARTIC Human Primer Mix 2
Nuclease free-Water
NEBNext Sample Purification Beads

Figure 3. Time to prepare the cDNA sequencing library



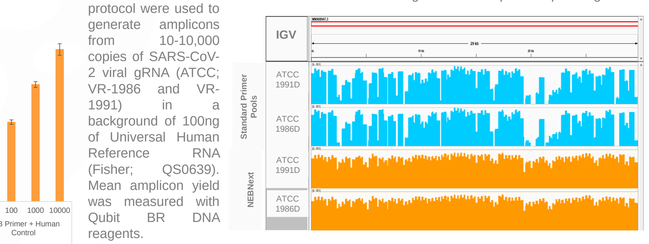
Results

Figure 4.



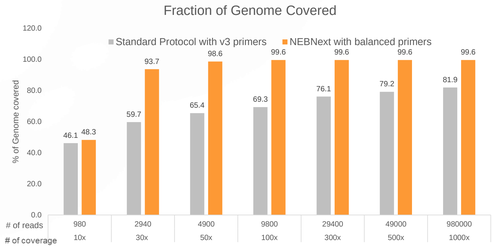
NEBNext ARTIC SARS-CoV-2 Kits RT-PCR reagents and protocol were used to generate amplicons from 10-10,000 copies of SARS-CoV-2 viral gRNA (ATCC; VR-1986 and VR-1991) in a background of 100ng of Universal Human Reference RNA (Fisher; QS0639). Mean amplicon yield was measured with Qubit BR DNA reagents.

Figure 5.



Amplicons were generated from 1,000 copies of SARS-CoV-2 viral gRNA (ATCC; VR-1986 and VR-1991) in 100ng of Universal Human Reference RNA (Fisher; QS0639) using commercially available ARTIC SARS-CoV-2 V3 primer pools and the standard RT-PCR protocol and balanced ARTIC SARS-CoV-2 primer pools. Libraries were subsequently generated using the NEBNext ARTIC SARS-CoV-2 Companion Kit for ONT. IGV visualization of reads aligned to the SARS-CoV-2 genome (minimap2).

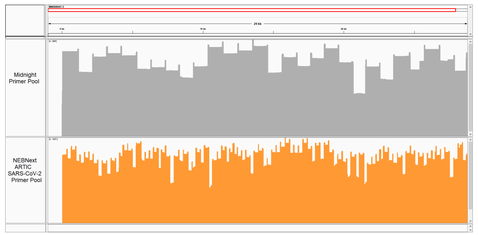
Figure 6.



Amplicons were generated from 1,000 copies of SARS-CoV-2 viral gRNA (ATCC; VR-1991) in 100ng of Universal Human Reference RNA (Fisher; QS0639) using commercially available ARTIC SARS-CoV-2 V3 primer pools and the standard RT-PCR protocol or NEBNext's protocol and balanced ARTIC SARS-CoV-2 primer pools. Libraries were subsequently generated using the NEBNext ARTIC SARS-CoV-2 Companion Kit for ONT and sequenced on GridION using R9.4.1 flow cells. Resulting data was downsampled to the data set with indicated reads number (in the figure) and nCoV-2019 novel coronavirus bioinformatics protocol was used to generate consensus genomes. Quast was used to compare the genome consensus to the reference SARS-CoV-2 sequence to obtain genome coverage percentage (as indicated by the numbers at the top of each bar). With >100x coverage, only the most 5' and 3' regions (~120 nucleotides) are missing from the consensus.

Figure 7.

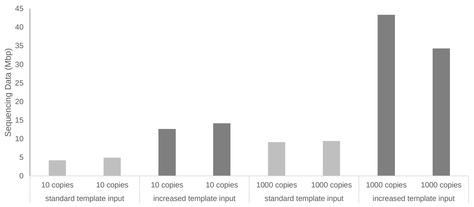
Adequate coverage of the SARS-CoV2 genome for both v3 NEB primers and Midnight primers



Midnight Primer set (IDT) was used with NEBNext ARTIC SARS-CoV-2 Companion kit to generate tiled amplicons with an average amplicon length of 1200bp. 1000 copies ATCC 1991 in 100ng Universal Human Reference RNA was used as the starting material. Libraries were prepared and sequenced. Sequencing data shows good coverage of the SARS-CoV2 genome for both the NEB balanced ARTIC primers and the Midnight primers.

Figure 8.

NEBNext ARTIC SARS-CoV-2 kit is compatible with nCov-2019 sequencing protocol v3 (LoCost)



NEBNext ARTIC SARS-CoV-2 companion kit was also optimized to be more inline with the LoCost method which omits the bead cleanup step after cDNA amplification. In this experiment, the RT-PCRs were performed on 10 copy or 1000 copies TWIST synthetic SARS-CoV2-2 RNA and cDNA libraries were prepared with no cDNA amplicon cleanup following v3 LoCost protocol. Our data (from 2hrs sequencing) demonstrated that LoCost protocol gave lower sequencing throughput (standard LoCost protocol recommended input amount). However, when LoCost was modified by increasing the amplicon input by 10-folder, there is a significant increase of data throughput.

Conclusions

- NEBNext ARTIC SARS-CoV-2 Companion Kits have a streamlined protocol and minimized costs by the reduction of reactions volumes
- The primer set in the NEBNext ARTIC SARS-CoV-2 kit has been balanced to achieve an improved genome coverage for SARS-CoV-2 genome comparing to v3 commercially available primer set
- Protocol has been optimized to have the best compatibility with nCoV-2019 sequencing protocol v3 LoCost and Nanopore PCR tiling of COVID-19 virus protocols

1. Josh Quick 2020. nCoV-2019 sequencing protocol v2 (Gunit). **protocols.io** https://dx.doi.org/10.17504/protocols.io.bdp7i5rn
2. Nick Loman, Will Rowe, Andrew Rambaut. ARTIC-nCoV-bioinformaticsSOP-v1.1.0. https://artic.network/ncov-2019/ncov2019-bioinformatics-sop.html
3. Rapid and Inexpensive Whole-Genome Sequencing of SARS-CoV-2 using 1200 bp Tiled Amplicons and Oxford Nanopore Rapid Barcoding ,Nikki E. Freed, Marketa Vlkova, ets. BioRxiv, https://doi.org/10.1101/2020.05.28.122648

sun@neb.com