Improving Multiplex Targeted Amplicon Sequencing of SARS-CoV-2 on Illumina Platforms

Kaylinnette Pinet, Lynne Apone, Luo Sun, Pingfang Liu, Andrew Gray, Julie Menin, Yanxia Bei, Jennifer Ong, Katherine Marks,
Kerri Sakowski, Nicole M. Nichols, Bradley W. Langhorst, Fiona J. Stewart, Eileen T. Dimalanta, and Theodore B. Davis | New England Biolabs, Inc.

Introduction

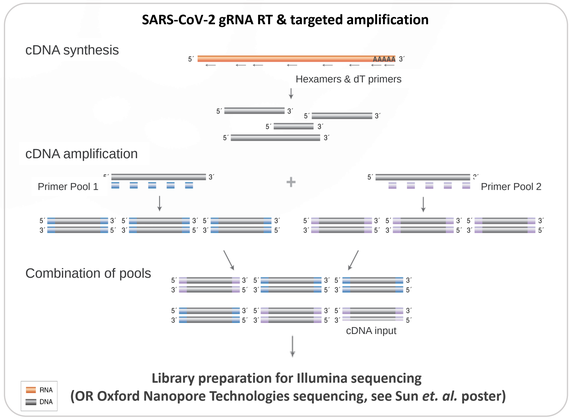
The impact of the novel coronavirus (severe acute respiratory syndrome coronavirus 2 [SARS-CoV-2]) on the global community has produced a critical need to develop reliable and accurate protocols for sequencing emergent pathogens. In response to this critical need, NEB applied our molecular expertise and resources to expand and improve upon ARTIC-based SARS-CoV-2 sequencing methods. The initial ARTIC SARS-CoV-2 sequencing protocol published in January of 2020 (Josh Quick 2020) is based on a multiplexed amplicon whole-viral-genome Oxford Nanopore Technologies® sequencing approach.[1] Here we present refined protocols and reagents for the sequencing of SARS-CoV-2 on Illumina® sequencing platforms.

The methodology demonstrated here incorporates optimized DNA polymerase master mixes that promote robust and high-throughput multiplex targeted amplicon sequencing on Illumina platforms. NEBNext® ARTIC SARS-CoV-2 Library Prep kits contain all reagents needed for RT-PCR and downstream library preparation. NEBNext ARTIC reagents for RT-PCR deliver ample amplicon yields from genomic RNA across a wide copy-number range. Furthermore, by employing a novel DNA polymerase formulation for library enrichment, we have eliminated the need to normalize amplicon concentration prior to library prep.

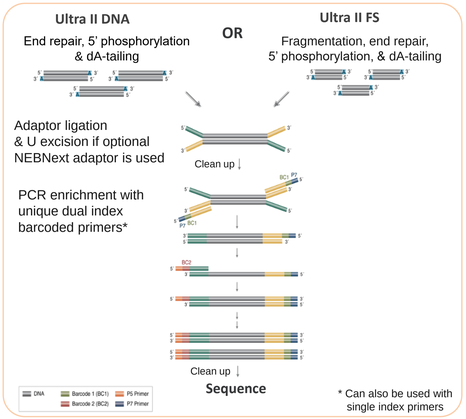
With these innovative reagents and techniques, we have developed two kits for SARS-CoV-2 sequencing on Illumina platforms. One kit follows a whole-amplicon sequencing approach, while the other utilizes enzymatic fragmentation and allows for sequencing on a 150-cycle MiSeq® flow cell. The same cost-effective RT-PCR protocol is used prior to library prep for both kits. The streamlined NEBNext ARTIC SARS-CoV-2 Library Prep protocols exhibit consistent SARS-CoV-2 genome coverage. Ultimately, these enhanced protocols and reagents enable the accurate sequencing and tracking of SARS-CoV-2 viral strains across the globe. These methods can also be readily applied to diverse multiplexed amplicon sequencing approaches, such as the identification and tracking of novel pathogens in the future.

Methods

NEBNext ARTIC SARS-CoV-2 genome amplification



Library preparation for Illumina sequencing

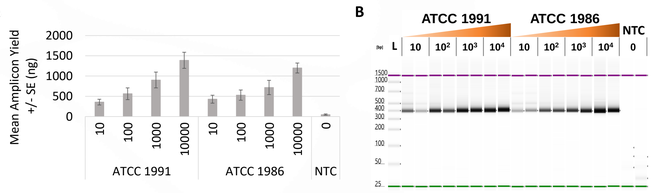


Sequencing and bioinformatics

- Illumina libraries were sequenced on a MiSeq instrument (2x250bp or 2x75bp)
- Sequencing reads were down-sampled with Seqtk and aligned to SARS-CoV-2 reference genome (NCBI, NC_045512) with Bowtie 2[2,3]

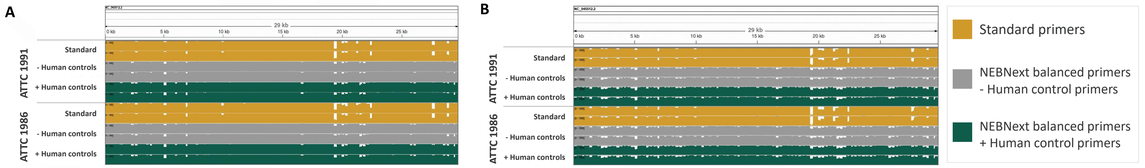
Results

Figure 1. Robust SARS-CoV-2 reverse transcription and amplification



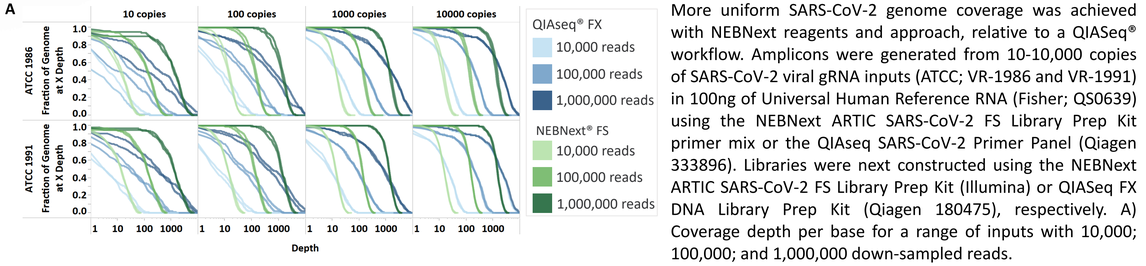
The NEBNext ARTIC SARS-CoV-2 Library Prep Kit RT-PCR reagents and protocol were used to generate amplicons from 0-10,000 copies of SARS-CoV-2 viral gRNA (ATCC; VR-1991 and VR-1986) in 100ng of Universal Human Reference RNA (Fisher; QS0639). A) Mean amplicon yield (ng) of 7 technical replicates per input across three trials measured with Qubit BR DNA reagents. B) Amplicon size confirmation with TapeStation DNA HS 1000 reagents.

Figure 2. Improved SARS-CoV-2 genome coverage depth with NEBNext ARTIC SARS-CoV-2 balanced primers



Amplicons were generated from 1,000 copies of SARS-CoV-2 viral gRNA (ATCC; VR-1991 and VR-1986) in 100ng of Universal Human Reference RNA (Fisher; QS0639) using commercially available (standard) ARTIC SARS-CoV-2 V3 primer pools or NEBNext's balanced ARTIC SARS-CoV-2 primer pools -/+ human control primers. Libraries were then constructed using NEBNext Ultra II DNA or Ultra II FS ARTIC SARS-CoV-2 Library Prep Kit. A) IGV visual of whole-amplicon library reads aligned to SARS-CoV-2 genome. B) IGV visual of fragmented amplicon library reads aligned to SARS-CoV-2 genome.

Figure 3. Optimized SARS-CoV-2 coverage depth per base with NEBNext ARTIC SARS-CoV-2 FS Library Prep Kit (Illumina)



More uniform SARS-CoV-2 genome coverage was achieved with NEBNext reagents and approach, relative to a QIASeq® workflow. Amplicons were generated from 10-10,000 copies of SARS-CoV-2 viral gRNA inputs (ATCC; VR-1986 and VR-1991) in 100ng of Universal Human Reference RNA (Fisher; QS0639) using the NEBNext ARTIC SARS-CoV-2 FS Library Prep Kit primer mix or the QIAseq SARS-CoV-2 Primer Panel (Qiagen 333896). Libraries were next constructed using the NEBNext ARTIC SARS-CoV-2 FS Library Prep Kit (Illumina) or QIASeq FX DNA Library Prep Kit (Qiagen 180475), respectively. A) Coverage depth per base for a range of inputs with 10,000; 100,000; and 1,000,000 down-sampled reads.

Conclusions

- NEBNext ARTIC SARS-CoV-2 Library Prep Kits for Illumina sequencing follow streamlined protocols with reagents specifically optimized for this workflow
- The NEBNext ARTIC SARS-CoV-2 RT-PCR reagents and protocol provide ample amplicon yields from a wide range of viral genome inputs, as well as versatile down-stream library prep options
- Amplicon normalization is not required prior to Illumina library preparation for uniform genome coverage
- Genome coverage depth has been optimized by developing a more balanced primer pool
- Libraries generated with NEBNext ARTIC SARS-CoV-2 FS Library Prep Kit (Illumina) shows improved and robust genome coverage relative to QIAseq SARS-CoV-2 Primer Panel and FX Library Prep Kit

References

1. Josh Quick 2020. nCoV-2019 sequencing protocol v2 (GunIt). **protocols.io** https://dx.doi.org/10.17504/protocols.io.bdp7i5rn
2. H. Li, Seqtk: a fast and lightweight tool for processing FASTA or FASTQ sequences, 2013. https://github.com/lh3/seqtk/
3. Langmead B, Salzberg S. Fast gapped-read alignment with Bowtie 2. *Nature Methods*. 2012, 9:357-359.