Next-generation sequencing

Improved protocol for SARS-CoV-2 genome sequencing with PCR enrichment

Introduction

The novel coronavirus SARS-CoV-2 continues to be one of the biggest global public health concerns of the modern era [1]. Faster and more sensitive analysis of mutations in the SARS-CoV-2 genome can enable tracking of worldwide transmission of the virus and provide a better understanding of infection mechanisms. Targeted amplification is a cost-effective approach in which PCR is used to generate amplicons to tile the entire SARS-CoV-2 genome prior to preparing the sequencing library. This allows generating enough coverage and depth to produce complete genome sequences even from samples that have partially degraded viral RNA genomes or low viral load.

Here we have described a robust workflow for preparing next-generation sequencing (NGS) libraries from samples that contain SARS-CoV-2 (Figure 1) using the primer panel from the ARTIC network [2], Thermo Scientific™ Phusion™ Plus PCR Master Mix, Invitrogen™ SuperScript™ IV VILO™ Master Mix, and the Invitrogen™ Collibri™ ES DNA Library Prep Kit for Illumina Systems.

Materials and methods

Samples

Clinical research specimens of bronchoalveolar lavage (BAL) were obtained with consent from the Santara Clinics Biobank in Lithuania. Each sample originated from a unique donor with qPCR-confirmed SARS-CoV-2. The quality and C_t values of the samples were within the acceptable ranges for further analysis (see “Results”).

RNA purification

The BAL samples were lysed, and RNA from the samples was purified using the Applied Biosystems™ MagMAX™ Viral/Pathogen Nucleic Acid Isolation Kit according to the standard protocol.

cDNA synthesis

The RNA from each sample was reverse-transcribed into cDNA using the SuperScript IV VILO Master Mix. Reactions of 10 µL volume were set up using 5.5 µL of purified RNA according to the standard protocol.

Figure 1. NGS workflow for the analysis of SARS-CoV-2 samples using a targeted amplification–based approach.

invitrogen

Enrichment PCR

Synthesized cDNA from the previous step was used directly for enrichment PCR using ARTIC v3 primer pools. Equal volumes of 100 µM primer stocks were combined in two pools according to the manufacturer's recommendation [2]. PCR reaction mixes (one per pool) were prepared with Phusion Plus Master Mix according to Table 1. The optimized enrichment PCR program is provided in Table 2.

Further, 12.5 µL of each enrichment PCR reaction were combined (25 µL in total for each sample) and purified with magnetic beads using the Collibri ES DNA Library Prep Kit following the "Remove EDTA from reaction" protocol.

Library preparation

Purified amplicons were converted to sequencing-ready libraries by enzymatic fragmentation using the Collibri ES DNA Library Prep Kit for Illumina Systems. The optimized fragmentation reaction setup and program are provided in Tables 3 and 4, respectively. After completion of the fragmentation program, the dual-indexed adaptors were ligated according to the standard protocol.

Table 1. Enrichment PCR reaction setup (two reactions per sample).

Reagent	Volume per 25 µL**
2X Phusion Plus PCR Master Mix	12.5 µL
Primer pool (1 or 2)*	1 µL
cDNA from previous reaction	2.5 µL
Water, nuclease free	9 µL

* Premixed pools are available from IDT (Cat. No. 10006788 was used for this study).
** Reaction volume can be reduced to 12.5 µL.

Table 2. Enrichment PCR program.

Temperature	Time	Cycles
98°C	30 sec	1
98°C	15 sec	30
63°C	5 min	
4°C	Hold	1

Table 3. Fragmentation reaction setup.

Reagent	Volume per 50 µL
Purified DNA	18 µL
Elution buffer	17 µL
10X Fragmentation and dA-tailing buffer	5 µL
Fragmentation and dA-tailing enzyme mix	10 µL

Table 4. Fragmentation program.

Step	Temperature	Duration
Block pre-cooling	4°C	As required
Fragmentation	37°C	15 min
dA-tailing	65°C	10 min
Hold	4°C	

The libraries were amplified using the Invitrogen™ Collibri™ Library Amplification Master Mix (optional step) and purified using the magnetic beads. Average library size was estimated on a Agilent™ 2100 Bioanalyzer™ instrument using the Agilent™ High Sensitivity DNA Kit. Library concentration was measured by qPCR using the Invitrogen™ Collibri™ Library Quantification Kit. Libraries were sequentially diluted to 1:10,000, and one of the dilutions was used to quantify 96 samples in two qPCR plates. Further details on the library preparation process can be found in the regular Collibri ES DNA Library Prep Kit user manual (MAN0018545) or previous SARS-CoV-2 sequencing protocol user guide, where the library preparation is performed in the same way (MAN0025323).

Sequencing and data analysis

The libraries were pooled and sequenced on an Illumina™ MiSeq™ instrument using the MiSeq™ Reagent Kit v2 Nano. Alignment with the reference genome NCBI ASM985889v3 was performed using the BWA-MEM algorithm (bwa mem – 32 -r 1.0 -k 19 -M -N 6 -v 1) included in the Burrows–Wheeler Alignment Tool.

Results

The 96 clinical research samples positive for SARS-CoV-2 were sequenced using the optimized amplicon-based workflow solution. The C_t values of the sequenced samples varied from 13.4 to 37.3, with a median value of 24.7. The average percentage of aligned reads in the collected samples was 97.6, providing mean coverage of 47.1x across the samples without amplicon failures (Figure 2). On average, less than 2% of the amplicons showed <20% sequencing depth. The average fraction of reads with ≥30x coverage was 95.3%, demonstrating high-quality results that can be used to identify SARS-CoV-2 strains in individual samples.

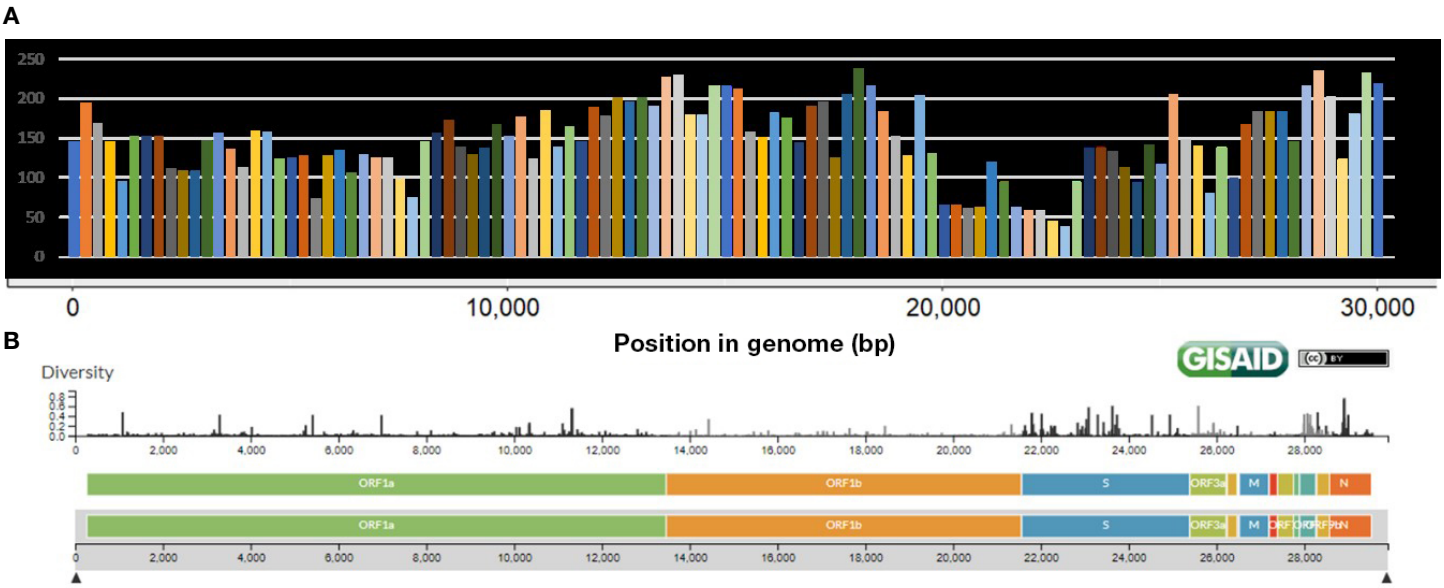

Figure 2. Average sequencing depth of 96 clinical samples positive for SARS-CoV-2 aligned against NCBI ASM985889v3 genome. **(A)** Each column shows sequencing depth of one amplicon (average from 96 samples) compared against a **(B)** snapshot of known positions and frequencies of known mutations from the GISAID database [3].

Conclusion

The amplicon-based approach using the primer panel from the ARTIC network [2], SuperScript IV VILO Master Mix, Phusion Plus PCR Master Mix, and the Collibri ES DNA Library Prep Kit for Illumina Systems provides sufficient sequencing coverage and depth for SARS-CoV-2 positive samples.

References

1. World Health Organization, Coronavirus, https://www.who.int/health-topics/coronavirus#tab=tab_1
2. ARTIC Network, https://artic.network/ncov-2019
3. GISAID, https://www.gisaid.org/phylodynamics/global/nextstrain/

Ordering information

Step	Kit and product	Quantity	Cat. No.
Purification of total RNA*	MagMAX Viral/Pathogen Nucleic Acid Isolation Kit	2,000 preps	A48310
		200 preps	A42352
cDNA synthesis	SuperScript IV VILO Master Mix	500 reactions	11756500
		50 reactions	11756050
PCR enrichment	ARTIC v3 Primer Pools	Choose a provider	
	Phusion Plus PCR Master Mix	2,000 reactions	F631XL
		500 reactions	F631L
		100 reactions	F631S
Library generation	Collibri ES DNA Library Prep Kit for Illumina Systems	≥96 preps	For the highest throughput with full Collibri kits, use A38607096 with A38607196 to support up to 192 preparations per run. For unlimited throughput, request Collibri ES DNA Library Prep Core Kit without indexes (A38607096W). For more recommendations on kits, check the Collibri webpages.
Library quantification**	Collibri Library Quantification Kit (preferred)	100 reactions	A38524100
		500 reactions	A38524500
	Qubit dsDNA HS Assay Kit	100 assays	Q32851
		500 assays	Q32854

* RNA can be purified using either the MagMAX Viral/Pathogen Nucleic Acid Isolation Kit or the MagMAX Viral/Pathogen II Nucleic Acid Isolation Kit. Optional protocol modifications to reduce sample and reagent use are described in the application note "MagMAX Viral/Pathogen kit protocol changes to enable increased SARS-CoV-2 testing throughput".

** Library quantification can be performed using the Collibri Library Quantification Kit (qPCR assay) or the Qubit dsDNA BR Assay Kit (fluorometric assay).

Learn more at **thermofisher.com/sarscov2-collibri**

invitrogen