HHS Public Access

Author manuscript

J Virol Methods. Author manuscript; available in PMC 2024 May 07.

Published in final edited form as:

J Virol Methods. 2022 January ; 299: 114346. doi:10.1016/j.jviromet.2021.114346.

Development and evaluation of a ligation-free sequence-independent, single-primer amplification (LF-SISPA) assay for whole genome characterization of viruses

Narayanan Jothikumar[a], **Theresa Cromeans**[b], **Jothikumar Shivajothi**[c], **Jan Vinjé**[d], **Jennifer Murphy**[a,*]

[a]Division of Foodborne, Waterborne and Environmental Diseases, Centers for Disease Control and Prevention, Atlanta, GA, 30329, USA

[b]CDC Foundation, USA Contracted to Division of Viral Diseases, Centers for Disease Control and Prevention, Atlanta, GA, 30329, USA

[c]COX, Atlanta, USA

[d]Division of Viral Diseases, National Center for Immunizations and Respiratory Diseases, Centers for Disease Control and Prevention, Atlanta, GA, 30329, USA

Abstract

Molecular identification and characterization of novel or re-emerging infectious pathogens is critical for disease surveillance and outbreak investigations. Next generation sequencing (NGS) using Sequence-Independent, Single-Primer Amplification (SISPA) is being used extensively in sequencing of viral genomes but it requires an expensive library preparation step. We developed a simple, low-cost method that enriches nucleic acids followed by a ligation-free (LF) 2-step Polymerase Chain Reaction (PCR) procedure for library preparation. A pan-chimeric universal primer (JS15N14) containing 15 nucleotides with a random tetradecamer (14N) attached to the 3′-end was designed. The complimentary primer (JS15) was used for nucleic acid enrichment in a first round PCR. A second PCR was designed to create Illumina sequencer-compatible sequencing-ready libraries for NGS. The new LF-SISPA protocol was tested using six RNA and DNA viral genomes (10.8–229.4 kilobases, kb) from an ATCC virome nucleic acid mix (ATCC® MSA-1008™) followed by analysis using One Codex, an online identification tool. In addition, a human stool sample known to be positive for norovirus GII was sequenced, and *de novo* assembly was performed using the Genome Detective Virus Tool allowing for near complete genome identification in less than 24 h. The LF-SISPA method does not require prior knowledge of target

*Corresponding author at: Waterborne Disease Prevention Branch, Division of Foodborne, Waterborne and Environmental Diseases, National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention, 1600 Clifton Road NE, Mailstop H23-9, Atlanta, GA, 30329, USA. njothik@gmail.com (N. Jothikumar), iod7@cdc.gov (J. Murphy).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.jviromet.2021.114346.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

sequences and does not require an expensive enzymatic library preparation kit, thereby providing a simple, fast, low-cost alternative for the identification of unknown viral pathogens.

Keywords

SISPA; Whole genome sequencing; Metagenomics; Norovirus; RNA and DNA viruses

1. Introduction

Development of PCR-based molecular methods to identify emerging and clinically relevant viruses requires prior knowledge of the genomic sequences. Alternatively, next generation sequencing (NGS) using random oligonucleotide primers can be applied to human samples or complex environmental samples to detect viruses without prior knowledge of their identity or genetic signatures. Rapid NGS-based genomic sequencing of viruses at an early stage of an outbreak or pandemic can be beneficial in that the etiological agent causing either event can quickly be identified and the appropriate public health measures initiated (Wu et al., 2020). In routine shotgun metagenomics protocols, commercial library preparation kits for detection of multiple viruses in a given sample are expensive (Glenn et al., 2019a) and time consuming (Glenn et al., 2019b).

Random amplification by sequence-independent single primer amplification (SISPA) was first reported by Reyes and Kim (1991). Initial nucleic acid enrichment was based on the endonuclease digestion of nucleic acids followed by adapter/primer ligation to target DNA fragments for PCR amplification using a single, specific complementary primer. More contemporary protocols combined SISPA with metagenomic sequencing for the identification of both known and unknown viruses. In our Ligation Free-SISPA (LF-SISPA) protocol, adapters and barcodes are directly added by PCR to obtain a sequencing ready library rather than utilizing a commercial kit.

The present study was undertaken as a proof-of-concept for a ligation-free library preparation workflow for the identification and characterization of clinically relevant viral pathogens. The LF-SISPA assay was evaluated using an ATCC virome nucleic acid mix containing RNA and DNA of six viral pathogens. The method was also applied to identify the complete genome sequence of a norovirus from the nucleic acid extract of a human stool sample.

2. Materials and methods

2.1. Nucleic acid virome standard (ATCC MSA-1008) and norovirus sample

The fully characterized ATCC virome nucleic acid mix (ATCC MSA-1008™) containing six viral nucleic acids with diverse genomic sizes (10.8 kb to 229.4 kb) and types of nucleic acid (RNA or DNA) was selected for assay development and testing. The ATCC virome nucleic acid mix contains human herpesvirus 5 (ds DNA, 229.4 kb), human mastadenovirus F (ds DNA, 34.2 kb), influenza B virus B/Florida/4/2006 (ss (−) RNA 8 segments, 14.2 kb), Zika virus (ss (+) RNA, 10.8 kb), human respiratory syncytial virus (ss (−) RNA, 15.2 kb), and reovirus 3 (ds RNA 10 segments, capsids, 23.5 kb). The estimated genome copy number for

each virus is 2×10^4 genome copies/μL. The clinical sample analyzed contained norovirus GII (ss (+) RNA, 7.5 kb; Ct 21.9, ~10^5 genome copies/μL) extracted from a UNEXP disk as previously described (Cromeans et al., 2019). Briefly, the UNEXP disk which was saturated with a nucleic acid (NA) extraction buffer was loaded with 60 μL of a clarified 10 % stool preparation; after 2 weeks NA was extracted with heat into a 10X volume of TE buffer.

2.2. cDNA synthesis

cDNA was synthesized using a SuperScript IV Reverse transcription kit (Applied Biosystems). Reverse transcription was performed in a volume of 10 μL, consisting of 2 μL of 5X RT buffer, 2 μL of 10 μM pan-chimeric primer JS15N14 (Table 1), 0.5 μL of 10 mM dNTP mix, 0.5 μL of 0.1 M DTT, 0.5 μL (200 U/μL) reverse transcriptase and 4.5 μL of nucleic acid extract (ATCC Virome-MSA-1008 or norovirus) and incubated at 23 °C for 20 min, 50 °C for 20 min followed by heat denaturation at 80 °C for 5 min.

2.3. Klenow reaction

Second strand DNA synthesis was performed using all 10 μL of the cDNA sample in a 25 μL reaction volume containing 2.5 μL of 10X Klenow buffer, 0.5 μL of 10 mM dNTP mix, 0.5 μL of Klenow fragment 3′ to 5′ exo- (Thermo Fisher Scientific) enzyme and 11.5 μL of nuclease free water. Reaction conditions included incubation at 37 °C for 30 min followed by heat denaturation at 80 °C for 5 min.

2.4. SISPA (First round PCR)

Nucleic acid enrichment was performed with the complimentary primer (JS15) without a tetradecamer at the 3′-end in a 50 μL reaction mixture consisting of 10 μL cDNA, 5 μL of 10 μM JS15 primer, 25 μL of TaqMan environmental master mix (version 2.0) PCR buffer (Life Technologies, Grand Island, NY) and 10 μL of nuclease free water. After an initial denaturation step at 95 °C for 10 min, PCR was performed for 30 s at 95 °C, 45 s at 60 °C, and 45 s at 72 °C for 45 cycles followed by a final extension step at 72 °C for 10 min. The amplification products were purified with AMPure SPRIselect beads (Beckman Coulter, Indianapolis, IN), and a 1:0.8 ratio was selected to obtain an average DNA fragment size of 392 bp as confirmed by the Agilent 4200 Bioanalyzer TapeStation using the Agilent HS D1000 tape. DNA concentration was determined using the Qubit dsDNA HS kit with the Qubit 3.0 Fluorimeter (ThermoFisher Scientific, Waltham, MA).

2.5. Sequence ready ligation-free library preparation for MiSeq illumina (Second round PCR)

Both forward (J1Fi5) and reverse (J1Ri7 or J2Ri7) barcoded primers (Table 1) were introduced during a second round PCR step. The barcoded primer (J1Fi5) included a sequencing adapter at the 5′ end and JS15 sequence at the 3′ end (Table 1). The barcoded primer (J1Ri7 or J2Ri7) included a sequencing adapter, a unique single index 8 base barcode at the 5′ end and JS15 sequence at the 3′ end (Table 1). Single-indexed PCR primers (i7) with unique indexes (that may be useful for multiplexing samples) were used. An 8-base single-indexed sequencing run workflow was adopted using indexes aaccgcgg or ggttataa (Table 1). These barcoded PCR primers were designed to be less than 60 bases.

PCR reactions were performed in a 50 μL reaction volume with 5 μL of amplified product from the PCR enrichment step, 1.25 μL of i5 primer at 10 μM, 1.25 μL of i7 primer at 10 μM, 25 μL of Environmental Master Mix (version 2.0) PCR buffer (Life Technologies, Grand Island, NY) and 17.5 μL of nuclease-free water. A short PCR amplification reaction was carried out with an initial denaturation step for 10 min at 95 °C followed by 25 cycles of 30 s at 95 °C, 30 s at 60 °C, 30 s at 68 °C. A final extension step at 68 °C for 5 min was included. The amplification products were purified with SPRIselect beads at a 1:0.8 ratio. After cleanup, the final barcoded DNA amplicons ranged between 300–700 bp.

2.6. Illumina NGS

Shotgun metagenomic sequencing was performed on the Illumina MiSeq platform (Illumina, USA). For testing of the method, only one sample per MiSeq run was sequenced. A PhiX v3 library was used as a balanced library control for Illumina sequencing runs to overcome the impact of low diversity. Read 1 and Read 2 sequencing primers (Table 1) were used on the MiSeq for paired-end sequencing (schematically illustrated in Fig. S1 in the supplementary data).

For ATCC virome mix sequencing, libraries prepared from the 2nd step PCR (J1Fi5 and J2Ri7) were sequenced using a MiSeq Reagent Kit (Illumina, USA) v3 kit (2 × 75 pb) that allowed for selection of 2 × 84 bp paired-end runs. Paired-end (2 × 84 bp) sequencing was performed by loading a final concentration of 5 pM ATCC virome library and 0.2 pM bacteriophage PhiX v3 Control library (Illumina, USA). An additional 5 μL of 100 μM Read 1 and Read 2 sequencing primers were added to well positions 12 and 14, respectively, of the MiSeq reagent cartridge. The index read sequencing primer was not included because only one sample per run was tested.

For norovirus library sequencing, the library prepared from the 2nd step PCR (J1Fi5 and J1Ri7) was sequenced and MiSeq Reagent Kit (Illumina, USA) v2 nano (2 × 250 bp) was used for its low cost when compared with MiSeq kit v3. The norovirus-positive sample library was sequenced with a paired-end run of 2 × 101 cycles at a final loading concentration of 2 pM and 0.8 pM for the norovirus and PhiX libraries, respectively. An additional 5 μL of 100 μM Read 1 and Read 2 sequencing primers were added, to well positions 12 and 14, respectively, of the MiSeq reagent cartridge. The index read sequencing primer was not included because only one sample per run was tested.

2.7. Bioinformatics analysis of ATCC virome using one codex tool

FASTQ files generated from sequencing the ATCC virome nucleic acid mix were submitted to the web-based One Codex bioinformatics tool (www.onecodex.com) on August 14, 2020 for NGS data analysis (Minot et al., 2015). The One Codex tool generated a report at the viral species level based on an examination of exact *k*-mer matches against the tool's internal viral databases. The One Codex analysis breaks each read into *k*-mers, and, in this case, the *k*-mers of each possible 31-base string in the read were compared to the database by exact alignment. The sequencing artefacts were filtered out of the sample based on the relative frequency of unique *k*-mers. The relative abundance of each virus was estimated based on the sequencing depth and coverage across every reference genome available in

the database. The One Codex analysis demonstrated the identification of viruses, and the substantial differences in virus identity percentage were not covered. This approach reduces the risk of misclassification and classifies only to the possible specificity. After confirmed identification of viral species, the sequences were aligned to the expected species using the Minimap2 add-on for One Codex, in order to provide a view of the coverage of each viral genome.

2.8. Bioinformatics analysis of norovirus using genome detective virus tool

Norovirus NGS FASTQ files were submitted for bioinformatic analysis with the Genome Detective Virus Tool (www.genomedetective.com) on April 04, 2021 to characterize the sequenced norovirus genomes (Vilsker et al., 2019). Paired-end data sets were assembled into nearly complete viral genomes by de novo assembly. The assembled sequences were used for identification and classification at the species level using reference genomes available in NCBI's RefSeq (O'Leary et al. (2016).

3. Results and discussion

The novel LF-SISPA method was evaluated using an ATCC virome nucleic acid mix representing highly divergent viral families that mimic a metagenomic sample and a diverse microbial community of both RNA and DNA viruses with genome size ranging from 7.5 to 229.4 kb. The complete sequences of all six viruses could be identified correctly with an average read depth ranging from 1390x to 16756x (Table 2).

The identification and characterization of the sequences using a bioinformatics approach, including *de novo* assembly for taxonomic classification, genotyping and data visualization was completed in one hour. The universal chimeric primer (JS15N14) containing the tetradecamer (14-mer) can potentially anneal to 268435456 different positions randomly throughout the target sequences to prime the first strand cDNA synthesis in the reverse transcription step and generate large pools of cDNA. Priming with the longer random primer provided sequence reads across the entire length of a viral genome (Zika virus) resulting in a genome coverage of 99.04 % (Table 2).

The genomic sequences of six viruses in the ATCC virome nucleic acid mix were successfully identified using the One Codex platform (the coverage plot for each virus is summarized in Fig. S2 in the supplemental data). The bioinformatic analysis performed using the One Codex tool efficiently separated host-associated reads from viral sequences. The percentage of viral reads from NGS mapped from 31.7–99.04% of six viral pathogen genomes included in the ATCC standard. The very low coverage for reovirus was attributed to the omission of a denaturation step for double-stranded RNA (dsRNA) before reverse transcription initiated. This resulted in reduced efficiency of reverse transcription and subsequent steps. Including a simple denaturation step to separate the complementary strands should improve the sequencing of dsRNA. The LF-SISPA NGS method successfully identified each genome from a virus mix of 10^4 nucleic acid copies per microliter.

Although the virome nucleic acid mix contained the same number of genomic copies for each virus in the mix, a difference in genome coverage due to the distribution of sequence

reads was observed. The difference may be attributed to the PCR amplification efficiency associated with highly structured viral genomes or repetitive regions at the 5′-UTR (Moser et al., 2016; Logan et al., 2014). The presence of cellular nucleic acids can also significantly impact the amplification of viral RNA, which may also reduce the sequencing depth of coverage and sensitivity of this method, especially for viruses at low initial concentrations; however, this is frequently a characteristic of unknown samples. Depletion of cellular host DNA has been shown to increase the depth of viral reads (Oyola et al., 2013).

The sequencing data analysis for norovirus started with 955,660 reads obtained from the UNEXP disk nucleic acid sample. After removal of non-viral reads approximately 275,736 reads remained for downstream analysis. Of these, 41,782 reads were mapped to a norovirus GII.3[P40] reference genome (NC_029646.1; Length: 7,518bp) with a genome depth of coverage of 502.3x (Fig. 2). A near complete genome sequence of 7074 bp with no gaps was obtained with the exception of the 5′ and 3′ ends of the genome. Reference genome coverage was 94.1 % and the genome sequence obtained by sequencing spans from 316 bp to 7389 bp in the reference. The genome of the norovirus reference isolate consists of 7518 nucleotides (nt) and contains three ORFs: ORF1 (5100 nt), ORF2 (1608 nt) and ORF3 (780 nt). The coverage obtained for all three ORFs was 93.9 %, 100 % and 89.6 %, respectively.

Yang et al. (Yang et al., 2017) observed that norovirus genome coverage was affected by the quality of RNA extraction, yield and library preparation. Several studies (Fonager et al., 2017; Hasing et al., 2016; Brown et al., 2016) found higher genome coverage for norovirus samples with lower Ct values. In another study, 13 strains were sequenced with Illumina MiSeq NGS and near complete genome sequences were obtained, ranging from 7204 to 7395 nucleotides of 7518 (Chhabra et al., 2018). Our LF-SISPA method correctly identified and recovered a near complete norovirus genomic sequence from a human clinical sample preserved on a solid matrix stored at room temperature for 14 days.

This study has certain limitations in read length and depth for some viral genomes. The random reverse transcription step that relies on 14 random nucleotides at the 3′ end (JS15N14) likely introduces bias in regions with high GC content or secondary structure associated with viral genomes and in subsequent random amplification steps. Despite the uneven high-coverage peaks obtained across the genome (Supplementary Fig. 2), the LF-SISPA is broadly applicable to detecting RNA and DNA viruses. Further study is needed to establish the sensitivity of the assay. Caution should be exercised in the 2-step PCR approach for the prevention of possible carry-over contamination; nucleic acid extraction followed by PCR must be performed in separate rooms. The possibility of error introduction by proof-reading PCR polymerases and resulting artefacts can be reduced by decreasing PCR cycles. Incorporating appropriate positive controls at different stages can reinforce confidence in the data analysis. The cost per run can be greatly reduced by parallel sequencing of multiple samples containing unique sample-specific barcode identifier sequences.

The major advantages of the novel LF-SISPA NGS protocol are the short turnaround time (approximately 1 day, Fig. 1) and minimal preparation cost associated with sequencing library preparation. The method requires less than one hour of “hands on” time for

sample library preparation. The majority of the remaining time is required for the reverse transcription, Klenow and PCR reactions (<5 h) and the sequence run (16 h) as shown in Fig. 1. Other types of clinical samples, as well as environmental samples, containing different viruses should be evaluated for the detection and characterization of virus sequences.

4. Conclusions

The LF-SISPA NGS protocol generated genome sequences for six viral genomes in an ATCC virome nucleic acid mix and norovirus without an enzymatic library preparation kit. The proposed ligation-free whole genome sequencing is less expensive than commercial kits and involves only a PCR-based library preparation, eliminating DNA shearing, end repair, A-tailing and ligation steps. This technology appears to be effective for identification and characterization of a variety of viral genomes and may offer the benefit of prokaryote and eukaryote identification, as well as testing for unknown pathogens in human clinical samples.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors would like to acknowledge Denise B. Lynch (One Codex) and Koen Deforche (Genome Detective) for bioinformatics support. We acknowledge Shatavia Morrison, Jeff Mercante, and Michael Arrowood for critical reading of the manuscript. The findings and conclusions in this article are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention. The use of trade names and names of commercial sources is for identification only and does not imply endorsement by the CDC or the U.S. Department of Health and Human Services.

References

Brown JR, et al. , 2016. Norovirus whole-genome sequencing by SureSelect target enrichment: a robust and sensitive method. J. Clin. Microbiol. 54 (10), 2530–2537. [PubMed: 27487952]

Chhabra P, et al. , 2018. Near-complete genome sequences of several new norovirus genogroup II genotypes. Genome Announc. 6 (6).

Cromeans T, et al. , 2019. A new solid matrix for preservation of viral nucleic acid from clinical specimens at ambient temperature. J. Virol. Methods 274, 113732.

Fonager J, et al. , 2017. A universal primer-independent next-generation sequencing approach for investigations of norovirus outbreaks and novel variants. Sci. Rep. 7 (1), 813. [PubMed: 28400558]

Glenn TC, et al. , 2019a. Adapterama I: universal stubs and primers for 384 unique dual-indexed or 147,456 combinatorially-indexed Illumina libraries (iTru & iNext). PeerJ 7, e7755. [PubMed: 31616586]

Glenn TC, et al. , 2019b. Adapterama II: universal amplicon sequencing on Illumina platforms (TaggiMatrix). PeerJ 7, e7786. [PubMed: 31616589]

Hasing ME, et al. , 2016. A next generation sequencing-based method to study the intrahost genetic diversity of norovirus in patients with acute and chronic infection. BMC Genomics 17, 480. [PubMed: 27363999]

Logan G, et al. , 2014. A universal protocol to generate consensus level genome sequences for foot-and-mouth disease virus and other positive-sense polyadenylated RNA viruses using the Illumina MiSeq. BMC Genomics 15, 828. [PubMed: 25269623]

Minot SS, Krumm N, Greenfield NB, 2015. One codex: a sensitive and accurate data platform for genomic microbial identification. bioRxiv. 10.1101/027607p. bioRxiv 027607.

Moser LA, et al. , 2016. A universal next-generation sequencing protocol to generate noninfectious barcoded cDNA libraries from high-containment RNA viruses. mSystems 1 (3).

O'Leary NA, et al. , 2016. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. 44 (D1), D733–D745. [PubMed: 26553804]

Oyola SO, et al. , 2013. Efficient depletion of host DNA contamination in malaria clinical sequencing. J. Clin. Microbiol. 51 (3), 745–751. [PubMed: 23224084]

Reyes GR, Kim JP, 1991. Sequence-independent, single-primer amplification (SISPA) of complex DNA populations. Mol. Cell. Probes 5 (6), 473–481. [PubMed: 1664049]

Vilsker M, et al. , 2019. Genome Detective: an automated system for virus identification from high-throughput sequencing data. Bioinformatics 35 (5), 871–873. [PubMed: 30124794]

Wu F, et al. , 2020. A new coronavirus associated with human respiratory disease in China. Nature 579 (7798), 265–269. [PubMed: 32015508]

Yang Z, et al. , 2017. Application of next generation sequencing toward sensitive detection of enteric viruses isolated from celery samples as an example of produce. Int. J. Food Microbiol. 261, 73–81. [PubMed: 28992517]

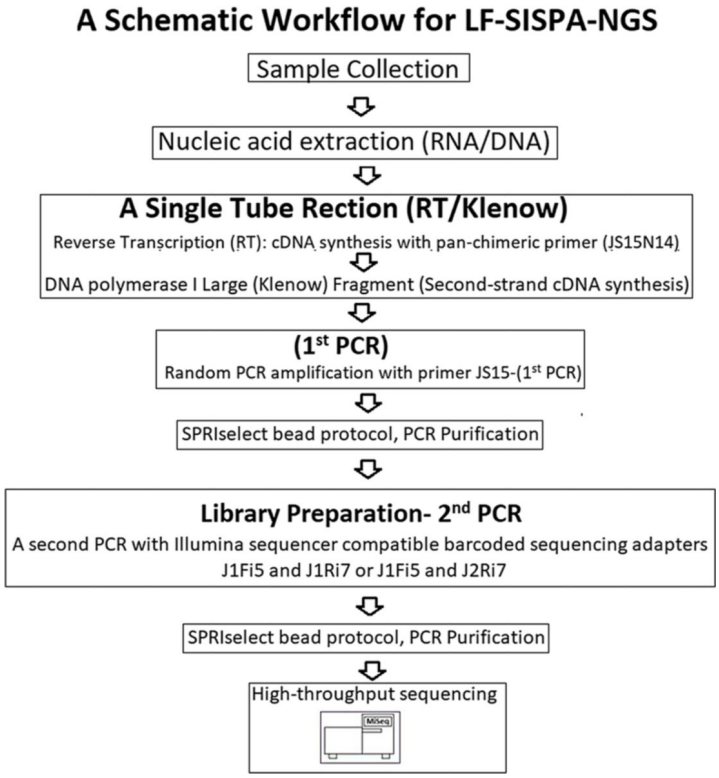

Fig. 1.

Stepwise overview of single-indexed library preparation workflow containing Illumina adapters and barcode without ligation for Illumina sequencing (MiSeq). The JS15 primer with 14 random nucleotides at the 3′ end (JS15N14) generates single-stranded and double-stranded DNA, respectively, by reverse transcription and a Klenow reaction. In the first-step PCR primer JS15 randomly amplifies and enriches the sample. In the second step-PCR using a pair of forward and reverse primers encoding Illumina adaptor (MiSeq) and JS15 primer sequences are incorporated into the amplicon from the first round PCR product. The reverse primer carries a unique barcoding index sequence embedded between the Illumina adaptor and JS15 primer sequences. The entire protocol is completed in approximately 1 day.

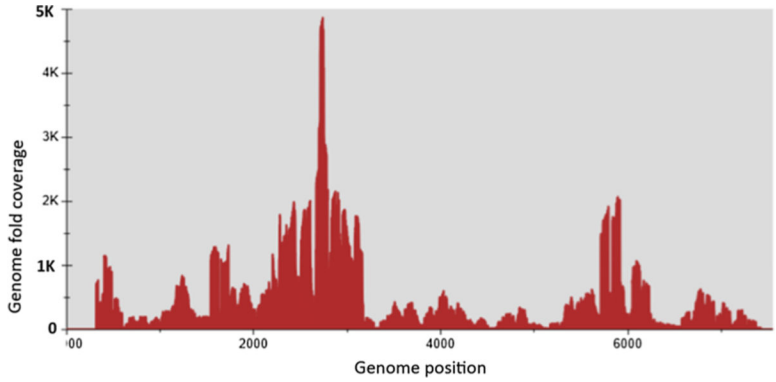

Fig. 2.

The genomic coverage map of norovirus GII based on the alignment of reads across the known reference sequence of norovirus GII (NC_029646). The Y and X axis represent the genome fold coverage and genome position respectively based on analysis by the Genome Detective Virus Tool.

Table 1

List of universal primers used in library preparation and sequencing primers for Illumina sequencing. (For interpretation of the references to colour in this Table legend, the reader is referred to the web version of this article)

Primer Name	Sequence (5'−3')
JS15N14	ACCGCCGCCCGTGAGNNNNNNNNNNNNNN
JS15	ACCGCCGCCCGTGAG
J1Fi5	AATGATACGGCGACCACCGAGATCTACACGCTTATGGTAATTGTACCGCCGCCCGTGAG
J1Ri7	CAAGCAGAAGACGGCATACGAGAT aaccgcgg AGTCAGCCAGCCACCGCCGCCCGTGAG
J2Ri7	CAAGCAGAAGACGGCATACGAGAT ggttataa AGTCAGCCAGCCACCGCCGCCCGTGAG
Read 1 sequencing primer	TATGGTAATTGTACCGCCGCCCGTGAG
Read 2 sequencing primer	AGTCAGCCAGCCACCGCCGCCCGTGAG

P7 barcoded primers J1Ri7 and J2Ri7 include unique 8-nt index sequences of aaccgcgg and ggttataa respectively and highlighted in red. JS15 sequence is highlighted in blue.

Table 2

Metagenomic profiling of the ATCC virome using the One Codex Viral Database.

	Genome length* (kb)	Number of Reads	Percent of classified reads	% Identity	% Coverage	Average depth x
Human herpesvirus 5 (ds DNA)	229.4	4690955	35.48	97.44	63.53	1652
Zika virus (positive-sense, ss RNA)	10.8	2022398	15.3	98.52	99.04	15589
Mammalian orthoreovirus 3 (segmented, ds RNA)	23.5	1548150	11.71	94.29	31.7	1390
Human mastadenovirus F (ds DNA)	34.2	1013610	7.67	99.35	76.08	2482
Influenza B virus (segmented, negative-sense, ss RNA)	14.2	649779	4.91	94.45	78.88	1551
Human orthopneumovirus (RSV) (negative-sense, ss RNA)	15.2	384724	2.91	99.39	79.19	16756

*Approximate genome length; human respiratory syncytial virus (RSV); single-stranded (ss); double-stranded (ds).