

## Article

# Library Preparation Based on Transposase Assisted RNA/DNA Hybrid Co-Tagmentation for Next-Generation Sequencing of Human Noroviruses

Zilei Zhang<sup>1,2,†</sup>, Danlei Liu<sup>1,2,†</sup>, Dapeng Wang<sup>1,\*</sup> and Qingping Wu<sup>1,2,\*</sup>

<sup>1</sup> Department of Food Science and Technology, School of Agriculture and Biology, Shanghai Jiao Tong University, Shanghai 200240, China; allyzhangzl@sjtu.edu.cn (Z.Z.); liudanlei@sjtu.edu.cn (D.L.)

<sup>2</sup> State Key Laboratory of Applied Microbiology Southern China, Guangdong Provincial Key Laboratory of Microbial Culture Collection and Application, Guangdong Open Laboratory of Applied Microbiology, Guangdong Institute of Microbiology, Guangzhou 510030, China

\* Correspondence: dapengwang@sjtu.edu.cn or norovirus@163.com (D.W.); wuqp@sjtu.edu.cn or wuqp203@163.com (Q.W.); Tel.: +86-21-34206918 (D.W.); +86-20-87688132 (Q.W.)

† These authors contribute equally to this work.

**Abstract:** Human noroviruses (HuNoVs) are one of the leading causes of foodborne illnesses globally. The viral genome is the most essential information for viral source tracing and viral transmission pattern monitoring. However, whole genome sequencing of HuNoVs is still challenging due to the sequence heterogeneity among different genotypes and low titer in samples. To address this need, in this study, the Transposase assisted RNA/DNA hybrid Co-tagmentation (TRACE-seq) method was established for next generation sequencing library preparation of HuNoVs. Our data demonstrated that almost the whole HuNoVs genome (>7 kb) could be obtained from all of the 11 clinical samples tested. Twelve genotypes including GI.3, GI.4, GI.5, GI.8, GII.2, GII.3, GII.4, GII.6, GII.12, GII.13, GII.14, and GII.21 were involved. Compared with the traditional method for viral metagenomics library preparation, optimized TRACE-seq greatly reduced the interference from the host's and bacterial RNAs. In addition, viral genome sequences can be assembled by using less raw data with sufficient depth along the whole genome. Therefore, for the high versatility and reliability, this method is promising for whole viral genome attainment. It is particularly applicable for the viruses with a low titer that are mixed with a complicated host background and are unable to be cultured in vitro, like the HuNoVs utilized in this study.

**Keywords:** human norovirus; next generation sequencing; library preparation; Tn5 transposase; TRACE-seq



**Citation:** Zhang, Z.; Liu, D.; Wang, D.; Wu, Q. Library Preparation Based on Transposase Assisted RNA/DNA Hybrid Co-Tagmentation for Next-Generation Sequencing of Human Noroviruses. *Viruses* **2021**, *13*, 65. <https://doi.org/10.3390/v13010065>

Academic Editor: Allan Brasier  
Received: 4 December 2020  
Accepted: 4 January 2021  
Published: 6 January 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

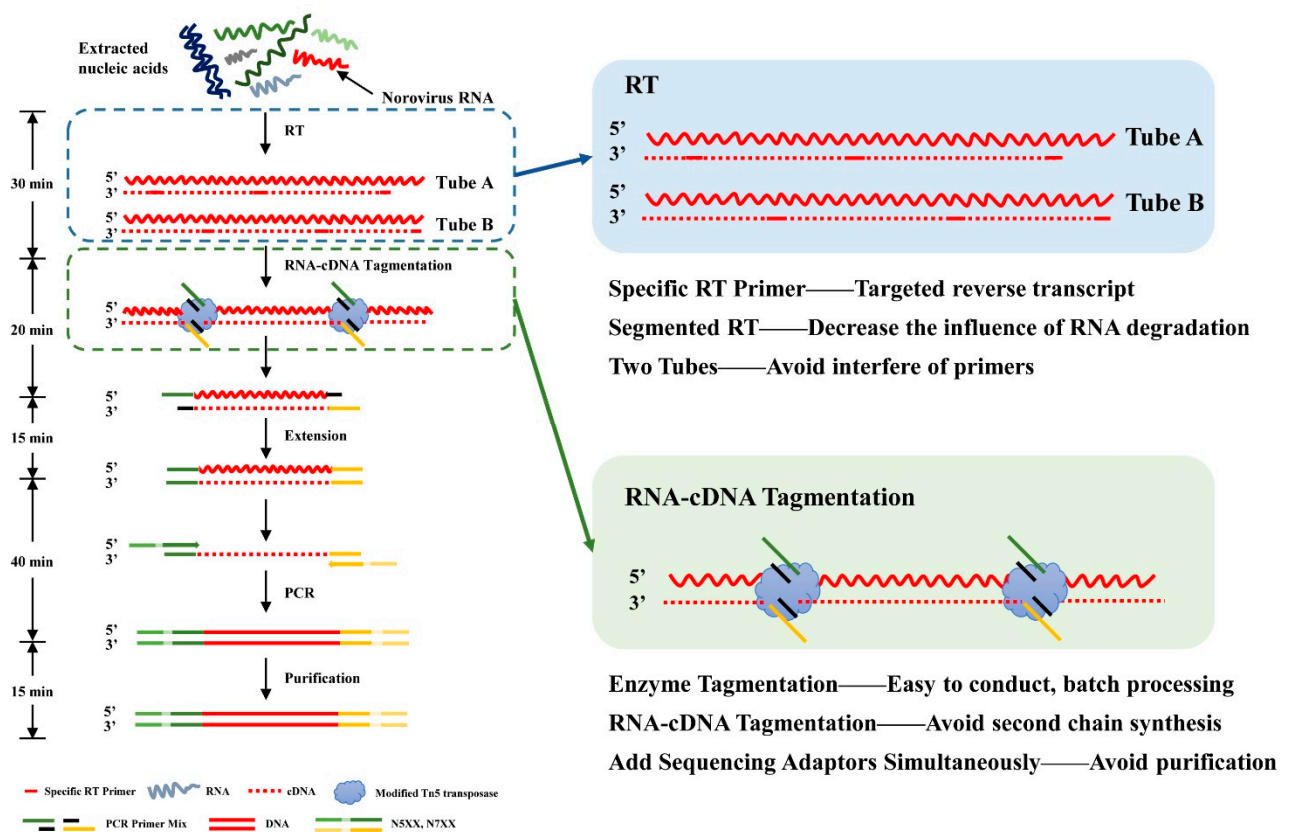
## 1. Introduction

Human noroviruses (HuNoVs) are the leading cause of non-bacterial epidemic gastroenteritis globally and cause an estimated 684 million illnesses of acute gastroenteritis worldwide, resulting in approximately \$60 billion in societal costs [1]. A norovirus (NoV) has a single-stranded, positive-sense genomic RNA of approximately 7.5 kb, that have a 5'-terminal virus protein, are genome-linked (VPg), and possess 3'-terminal poly (A) [2]. Based on phylogenetic analysis of major capsid protein (VP1) amino acid sequences, at least 10 norovirus genogroups (GI-GX) and 48 genotypes have been recognized [3]. The uncorrected pairwise distance ranges for strain, genotype, and genogroup were 0–14.1%, 14.3–43.8%, and 44.9–61.4%, respectively [4]. No routine cultivation method in vitro was available for different genotypes of HuNoVs isolation and replication [5]. To prevent and control the HuNoVs outbreaks, the key is to identify the viral pollution source and monitor viral transmission patterns based on the viral genome sequences [6].

Metagenomic next-generation sequencing (NGS) is an effective approach for broad-spectrum pathogen identification in clinical samples, as nearly all potential pathogens (viruses, bacteria, fungi, and parasites) can be detected by identified unique DNA and/or RNA sequences [7]. Viral genome sequencing is a branch of RNA sequencing (RNA-seq), and relies on second strand complementary DNA (cDNA) synthesis to generate initial material for library preparation. The common protocols contain a few key steps, including RNA extraction, poly-A selection or ribosomal RNA depletion, reverse transcription, second strand complementary DNA (cDNA) synthesis, adapter ligation, and PCR amplification [8–10]. Viral metagenomics can be realized in two ways: one is to enrich viral RNA by the poly A-capture method to remove as many hosts and bacterial rRNAs as possible (mRNA-seq), the other is to use ribosomal random primers for background sequences deletion, i.e., primers selected among random hexamers that match rRNA genes from the host or the matrix [11,12].

Traditional methods still have limited capacity to remove background data for viral genome sequencing, though host or background nucleic acids can account for up to 99% of the raw reads [12]. A crucial part of NGS is to increase the minimal required number of sequencing reads per sample (SRPS) in order to get a reliable in-depth analysis. Underestimation of SRPS results in a low yield of target sequences. The known minimum required number of SRPS was 0.008% for the Epstein-Barr virus (EBV) [13] and 0.3% for the Zika virus [14]. The existence of background data increases the requirements for sequencing data and the cost of each sample. Meanwhile, it also increases the difficulty of subsequent bioinformatics analysis, sample analysis times, and data storage costs. In addition, it is difficult to design primers that work for a broad range of strains with high sensitivity and efficiency, because of sequence heterogeneity among HuNoVs genomes [15]. Therefore, it is a big challenge to obtain whole genome sequences for HuNoVs compared with other viruses due to their sequence heterogeneity and complicated background data.

Tn5-mediated transposition of double-strand DNA has been widely utilized in various NGS tagmentation. DNA sequencing library preparation based on Tn5 transposase could be processed conveniently and has low requirements for sample inputs [16,17]. Tn5 transposase has been altered to be capable of direct tagmentation of RNA/DNA hybrids *in vitro* in 2020, named Transposase-assisted RNA/DNA hybrids Co-tagmentation ('TRACE-seq') [18,19]. TRACE-seq offers a greatly simplified protocol and produces results with higher reproducibility and GC uniformity compared with prevailing RNA-seq methods [18]. However, this method is still at the theoretical stage and has not been applied to clinical samples. In this study, HuNoVs clinical samples were used to evaluate the strategy for establishing a library preparation method for specific viruses based on the theory of TRACE-seq. Each sample was divided into tube A and B for reverse transcription using the TRACE-seq method, which optimized the method of library preparation for NGS of HuNoVs (Figure 1). The RNA/DNA hybrid was tagmented by the modified Tn5 transposase and partial sequencing adaptors were added to fragment ends simultaneously. By achieving specific recognition of HuNoVs viral genomes, the percentage of effective reads could be increased, and the difficulty of complete reverse transcription caused by high GC regions can be overcome.



**Figure 1.** Workflow of the next-generation sequencing (NGS) library preparation strategy based on Transposase-assisted RNA/DNA hybrids Co-tagmentation (termed TRACE-seq) Library construction processes are listed in the left panel. The main optimizations of this method are marked with two boxes and zoomed to show details with advantages listed in the right panel. The process consists of four components: RNA reverse transcription, RNA/cDNA hybrid tagmentation, PCR amplification, and purification. Specifically, RNA is divided into tube A and B for reverse transcription using primers designed for human noroviruses (HuNoVs). The RNA/DNA hybrid is then tagmented by the modified Tn5 transposase directly. High-fidelity DNA polymerase then amplifies the cDNA into a sequencing library after initial end extension. The purified product is an indexed library which is ready for sequencing.

## 2. Materials and Methods

### 2.1. RNA Extraction and RT-qPCR Identification of Clinical Samples

HuNoVs-positive clinical samples were kindly provided by Dr. Zhiyong Gao at the Beijing Center for Disease Control and Prevention (CDC), Dr. Ningbo Liao at the Zhejiang CDC, and Dr. Huiying Li at the Chinese CDC [20]. Viral RNAs were extracted from 140  $\mu$ L of 10% stool suspension with a HiPure Viral RNA Kit (Magen, China) according to the manufacturer's instructions without carrier RNA. RNA was eluted with 30  $\mu$ L RNase-free water. Duplicate one-step RT-qPCR assays for GI and GII HuNoVs were detected in a single tube using a Biorad CFX96 qPCR machine [20,21].

### 2.2. Primer Sets Design

The complete sequences of GI and GII HuNoVs before Oct 2019 ( $n = 1137$ ) were downloaded from the National Center for Biotechnology Information (NCBI, <https://www.ncbi.nlm.nih.gov/>, accessed on Oct 2019) and aligned using mafft to identify conserved regions for primers design separately [22]. Principles for reverse transcription primers selection including (1) high conservation of the fragment region, (2) length with 12–14 bp, and (3) interval with 1000–1500 bp. Primers designed in this study are listed in Table S1. Primers were synthesized by Sangon Biotech Co., Ltd. (Shanghai, China) and diluted to

200  $\mu$ M as stock-solutions. They were further configured into primer sets with a final concentration of 10  $\mu$ M according to the requirements of tube A or B.

### 2.3. TRACE-Seq Library Preparation and Sequencing

Libraries were prepared using TruePrep<sup>®</sup> RNA Library Prep Kit for Illumina (Vazyme Biotech Co., Ltd., Nanjing, China) with some modification. RNA (8.0  $\mu$ L) was heated at 65 °C for 5 min and quickly placed on ice for 2 min. It was then mixed with 2.0  $\mu$ L 5  $\times$  genomic DNA wiper Mix and incubated at 42 °C for 2 min to remove genomic DNA. The solution was then divided into tube A / tube B equally and added to the different primer sets described in Section 2.2. Then 1.0  $\mu$ L 10  $\times$  RT Mix, 0.5  $\mu$ L HiScript III Enzyme Mix, 1.0  $\mu$ L Primer sets, and 2.5  $\mu$ L Nuclease-free ddH<sub>2</sub>O were added into each tube for reverse transcription. Reaction conditions were as follow: 25 °C and 5 min, 37 °C and 30 min, and 85 °C and 5 s. Transcription products in tube A and tube B were combined, then we added 10.0  $\mu$ L tagment buffer and 2.0  $\mu$ L Tn5 VR1, then pipetted 20 times gently to mix well. The tagmentation conditions were 55 °C and 15 min and the solution was then held held at 10 °C. After the reaction was completed, 2.0  $\mu$ L TStop Solution was immediately added to the reaction solution. After using a vortex for mixing, the solution was then incubated at room temperature for 5 min. Then we added 50  $\mu$ L VAHTS HiFi Amplification Mix, 5.0  $\mu$ L PCR Primer Mix 3 for Illumina, and 1.0  $\mu$ L TSE into the solution (34  $\mu$ L) and incubated it at 55 °C for 5 min, 60 °C for 5 min, and 95 °C for 5 min for extension. TruePrep<sup>®</sup> Index Kit V2 for Illumina (Vazyme, Nanjing, China) N5XX (5.0  $\mu$ L) and N7XX (5.0  $\mu$ L) were added according to the number of samples and index matching strategy. The cycling conditions were as follows: initial denaturation at 95 °C for 3 min, 20 cycles of PCR with 98 °C for 20 s, 60 °C for 15 s, 72 °C for 30 s, and final extension at 72 °C for 5 min. The library was purified using 80  $\mu$ L (0.8 $\times$ ) VAHTS<sup>®</sup> DNA Clean Beads (Vazyme, Nanjing, China). Sequencing was performed using the Illumina X-10 platform to generate 2  $\times$  150 bp reads.

### 2.4. Bioinformatics

Reads from the Illumina X-10 were trimmed using fastp to remove sequences with low quality [23]. Clean reads were de novo assembled by SPAdes [24] or MEGAHIT [25]. Bowtie2 was used to extract the sequence of HuNoVs before assembling when the result of a de novo sequence was not ideal [26]. Contigs were identified using BLASTn based on a curated reference list that consists of all HuNoVs complete-genome sequences in GenBank that were uploaded before Oct 2019 ( $n = 1137$ ) [27]. Sequences were then validated and genotyped using a web-based genotyping tool (<http://www.rivm.nl/mpf/norovirus/typingtool>, accessed on Dec 2020, Version2.0) [28,29]. Mapped reads to contigs (CLC genome workbench) was used to identify coverage reads of HuNoVs per sample. After mapping reads to the foodborne pathogen database, sequence taxonomy was assigned by Kraken2 [30,31]. Data unification and heat map drawing were realized by using TBtools [32]. The raw sequence data reported in this paper have been deposited in the Genome Sequence Archive [33] in National Genomics Data Center [34], Beijing Institute of Genomics (China National Center for Bioinformatics), Chinese Academy of Sciences, under the accession number CRA003583 that is publicly accessible at <https://bigd.big.ac.cn/gsa>, accessed on 4 Dec 2020.

## 3. Results

### 3.1. Sequencing Results

Eleven HuNoVs positive clinical samples were used for method development (Figure 2). Raw data for each sample was 6–240 Mb. A HuNoVs genome above 7 kb could be obtained for all tested samples. The BLAST score is a number used to assess the biological relevance of a sequence. All the sequences were above 70.0, indicating that the assembled sequences are of high similarity with the norovirus sequences in the database. BLAST scores of sequences belonging to the rarely reported genotypes including GI.3[P10], GI.4[P4], GI.5[P5], GI.8[P8] were relatively low. This situation was mainly caused by the lack of sufficient

co-responding genotyped whole genome sequences in the database. Most of the sequences cover the entire three ORF regions, while the 5' or 3' end sequence was missing in a few samples, which has no effect on the genotyping or evolution analyzing of the main coding protein region. Genotypes involved in this study including GI.3, GI.4, GI.5, GI.8, GII.2, GII.3, GII.4, GII.6, GII.12, GII.13, GII.14, and GII.21, indicated that this method is suitable for a wide range of HuNoVs genotypes. Samples SJTU-1030 and SJTU-7565 were found to contain two different genotypes of HuNoVs strains, respectively. Reads coverage were listed in Figure 2. No special location was found with insufficient sequencing depth.

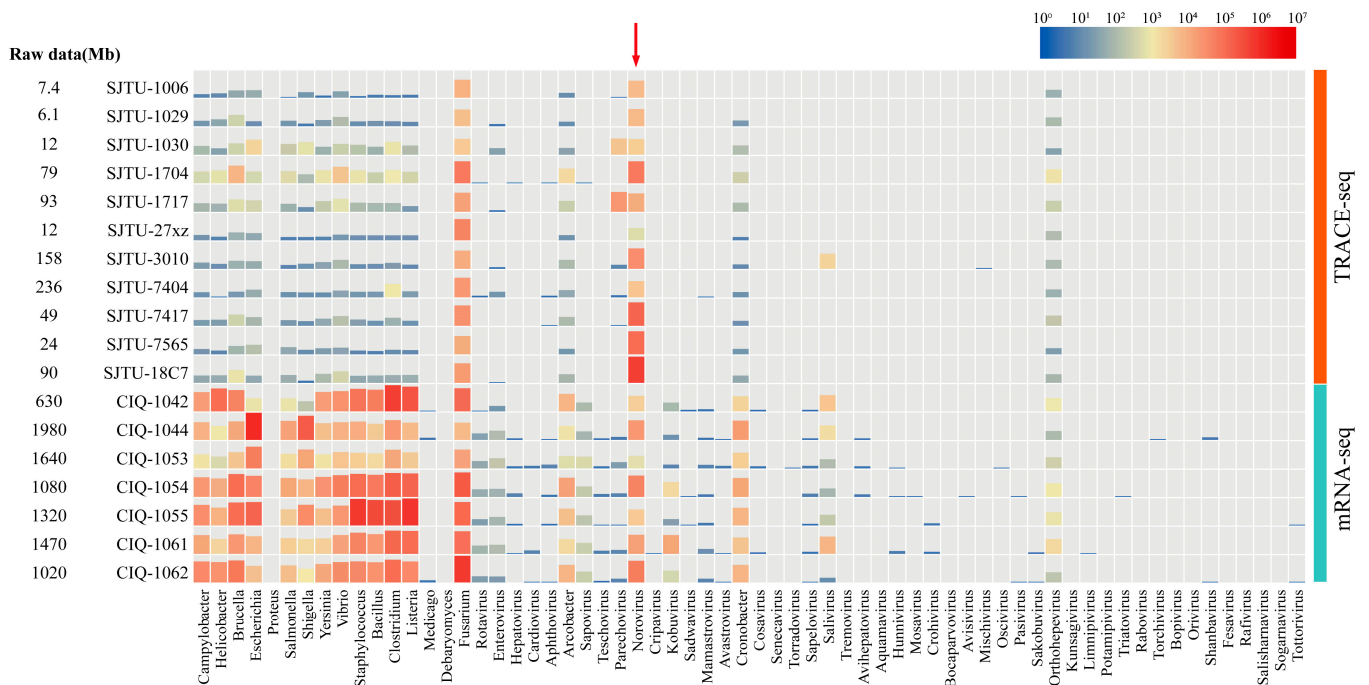
Sample	Raw data (Mb)	Length (bp)	BLAST score	Polymerase genotype	Capsid genotype	Genome coverage	Read coverage
SJTU-1006	7.4	7494	89.3556	GII.P12	GII.3		
SJTU-1029	6.1	7433	84.986725	GII.P31 (GII.Pe)	GII.4_ Sydney_2012		
SJTU-1030	12	7298	87.249245	GII.P7	GII.6		
		7305	89.63404	GII.P12	GII.3		
SJTU-1704	79	7502	85.12838	GII.P31 (GII.Pe)	GII.4_ Sydney_2012		
SJTU-1717	93	7441	89.6704	GII.P12	GII.3		
SJTU-27xz	12	7279	76.55752	GII.P16	GII.2		
SJTU-3010	158	7761	71.1978	GLP10 (GI.Pa)	GI.3		
SJTU-7404	236	7211	81.07364	GII.P21 (GII.Pb)	GII.21		
		6147	74.1541	GLP4	GI.4		
SJTU-7417	49	7566	85.35542	GII.P33 (GII.Pg)	GII.12		
SJTU-7565	24	7908	73.544975	GLP5	GI.5		
		7407	87.99723	GII.P7	GII.14		
SJTU-18C7	90	7504	81.50328	GII.P21 (GII.Pb)	GII.13		
		5034	81.11294	GII.P21 (GII.Pb)	GII.21		
		3124	73.56948	GLP8	GI.8		

**Figure 2.** Detailed sequencing results of HuNoVs positive samples using TRACE-seq. Results were analyzed by each sample. Sequences which could be genotyped by polymerase and capsid are listed. In the genome coverage column, sequences obtained in this study that matched the reference sequence of HuNoVs are shown in green. In the last column, altitude indicates the reads coverage in each position of obtained sequences.

Three different genotypes of HuNoVs were identified in sample SJTU18C7, while only one full-length viral genome was obtained. A full-length genome of GI.13 and two genome fragments with 3 kb of GI.8 and GI.21 were assembled. According to our previous study, one strain was always found to be dominant in multiple infections in one patient with more than 65% mapped reads [35]. This situation might lead to underestimation of SRPS for the minor strains in the multiple infection sample, thus whole genomes of all infected strains are unable to be obtained. Since this method is highly targeted and the amount of raw data for different samples varies greatly, it is recommended that the library should be quantified by qPCR library quantification to effectively evaluate the sequencing volume distribution of samples before loading the sample in the chip.

### 3.2. Effective Data Distribution

In the traditional method for viral metagenomic, mRNA in extracted RNA solution is captured for library preparation while the mRNA of host and bacteria are also involved, thus interfering with the analysis of viral genome (Figure 3). Meanwhile in the TRACE-seq method, the proportion of HuNoVs reads strongly increased through the reverse transcription by specific primers, while reads of food-borne pathogenic bacteria are reduced from  $10^5$  to  $10^2$  (Figure 3). However, other sequence interferences still exist in this method. *Fusarium* was found to be the main interfering pathogen in tested samples. This may be related to the undesired spurious priming between designed the primers and genome of *Fusarium*. The situation could be resolved by increasing the amount of sequencing data. It can also be seen from the distribution map that this method is highly pertinent and cannot achieve metagenome analysis of poly-A microorganisms in the same way as the traditional mRNA method, thus this technique lacks information on other pathogenic microorganisms in the sample.



**Figure 3.** Species distribution heatmap of effective reads of the TRACE-seq method and mRNA-seq method. Samples are listed by the left side, the main food-borne pathogenic microorganisms are listed in abscissa. The effective reads number of each pathogen are shown by the height of column and color from blue to red by the increase of the reads number. The relationship between the effective reads number and the related color is shown on the upper right site. Reads of norovirus are pointed to with a red arrow.

#### 4. Discussion

Sequencing of viral genomes has become a pivotal part of virological research [36]. Whole genome or long sequence is essential for either tracing the contaminating event or identifying emerging variants or viruses. Therefore, efforts have been paid towards exploring a method for the specific enrichment of HuNoVs sequences. Sureselect is a target enrichment method using a panel of custom-designed 120-mer RNA baits that are complementary to all publicly available HuNoVs sequences [37]. Metagenomic sequencing with spiked primer enrichment (MSSPE) enriches targeted viral RNA sequences by using spiked primers [38]. Another strategy to enhance the recovery of HuNoVs in complex matrices is the capture-based metagenomics (ViroCap), where nucleic acid libraries are enriched in viral sequences using probes [39]. Although many target enrichment arrays were designed to capture virus sequences, the target sequences still represent a small proportion of the total reads, thereby suggesting non-specific capturing [12]. Therefore, it is necessary to explore library preparation methods aiming for the target virus.

Our method successfully combined the advantages of transposase with the genome characteristics of HuNoVs, and verified that the TRACE-seq theory can be applied to the whole genome sequencing of RNA virus samples. Due to the high genetic diversity of the HuNoVs genus, it is difficult to design primers that work for a broad range of strains with high sensitivity and efficiency [15]. The TRACE-seq method does not need to design PCR primers of about 20 bp, as short-chain reverse transcription primers of only 12–14 bp work efficiently with high specificity. Therefore, it is very suitable for the whole genome sequencing of HuNoVs. In this study, segmented design of the reverse transcription primers was implemented to avoid the degradation of the target viral RNA and the inability of full-length reverse transcription due to the existence of a high GC region. In addition, in order to maximize the coverage of all regions of the genome and avoiding the genome sequencing errors caused by primer addition, two tubes A and B were operated simultaneously in the reverse transcription step with different primer sets.

This method is aimed towards whole genome analysis of a specific virus sequence for HuNoVs, and thus information of other microorganisms in the sample would be missed. It is suitable for samples whose pathogens have been confirmed through nucleic acid or immune assays that have been previously identified but are not suitable for unknown pathogen metagenomics. The method reduces the operating steps of the process, which minimizes the consumption and loss caused by the library preparation process. The magnetic bead purification step only needs to be done after the final PCR amplification, which avoids cDNA second-strand synthesis and ssDNA purification operations, and therefore minimized the loss of target fragments. The cost of the library for each sample is only approximately \$50. By excluding a large amount of metagenomic information, and only analyzing specific viruses, the probability of obtaining a whole genome of the target virus is improved, thereby potentially reducing sequencing costs substantially. In addition, since only a small amount of data is needed for subsequent sequencing, more samples could be sequenced on a chip by labeling samples with different indexes.

#### 5. Conclusions

The method developed in this study provides a promising path for viral genome acquisition, especially for viruses with a low titer, that are mixed with a complicated host background, and that cannot be cultured *in vitro*. Obtaining whole genome information is the best preservation and analysis method for these viruses. Optimized TRACE-seq processes provide high versatility and reliability, which is very suitable for obtaining information for a specific viral whole genome.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/1999-4915/13/1/65/s1>, Table S1: Primers designed in this study.

**Author Contributions:** Conceptualization, D.L.; methodology, D.L.; software, D.L.; validation, D.L.; formal analysis, D.L.; resources, D.W.; data curation, Z.Z.; writing—original draft preparation,

Z.Z.; writing—review and editing, D.L. and D.W.; visualization, Z.Z.; supervision, Q.W.; project administration, Q.W.; funding acquisition, D.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China (31772078).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data is contained within the article or supplementary material.

**Acknowledgments:** We are very grateful for the technical support provided by Liyan Tao and Jinming Wang from Nanjing Vazyme Biotechnology Co., Ltd. We would like to thank Zhiyong Gao at the Beijing Center for Disease Control and Prevention (CDC), Ningbo Liao at the Zhejiang CDC, and Huiying Li at the Chinese CDC for providing HuNoVs-positive clinical samples.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Cates, J.E.; Vinjé, J.; Parashar, U.; Hall, A.J. Recent advances in human norovirus research and implications for candidate vaccines. *Expert Rev. Vaccines* **2020**, *19*, 539–548. [[CrossRef](#)]
2. Vinjé, J.; Estes, M.K.; Esteves, P.; Green, K.Y.; Katayama, K.; Knowles, N.J.; L’Homme, Y.; Martella, V.; Vennema, H.; White, P.A. ICTV Virus Taxonomy Profile: Caliciviridae. *J. Gen. Virol.* **2019**, *100*, 1469–1470. [[CrossRef](#)] [[PubMed](#)]
3. Chhabra, P.; de Graaf, M.; Parra, G.I.; Chan, M.C.; Green, K.; Martella, V.; Wang, Q.; White, P.A.; Katayama, K.; Vennema, H.; et al. Updated classification of norovirus genogroups and genotypes. *J. Gen. Virol.* **2019**, *100*, 1393–1406, Corrigendum in **2020**, *101*, 893. [[CrossRef](#)] [[PubMed](#)]
4. Zheng, D.; Ando, T.; Fankhauser, R.L.; Beard, R.S.; Glass, R.I.; Monroe, S.S. Norovirus classification and proposed strain nomenclature. *Virology* **2006**, *346*, 312–323. [[CrossRef](#)]
5. Estes, M.K.; Ettayebi, K.; Tenge, V.R.; Murakami, K.; Karandikar, U.; Lin, S.; Ayyar, B.V.; Cortes-Penfield, N.W.; Haga, K.; Neill, F.H.; et al. Human norovirus cultivation in nontransformed stem cell-derived human intestinal enteroid cultures: Success and challenges. *Viruses* **2019**, *11*, 638. [[CrossRef](#)] [[PubMed](#)]
6. Gardy, J.L.; Loman, N.J. Towards a genomics-informed, real-time, global pathogen surveillance system. *Nat. Rev. Genet.* **2018**, *19*, 9–20. [[CrossRef](#)]
7. Wilson, M.R.; Sample, H.A.; Zorn, K.C.; Arevalo, S.; Chiu, C.Y. Clinical metagenomic sequencing for diagnosis of meningitis and encephalitis. *N. Engl. J. Med.* **2019**, *380*, 2327–2340. [[CrossRef](#)]
8. Cui, P.; Lin, Q.; Ding, F.; Xin, C.; Gong, W.; Zhang, L.; Geng, J.; Zhang, B.; Yu, X.; Yang, J.; et al. A comparison between ribo-minus RNA-sequencing and polyA-selected RNA-sequencing. *Genomics* **2010**, *96*, 259–265. [[CrossRef](#)]
9. Gubler, U.; Hoffman, B.J. A simple and very efficient method for generating cDNA libraries. *Gene* **1983**, *25*, 263–269. [[CrossRef](#)]
10. Mortazavi, A.; Williams, B.A.; McCue, K.; Schaeffer, L.; Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **2008**, *5*, 621–628. [[CrossRef](#)] [[PubMed](#)]
11. Endoh, D.; Mizutani, T.; Kirisawa, R.; Maki, Y.; Saito, H.; Kon, Y.; Morikawa, S.; Hayashi, M. Species-independent detection of RNA virus by representational difference analysis using non-ribosomal hexanucleotides for reverse transcription. *Nucleic Acids Res.* **2005**, *33*, e65. [[CrossRef](#)] [[PubMed](#)]
12. Strubbia, S.; Schaeffer, J.; Besnard, A.; Wacrenier, C.; Le Mennec, C.; Garry, P.; Desdouts, M.; Le Guyader, F.S. Metagenomic to evaluate norovirus genomic diversity in oysters: Impact on hexamer selection and targeted capture-based enrichment. *Int. J. Food Microbiol.* **2020**, *323*, 108588. [[CrossRef](#)] [[PubMed](#)]
13. Allen, U.D.; Hu, P.; Pereira, S.L.; Robinson, J.L.; Paton, T.A.; Beyene, J.; Khodai Booran, N.; Dipchand, A.; Hébert, D.; Ng, V. The genetic diversity of Epstein-Barr virus in the setting of transplantation relative to non-transplant settings: A feasibility study. *Pediatr. Transplant.* **2016**, *20*, 124–129. [[CrossRef](#)] [[PubMed](#)]
14. Calvet, G.; Aguiar, R.S.; Melo, A.S.; Sampaio, S.A.; De Filippis, I.; Fabri, A.; Araujo, E.S.; de Sequeira, P.C.; de Mendonça, M.C.; de Oliveira, L. Detection and sequencing of Zika virus from amniotic fluid of fetuses with microcephaly in Brazil: A case study. *Lancet Infect. Dis.* **2016**, *16*, 653–660. [[CrossRef](#)]
15. Desdouts, M.; de Graaf, M.; Strubbia, S.; Oude Munnink, B.B.; Kroneman, A.; Le Guyader, F.S.; Koopmans, M.P.G. Novel opportunities for NGS-based one health surveillance of foodborne viruses. *One Health Outlook* **2020**, *2*, 1–8. [[CrossRef](#)]
16. Adey, A.; Morrison, H.G.; Asan, X.; Kitzman, J.O.; Turner, E.H.; Stackhouse, B.; MacKenzie, A.P.; Caruccio, N.C.; Zhang, X.; et al. Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition. *Genome Biol.* **2010**, *11*, R119. [[CrossRef](#)]
17. Picelli, S.; Björklund, A.K.; Reinius, B.; Sagasser, S.; Winberg, G.; Sandberg, R. Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome Res.* **2014**, *24*, 2033–2040. [[CrossRef](#)]
18. Di, L.; Fu, Y.; Sun, Y.; Li, J.; Liu, L.; Yao, J.; Wang, G.; Wu, Y.; Lao, K.; Lee, R.W.; et al. RNA sequencing by direct tagmentation of RNA/DNA hybrids. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 2886–2893. [[CrossRef](#)]



19. Lu, B.; Dong, L.; Yi, D.; Zhang, M.; Zhu, C.; Li, X.; Yi, C. Transposase-assisted tagmentation of RNA/DNA hybrid duplexes. *Elife* **2020**, *9*, e54919. [[CrossRef](#)]
20. Liu, D.; Zhang, Z.; Wu, Q.; Tian, P.; Geng, H.; Xu, T.; Wang, D. Redesigned duplex RT-qPCR for the detection of GI and GII human noroviruses. *Engineering* **2020**, *6*, 442–448. [[CrossRef](#)]
21. Zhang, H.; Liu, D.; Zhang, Z.; Hewitt, J.; Li, X.; Hou, P.; Wang, D.; Wu, Q. Surveillance of human norovirus in oysters collected from production area in Shandong Province, China during 2017–2018. *Food Control* **2021**, *121*, 107649. [[CrossRef](#)]
22. Katoh, K.; Kuma, K.; Toh, H.; Miyata, T. MAFFT version 5: Improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* **2005**, *33*, 511–518. [[CrossRef](#)] [[PubMed](#)]
23. Chen, S.; Zhou, Y.; Chen, Y.; Gu, J. Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinform. Oxf. Engl.* **2018**, *34*, i884–i890. [[CrossRef](#)] [[PubMed](#)]
24. Nurk, S.; Bankevich, A.; Antipov, D.; Gurevich, A.A.; Korobeynikov, A.; Lapidus, A.; Prjibelski, A.D.; Pyshkin, A.; Sirotkin, A.; Sirotkin, Y.; et al. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. *J. Comput. Biol. A J. Comput. Mol. Cell Biol.* **2013**, *20*, 714–737. [[CrossRef](#)]
25. Li, D.; Liu, C.; Luo, R.; Sadakane, K.; Lam, T. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinform. Oxf. Engl.* **2015**, *31*, 1674–1676. [[CrossRef](#)]
26. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [[CrossRef](#)]
27. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. BLAST+: Architecture and applications. *BMC Bioinform.* **2009**, *10*, 421. [[CrossRef](#)]
28. Kroneman, A.; Vennema, H.; Deforche, K.; Avoort, H.V.D.; Peñaranda, S.; Oberste, M.S.; Vinjé, J.; Koopmans, M. An automated genotyping tool for enteroviruses and noroviruses. *J. Clin. Virol.* **2011**, *51*, 121–125. [[CrossRef](#)]
29. Vilsker, M.; Moosa, Y.; Nooij, S.; Fonseca, V.; Ghysens, Y.; Dumon, K.; Pauwels, R.; Alcántara, L.C.; Vanden Eynden, E.; Vandamme, A.; et al. Genome Detective: An automated system for virus identification from high-throughput sequencing data. *Bioinformatics* **2019**, *35*, 871–873. [[CrossRef](#)]
30. Wood, D.E.; Lu, J.; Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biol.* **2019**, *20*, 257. [[CrossRef](#)] [[PubMed](#)]
31. Labbé, R.G.; García, S. Foodborne viruses. In *Guide to Foodborne Pathogens*, 2nd ed.; Wiley Blackwell: Hoboken, NJ, USA, 2013; pp. 352–376.
32. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol. Plant* **2020**, *13*, 1194–1202. [[CrossRef](#)] [[PubMed](#)]
33. Wang, Y.; Song, F.; Zhu, J.; Zhang, S.; Yang, Y.; Chen, T.; Tang, B. GSA: Genome Sequence Archive. *Genom. Proteom. Bioinform.* **2017**, *15*, 14–18. [[CrossRef](#)] [[PubMed](#)]
34. Zhang, Z.; Ma, L.; Abbasi, A.A.; Raza, R.Z.; Gao, F. Database resources of the National Genomics Data Center in 2020. *Nucleic Acids Res.* **2020**, *48*, D24–D33.
35. Liu, D.; Zhang, Z.; Li, S.; Wu, Q.; Tian, P.; Zhang, Z.; Wang, D. Fingerprinting of human noroviruses co-infections in a possible foodborne outbreak by metagenomics. *Int. J. Food Microbiol.* **2020**, *333*, 108787. [[CrossRef](#)]
36. Radford, A.D.; Chapman, D.; Dixon, L.; Chantrey, J.; Darby, A.C.; Hall, N. Application of next-generation sequencing technologies in virology. *J. Gen. Virol.* **2012**, *93*, 1853. [[CrossRef](#)]
37. Brown, J.R.; Roy, S.; Ruis, C.; Romero, E.Y.; Shah, D.; Williams, R.; Breuer, J. Norovirus whole-genome sequencing by SureSelect target enrichment: A robust and sensitive method. *J. Clin. Microbiol.* **2016**, *54*, 2530–2537. [[CrossRef](#)]
38. Deng, X.; Achari, A.; Federman, S.; Yu, G.; Somasekar, S.; Bártolo, I.; Yagi, S.; Mbala-Kingebeni, P.; Kapetshi, J.; Ahuka-Mundeke, S.; et al. Metagenomic sequencing with spiked primer enrichment for viral diagnostics and genomic surveillance. *Nat. Microbiol.* **2020**, *5*, 443–454. [[CrossRef](#)]
39. Briese, T.; Kapoor, A.; Mishra, N.; Jain, K.; Kumar, A.; Jabado, O.J.; Lipkin, W.I. Virome capture sequencing enables sensitive viral diagnosis and comprehensive virome analysis. *mBio* **2015**, *6*, e1415–e1491. [[CrossRef](#)]