NOTE *Virology*

Use of S1 nuclease in deep sequencing for detection of double-stranded RNA viruses

Saya SHIMADA[1)], Makoto NAGAI[1,2)]*, Hiromitsu MORIYAMA[3)], Toshiyuki FUKUHARA [3)], Satoshi KOYAMA[4)], Tsutomu OMATSU[2)], Tetsuya FURUYA[2,5)], Junsuke SHIRAI[1,2)] and Tetsuya MIZUTANI[2)]

[1)]*Laboratory of Epizootiology, Cooperative Department of Veterinary Medicine Faculty and Agriculture, Tokyo University of Agriculture and Technology, Saiwai, Fuchu, Tokyo 183–8509, Japan*
[2)]*Research and Education Center for Prevention of Global Infectious Diseases of Animal, Tokyo University of Agriculture and Technology, Saiwai, Fuchu, Tokyo 183–8509, Japan*
[3)]*Laboratory of Molecular and Cellular Biology, Department of Applied Biological Science Faculty and Agriculture, Tokyo University of Agriculture and Technology, Saiwai, Fuchu, Tokyo 183–8509, Japan*
[4)]*Laboratory of Ethology, Cooperative Department of Veterinary Medicine Faculty and Agriculture, Tokyo University of Agriculture and Technology, Saiwai, Fuchu, Tokyo 183–8509, Japan*
[5)]*Laboratory of Microbiology, Department of Veterinary Medicine Faculty and Agriculture, Tokyo University of Agriculture and Technology, Saiwai, Fuchu, Tokyo 183–8509, Japan*

(Received 20 November 2014/Accepted 25 March 2015/Published online in J-STAGE 6 April 2015)

ABSTRACT. Metagenomic approach using next-generation DNA sequencing has facilitated the detection of many pathogenic viruses from fecal samples. However, in many cases, majority of the detected sequences originate from the host genome and bacterial flora in the gut. Here, to improve efficiency of the detection of double-stranded (ds) RNA viruses from samples, we evaluated the applicability of S1 nuclease on deep sequencing. Treating total RNA with S1 nuclease resulted in 1.5–28.4- and 10.1–208.9-fold increases in sequence reads of group A rotavirus in fecal and viral culture samples, respectively. Moreover, increasing coverage of mapping to reference sequences allowed for sufficient genotyping using analytical software. These results suggest that library construction using S1 nuclease is useful for deep sequencing in the detection of dsRNA viruses.

KEY WORDS: deep sequencing, double-stranded RNA viruses, group a rotavirus, next-generation sequencer, S1 nuclease

doi: 10.1292/jvms.14-0607; *J. Vet. Med. Sci.* 77(9): 1163–1166, 2015

Non-Sanger-based next-generation sequencing (NGS) is a comprehensive technology for the characterization of nucleic acid sequences, without prior knowledge of genetic background [12, 14]. NGS has facilitated the discovery of many pathogenic viruses from a broad range of samples [7, 10]. However, numerous sequence reads originating from bacterial species and hosts are recovered when using conventional metagenomic approaches. Thus, to detect viral genome sequence reads from minor virus populations efficiently, it is important to exclude contaminating nucleic acids, including those of bacteria and hosts. S1 nuclease was first identified in *Aspergillus oryzae* as an endonuclease specific for single-stranded (ss) polynucleotides [1, 2]. S1 nuclease hydrolyzes ssRNA, ssDNA and ss regions of double-stranded (ds) polynucleotide, but it does not degrade dsRNA [3, 13, 15]. S1 nuclease was used in the present study, because it is inexpensive and easily inactivated by ethylenediaminetetraacetic acid.

Group A rotaviruses (RVAs) are major etiological agents of acute gastroenteritis, particularly in neonatal animals; RVAs can cause economic losses to the livestock industry [4, 11]. RVAs possess a dsRNA genome composed of 11 genome segments [5]. Whole genome analyses of RVA based on their nucleotide sequences are important for comprehensive understanding of the evolution of RVAs, which involves genetic re-assortment events and interspecies transmission [6, 8]. In a previous study, applying DNase I treatment to total RNA after RNA extraction markedly reduced the number of extra sequence reads in preliminary deep sequence analysis (data not shown). However, a considerable number of sequence reads from the RNA of bacterial species and hosts were detected in fecal and viral culture samples. In this study, to improve the efficiency of detection of dsRNA viruses from fecal and viral culture samples, we evaluated the applicability of S1 nuclease-treatment in deep sequencing.

Four fecal samples [2 from calves (calf no. 1: from 6 days old calf with diarrhea collected in 2013 in Japan and calf no. 2: from 26 days old calf with diarrhea collected in 2013 in Japan) and 2 from piglets (pig no.1: from healthy piglet collected in 2014 in Japan and pig no. 2: from piglet with diarrhea in 2014 in Japan)] and 3 viral culture samples were evaluated using the rapid antigen detection kit, Dipstick "Eiken" Rota immunochromatographic assay (Eiken Chemical Co., Ltd., Tokyo, Japan), for the presence of RVA. The results revealed that the samples from calf nos. 1 and 2 were weakly positive and positive, respectively, for RVA, while those from pig nos. 1 and 2 were strongly positive for RVA. The fecal samples were diluted (1:9 [v/v]) in sterile phosphate-buffered saline, centrifuged at 8,000 × *g* for 10 min

*CORRESPONDENCE TO: NAGAI, M., Research and Education Center for Prevention of Global Infectious Diseases of Animal, Tokyo University of Agriculture and Technology, Saiwai, Fuchu, Tokyo 183–8509, Japan. e-mail: m-nagai@cc.tuat.ac.jp

Table 1. Comparison of RVA sequence read counts and percentages of RVA sequence reads between non-treated and S1nuclease treated samples

		Fecal samples				Viral culture samples		
		calf No.1	calf No.2	pig No.1	pig No.2	bovine RVA (G6P[5])	swine RVA (G9P[23])	swine RVA (G4P[23])
RVA sequence read counts	non-treated	2,190	24,737	30,768	366,339	1,901	975	34,499
	S1 nuclease treated	62,260	140,402	401,860	539,964	398,850	181,870	347,478
percentage of RVA sequence reads (%) [a)]	non-treated	0.9	9.4	5.4	75.0	0.5	0.2	7.9
	S1 nuclease treated	21.1	47.3	98.3	98.6	93.0	33.1	99.4

a) percentage of RVA sequence reads (%): RVA sequence reads/total reads.

at 4°C and stored at −80°C until further use. A bovine RVA (RVA/Cow-tc/JPN/Hori-No.14/1997/G6P[5]) and two swine RVA (RVA/Pig-tc/JPN/BU9/2014/G9P[23] and RVA/Pig-tc/JPN/Miyamoto/1997/G4P[23]) were inoculated into MA-104 cells. Supernatants were collected at post-inoculation day 5 and stored at −80°C, without centrifugation, until further use. Total RNA was extracted from all samples using TRIzol® LS Reagent (Life Technologies, Carlsbad, CA, U.S.A.), following which the RNA samples were treated with DNase I (0.5 U/μl; TaKaRa Bio Inc., Otsu, Japan). Each RNA sample was divided into 2 equal volumes, of which one was left untreated and the other was treated with S1 nuclease (27 U/μl; TaKaRa Bio Inc.) at 23°C for 15 min. After purification by ethanol precipitation, both the non-treated and S1 nuclease-treated samples were normalized to 50 *n*g/reaction using a Qubit® 2.0 Fluorometer (Invitrogen, Carlsbad, CA, U.S.A.). A cDNA library was constructed using the NEBNext® Ultra™ RNA Library Prep Kit to facilitate sequencing with Illumina version 2.0 (New England Biolabs, Ipswich, MA, U.S.A.), according to the manufacturer's guidelines. Briefly, the RNA samples were fragmented and used to synthesize ds-cDNA, which was then purified using Agencourt® AMPure® XP Beads (Beckman Coulter, Pasadena, CA, U.S.A.). To select fragments of approximately 200-bp in length after A-Tailing and adaptor ligation, 2 clean-up steps were performed using the beads. The library was then enriched by polymerase chain reaction (PCR), quantified using a Qubit® 2.0 Fluorometer (Invitrogen) and sequenced using a MiSeq bench-top sequencer (Illumina, San Diego, CA, U.S.A.) with 51 single-end reads. The sequence data were analyzed using the MiSeq Reporter program (Illumina) to generate the reads in FASTQ format. Trimmed reads were assembled into contigs by *de novo* assembly with default parameters (automatic word and bubble size), using the CLC Genomics Workbench 6.0 (CLC; CLC bio, Aarhus, Denmark). Using the assembled contigs as references, consensus sequences for all the RVA segments were obtained. These consensus sequences were in turn used as references by the read mapper tool in CLC. The RVA sequence reads from the non-treated and S1 nuclease-treated samples were compared using mapping results of the read mapper tool with default mapping parameters (mismatch cost, 2; insertion cost, 3; deletion cost, 3; length function, 0.5; and similarity function, 0.8).

The results of the comparison of total RVA sequence read counts and percentage of RVA sequence reads (RVA sequence reads/total reads) between the non-treated and S1 nuclease-treated samples are listed in Table 1. The number of RVA sequence reads in all the genome segments was higher in the S1 nuclease-treated samples than in the non-treated samples. The percentage of RVA sequence reads was also higher in the S1 nuclease-treated samples than in the non-treated samples, particularly in the viral culture samples. Moreover, S1-treatment did not show any strong bias in relative numbers homologous to viral segments (Table 2).

Furthermore, RVA genotyping was performed by mapping the reads of the samples to the RVA reference sequences by using CLC[9]. Figure 1 illustrates the mapping results of the bovine fecal (calf no. 1) and swine fecal (pig no. 1) samples, which represent the VP4 bovine (P[1], P[5], P[11] and P[14]; A) and VP7 swine genotypes (G2, G4, G5 and G9; B), respectively. Mapping sequence reads from the non-treated and S1 nuclease-treated samples to the reference sequences of representative strains revealed that the coverage of sequence reads of the S1 nuclease treated-samples against the reference sequences (calf no.1: 98.1%, swine no.1: 97.7%) was higher than that of the non-treated samples (calf no.1: 37.0%, swine no.1: 39.1%).

In the present study, treating total RNA with S1 nuclease prior to NGS reduced the number of extra sequence reads and increased the number of reads of RVA obtained for the bovine fecal, swine fecal and viral culture samples. Furthermore, the increased number of sequence reads of RVA facilitated genotyping with mapping to reference sequences on the genome analysis software. These results suggest that S1 nuclease-treatment during preparation of viral RNA for NGS is more useful than DNase I treatment alone, for the detection of RVA and dsRNA viruses. S1 nuclease-treatment facilitates the detection of not only viruses belonging to the *Reoviridae* and *Picobirnaviridae* families, but also non-identified novel viruses.

ACKNOWLEDGMENT. This study was supported by the Research Project for Improving Food Safety and Animal Health of the Ministry of Agriculture, Forestry and Fisheries of Japan.

REFERENCES

1. Ando, T. 1966. A nuclease specific for heat-denatured DNA in isolated from a product of *Aspergillus oryzae*. *Biochim. Biophys. Acta* **114**: 158–168. [Medline] [CrossRef]

Table 2. Relative number of RVA sequence reads homologous to each segment

	Calf No.1				Calf No.2				Pig No.1				Pig No.2			
	non-treated		S1 nuclease treated		non-treated		S1 nuclease treated		non-treated		S1 nuclease treated		non-treated		S1 nuclease treated	
	read counts	%[a)]	read counts	%	read counts	%	read counts	%	read counts	%	read counts	%	read counts	%	read counts	%
VP1	402	18.4	12,357	19.8	5,041	20.4	29,769	21.2	7,662	24.9	92,178	22.9	73,770	20.1	108,149	20.0
VP2	367	16.8	10,527	16.9	4,076	16.5	26,013	18.5	5,480	17.8	71,964	17.9	67,888	18.5	106,540	19.7
VP3	424	19.4	12,568	20.2	3,791	15.3	23,040	16.4	4,333	14.1	58,400	14.5	60,925	16.6	82,565	15.3
VP4	76	3.5	3,226	5.2	1,015	4.1	7,899	5.6	3,325	10.8	45,288	11.3	29,090	7.9	57,786	10.7
VP6	108	4.9	2,832	4.5	1,459	5.9	5,720	4.1	778	2.5	15,830	3.9	13,229	3.6	22,394	4.1
VP7	92	4.2	1,742	2.8	1,674	6.8	6,250	4.5	1,263	4.1	20,731	5.2	14,416	3.9	21,563	4.0
NSP1	324	14.8	5,453	8.8	3,001	12.1	13,421	9.6	2,126	6.9	26,172	6.5	27,043	7.4	35,353	6.5
NSP2	149	6.8	5,140	8.3	1,227	5.0	10,132	7.2	1,986	6.5	23,917	6.0	28,073	7.7	36,870	6.8
NSP3	130	5.9	6,153	9.9	2,327	9.4	11,510	8.2	2,791	9.1	34,164	8.5	33,953	9.3	44,243	8.2
NSP4	85	3.9	1,198	1.9	742	3.0	3,557	2.5	814	2.6	10,751	2.7	9,188	2.5	12,441	2.3
NSP5	33	1.5	1,065	1.7	386	1.6	3,090	2.2	201	0.7	2,465	0.6	8,764	2.4	12,060	2.2
total	2,190	100	62,261	100	24,739	100	140,401	100	30,759	100	401,860	100	366,339	100	539,964	100

	RVA/Cow-tc/JPN/Hori-No.14/1997/G6P[5]				RVA/Pig-tc/JPN/BU9/2014/G9P[23]				RVA/Pig-tc/JPN/Miyamoto/1997/G4P[23]			
	non-treated		S1 nuclease treated		non-treated		S1 nuclease treated		non-treated		S1 nuclease treated	
	read counts	%	read counts	%	read counts	%	read counts	%	read counts	%	read counts	%
VP1	470	24.7	97,019	24.3	192	19.7	37,413	20.6	7,037	20.4	73,527	21.2
VP2	415	21.8	79,098	19.8	154	15.8	34,893	19.2	6,418	18.6	63,481	18.3
VP3	364	19.1	88,677	22.2	118	12.1	27,585	15.2	4,958	14.4	56,224	16.2
VP4	58	3.1	17,853	4.5	134	13.7	21,674	11.9	4,475	13.0	40,859	11.8
VP6	32	1.7	8,742	2.2	32	3.3	6,263	3.4	715	2.1	8,704	2.5
VP7	43	2.3	8,810	2.2	19	1.9	4,174	2.3	1,779	5.2	18,166	5.2
NSP1	108	5.7	27,380	6.9	85	8.7	16,127	8.9	2,112	6.1	21,856	6.3
NSP2	182	9.6	30,286	7.6	67	6.9	10,641	5.9	2,275	6.6	23,188	6.7
NSP3	155	8.2	28,768	7.2	114	11.7	13,912	7.6	3,625	10.5	31,816	9.2
NSP4	55	2.9	8,621	2.2	37	3.8	5,451	3.0	925	2.7	8,117	2.3
NSP5	19	1.0	3,596	0.9	23	2.4	3,737	2.1	180	0.5	1,540	0.4
total	1,901	100	398,850	100	975	100	181,870	100	34,499	100	347,478	100

a) RVA sequence reads of each viral segment/total RVA sequence reads of the sample.

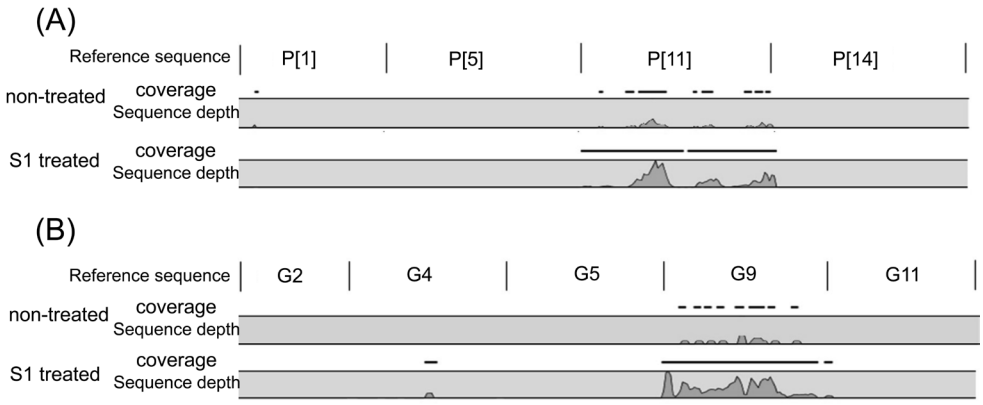

Fig. 1. Mapping of RVA sequence reads of non-treated and S1 treated samples of calf No.1 and swine No.1 against bovine reference sequences of (A) VP4 and (B) swine reference of VP7. Positional sequence coverage and sequencing depth of VP4 (P[1], P[5], P[11] and P[14]) and VP7 (G2, G4, G5, G9 and G11) of calf No.1 and swine No.1, based on reference assembly to P[1]: NCDV Lincoln (AB119636), P[5]: WC3 (AY05071), P[11]: B223 (D13394), P[14]: RVA/Cow-wt/JPN/Tottori-SG/2013/G15P[14] (AB853893), G2: RVA/Vaccine/USA/RotaTeq-SC2-9/G2P7 [5] (GU565068), G4: RVA/Vaccine/USA/RotaTeq-BrB-9/1996/G4P7[5] (GU565090), G5: RVA/Pig-tc/ESP/OSU-C5111/2010/G5P[7] (KJ450849), G9: JP3-6 (AB176678) and G11: HLJhg7 (JX498964).

2. Berk, A. J. and Sharp, P. A. 1978. Structure of the adenovirus 2 early mRNAs. *Cell* **14**: 695–711. [Medline] [CrossRef]
3. Desai, N. A. and Shankar, V. 2003. Single-strand-specific nucleases. *FEMS Microbiol. Rev.* **26**: 457–491. [Medline] [CrossRef]
4. Dhama, K., Chauhan, R. S., Mahendran, M. and Malik, S. V. 2009. Rotavirus diarrhea in bovines and other domestic animals. *Vet. Res. Commun.* **33**: 1–23. [Medline] [CrossRef]
5. Estes, M. K. and Greenberg, H. B. 2013. Rotaviruses. pp. 1347–1401. *In* Fields Virology, 6th ed, (Knipe, D. M., Howley, P. M., Cohen, J. I., Griffin, D. E., Lamb, R. A., Martin, M. A., Racaniello, V. R. and Roizman, B. eds.), Lippincott, Philadelphia.
6. Ghosh, S. and Kobayashi, N. 2011. Whole-genomic analysis of rotavirus strains: current status and future prospects. *Future Microbiol.* **6**: 1049–1065. [Medline] [CrossRef]
7. Marston, D. A., McElhinney, L. M., Ellis, R. J., Horton, D. L., Wise, E. L., Leech, S. L., David, D., de Lamballerie, X. and Fooks, A. R. 2013. Next generation sequencing of viral RNA genomes. *BMC Genomics* **14**: 444. [Medline] [CrossRef]
8. Matthijnssens, J., Ciarlet, M., Heiman, E., Arijs, I., Delbeke, T., McDonald, S. M., Palombo, E. A., Iturriza-Gómara, M., Maes, P., Patton, J. T., Rahman, M. and Van Ranst, M. 2008. Full genome-based classification of rotaviruses reveals a common origin between human Wa-Like and porcine rotavirus strains and human DS-1-like and bovine rotavirus strains. *J. Virol.* **82**: 3204–3219. [Medline] [CrossRef]
9. Minami-Fukuda, F., Nagai, M., Takai, H., Murakami, T., Ozawa, T., Tsuchiaka, S., Okazaki, S., Katayama, Y., Oba, M., Nishiura, N., Sassa, Y., Omatsu, T., Furuya, T., Koyama, S., Shirai, J., Tsunemitsu, H., Fujii, Y., Katayama, K. and Mizutani, T. 2013. Detection of bovine group a rotavirus using rapid antigen detection kits, rt-PCR and next-generation DNA sequencing. *J. Vet. Med. Sci.* **75**: 1651–1655. [Medline] [CrossRef]
10. Mokili, J. L., Rohwer, F. and Dutilh, B. E. 2012. Metagenomics and future perspectives in virus discovery. *Curr. Opin. Virol.* **2**: 63–77. [Medline] [CrossRef]
11. Papp, H., László, B., Jakab, F., Ganesh, B., De Grazia, S., Matthijnssens, J., Ciarlet, M., Martella, V. and Bányai, K. 2013. Review of group A rotavirus strains reported in swine and cattle. *Vet. Microbiol.* **165**: 190–199. [Medline] [CrossRef]
12. Radford, A. D., Chapman, D., Dixon, L., Chantrey, J., Darby, A. C. and Hall, N. 2012. Application of next-generation sequencing technologies in virology. *J. Gen. Virol.* **93**: 1853–1868. [Medline] [CrossRef]
13. Roberts, T. M., Bikel, I., Yocum, R. R., Livingston, D. M. and Ptashne, M. 1979. Synthesis of simian virus 40 t antigen in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **76**: 5596–5600. [Medline] [CrossRef]
14. Schuster, S. C. 2008. Next-generation sequencing transforms today's biology. *Nat. Methods* **5**: 16–18. [Medline] [CrossRef]
15. Vogt, V. M. 1973. Purification and further properties of single-strand-specific nuclease from *Aspergillus oryzae*. *Eur. J. Biochem.* **33**: 192–200. [Medline] [CrossRef]