



Genome Sequence of *Dengue virus 3* from the *Pythium insidiosum* Transcriptomes

Yeonhwa Jo¹, Hoseong Choi¹ and Won K. Cho^{1,2*}

¹ Department of Agricultural Biotechnology, College of Agriculture and Life Sciences, Seoul National University, Seoul, Republic of Korea, ² The Taejin Genome Institute, Hoengseong, Republic of Korea

Keywords: genome, *Dengue virus 3*, *Pythium insidiosum*, *de novo* assembly, single nucleotide polymorphism

BACKGROUND

Pythium species are a kind of the fungus-like oomycetes presenting in soil and aquatic environments (Latijnhouwers et al., 2003). They are regarded as serious plant pathogens, resulting in economic losses of many crops. Of known *Pythium* species, *P. insidiosum* is the only oomycete that can infect both humans and animals (Gaastra et al., 2010). *P. insidiosum* causes a life-threatening infectious diseases known as pythiosis (Mendoza et al., 1993; Gaastra et al., 2010). So far, infection of *P. insidiosum* has been reported in horses, dogs, and humans in tropical and subtropical regions. *P. insidiosum* inhabits swampy areas and produces motile zoospores colonizing surfaces of water plants (Mendoza et al., 1993; Gaastra et al., 2010). Contact of zoospores with tissues of human and animal causes serious infection (Mendoza et al., 1993).

Dengue virus (DENV) causes severe visceral and central nervous system disease in humans (Martina et al., 2009). DENV belongs to the genus *Flavivirus* and consists of single-stranded positive-sense RNA (Chambers et al., 1990; Bai et al., 2013). Infection of DENV is mediated by two major mosquito vectors: *Aedes albopictus* and *Aedes aegypti* (Martina et al., 2009). So far, four different serotypes of DENV have been identified in tropical and subtropical regions (Bai et al., 2013). The genome of DENV is composed of a single-stranded positive-sense RNA about 10,700 nucleotides (nt) in length. The DENV genome consists of a single open reading frame (ORF) encoding three structural (C, prM, and E) and seven nonstructural (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) proteins (Chambers et al., 1990).

Next generation sequencing (NGS) facilitates the identification of known viruses in various organisms (Radford et al., 2012). NGS can be applied not only in detection or diagnosis of viruses but also in metagenome-based approaches to detect unexpected disease-associated viruses and novel viruses (Adams et al., 2009; Radford et al., 2012). In addition, complete or draft viral genomes can be assembled from NGS data (Jo et al., 2015).

In this study, we identified DENV3 from *P. insidiosum* transcriptome data and assembled a nearly complete genome of DENV3. In addition, we examined single nucleotide variations (SNVs) in the *P. insidiosum* transcriptome, demonstrating quasispecies of DENV3.

METHODS

Oomycete Material, Growth Condition, and Library Preparation for 454 Sequencing

The *P. insidiosum* transcriptome data was obtained from the previous study (Krajaejun et al., 2014). In brief, *P. insidiosum* strain Pi-S was isolated from a Thai patient showing vascular pythiosis. The obtained *P. insidiosum* was cultured in Sabouraud dextrose broth and incubated at 28°C and 37°C for 1 week. Mycelia from each condition was harvested and washed with sterile water. Again, the

OPEN ACCESS

Edited by:

Ayae Honda,
Hosei University, Japan

Reviewed by:

Takeshi Kurosu,
National Institute of Infectious
Diseases, Japan
Shigeo Sugita,
Equine Research Institute,
Japan Racing Association, Japan

*Correspondence:

Won K. Cho
wonkyong@gmail.com

Specialty section:

This article was submitted to
Virology,
a section of the journal
Frontiers in Microbiology

Received: 24 January 2016

Accepted: 31 May 2016

Published: 15 June 2016

Citation:

Jo Y, Choi H and Cho WK (2016)
Genome Sequence of Dengue virus 3
from the *Pythium insidiosum*
Transcriptomes.
Front. Microbiol. 7:926.
doi: 10.3389/fmicb.2016.00926

mycelia samples were transferred to 2 ml of microcentrifuge tube and incubated at 28°C and 37°C, respectively, for 24 h. Harvested mycelia was immediately frozen in liquid nitrogen for further study. Poly-A tailed mRNAs were isolated from total RNAs and used for cDNA library preparation. Sequencing was performed using the Genome Sequencer (GS) FLX Titanium platform (Roche Applied Sciences, Penzberg, Germany).

De novo Transcriptome Assembly

We downloaded two raw data with following accession numbers, DRR004443 and DRR004444, from DDBJ (DNA Data Bank of Japan) (https://trace.ddbj.nig.ac.jp/index_e.html). We used a workstation (two six-core CPUs and 256-GB RAM) operated using the Ubuntu 12.04.5 LTS operation system for all bioinformatics analyses in this study. Each individual FASTQ file for each library was first subjected to *de novo* transcriptome assembly using the Trinity program (version 2.0.2, released 22nd January 2015) with default parameters (Grabherr et al., 2011). We obtained two independent transcriptomes from the two samples (Table 1). For *de novo* genome assembly of DENV3, we combined two raw data sets and performed transcriptome assembly again using Trinity, Velvet programs (Zerbino and Birney, 2008), and iAssembler program (Zheng et al., 2011). After that, contigs associated with DENV3 were selected by blast search against known DENV3 reference sequences. Genome of DENV3 was manually assembled.

Identification of Viruses in the *P. insidiosum* Transcriptomes

To identify viruses in the *P. insidiosum* transcriptomes, two different approaches were used. For the first, *de novo* assembled contigs were blasted against complete viral reference sequences (<http://www.ncbi.nlm.nih.gov/genome/viruses/>). For the BLAST search, the MEGABLAST algorithm with a cut-off *e*-value of 1e-5 was applied. MEGABLAST is much faster than other sequence similarity programs and provides very reliable sequence similarities for the identification of viruses. For the second approach, all the FASTQ files were converted into FASTA files using the FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_

toolkit/) and FASTA files from each library were directly blasted against complete viral reference sequences.

Examination of Single Nucleotide Variations

To identify SNVs, all the raw data (FASTQ files) were aligned on the obtained DENV3 genome sequence using the Burrows–Wheeler Aligner (BWA) with default parameters (Li and Durbin, 2009). The obtained SAM files were converted to BAM files for sorting using SAMtools (Li et al., 2009). The BAM files were sorted for SNV calling using SAMtools. Afterward, mpileup was conducted to generate the VCF format. The SNVs were then called using BCFtools implemented in SAMtools, and finally, SNVs were filtered using BCFtools. For the visualization of mapped reads on the genome, aligned SAM files were imported to the Tablet program (Milne et al., 2010).

RESULTS

Identification of DENV3 in the *P. insidiosum* transcriptomes

In a search of viruses infecting fungi in the publicly available transcriptome data, we found DENV3-associated sequences from *P. insidiosum* transcriptomes. The transcriptomes were composed of two different libraries, which were generated from two different growth conditions, 28°C and 37°C. BLAST search revealed that 0.36% (638 reads) and 0.11% (229 reads) at 28°C and 37°C, respectively, were sequences associated with DENV3 (Figure 1A). The portion of DENV3 associated reads at 28°C was more than three times that at 37°C. This result indicates that replication of DENV3 was reduced two-fold when the temperature was increased.

De novo Viral Genome Assembly for DENV3

Many sequenced reads were associated DENV3. Therefore, we *de novo* assembled transcriptome of *P. insidiosum* to identify contigs associated with the DENV3 genome. BLAST search identified 116 (28°C) and 20 (37°C) contigs associated with DENV3 (Figure 1B; Supplementary Table S1). The longest contig was 2332 bp, while the shortest contig was 177 bp. We assembled the genome of DENV3. The genome of DENV3, referred to as isolate Pythium, was 10,671 nucleotides (nt) in length (accession number: KT424097). By BLASTN search against the NCBI nucleotide database, the DENV3 isolate Pythium was found to be highly matched to the known DENV3 isolate HN/2013/20 (KJ622192.1) with 98% sequence identity. The DENV3 isolate HN/2013/20 was isolated from Henan, China, in 2013, indicating the possible origin of DENV3 isolate Pythium.

We mapped raw data on the genome of the obtained DENV3 isolate Pythium. As shown in Figure 1C, the number of reads mapped on the genome was much higher in the transcriptome from 28°C as compared to that of 37°C (Figure 1C).

Analysis of Single Nucleotide Variations

We examined SNVs for DENV3 in two different transcriptomes (Supplementary Table S2). From the transcriptome at 28°C, we identified 73 SNVs including 11 InDels (Insertion and Deletions), while 23 SNVs harboring nine InDels were detected from the

TABLE 1 | Summary of *de novo* transcriptome assembly and the number of identified DENV3 associated reads and contigs.

Index	28°C	37°C
SRA accession No.	DRR004443	DRR004444
Total genes	14,327	17,761
Total transcripts	14537	18049
Percent GC	59.23	60.29
Contig N50	794	827
Median contig length	697	726
Average contig	751.7	797.18
Total assembled bases	10927464	14388240
Total No. of reads	177947	204241
No. of DENV3 reads	638	229
No. of DENV3 contigs	116	20

clearly identified a virus in the human pathogenic fungus so far.

To clearly demonstrate the presence of DENV3 in *P. insidiosum* transcriptomes, we assembled the genome of DENV3 isolate *Pythium* and examined its SNVs using *Pythium* transcriptome data. However, we are not sure whether the *P. insidiosum* was really infected by DENV3. It seems that the contamination rate of DENV3 in *P. insidiosum* culture should be very high. For example, a patient was infected with dengue virus and had viremia. Total RNAs from the viremia were transferred into *P. insidiosum* culture medium and slowly degraded during culture. Libraries were prepared from the culture medium containing dengue virus and sequenced by a next-generation sequencer. Moreover, the low amount of DENV3 at 37°C as compared to that at 28°C indicates the instability of the DENV3 RNA genome at higher temperatures. If DENV3 can replicate in *P. insidiosum*, we can get a high level of DENV3 RNA genome at 37°C, because flavivirus polymerase usually works better at higher temperatures (Simon et al., 2015).

Many previous studies have also reported microbial contamination in clinical samples by NGS. For example, substantial bacterial contamination has routinely been identified in RNA-Seq data (Strong et al., 2014) as well as DNA data (Laurence et al., 2014). Due to the high level of sensitivity of the NGS technique, NGS is currently used to discover new viruses associated with diseases in clinical virology (Datta et al., 2015; Perlejewski et al., 2015). Taken together, this study provides the complete genome of DENV3 from *P. insidiosum* transcriptomes that might have been contaminated during sample preparation.

REFERENCES

- Adams, I. P., Glover, R. H., Monger, W. A., Mumford, R., Jackeviciene, E., Navalinskiene, M., et al. (2009). Next-generation sequencing and metagenomic analysis: a universal diagnostic tool in plant virology. *Mol. Plant Pathol.* 10, 537–545. doi: 10.1111/j.1364-3703.2009.00545.x
- Bai, Z., Liu, L.-C., Jiang, L.-Y., Liu, Q., Cao, Y.-M., Xu, Y., et al. (2013). Complete genome sequence of dengue virus serotype 3 from Guangzhou, China. *Genome Announc.* 1:e0020812. doi: 10.1128/genomeA.00208-12
- Cai, G., and Hillman, B. I. (2013). Phytophthora viruses. *Adv. Virus Res.* 86, 327–350. doi: 10.1016/B978-0-12-394315-6.00012-X
- Chambers, T. J., Hahn, C. S., Galler, R., and Rice, C. M. (1990). Flavivirus genome organization, expression, and replication. *Annu. Rev. Microbiol.* 44, 649–688. doi: 10.1146/annurev.mi.44.100190.003245
- Datta, S., Budhaujiya, R., Das, B., and Chatterjee, S. (2015). Next-generation sequencing in clinical virology: discovery of new viruses. *World J. Virol.* 4:265. doi: 10.5501/wjv.v4.i3.265
- Gaastera, W., Lipman, L. J., De Cock, A. W., Exel, T. K., Pegge, R. B., Scheurwater, J., et al. (2010). *Pythium insidiosum*: an overview. *Vet. Microbiol.* 146, 1–16. doi: 10.1016/j.vetmic.2010.07.019
- Gillings, M., Tesoriero, L., and Gunn, L. (1993). Detection of double-stranded RNA and virus-like particles in Australian isolates of *Pythium irregulare*. *Plant Pathol.* 42, 6–15. doi: 10.1111/j.1365-3059.1993.tb01466.x

DATA ACCESS

The two raw data in this study can be available from DDBJ (DNA Data Bank of Japan) (https://trace.ddbj.nig.ac.jp/index_e.html) with following accession numbers, DRR004443 and DRR004444. Genome sequence of DENV3 isolate *Pythium* was deposited in GenBank with accession number KT424097.

AUTHOR CONTRIBUTIONS

WC designed the research; YJ, HC, and WC performed the research; YJ, HC, and WC analyzed the data; and YJ, HC, and WC wrote the paper.

ACKNOWLEDGMENTS

This work was carried out with the support of the “Cooperative Research Program for Agriculture, Science and Technology Development (Project No. PJ01186102)” conducted by the Rural Development Administration, Republic of Korea. This work is dedicated to the memory of my father, Tae Jin Cho (1946–2015).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2016.00926>

Supplementary Table S1 | Detailed information of megablast results to identify contigs associated with DENV3.

Supplementary Table S2 | Information of identified single nucleotide variations for DENV3 in two samples grown at 28°C and 37°C, respectively.

- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883
- Jo, Y., Choi, H., and Cho, W. K. (2015). De novo assembly of a bell pepper endornavirus genome sequence using RNA sequencing data. *Genome Announc.* 3, e00061–e00015. doi: 10.1128/genomeA.00061-15
- Kohno, S., Fujimura, T., Rulong, S., and Kwon-Chung, K. (1994). Double-stranded RNA virus in the human pathogenic fungus *Blastomyces dermatitidis*. *J. Virol.* 68, 7554–7558.
- Krajaeun, T., Lerksuthirat, T., Garg, G., Lowhnoo, T., Yingyong, W., Khositnithikul, R., et al. (2014). Transcriptome analysis reveals pathogenicity and evolutionary history of the pathogenic oomycete *Pythium insidiosum*. *Fungal Biol.* 118, 640–653. doi: 10.1016/j.funbio.2014.01.009
- Kurosu, T. (2011). Quasispecies of dengue virus. *Trop. Med. Health* 39, 29. doi: 10.2149/tmh.2011-S02
- Latijnhouwers, M., De Wit, P. J., and Govers, F. (2003). Oomycetes and fungi: similar weaponry to attack plants. *Trends Microbiol.* 11, 462–469. doi: 10.1016/j.tim.2003.08.002
- Laurence, M., Hatzis, C., and Brash, D. E. (2014). Common contaminants in next-generation sequencing that hinder discovery of low-abundance microbes. *PLoS ONE* 9:e97876. doi: 10.1371/journal.pone.0097876
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352
- Martina, B. E., Koraka, P., and Osterhaus, A. D. (2009). Dengue virus pathogenesis: an integrated view. *Clin. Microbiol. Rev.* 22, 564–581. doi: 10.1128/CMR.00035-09
- Mendoza, L., Hernandez, F., and Ajello, L. (1993). Life cycle of the human and animal oomycete pathogen *Pythium insidiosum*. *J. Clin. Microbiol.* 31, 2967–2973.
- Milne, I., Bayer, M., Cardle, L., Shaw, P., Stephen, G., Wright, F., et al. (2010). Tablet—next generation sequence assembly visualization. *Bioinformatics* 26, 401–402. doi: 10.1093/bioinformatics/btp666
- Perlejewski, K., Popiel, M., Laskus, T., Nakamura, S., Motooka, D., Stokowy, T., et al. (2015). Next-generation sequencing (NGS) in the identification of encephalitis-causing viruses: unexpected detection of human herpesvirus 1 while searching for RNA pathogens. *J. Virol. Methods* 226, 1–6. doi: 10.1016/j.jviromet.2015.09.010
- 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. doi: 10.1099/vir.0.043182-0
- Sharma, S., Gupta, S., and Shrivastava, J. (2011). Presence of virus like particles in human pathogenic fungi: *chrysosporium* sps and *Candida albicans*. *Indian J. Virol.* 22, 104–110. doi: 10.1007/s13337-011-0048-x
- Simon, A. Y., Sutherland, M. R., and Prydzial, E. L. (2015). Dengue virus binding and replication by platelets. *Blood* 126, 378–385. doi: 10.1182/blood-2014-09-598029
- Strong, M. J., Xu, G., Morici, L., Bon-Durant, S. S., Baddoo, M., Lin, Z., et al. (2014). Microbial contamination in next generation sequencing: implications for sequence-based analysis of clinical samples. *PLoS Pathog.* 10:e1004437. doi: 10.1371/journal.ppat.1004437
- Zerbino, D. R., and Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18, 821–829. doi: 10.1101/gr.074492.107
- Zheng, Y., Zhao, L., Gao, J., and Fei, Z. (2011). iAssembler: a package for de novo assembly of Roche-454/Sanger transcriptome sequences. *BMC Bioinformatics* 12:453. doi: 10.1186/1471-2105-12-453

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Jo, Choi and Cho. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.