



# Nearly Complete Genome Sequences of Human Norovirus Belonging to Several Genotypes from Valencia, Spain

Cristina Santiso-Bellón,<sup>a</sup> Carolina Monzó,<sup>b</sup> Azahara Fuentes-Trillo,<sup>b</sup> Susana Vila-Vicent,<sup>a</sup> Juliana da Silva Ribeiro de Andrade,<sup>c</sup>  Roberto Gozalbo-Rovira,<sup>a</sup> Javier Buesa,<sup>a</sup> Felipe J. Chaves,<sup>b,d</sup>  Jesús Rodríguez-Díaz<sup>a</sup>

<sup>a</sup>Department of Microbiology, School of Medicine, University of Valencia, Valencia, Spain

<sup>b</sup>Genomic and Genetic Diagnosis Unit, Hospital Clínico Research Foundation (INCLIVA), Valencia, Spain

<sup>c</sup>Laboratory of Comparative and Environmental Virology, Oswaldo Cruz Institute, Rio de Janeiro, Brazil

<sup>d</sup>CIBER of Diabetes and Associated Metabolic Diseases (CIBERDEM), Madrid, Spain

**ABSTRACT** Human noroviruses are responsible for most nonbacterial acute gastroenteritis cases. The GII.2, GII.4, and GII.17 genotypes of human noroviruses have recently arisen as the most frequent genotypes found in humans worldwide. We report here seven nearly complete genomes of these genotypes from patients with acute gastroenteritis in Valencia, Spain.

Human noroviruses are responsible for 18% of all cases of acute gastroenteritis worldwide (1) and are the third most common mortality cause in children under 5 years old (2). The *Norovirus* genus belongs to the *Caliciviridae* family and is classified into six distinct genogroups (GI to GVI), which are further subdivided into different genotypes (3). Norovirus strains are subjected to continuous processes of punctual mutations and/or recombination. Therefore, new strains emerge periodically (4, 5). In the last few years, a new recombinant strain has emerged (GII.P4.NewOrleans2009\_GII.4.Sydney2012). Together with this variant, GII.17 and GII.2 are responsible for the majority of outbreaks worldwide (6, 7).

We report here seven nearly complete genome sequences of human GII noroviruses (GII.2, GII.4, and GII.17) initially detected by conventional reverse transcriptase PCR (RT-PCR) in feces collected from both adults and children. Samples were from sporadic cases and outbreaks of acute gastroenteritis from November 2015 to September 2017. All patients were treated in different hospitals in Valencia, Spain.

Viral RNA was extracted from 250  $\mu$ l of 10% fecal suspension in phosphate-buffered saline using TRIzol reagent (Invitrogen). Total RNA was sent to Macrogen, Inc., for library construction using a TruSeq mRNA library prep kit (Illumina) and sequencing on a HiSeq platform (Illumina). Raw read quality control was assessed with FastQC version 0.11.5 (8). Reads were quality trimmed using seqtk trimfq version 1.2-r101-dirty (9), and genomes were assembled using SPAdes version 3.11.1 (10) and MEGAHIT version 1.2.4 (11). The obtained contigs were annotated using BLASTn against the full viral genome database from NCBI. Contigs assigned as norovirus were further genotyped using the norovirus genotyping tool (12). All genome sequences contained three open reading frames (ORFs).

Table 1 summarizes the most relevant information of each of the sequences. The isolate 3453 sequence had a 99.22% (7,468/7,527 [number of nucleotides identical to the reference sequence compared to the total number of nucleotides of the obtained sequence]) nucleotide similarity with the isolate NORO\_187\_01\_10\_2015 GII.2 strain (MH218655). The isolate 3351 sequence had a 99.63% (7,505/7,533) nucleotide similarity with the Hu/GII.P17\_GII.17/KR/2014/CAU-265 strain (KU561253). The isolate 3587 sequence shared 99.5% (7,520/7,558) nucleotide similarity with the Hu/GII.17/142700/

**Citation** Santiso-Bellón C, Monzó C, Fuentes-Trillo A, Vila-Vicent S, da Silva Ribeiro de Andrade J, Gozalbo-Rovira R, Buesa J, Chaves FJ, Rodríguez-Díaz J. 2019. Nearly complete genome sequences of human norovirus belonging to several genotypes from Valencia, Spain. *Microbiol Resour Announc* 8:e00641-19. <https://doi.org/10.1128/MRA.00641-19>.

**Editor** Jelle Matthijnssens, KU Leuven

**Copyright** © 2019 Santiso-Bellón et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Jesús Rodríguez-Díaz, [jesus.rodriguez@uv.es](mailto:jesus.rodriguez@uv.es).

C.S.-B. and C.M. contributed equally to this work.

**Received** 29 May 2019

**Accepted** 2 August 2019

**Published** 22 August 2019

**TABLE 1** Assembly characteristics for the norovirus contigs obtained<sup>a</sup>

Isolate	GenBank accession no.	Polymerase genotype <sup>b</sup>	Capsid genotype <sup>b</sup>	ORF1 nt <sup>c</sup>	ORF2 nt <sup>c</sup>	ORF3 nt <sup>c</sup>	No. of norovirus reads/ raw reads (%)	Coverage (x)	Genome coverage (%) (genome length [bp]) <sup>d</sup>	GC content (%)	GenBank accession no. of most similar strain (nt lacking in the 5' end)
3453	MK789430	GII.P2	GII.2	1-5094	5075-6703	6703-7482	4,529,781/31,181,056 (14.5)	2,675.12	99.9 (7,536/7,546)	49.8	MH218655 (GTGAATGAAG)
3351	MK789431	GII.P17	GII.17	1-5097	5078-6700	6700-7499	1,360,087/94,666,148 (1.4)	2,587.67	99.9 (7,533/7,549)	48.2	KU561253 (GTGAATGAAGATGCG)
3587	MK789432	GII.P17	GII.17	5-5113	5094-6716	6716-7495	67,237,010/70,923,098 (94.8)	7,674.57	100 (7,558/7,558)	48.1	KT380915
3618	MK789433	GII.P17	GII.17	1-5099	5080-6702	6702-7481	15,904,593/46,226,582 (34.4)	6,316.85	99.8 (7,542/7,556)	48.2	KU561253 (GTGAATGAAGATGG)
3546	MK789434	GII.P4 New Orleans 2009	GII.4 Sydney 2012	1-5093	5074-6696	6696-7502	39,211,847/90,843,600 (46.1)	7,716.42	99.9 (7,540/7,551)	49.9	MH218612 (GTGAATGAAGA)
3549	MK789435	GII.Pe	GII.4 Sydney 2012	5-5104	5085-6707	6707-7513	90,198,955/90,548,480 (99.6)	7,792.28	100 (7,569/7,569)	49.6	MH218701
3552	MK789436	GII.P4 New Orleans 2009	GII.4 Sydney 2012	1-5090	5071-6693	6693-7499	68,287,930/88,189,676 (77)	7,755.68	99.8 (7,536/7,550)	49.8	MH218612 (GTGAATGAAGATGG)

<sup>a</sup> nt, nucleotides.<sup>b</sup> Polymerase and capsid genotypes were assigned using the norovirus genotyping tool (12).<sup>c</sup> The ORFs were automatically assigned by GenBank after sequence submission.<sup>d</sup> The total expected genome length is determined relative to the sequence annotated as "full genome" with the higher similarity after BLAST search.

Shanghai/2014 strain (KT380915). The sequence of sample 3618 had a 99.11% (7,475/7,542) nucleotide similarity with the Hu/GII.P17\_GII.17/KR/2014/CAU-265 strain (KU561253). The isolate 3546 sequence had a 98.55% (7,430/7,539) nucleotide similarity with the GII.P4 New Orleans 2009\_GII.4 Sydney 2012 GII isolate NORO\_141\_02\_11\_2014 (MH218612). The sequence of sample 3549 shared a 99.55% (7,512/7,546) nucleotide similarity with the GII.Pe\_GII.4 Sydney 2012 NORO\_42\_01\_09\_2014 strain (MH218701). The isolate 3552 sequence had a 98.45% (7,419/7,536) nucleotide similarity with the GII.P4 New Orleans 2009\_GII.4 Sydney 2012 NORO\_141\_02\_11\_2014 strain (MH218612).

In summary, our study describes seven nearly complete genome sequences of different genotypes of norovirus isolates, providing additional important information to use as reference for phylogenetic and evolutionary studies.

**Data availability.** The GenBank accession numbers for the norovirus genomes are MK789430, MK789431, MK789432, MK789433, MK789434, MK789435, and MK789436. The sequence data are available in the Sequence Read Archive (SRA) under BioProject number PRJNA497363.

## ACKNOWLEDGMENTS

This work was supported by Spanish government (Ministerio de Economía y Competitividad) grants AGL2014-52996-C2-2-R and RYC-2013-12442 to J.R.-D. R.G.-R. is the recipient of a postdoctoral grant from the Valencian government (APOST/2017/037), and S.V.-V. is the recipient of a predoctoral fellowship from the Valencian government (ACIF/2016/437). A.F.-T. is the beneficiary of a predoctoral fellowship from the Valencian government (ACIF/2018/303). J.D.S.R.D.A. received financial support from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brasil (CAPES), finance code 001.

## REFERENCES

- Ahmed SM, Hall AJ, Robinson AE, Verhoef L, Premkumar P, Parashar UD, Koopmans M, Lopman BA. 2014. Global prevalence of norovirus in cases of gastroenteritis: a systematic review and meta-analysis. *Lancet Infect Dis* 14:725–730. [https://doi.org/10.1016/S1473-3099\(14\)70767-4](https://doi.org/10.1016/S1473-3099(14)70767-4).
- Lanata CF, Fischer-Walker CL, Olascoaga AC, Torres CX, Aryee MJ, Black RE. 2013. Global causes of diarrheal disease mortality in children <5 years of age: a systematic review. *PLoS One* 8:e72788. <https://doi.org/10.1371/journal.pone.0072788>.
- Kroneman A, Vega E, Vennema H, Vinjé J, White PA, Hansman G, Green K, Martella V, Katayama K, Koopmans M. 2013. Proposal for a unified norovirus nomenclature and genotyping. *Arch Virol* 158:2059–2068. <https://doi.org/10.1007/s00705-013-1708-5>.
- Bull RA, Eden JS, Rawlinson WD, White PA. 2010. Rapid evolution of pandemic noroviruses of the GII.4 lineage. *PLoS Pathog* 6:1–10. <https://doi.org/10.1371/journal.ppat.1000831>.
- Siebenga JJ, Vennema H, Renckens B, de Bruin E, van der Veer B, Siezen RJ, Koopmans M. 2007. Epochal evolution of GII.4 norovirus capsid proteins from 1995 to 2006. *J Virol* 81:9932–9941. <https://doi.org/10.1128/JVI.00674-07>.
- Chan MCW, Hu Y, Chen H, Podkolzin AT, Zaytseva EV, Komano J, Sakon N, Poovorawan Y, Vongpunswad S, Thanusuwannasak T, Hewitt J, Croucher D, Collins N, Vinjé J, Pang XL, Lee BE, De Graaf M, Van Beek J, Vennema H, Koopmans MPG, Niendorf S, Poljsak-Prijatelj M, Steyer A, White PA, Lun JH, Mans J, Hung TN, Kwok K, Cheung K, Lee N, Chan PKS. 2017. Global spread of norovirus gII.17 Kawasaki 308, 2014–2016. *Emerg Infect Dis* 23:1354–1359. <https://doi.org/10.3201/eid2308.161138>.
- Ao Y, Cong X, Jin M, Sun X, Wei X, Wang J, Zhang Q, Song J, Yu J, Cui J, Qi J, Tan M, Duan Z. 2018. Genetic analysis of re-emerging GII.P16–GII.2 noroviruses in 2016–2017 in China. *J Infect Dis* 218:133–143. <https://doi.org/10.1093/infdis/jiy182>.
- Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27:863–864. <https://doi.org/10.1093/bioinformatics/btr026>.
- Li H. 2015. Seqtk. <https://github.com/lh3/seqtk>.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
- Li D, Liu CM, Luo R, Sadakane K, Lam TW. 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31:1674–1676. <https://doi.org/10.1093/bioinformatics/btv033>.
- Kroneman A, Vennema H, Deforche K, Avoort HVD, Peñaranda S, Oberste MS, Vinjé J, Koopmans M. 2011. An automated genotyping tool for enteroviruses and noroviruses. *J Clin Virol* 51:121–125. <https://doi.org/10.1016/j.jcv.2011.03.006>.