

# Complete nucleotide sequence of a norovirus GII.4 genotype: evidence for the spread of the newly emerged pandemic Sydney 2012 strain to China

Yuqi Huo · Ailing Cai · Hui Yang · Mingli Zhou ·  
Jiaxin Yan · Dingxiang Liu · Shuo Shen

Received: 19 July 2013 / Accepted: 16 November 2013 / Published online: 27 November 2013  
© Springer Science+Business Media New York 2013

**Abstract** A newly emerged pandemic Sydney GII.4-like norovirus (NoV) (Jingzhou GII.4) was isolated in Jingzhou, China in April, 2013, demonstrating the rapid spread of the variant to China. The complete nucleotide sequence was compared with the prototype Sydney 2012 variant and its VP1 gene with that of Huzhou strain (isolated in January 2013 in Huzhou, China). The result demonstrates that the new variant has evolved rapidly, including mutations in the hypervariable P2 domain of the major capsid protein VP1. Our study also shows that the new Jingzhou GII.4 variant co-circulated with GII.3 and GI.2 at the same time, supporting further monitoring of the evolution of the new NoV variant in China.

**Keywords** Norovirus · Pandemic Sydney 2012-like GII.4 · Chinese variants Jingzhou2013GII.4 and Huzhou2013GII.4 · Complete nucleotide sequence

Norovirus (NoV) is one of the leading pathogens of acute viral gastroenteritis worldwide [1–3] in the last 20 years, causing infections of humans of all ages. Following the introduction of Rotavirus vaccines (Rotarix<sup>®</sup> and RotaTeq<sup>®</sup>) into national immunization programs, NoV has

become the leading cause of acute gastroenteritis in children in some countries [4–6]. The dominant strains are of genogroup II, genotype 4 (GII.4), and emerged rapidly with new variants detected every 2 or 3 years, associated with pandemics or prevailing in particular regions [7]. A newly emerged GII.4 isolate was identified in March 2012 during an epidemic in Sydney. The variant spread quickly to North America [8], Europe [9], and other regions [10] in the world. For example, from September to December 2012, in USA, 19–58 % of all cases of acute gastroenteritis were caused by the Sydney GII.4 variant [5]. In an epidemic in Huzhou, China, a Sydney GII.4-like variant was identified, and the sequence of VP1 gene was determined and submitted to Genbank.

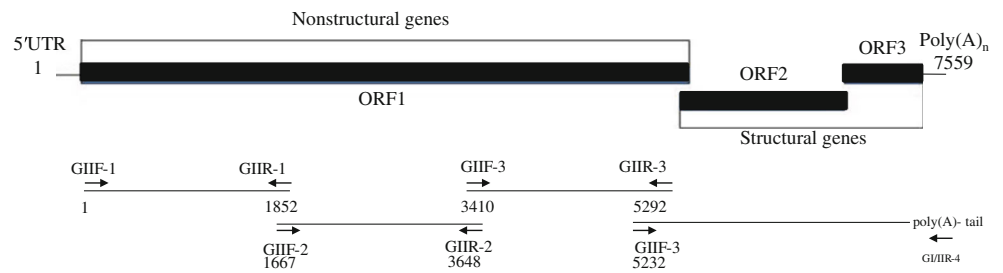
In April 2013, a new Sydney 2012-like GII.4 variant was isolated in Jingzhou, Hubei, China from a 2-year-old boy hospitalized with acute gastroenteritis. At the same time GII.3 and GI.2 NoVs were isolated. The complete nucleotide sequences of the three strains were determined and analyzed. Experiment procedures are briefly as follows. Stool samples from inpatient and outpatient with acute gastroenteritis were collected and diluted with 0.01 M PBS, pH 7.2, to make a 20 % of suspension (w/v). The suspension was cleared by low-speed centrifugation and 140  $\mu$ l of supernatant was used for RNA extraction with QIAamp Viral RNA mini kit following manufacturer's instructions. Extracted RNA was used for the first strand cDNA synthesis using PrimeScript 1st Strand cDNA Synthesis Kit. The synthesized cDNA was used to screen for the presence of NoV genome by using detection kit for Norwalk-like virus (PCR-Fluorescence Probing method). The mixed degenerate primer sets were used for amplification of the ORF1–ORF2 junction of GI and GII genogroups from positive samples [11]. PCR-amplified fragments were sequenced and then blasted to find the

Y. Huo · J. Yan · S. Shen (✉)  
Wuhan Institute of Biological Products Co., Ltd., 9 Linjiang  
Avenue, Wuchang, Wuhan 430060, China  
e-mail: 2628970324@qq.com

A. Cai · H. Yang · M. Zhou  
The Third People's Hospital of Jingzhou, Jingzhou, Hubei,  
China

D. Liu  
Nanyang Technological University, Singapore, Singapore

**Fig. 1** Cloning strategy. Four overlapping fragments covering the whole genome were cloned into the pGEM-T vector. Arrows indicate the orientation of primers used for cloning. The nt positions of individual fragments are shown below each primer



**Table 1** Differences in nt and predicted aa of three ORFs between Sydney 2012 and Jingzhou 403 strains

Position <sup>a</sup>	ORF1			Position <sup>b</sup>	ORF2			Position <sup>c</sup>	ORF3		
	Sydney 2012	Jingzhou403	aa changes <sup>d</sup>		Sydney 2012	Jingzhou403	aa changes <sup>d</sup>		Sydney 2012	Jingzhou403	aa changes <sup>d</sup>
85	C	T	L(29) → F	928	G	A	<i>D(310) → N</i>	218	A	G	K(73) → R
164	G	T	R(55) → I	968	T	A	<i>T(323) → A</i>	275	T	C	F(92) → S
584	C	T	S(195) → L	1118	G	A	<i>R(373) → H</i>	548	T	G	V(183) → G
1201	G	A	A(401) → T	1130	C	T	<i>A(377) → V</i>	632	C	T	S(211) → F
1487	A	G	K(496) → R	1241	A	C	H(414) → P				
2173	A	G	M(725) → V	1505	C	A	T(501) → N				
2284	A	G	S(762) → G	1618	G	C	V(540) → L				
2348	T	C	I(783) → T								
3127	A	G	S(1043) → G								
4034	G	A	S(1345) → N								

Letters in italic indicate aa changes located in the hypervariable P2 region of VP1. Synonymous mutations are not included. Numbers in parentheses are aa positions (from the N terminal)

<sup>a</sup> Numbering from the 5' end of the nt sequence of ORF1

<sup>b</sup> Numbering from the 5' end of the nt sequence of ORF2

<sup>c</sup> Numbering from the 5' end of the nt sequence of ORF3

<sup>d</sup> aa changes from Sydney 2012 to Jingzhou 403

complete genome sequences with closest similarity to Jingzhou isolates for design of sequencing primers. Four primer sets for each of two genotypes were used for sequencing of the whole genome of isolated strains, with the exception of the first 24 nucleotides of the 5'-end of the viral genomes which are derived from primers.

The whole genome of Jingzhou GII.4-2013 strain (Jingzhou 403) (Fig. 1) was cloned and sequenced (Genbank accession number, KF306214), representing the first complete nucleotide sequence of GII.4 Sydney 2012-like Chinese strain. The genome of GII.4 Jingzhou 2013 is 7,559 nucleotides (nt) in length with a polyadenylated tail. ORF1 (5,100 nt), ORF2 (1,623 nt), and ORF3 (807 nt) are overlapped with 20 and 1 nt, respectively. In comparison with the original Sydney strain (Genbank accession number, JX459908.1), there are 54, 24, and 9 differences in nucleotide sequences in ORF1, ORF2, and ORF3, causing 10, 7, and 4 amino acid (aa) changes, respectively (Table 1). Seven mutations leading to aa changes are located at ORF2, coding for the capsid protein VP1, and 4

are within the P2 hypervariable domain. Of the four aa changes located in hypervariable P2 region of GII.4 Jingzhou 2013 strain, residues 373 and 377 are located adjacent to putative epitope C (consisting of aa residues 296 and 372) and epitope B (consisting of aa residues 340 and 376), respectively [12]. Such changes might enable new strains evade human immune responses by changing antibody recognition sites and/or receptor binding sites resulting in varied binding patterns to HBGA receptors [13]. Compared with the VP1 gene of Huzhou 2013 strain, 14 nt differences were identified, resulting in three aa changes (Table 2). The changes may correlate with escape from herd immunity [14]. The new GII.4 variant co-circulated with NoVs GII.3 and GI.2 in Jingzhou. The complete nt sequences of the representative strains of Jingzhou GII.3 and GI.2 were determined (Genbank accession numbers KF306213 and KF306212).

The phylogenies (Fig. 2) show the evolution of GII.4 lineage of major pandemic isolates [15]. To further clarify the evolution of GII.4 lineages, 76 GII.4 VP1 sequences of

**Table 2** Differences in nt and predicted aa of ORF2 between Huzhou128 and Jingzhou403 strains

Position <sup>a</sup>	ORF2		aa changes <sup>b</sup>
	Huzhou128	Jingzhou403	
967	A	G	T(323) → A
1130	C	T	A(377) → V
1505	C	A	T(502) → N

Synonymous mutations are not included

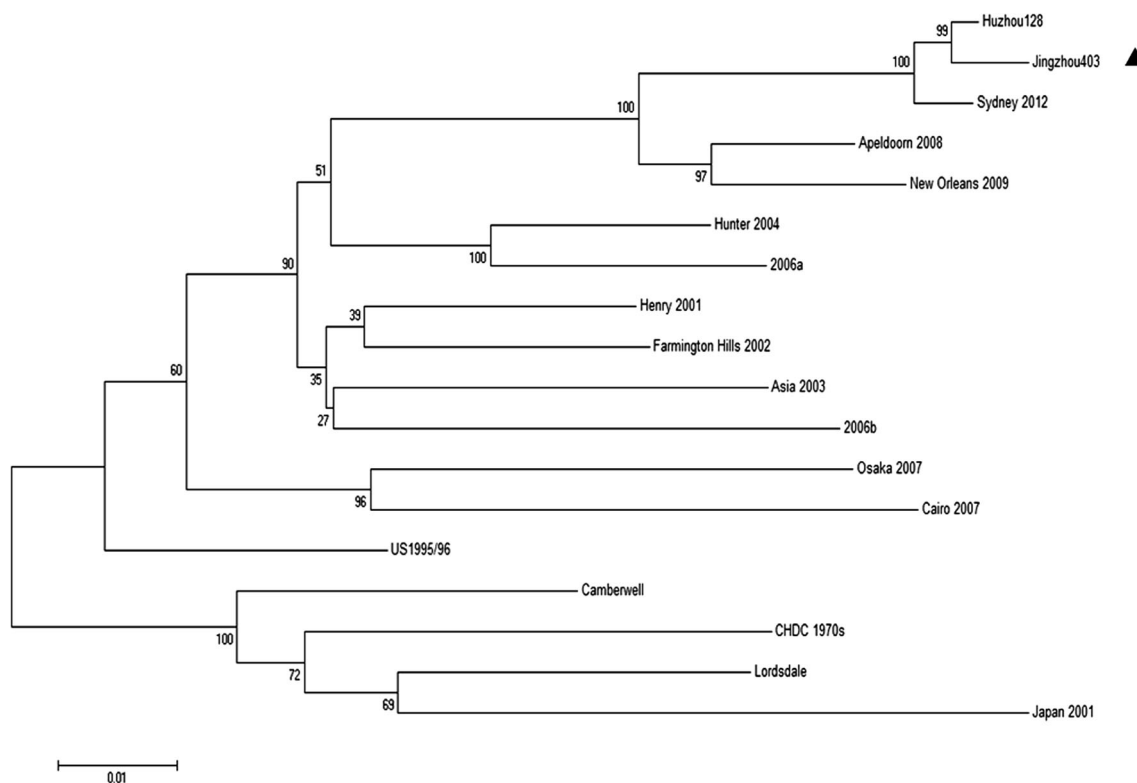
<sup>a</sup> Corresponding to the 5' end of the nt sequence of ORF2

<sup>b</sup> Numbers in parentheses are aa positions (from the N terminal)

NoVs spanning a 40-year period (1974–2013) were submitted to an evolutionary analysis (sequence accession numbers are available upon request). Rates of nucleotide substitution per site per year and the time to the most recent common ancestor were estimated by using the Bayesian Markov Chain Monte Carlo algorithm as implemented in

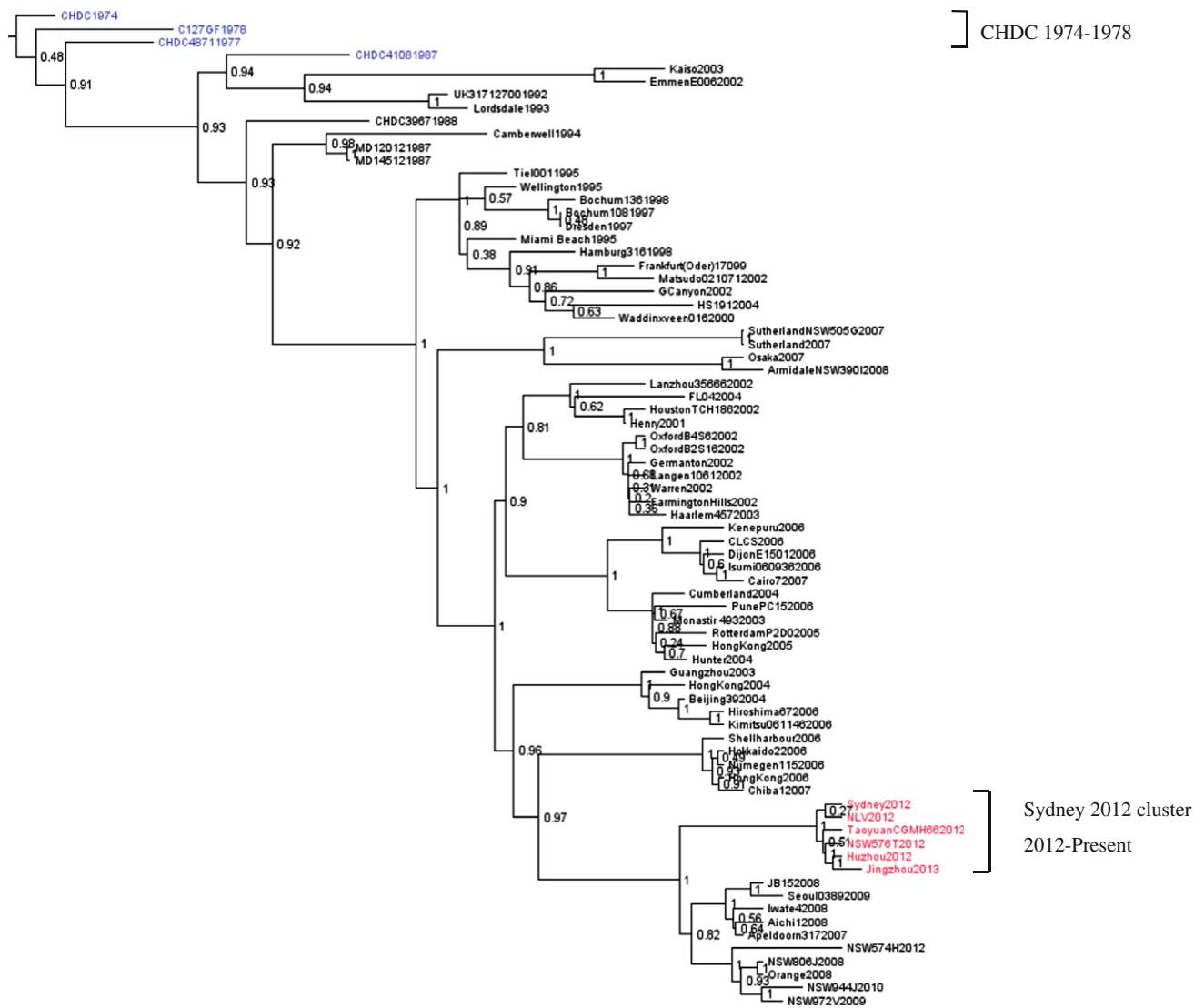
the BEAST software package [16]. The rate of evolution, based on the most conservative molecular clock model (strict clock) was  $4.80 \times 10^{-3}$  nucleotide/site/year ( $4.33\text{--}5.28 \times 10^{-3}$ ) and the mean age of the population was 40 years. Based on above evolutionary analysis, a maximum clade credibility tree was inferred (Fig. 3), showing CHDC GII.4 sequences from the 1970s forming a cluster that likely evolved into the present Sydney 2012 cluster.

The result explain the rapid spread of this GII.4 variant to China and the rest of the world, and rapid molecular evolution of NoVs, attributing to mutations in response to herd immunity. The impact of amino acid changes in VP1 of variants on the antigenicity or immunogenicity needs to be investigated as multiple reports have concluded that the major capsid protein of GII.4 are evolving rapidly [17, 18]. Understanding of the relationship among antigenic variations in the major capsid protein of pandemic strains will certainly shed new light on the development of effective



**Fig. 2** Phylogenetic analysis of dominant GII.4 strains circulating during 1974–2013. Phylogenetic tree was constructed based on the complete sequences of ORF2 of NoVs. The evolutionary distances were computed using the Maximum Composite Likelihood method, and the tree was plotted by using the neighbor-joining method. Numbers at each branch point are bootstrap values for that branch supported cluster. The newly identified strain Jingzhou 403 was marked with a triangle. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. Strains and Genbank accession No. are as follows:

Huzhou128, KC473546.1; CHDC 1970s (1974–1977), FJ537134; Lordsdale (1987–1993), X86557; Camberwell (1987–1994), AF145896; US1995/96 (1995–2002), AY741811; Henry 2001 (2000–2004), EU310927; Japan 2001 (2002–2003), AB294779; Farmington Hills 2002 (2002–2004), AY502023; Asia 2003 (2003–2006), DQ369797; Hunter 2004 (2004–2006), DQ078814; 2006a (2006–2008), EF187497; 2006b (2006–2012), EF684915; Osaka 2007 (2007–2008), AB541319; Cairo 2007 (2007), GQ845368; Apeldoorn 2008 (2007–2011), HQ009513; New Orleans 2009 (2008–2012), GQ845367; Sydney 2012 (2012), JX459908



**Fig. 3** Maximum clade credibility tree of GII.4 NoVs VP1 sequence dataset (76 sequences). Each internal node is labeled with the posterior probability monophyly of the corresponding clade. GII.4

sequences from the CHDC1974 cluster and Sydney 2012 cluster are indicated. Branch length was calibrated to reflect temporal patterns

vaccines [19] and therapeutic monoclonal antibodies [20]. Considering the large population in China and NoVs becoming the second leading pathogen of non-bacterial acute gastroenteritis [21, 22], further spreading and evolution of NoVs need to be closely monitored in China.

**References**

1. B.A. Lopman, M.H. Reacher, Y. Van Duynhoven, F.X. Hanon, D. Brown, M. Koopmans, *Emerg. Infect. Dis.* **9**, 90–96 (2003)
2. A.J. Hall, V.G. Eisenbart, A.L. Etingue, L.H. Gould, B.A. Lopman, U.D. Parashar, *Emerg. Infect. Dis.* **18**, 1566–1573 (2012)
3. M.D. Kirk, K.E. Fullerton, G.V. Hall, J. Gregory, R. Stafford, M.G. Veitch, N. Becker, *Clin. Infect. Dis.* **51**, 907–914 (2010)

4. M. Hemming, S. Räsänen, L. Huhti, M. Paloniemi, M. Salminen, T. Vesikari, *Eur. J. Pediatr.* **172**(6), 739–746 (2013)
5. F. Bucardo, P.E. Lindgren, L. Svensson, J. Nordgren, *PLoS One* **6**(10), e25962 (2011)
6. H.L. Koo, F.H. Neill, M.K. Estes, F.M. Munoz, A. Cameron, H.L. DuPont, R.L. Atmar, *J. Ped. Infect. Dis.* **2**(1), 57–60 (2013)
7. J.J. Siebenga, H. Vennema, D.P. Zheng, J. Vinje, B.E. Lee, X.L. Pang, E.C. Ho, W. Lim, A. Choudekar, S. Broor et al., *J. Infect. Dis.* **200**, 802–812 (2009)
8. *MMWR, Morb. Mortal. Wkly Rep.* **62**, 55 (2013)
9. J. Fonager, L.S. Hindbaek, T.K. Fischer, *Euro. Surveil.* **18**, 9 (2013)
10. M. Rahman, S. Nahar, M.H. Afrad, A.S. Faruque, T. Azim, *Emerg. Infect. Dis.* (2013). doi:10.3201/eid1908.130227
11. T. Kageyama, S. Kojima, M. Shinohara, K. Uchida, S. Fukushi, F.B. Hoshino, N. Takeda, K. Katayama, *J. Clin. Microbiol.* **41**, 1548–1557 (2003)
12. E.F. Donaldson, L.C. Lindesmith, A.D. Lobue, R.S. Baric, *Immunol. Rev.* **225**, 190–221 (2008)

13. L.C. Lindesmith, V. Costantini, J. Swanstrom, K. Debbink, E.F. Donaldson, J. Vinjé, R.S. Baric, *J. Virol.* **87**(5), 2803–2813 (2013)
14. K. Debbink, L.C. Lindesmith, E.F. Donaldson, V. Costantini, M. Beltramello, D. Cortj, J. Swanstrom, A. Lanzavecchia, J. Vinjé, R.S. Baric, *J. Infect. Dis.* **208**(11), 1877–1887 (2013)
15. J.S. Eden, M.M. Tanaka, M.F. Boni, W.D. Rawlinson, P.A. White, *J. Virol.* **87**(11), 6270–6282 (2013)
16. A.J. Drummond, A. Rambaut, *BMC. Evol. Biol.* **7**, 214 (2007)
17. L.C. Lindesmith, E.F. Donaldson, R.S. Baric, *J. Virol.* **85**, 231–242 (2011)
18. L.C. Lindesmith, M. Beltramello, E.F. Donaldson, D. Corti, J. Swanstrom, K. Debbink, A. Lanzavecchia, R.S. Baric, *PLoS Pathog.* (2012). doi:[10.1371/journal.ppat.1002705](https://doi.org/10.1371/journal.ppat.1002705)
19. D.I. Bernstein, R.L. Atmar, M. Lyon, J.J. Treanor et al., IDWeek 2013, San Francisco, CA, Oct. 2–6, 2013
20. Z. Chen, S.V. Sosnovtsev, K. Bok, G.I. Parra, M. Makiya, L. Agulto, K.Y. Green, R.H. Purcell, *J. Virol.* **87**(17), 9547–9557 (2013)
21. L.J. Liu, W. Liu, Y.X. Liu, H.J. Xiao, N. Jia, G. Liu, Y.G. Tong, W.C. Cao, *Am. J. Trop. Med. Hyg.* **82**, 717–722 (2010)
22. H. Yan, F. Yagyu, S. Okitsu, O. Nishio, H. Ushijima, *J. Virol. Methods* **114**(1), 37–44 (2003)