Infection, Genetics and Evolution 53 (2017) 199–205

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Research paper

Variation analysis of norovirus among children with diarrhea in rural Hebei Province, north of China

Niu Qiao [a,f,1], Song-Mei Wang [b,1], Jin-Xia Wang [c,d,1], Bin Kang [a,f], Shan-Shan Zhen [a,f], Xin-Jiang Zhang [e], Zhi-Yong Hao [e], Jing-Chen Ma [d], Chao Qiu [a,f], Yu-Liang Zhao [d], Lei Liu [a,f,*], Xuan-Yi Wang [a,f,**]

[a] Key Laboratory of Medical Molecular Virology of MoE & MoH, Fudan University, Shanghai 200032, People's Republic of China
[b] Laboratory of Molecular Biology, Training Center of Medical Experiments, School of Basic Medical Sciences, Fudan University, Shanghai 200032, People's Republic of China
[c] College of Public Health, North China University of Science and Technology, Tangshan 063009, People's Republic of China
[d] Hebei Province Center for Disease Control and Prevention, Shijiazhuang 050021, People's Republic of China
[e] Zhengding County Center for Disease Control and Prevention, Zhengding 050800, People's Republic of China
[f] Institutes of Biomedical Sciences, Fudan University, Shanghai 200032, People's Republic of China

ARTICLE INFO

Article history:
Received 12 March 2017
Received in revised form 3 June 2017
Accepted 6 June 2017
Available online 7 June 2017

Keywords:
Norovirus
Genotyping
Phylogenetic tree
Haplotype network
Evolutionary rates
China

ABSTRACT

To understand the distribution of genotyping, as well as evolution of norovirus circulating among children < 5 yrs., a population-based diarrhea surveillance targeted children < 5 yrs. was conducted in rural Zhengding County, Hebei Province, China between October 2011 and March 2012. RT-PCR was used to amplify the capsid-encoding region of GI and GII norovirus to identify norovirus infection. All PCR products were sequenced and analyzed for genotyping and constructing phylogenetic tree. Dynamic distribution network was constructed by TempNet to illustrate the genetic relationships at two different time points. Bayesian evolutionary inference techniques were applied by BEAST software to study the norovirus evolution rate. During the 6-month surveillance period, 1091 episodes of diarrhea were reported from 5633 children under 5 years of age lived in catchment area. 115 of 1091 stool specimens were detected as norovirus positive (10.54%). Five genotypes based on capsid gene sequences were identified, including GII.2 (11), GII.3 (52), GII.4 (47), GII.6 (4) and GII.7 (1). An identical haplotype of GII.4 circulated between 2006 and 2011 in Hebei Province. A mean rate of 6.29×10^{-2} nucleotide substitutions/site/year (s/s/y) was obtained for GII.3 viruses in Hebei, while the GII.4 viruses evolved at a mean rate of 3.67×10^{-2} s/s/y. In conclusions, GII.3 (45.22%) and GII. 4(40.87%) are the predominant strain in Hebei Province in the winter season of 2011 and 2012. Different from the current consensus, our study shows that GII.3 noroviruses in Hebei Province evolved at a faster rate than GII.4 viruses.

1. Introduction

Acute gastroenteritis is a major public health problem among children worldwide, and noroviruses (NoVs) is recognized as the major cause of viral gastroenteritis among all age groups (Ahmed et al., 2014). NoVs are genetically classified into seven genogroups (GI-GVII) that further subdivided into genotypes. GI and GII NoVs are the genogroups primarily responsible for human illness, although GIV has also been detected in humans (Ao et al., 2014; Eden et al., 2014; Kroneman et al., 2013; Martella et al., 2013; Baehner et al., 2016). Based on viral protein 1 (VP1) sequences, the GI and GII genogroups can be further divided into 9 and 22 genotypes, respectively. In addition, GIV.1 from GIV group is found also infected humans (Zheng et al., 2006; Mesquita et al., 2010). Despite the broad genetic diversity, the majority of acute gastroenteritis outbreaks due to NoVs infection is caused by GII.4 NoVs (Siebenga et al., 2009; Lindesmith et al., 2008), while GII.3 NoVs are one of the most common genotypes associated with sporadic NoV infection, particularly in children, where they often are identified as the dominant genotype (Siebenga et al., 2009; Boon et al., 2011; Barreira et al., 2010). It is generally believed that sequence mutations and homologous genome recombination are the two main mechanisms for the current NoV variations (Bull and White, 2011).

NoVs are highly genetically and antigenically diverse and their epidemiology and transmission patterns are similar to that of influenza viruses. Epidemiologic studies have shown that GII.4 NoVs has a rapid local transmission and that novel epidemic strains emerged every 2 to 3 years and spread globally in months (Siebenga et al., 2009; Lindesmith et al., 2008). The surface-exposed host ligand binding site on the NoV capsid is under heavy immune selection and likely evolves

* Corresponding author at: Institutes of Biomedical Sciences, Shanghai Medical College, Fudan University, 138 Yi Xue Yuan Rd., Shanghai, People's Republic of China.
** Correspondence to: X.-Y. Wang Key Laboratory of Medical Molecular Virology of MoE & MoH, Shanghai Medical College, Fudan University, 138 Yi Xue Yuan Rd., Shanghai, People's Republic of China.
E-mail addresses: liulei@fudan.edu.cn (L. Liu), xywang@shmu.edu.cn (X.-Y. Wang).
[1] Niu Qiao, Song-Mei Wang, and Jin-Xia Wang contributed equally.

http://dx.doi.org/10.1016/j.meegid.2017.06.007
1567-1348/

by antigenic drift in the face of human herd immunity (Lindesmith et al., 2008; Bull and White, 2011). Fast mutations, like genetic drift, in the surface proteins allow virus to elude host immunity, resulting in an ineffective immune protection produced by previous infection. As influenza viruses, vaccines could be targeted to protect against NoV infections (Lindesmith et al., 2008). A thorough understanding of the evolutionary pattern, evolutionary rate, genetic diversity, and epidemic cycle of NoVs may help interpret how these viruses change, evade the host immune response, and adapt to the host environment.

NoV genome consists of three open reading frames (ORFs) (Glass et al., 2009). Current classification of NoV genotypes is based on the genetic diversity within the polymerase (regions A and B) and the major capsid (VP1; regions C, D, and E) gene (Eden et al., 2014; Martella et al., 2013; Ando et al., 2000; Kirsten et al., 2009) Based on overall performance, the region C is recommended for routine genotyping of NoVs, while the region D may be useful for identifying new GII.4 variants (Kirsten et al., 2009).

In this article, the genetic characterization and phylogenetic analyses of the NoVs detected in Hebei Province were conducted to determine their genogroup and genotypes and constructed the corresponding phylogenetic tree. The genetic relationships of these NoVs in Hebei Province at two different time points were also illustrated by constructing a haplotype network, which is a useful tool for understanding the number, the relative frequency and dissimilarity between haplotypes within a single population. Furthermore, Bayesian genealogical inference of time-measured trees was performed using Markov chain Monte Carlo (MCMC (Green, 1995)) sampling to calculate the evolutionary rates of the dominant NoV strains.

2. Material and methods

2.1. Study population and surveillance

In total, 34 villages located in five townships in Zhengding County, Hebei Province, China, were selected as the catchment area for the population- and health care facility-based viral diarrhea surveillance targeted children < 5 years of age. All health-care providers that offered health care for the children in the catchment area constitute the surveillance system, including 101 village clinics, 5 township hospitals, and 1 county hospital. The surveillance was conducted from October 1, 2011 through March 31, 2012, the peak season for viral diarrheal illness in children. Bulk stool samples were collected from children with diarrhea by health care providers during their visits to the hospitals or clinics. The total number of children who contributed to the cohort were 6441. Of these, 1211 diarrhea episodes were reported, and 1091 (90.1%) children provided stool samples for NoV test through surveillance system, resulting in a diarrhea incidence rate of 188.0/1000 person/year (Zhen et al., 2015). This study was reviewed and approved by the Institutional Review Board of Hebei Center for Disease Control and Prevention. Written inform consent was obtained from the parent/guardian of each child. The study was performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments.

2.2. Laboratory diagnosis of NoVs

Stool samples were prepared as 20% fecal suspensions, and total viral RNA was extracted using the immunomagnetic virus DNA/RNA extraction kit (Xi'an Tianlong Science & Technology Co. Ltd), and reverse transcribed into cDNA using GoScript Reverse Transcription System A5001 (Promega Corporation) according to the manufacturer's instructions.

GI and GII NoVs were detected in separate reactions by conventional Reverse Transcriptase Polymerase Chain Reaction (RT-PCR) using consensus primers, G1SKF/G1SKR and COG2F/G2SKR (Kojima et al., 2002) for both GI and GII NoVs, respectively. These primers amplify a region spanning from position 5342 to 5671 (330 bp) of the GI NoV genome (Norwalk/68, GenBank accession No. M87661) containing 17 bp of 3′ end ORF1 and 313 bp of 5′ end ORF2, or spanning from position 5003 to 5389 (387 bp) of GII NoV genome (Lordsdale/93, GenBank accession No. X86557) containing 83 bp of 3′ end ORF1 and 304 bp of 5′ end ORF2. The GoTaqDNA polymerase (including 5× buffer) was purchased from Promega Company. Sanger dideoxy termination sequencing was applied based on the amplified PCR products representing corresponding norovirus genomic fragments. The sequencing was carried out by the Biosune Co., Ltd. in Shanghai, China. All obtained sequences were in high quality. Manual editing was performed to delete indels that occurred in >50% of the sequences, to allow better and reliable outcomes of the sequence comparison.

2.3. Reference sequences collection

A total of 25 NoV sequences that were originally isolated in Hebei Province, China, are available from GenBank Database. 24 of them that were isolated from 2006 to 2007 were downloaded for the analysis of the temporal network. In addition, 11 GII NoV sequences were obtained from the Database of the Netherlands National Institute for Public Health and the Environment (RIVM) to construct phylogenetic tree. The detailed accession numbers of these reference sequences download from GeneBank was shown in the Supplemental materials (Table S1).

2.4. Sequence analysis

The nucleotide sequences obtained from our study were firstly genotyped by NCBI BLAST (Pevsner, 2013) tool. The resulting NoV genotypes were then confirmed by the online NoV genotyping tool (http://www.rivm.nl/mpf/norovirus/typingtool) (Kroneman et al., 2011) based on ORF2 sequences offered by RIVM.

The nucleotide sequences were aligned using the multiple sequence alignment program MUSCLE (Edgar, 2004) implemented by Molecular Evolutionary Genetics Analysis (MEGA) software version 6.06 (Tamura et al., 2013). Manual editing was carried out to delete indels occurring in >50% of the sequences. Some of the sequences were truncated at both the 5′ and 3′ ends, possibly confounding the phylogeny of local clades.

2.5. Recombination analysis

The Genetic Algorithms for Recombination Detection (GARD (Kosakovsky Pond et al., 2006)) and the Single Break-Point (SBP (Kosakovsky Pond, 2006)) methods from Datamonkey Web Site (Delport et al., 2010; Pond and Frost, 2005) were utilized to determine if recombination occurred within the capsid protein-encoding sequences, which can therefore be used to determine if recombination was a factor in the evolution of the capsid protein-encoding sequences.

2.6. Phylogenetic analyses

To infer the evolutionary relationships among the NoV strains, Maximum Likelihood (ML) phylogenetic trees were constructed by using MEGA software v6.06 (Tamura et al., 2013; Hall, 2013). A separate Hasegawa-Kishino-Yano (HKY (Hasegawa et al., 1985)) substitution model with gamma-distributed rate variation among sites (Beth et al., 2006) and the complete deletion for gaps/missing data treatment (Hall, 2013) were applied. In the search process for the optimal evolutionary tree, Neighbor-Joining (NJ) tree as initial tree with Nearest-Neighbor Interchange (NNI) for ML heuristic method was chosen. Besides, a ML bootstrap analysis (1000 replicates) was used to evaluate the robustness of the phylogenetic grouping.

2.7. Temporal network analyses

To compare the genetic distributions of the NoVs isolated in Hebei Province, a haplotype network was constructed at two different time