Journal of Clinical Virology 56 (2013) 286–292

Contents lists available at SciVerse ScienceDirect

Journal of Clinical Virology

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

P2 domain profiles and shedding dynamics in prospectively monitored norovirus outbreaks

Faizel H.A. Sukhrie[a,b,*], Peter Teunis[c,d], Harry Vennema[a], Jolanda Bogerman[e], Sebastian van Marm[a], Matthias F.C. Thijs Beersma[b], Marion Koopmans[a,b]

[a] Laboratory for Infectious Diseases and Perinatal Screening, Centre for Infectious Disease Control (RIVM), Bilthoven, The Netherlands
[b] Department of Virology, Erasmus Medical Center, Rotterdam, The Netherlands
[c] Epidemiology and Surveillance Unit, Centre for Infectious Disease Control (RIVM), Bilthoven, The Netherlands
[d] Hubert Department of Global Health, Rollins School of Public Health, Emory University, Atlanta, USA
[e] Municipal Health Service, Rotterdam, The Netherlands

ARTICLE INFO

Article history:
Received 6 September 2012
Received in revised form 3 December 2012
Accepted 7 December 2012

Keywords:
Norovirus transmission
Shedding
P2 domain

ABSTRACT

Background: Norovirus P2 domain is commonly used to extrapolate transmission within an outbreak (OB) setting. The current definition is that transmission among cases is considered to be proven when no sequence variation is found.
Objectives: Previous studies have shown a high mutation rate and errors during replication of the norovirus genome, therefore the validity of this criterion must be evaluated.
Study design: Sequences of the P2 domain were obtained from patients and health care workers sampled during 4 prospectively GII.4 outbreaks. Fecal samples were tested by RT-PCR for presence of norovirus RNA against a standard control preparation to allow quantification. Estimated time of onset of shedding was derived from shedding kinetics modeled on data from sequential sampling. Thereby P2 sequence variation could be linked to estimated total virus excretion in individual subjects.
Results: In all the outbreaks, P2 domain variation was found that resulted in unique codon changes in some patients. Mutations were found in 14% of initial samples and >50% of follow-up samples taken from patients involved in an outbreak. In three patients, aa mutations was observed in or near sites involved in host or antigen binding.
Conclusions: We concluded that P2 domain variation increases with duration of virus shedding, but was unrelated to total amounts of virus shed. Therefore, we propose that cluster identification based on identical sequences should be relaxed to accommodate minor sequence variation. When using sequence data to support outbreak investigations, sequence diversity should be interpreted in relation to timing of sampling since onset of illness.

1. Background

Noroviruses (NoV) are a major cause of gastroenteritis worldwide and are most commonly associated with outbreaks in health care settings.[1] Onward transmission of noroviruses is common when guidelines for outbreak control are not applied rigorously.[2] For developing effective control measures, a proper understanding of the transmission patterns during outbreaks is needed, including the role of healthcare workers and asymptomatic shedders. Molecular typing of NoV-positive stool samples can be used to determine links between individual cases. A systematic analysis of genome diversity in a large dataset collected through the Foodborne viruses of Europe network concluded that the optimal target for sequence-based linking of cases was the capsid gene.[3] For practical reasons, currently the P2 domain of the NoV is used.[4–7] This genome region is considered to be the most variable part of the genome since it codes for the protruding domain of the capsid protein, which contains the receptor binding domain and important epitopes targeted by antibodies that inhibit binding.[8,9] In GII.4 NoV, the P2 domain evolves by accumulation of mutations under selective pressure from host immunity.[10–12] Accumulation of mutations in this domain was also shown in immunocompromised patients

Abbreviations: NoV, norovirus; HCW, health care workers; OB, outbreaks; EMC, university hospital.

* Corresponding author at: Center for Infectious Disease and Screening, Department of Virology, P.O. Box 1, 3720 BA, A. van Leeuwenhoeklaan 9, Bilthoven, The Netherlands. Tel.: +30 2742909; fax: +30 2744418.
E-mail address: faizel.sukhrie@rivm.nl (F.H.A. Sukhrie).

1386-6532/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jcv.2012.12.006

with prolonged shedding of viruses.[13] Recently the use of next generation sequencing identified minority variants present during transmission events.[14]

2. Objectives

P2 domain sequencing has been used for identifying the transmission pathways and links during outbreaks[4–7] and identical P2 domain sequences are considered evidence for a cluster. However, with the high mutation rate of norovirus,[15] nucleotide changes may occur within a short time interval, raising the question what would be an appropriate minimum number of nucleotide changes for defining a cluster of cases connected by direct transmission links. This question is relevant because outbreaks may be missed with common cluster detection algorithms that use time and space, or pseudo-outbreaks may occur when many patients are hospitalized during peak season.[6,7] Therefore, we set out to quantify P2 domain variation, during four prospectively monitored outbreaks in three nursing homes and a tertiary care hospital.[16] We sampled NoV-positive patients and health care workers (HCW), identified through an enhanced outbreak protocol irrespective of their symptom status. Variation in sequence data between and within outbreaks, as well as variation between and within infected subjects was analyzed and correlated with the estimated number of viruses shed by each individual. The results can be used to redefine criteria for linking of cases to outbreaks.

3. Study design

We prospectively monitored four GII.4 outbreaks starting from January 2009 until March 2011 in the region of Rotterdam in the University Hospital (EMC) and in nursing homes.[16] Sampling was based on an enhanced outbreak investigation protocol focusing on the identification of possible sources and modes of transmission of NoV.[16] The study protocol included random sampling irrespective of symptoms of all patients and HCW on affected wards with NoV. Patients and their contacts involved in the outbreaks who met the inclusion criteria of the study protocol were tested weekly until a negative test was returned. Each case was confirmed by sequencing region A (genotyping) followed by P2 domain sequencing.[6,17] The amount of virus shed by each subject was estimated from real time RT-PCR analysis of all fecal samples, using RNA transcribed from a plasmid containing a sequence spanning all commonly used diagnostic targets as a reference template,[18] allowing us to investigate correlation between virus excretion and P2 domain changes over time. Background sequences from the same geographic region were collected from patients outside the studied outbreaks, with NoV infection detected $\leq$2 days after admission.

3.1. Sequence analysis

RNA fragments were reverse transcribed with random hexamers (Invitrogen), yielding cDNA that was amplified by a nested PCR and subsequently sequenced using the ABI-PrismBigDye Terminator v3.0 Ready Reaction Cycle kit. The same primers were used for amplification and sequencing the P2 domain (primers 1st PCR: F: 5′ gangatgtcttcacagtctctt 3′, R: 5′ cattcctgggggagtagaca 3′,[4] Nested primers: F: 5′ gtgccacccacagttgag 3′, R: 5′ gggagtagacagtccaa 3′). DNA Sequences were entered and assembled in bionumerics 6.6.2 (Applied Maths, Sint-Martens-Latem, Belgium) and evaluated manually for their quality by looking for the number of ambiguities, errors, mismatches and deletions.

Sequences were aligned; genotype and variant assignment was based on the RdRp region[16] using the norovirus typing tool (http://www.rivm.nl/mpf/norovirus/typingtool).[19] Full-length P2 domain sequences (600 nucleotide) were then subjected to pairwise analysis (UPGMA) to identify strains linked to the same outbreak, and by advanced cluster analysis (maximum parsimony), to compare diversity within and between outbreaks and robustness of clustering. Sequence diversity within patient and between patients within an outbreak was assessed by comparing the minimum and maximum number of mismatches for each outbreak separately. Translated sequences were reviewed to look for possible directional amino acid mutations.

3.2. Sampling and virus shedding

To study the effect of sampling delay, the time of onset of shedding was estimated by extrapolation from modeled shedding kinetics, based on data from all subjects with follow up samples (Teunis et al., submitted for publication). An RNA standard template was used to translate CT values in fecal samples into an estimated viral load. PCR based estimates of NoV shedding were then used to calculate total numbers of viruses shed by sampled subjects, allowing analysis of sequence variation against viral load, clinical symptoms (symptomatic or asymptomatic), and occupational status (HCW/patients). To characterize the rate of sequence variation all data were pooled and the survivor function for sequence change was calculated, using a Kaplan–Meier estimator[20] describing the probability of any nucleotide changes versus time from onset of shedding.

4. Results

4.1. Strain typing and clustering

The four GII.4 outbreaks occurred in 3 nursing homes and 1 university hospital. The ages of the included HCW/patients from the hospital setting ranged from 25 to 77 and 54 to 77 years, respectively. In the nursing homes, ages for the residents were high (72–95 years), while for the HCW this ranged from 20 to 63 years. Details of the outbreaks have been described elsewhere (15). In total 175 HCW and 77 patients consented to enhanced case finding, of which 50 HCW and 47 patients tested positive for NoV infection (Table 1). Capsid gene sequencing was successful in 109 NoV positive stool samples from a total of 252 sampled cases, comprising 44 HCW and 37 patients.[16] OB 4 yielded 48 sampled cases but the data is not published yet. Failed sequences were repeatedly tested but persistently failed to produce sequence information. The success of sequencing was unrelated to the levels of virus shedding (data not shown).

Phylogenetic analysis of all P2 domain sequences showed a clear discrimination of the four GII.4 outbreak clusters, but with mixed results for OB 1: here, the outbreak strains segregated into three different clusters: GII.4 2008 (17 cases) and 2 clusters belonging to the GII.4 2006b variant lineage (2 cases each) (Fig. 1). As this suggests that a minority of the patients was from a different, unrelated cluster, detailed molecular analysis was not performed for these strains. Data retrieved from the hospital database showed that one of the samples belongs to a nurse who was involved in patient interviews. The other three subjects were patients who had been admitted into the hospital. From the epidemiological data, it was clear that one patient developed diarrhea after admission, indicating a nosocomial infection.

Comparison of these results against the strain diversity observed in the background dataset (defined as sequences from patients admitted with norovirus infection) showed that these were unique and distinct from the outbreak sequences with few exceptions (4%) (Fig. 1). In OB 4 a unique single case was observed who showed