Molecular epidemiology of Norovirus strains circulating in Ireland from 2003 to 2004

A. WATERS1*, S. COUGHLAN2, L. DUNFORD2 AND W. W. HALL1,2
1 Centre for Research in Infectious Diseases (CRID), University College Dublin, Belfield, Dublin 4, Ireland
2 National Virus Reference Laboratory (NVRL), University College Dublin, Belfield, Dublin 4, Ireland

(Accepted 11 January 2006, first published online 29 March 2006)

SUMMARY
Since 2002, the burden of norovirus (NoV) infection in Ireland has increased. Outbreaks in institutional settings are the most common causing widespread disruption to health service delivery. This is the first national study of NoV in the Republic of Ireland and its aim was to identify the major NoV strains circulating in Ireland over a 13-month period between November 2003 and November 2004, inclusive. A prospective study screened faecal samples (n = 478) for NoV RNA. Positive samples (n = 116) were further analysed by a second PCR, targeted to the orf1/orf2 junction of the virus. Phylogenetic analysis was based on sequence alignments of this domain. GII/4 viruses represented 92.2% of sequences, 2.7% were GII/2, GII/3 and GGIIB cluster-like strains. The remaining 5.2% were of GI origin. NoV was detectable throughout the study period, although two peaks of infection were observed. The majority of infections were caused by a range of closely related GII/4 NoV strains.

INTRODUCTION
Norovirus (NoV) is recognized as a worldwide cause of acute epidemic non-bacterial gastroenteritis. Infection is characterized by severe projectile vomiting, usually accompanied by profuse diarrhoea [1]. NoV infection is common in all age groups and is known for its low infectious dose and high attack rate [2]. A distinct winter/spring seasonal peak of infectious outbreaks is associated with the virus [3]. Transmission can be waterborne or foodborne, although person-to-person transmission through direct contact and aerosols is mainly responsible for large outbreaks [4, 5].

NoV are members of the family Caliciviridae [6] and are single-stranded positive-sense RNA viruses. The viral genome is divided into three open reading frames (orf) encoding the non-structural proteins including the viral RNA-dependent RNA polymerase (RdRp) (orf1), the capsid protein (orf2) and a small basic protein (orf3) [7]. They are divided into three distinct genogroups, two of which, genogroup I (GI) and genogroup II (GII), are pathogenic to humans [8]. GI includes the prototype Norwalk virus, Desert Shield virus and Southampton virus and GII includes Lordsdale virus, Grimsby virus, Snowy Mountain virus and Hawaii virus [7, 9–11]. Each genogroup may be divided into several genetic clusters which share approximately 75% nucleotide sequence identity within a genogroup and <60% nucleotide similarity between genogroups [12]. This genetic diversity is a direct result of both the error-prone viral RdRp and strain recombination [13].

Genogroup II/genotype 4 (GII/4) virus strains or Lordsdale-like strains are recognized as the most predominant NoV strains globally [14]. The United
Kingdom recently characterized 21 outbreaks during peak infection in January 2002, all of which were attributable to Lordsdale-like viruses [15]. Analysis of GII sequences, isolated from an Irish hospital, also confirms the high prevalence of GII/4 strains, especially within the health-care setting [16]. In 2002, across Europe, a striking increase in the number of outbreaks coincided with the emergence of a ‘new variant’ GII/4-2002 strain, which contained two consistent nucleotide point mutations in the RdRp gene [17]. Subsequent outbreaks in 2004 at an international scout jamboree in The Netherlands were caused by strains that were more similar in this region to those strains dated pre-2002 ([18], Foodborne Virus Network, FBVN, personal communication).

‘Winter vomiting disease’ has emerged as the principal cause of gastroenteritis outbreaks in Ireland, with the number of outbreaks attributed to NoV increasing from 37 (out of 64) in 2001 to 154 (out of 188) in 2002 [19]. This dramatic increase resulted in closure of wards, widespread disruption to health-care services and growing economic costs to the Department of Health and Children. In 2002, 82% of all outbreaks occurred within the health-care setting [20]. Diagnostic testing, in the Republic of Ireland, is largely the responsibility of the National Virus Reference Laboratory (NVRL). The NVRL investigates the majority of cases of suspected viral gastroenteritis and as such, receives samples from public health, general practitioners, créches, schools, hotels, hospitals and other institutions. The NVRL is solely responsible for the molecular diagnosis of NoV. The number of faecal samples, submitted for testing, from patients presenting with symptoms of viral gastroenteritis, increased by 600% from 2001 to 2002, and in March 2003 reverse transcriptase–polymerase chain reaction (RT–PCR) was implemented as the main diagnostic method for NoV infection [21].

Community-based infections exhibit much larger genetic diversity [22] whereas GII/4 strains are endemic in hospitals and long-term care facilities [23, 24]. Thus, the characterization of outbreaks alone can result in an underestimation of the true number of strains circulating at a particular time. Previous publications detailing sporadic infection in an Irish hospital included Grimsby-like, Southampton-like and Mexico-like NoV strains [16].

In light of recent changes in relation to the burden of disease caused by NoV, molecular characterization of all Irish noroviral strains is crucial. The objectives of this study were, for the first time, to isolate a representative sample of all circulating NoV strains in Ireland and elucidate, using phylogenetics and sequential alignments, the genetic diversity existing within these strains. This prospective study was designed to include representative samples from both NoV outbreaks and those isolated from sporadic community-based infection in order to report on the variability of strains originating from all settings. Molecular characterization of Irish NoV strains has never been performed previously on a national level and the results generated may contribute to a better understanding of the evolution of infection in the Republic of Ireland.

METHODS

Study design and sampling method

A prospective study was carried out between November 2003 and November 2004 on anonymized faecal samples submitted to the NVRL for viral gastroenteritis investigation, following initial clinical diagnosis. The national guidelines for the management of outbreaks of NoV infection in healthcare settings, outlined by the Health Protection Surveillance Centre (HPSC), state if two out of six samples are positive, an outbreak should be confirmed and reported [20]. As such, every eighth faecal sample from patients suffering from the sudden onset of vomiting and diarrhoea, received by the NVRL, were chosen for participation in this study. This was done in order to considerably reduce the possibility of repeatedly isolating viral strains from the same outbreak. The initial diagnostic result of the NVRL was unknown at the time of sampling.

Specimen collection

A total of 3824 faecal samples were collected by the NVRL during the study period, of which 478 were analysed. Limited demographic data was included with each sample to limit traceability, with only the age of the patient and source of sample accompanying each specimen and as such, full epidemiological analysis of each outbreak of infection could not be carried out. In addition, 10 sequences from 2002 were also included in the phylogenetic analysis and each was chosen from 10 separate outbreaks and were used as a retrospective ‘snapshot’ of strains circulating in Ireland in 2002. No demographic details accompanied these samples. Approximately 2–4 GII/4 strains per
month, were chosen for characterization of the ‘new variant’ region located in the gene encoding the RdRp. A total of 22 specimens were included in total.

RNA extraction and screening RT–PCR

Approximate 10% faecal suspensions were prepared in MEM (Minimum Essential Medium, Life Technologies, Paisley, Scotland) and stored at 4 °C. RNA was extracted using the Boom extraction method and stored at −70 °C [25]. Reverse transcription was carried out on 20 ml extracted nucleic acid, using random primers (Promega, Southampton, UK) and Moloney Murine Leukaemia Virus (MMLV)–reverse transcriptase (Promega) according to the protocol outlined by Maguire et al. [26]. All cDNA was stored at −40 °C. A screening PCR, using the modified JV12Y and JV13I primer set, was carried out on all cDNA [27, 28].

Amplification of the capsid non-structural (N/S) domain

Genogroup specific RT–PCR amplified the orf1/orf2 junction or the capsid N/S domain of the virus. The PCR reaction was carried out under the following conditions; 200 μM each dNTP (Roche Diagnostics, Lewes, UK), 1·5 mM MgCl₂, 500 μM each primer, 2·5 U Taq polymerase (Promega), 50 mM KCl, 10 mM Tris–HCl and 0·1% Triton X-100. Amplification was as follows; denaturation at 95 °C for 5 min, followed by 40 cycles of 94 °C for 30 s; annealing of primers was at 57 °C and 55 °C for 30 s for the GI primer (GIICapN/S F GIICapN/S R) and the GII primer sets (GIICapN/S F/GII R) respectively; elongation at 72 °C for 1 min followed by a final step of 72 °C for 10 min. The capsid N/S primers are listed in Table 1. Samples requiring further sequencing analysis were amplified as above, using the GGIIBpol F/GII R primer set. Annealing temperature was 52 °C.

Nucleotide sequencing

Amplicons (50 μl for each sample) were purified using the High Pure PCR product purification kit (Roche Diagnostics). Sequencing was carried out according to the manufacturer’s protocol using an ABI Prism 3.10 automated genetic analyser and the BigDye terminator V1.1 cycle sequencing kit (Applied Biosystems, Warrington, UK). Chromatograms were analysed using Chromas software, version 2.24 (http://www.technelysium.com.au/chromas.html).

Phylogenetic analysis

Sequences were aligned using Clustal W (http://www.ebi.ac.uk/clustalw) and DAMBE (http://aix1.uottawa.ca/~xxia/software/software.htm) software. Modeltest (http://bioag.byu.edu/zoology/crandall_lab/modeltest.htm) was used to compare the likelihood of the different parameters for all models of substitution and select the ‘best-fit’ model for the aligned dataset. All identical sequences were grouped together to reduce the computational time of the phylogenetic analysis and relabelled with an alphabetic letter (B–M). Maximum likelihood trees were constructed, with PAUP* version 4.0 using the HKY model of substitution based on a gamma distribution and equal base frequencies. Bootstrap resampling was carried out for 1000 replicates of the dataset using PAUP*. Values of <70% are not shown.

Accession numbers. The Genbank nucleotide sequence accession numbers for all the Irish phylogenetic sequences are: DQ157930-DQ157956.

RESULTS

Faecal samples were prospectively analysed throughout the study period. Of the total of 478 samples collected for this study, 116 were identified as positive for NoV RNA. The screening PCR detected samples...
of both GI and GII origin. Faecal samples were received from all regions of the Republic of Ireland, with 24.3% testing positive for NoV RNA. The median age of patients confirmed positive for NoV infection was 75.5 years (range 6 months to 95 years). Specimens received from health-care institutions, including hospitals (98/116) and long-term care facilities (12/116), contributed to the vast majority of all positives identified (95%) (Table 2).

NoV infection was detectable throughout the study period, although a noticeable seasonal distribution was identified, with two distinct outbreak peaks of infection observed. The largest number of positive samples was identified between February and April 2004 (49%, 57/116) and a less significant autumn/winter peak was evident between October and November 2004 (24%, 28/116) (Fig. 1). This peak of infection has continued until April 2005 (data not shown). This pattern of infection was reflected in the national epidemiology figures for viral gastroenteritis in Ireland in 2004 [19].

Phylogenetic analysis was carried out on nucleotide alignments of the capsid N/S domain. Sequenced amplicons were 289 nucleotides in length. A total of 144 sequences were used to compile the phylogram; 114 sequences identified during the study period, 10 from 2002 and 20 standard strain sequences (obtained from Genbank) (Fig. 2). GI sequences represented 5.2% of all strains identified and the remaining 94.8% of all positive samples were GII NoV, of which 92.2% branched as GII/4 strains. A complete classification of all the Irish strains identified is shown in Table 3.

GII/4 sequences were the most common strain type identified. Identical sequences were grouped together and relabelled with an alphabetic letter (Table 3). Of all the Irish GII/4 sequences, 22.4% of sequences analysed clustered as group H strains and this nucleotide sequence exhibited 100% identity with the Farmington Hills strain. The second most common set of sequences group B represented 18.7% of the GII/4 strains sequences and showed 98% identity with the Oxford strain [5]. Group B and group H sequences were only identified as circulating between December 2003 and July 2004 and thus represent the main strains responsible for the large number of outbreaks during the springtime peak of infection.

Group J represented 13.1% of all GII/4 sequences and is seen to be largely responsible for the second autumnal peak of infection.

Of particular interest is the GII/4 group F. This consisted of three GII/4 sequences, all of which contained two consistent nucleotide mutations, never previously identified in any published GII/4 strain. Sequence analysis of the capsid gene revealed a novel silent mutation (AAT to AAC) and newly identified non-synonymous mutation in the polymerase gene. Lopman et al. recently outlined two significant silent mutations (AACCTTG to AATCTG) in the polymerase

Table 2. Demographic data received for each sample positive for NoV RNA. This is the first national study and included patient samples from all four provinces within the Republic of Ireland. Setting of infection was divided into three categories (hospital, long-term care facility and community-based/sporadic). The largest number of samples were received from Dublin hospitals, reflecting the larger population size and the higher number of hospitals in situ
gene of GII/4 noroviral strains, called the ‘new variant’ or GII/4-2002 mutations [17]. The polymerase mutation, in this Irish sequence, is found in the same sequence motif changing the asparagine residue (AAT/AAC) to a serine (AGC).

Twenty-two Irish GII/4 polymerase sequences were aligned with respect to the 6-bp polymerase gene motif and a distinct pattern of infection was observed (Fig. 3). This study revealed nine ‘new variant’ strains were circulating between November 2003 and June
2004. Only one ‘new variant’ strain (Irl04N298) was identified between April 2004 and June 2004, and during this period the emergence of the ‘GII/4-2004’ strain was observed. Ten ‘GII/4-2004’ sequences were identified in the subset analysed and were circulating between April 2004 and November 2004. This is the first account molecularly characterizing the GII/4 strain population in Ireland since these mutations were first described.

The remaining GII sequences represented 3 out of 116 positive specimens. One sequence clustered most closely with the MX virus (GII/3) standard strain with a bootstrap value of 81%. The two other GII sequences clustered with the Oberhausen & Satima [29] standard noroviral strains (GII/3 and GII/2 respectively). Further upstream sequence analysis revealed all three strains to belong to the newly identified GGIIB recombinant cluster of NoV, commonly associated with paediatric cases of viral gastroenteritis. These sequences were all received from different locations throughout the country over a period of 4 months (November 2003–March 2004). Two sequences were isolated from confirmed community-based infection. Only one of the sequences was isolated from a person <5 years [30]. This is the first time the presence of the GGIIB cluster in Ireland has been described and all three sequences isolated were different. Analysis of the phylogenetic tree (Fig. 2) potentially demonstrates that the GGIIB sequences isolated are either recombinants of Mexico-like or Oberhausen-like viral genetic clusters. All three sequences exhibited 98% identity with the French GGIIB Vannes strain (AY773210).

All six GI sequences clustered most closely with the Southampton virus (GI/2) [10], with a bootstrap value of 99% (Fig. 2, Table 3). Group N consists of three sequences all received from the same hospital. The remaining three GI strains were identified from community-based cases of NoV. Only four GI sequences were included in the phylogenetic analysis, as two samples would not amplify with the GICapN/S F/GICapN/S R primer set.

A retrospective subset of 10 positive samples from 2002 was included in the phylogenetic analysis and all 2002 sequences were labelled as Irl02, followed by the sequence number. No additional demographic data

---

**Table 3. Molecular characterization of NoV strains circulating in Ireland throughout the study period.** All identical sequences were grouped together and relabelled with an alphabetic letter. The GII/4 sequences represent 97.3% of all GII sequences detected and consisted of 11 lineages, defined as B-M. Four different genotypes were detected overall, GI/2, GII/2, GII/3 and GII/4. Lineage group H is identical to the Farmington Hills standard strain.

<table>
<thead>
<tr>
<th>Geno-group</th>
<th>GI (n=6)</th>
<th>GII (n=110)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5.2%</td>
<td>94.8%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Genotype</th>
<th>GII 2</th>
<th>GII 3</th>
<th>GII 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100%</td>
<td>1-7%</td>
<td>0-9%</td>
</tr>
<tr>
<td></td>
<td>97.3%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lineage</th>
<th>Putative GGIIB</th>
<th>Putative GGIIB</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>18-7%</td>
<td>2-8%</td>
</tr>
</tbody>
</table>

* Newly identified sequence.
accompanied these samples. 50% of sequences clustered as GII/4 strains and 50% of sequences (Irl02GroupA) clustered as Oberhausen-like strains (GII/3). This is in contrast to those isolated over the study period, where 92% of all sequences clustered as GII/4 strains.

DISCUSSION

This is the first report detailing the molecular characterization of NoV in Ireland on a national scale. The NVRL receives samples from all settings of infection and as such, samples originating from both outbreak and community-based cases were included in order to give a comprehensive assessment of the burden of norovirus disease in Ireland. Most samples received for analysis were derived from the hospital setting, reflecting the epidemiology of infection in Ireland, where in 2002, 93% of outbreaks and in 2004, 82% of outbreaks occurred within the health-care setting. This situation is very similar to that documented for the United Kingdom with 93% of all outbreaks in 2000 reported to have occurred in residential homes and hospitals [24, 31, 32]. The high prevalence of GII/4 strains in Ireland is, however, not confined to institutional outbreaks, as many general practitioners send their samples to the NVRL via their local hospital. Identifying the exact source of the specimen is, therefore, a very difficult task and a limitation of the study. As previously documented, for other countries, the majority of people affected by NoV were elderly [33, 34] and Ireland’s aging population is likely to exacerbate this situation into the future.

Seasonal distribution of NoV is well documented and the pattern of infection in the United Kingdom has demonstrated outbreaks starting as early as October, continuing through the winter months and finishing in late spring [3]. The seasonal variance of the virus was very apparent in the present study with two peaks of large numbers of outbreaks occurring in both spring and early winter. The virus was detectable all year round, illustrating how the ‘winter vomiting disease’ is a misnomer. NoV became a notifiable disease in January 2004 and the HPSC are now monitoring all outbreaks through the national notifiable disease in January 2004 and the HPSC are now monitoring all outbreaks through the national guidelines [20].

Lordsdale-like viruses accounted for the majority of infections and represented 92.2% of all isolates. GII/4 predominance is well documented and is endemic in many areas such as the United Kingdom, Spain, The Netherlands, New Zealand, continental Europe and the United States [14, 15, 17, 23, 33, 35]. Previous publications have reported the dominance of Grimsby-like and Bristol-like GII/4 strains in Ireland [16], although this is the first account detailing both the presence and high prevalence of the Farmington Hills and Oxford-like strains in Ireland [31].

In 2002, atypical summertime outbreak peaks were recorded for many European countries including The Netherlands and the United Kingdom and coincided with the emergence of a ‘new variant’ mutant GII/4-2002 strain [17]. In 2004 a large outbreak at an international jamboree was caused by several strains, including a GII/4 strain, which did not contain the 2002 mutations. This was named the GII/4-2004 strain. The earliest detected outbreak of the GII/4-2002 sequences in Ireland was in February 2002 [21]. Our study suggests that the two point mutations have only recently become less common within the Irish sample population, as sequences containing these mutations were detected as late as June 2004. After April 2004 the ‘GII/4-2004’ strain mutations became dominant in all subsequent GII/4 isolates and may be a factor responsible for the second peak of NoV infection in the latter months of the study period.

Of particular interest is a non-synonymous mutation, in an Irish GII/4 sequence, located in the same two-codon region of polymerase gene that resulted in the amino-acid change of an asparagine residue to a serine residue. The significance of this mutation is not clear. Although both serine and asparagine contain the same functional group and enzyme activity is probably little affected, we did not investigate this possibility any further. This sequence also contained a previously undetected silent nucleotide mutation in the capsid gene. BLAST searches on the NCBI network revealed no similar strains with these mutations. Full sequence analysis may confirm whether this represents a new GII/4 strain, or merely reflects NoV tendency to use antigenic drift as a mechanism of immune evasion.

Phylogenetic analysis was based on orf1/orf2 junction or capsid N/S alignments. Small amplicons of the polymerase region are of limited value in phylogeny studies, especially within the health-care setting, as over short periods of time many sequences are identical [26]. Extensive phylogenetic analysis and pairwise similarity tests by Katayama et al. [13] concluded that the capsid N/S domain is most capable of distinguishing between genetic clusters. In the present study, we designed genogroup specific primers to amplify the capsid N/S domain, for use in the
phylogenetic analysis of the newly identified Irish sequences. One GI cluster and three GII clusters of NoV were identified and were defined as Southampton-like (GI/2), Saitama-like (GII/2), MX-like (GII/3) and Lordsdale-like (GII/4). Further sequence analysis led to the first report of the GGIIB recombinant cluster in Ireland. This strain has been described in outbreaks in Sweden, especially within the paediatric population [30, 36]. Isolation of this sequence, within our sample population was not confined to paediatric patients.

NoV is clearly shown as an extremely important pathogen in Ireland, causing widespread disruption to public health. Although the sample population included community-based and institutional outbreak infection, the overwhelming predominance of GII/4 sequences very much reflects the growing burden of noroviral disease in Ireland within the health-care system. This situation is distinct from other countries, such as the United States where non-health-care outbreaks, mainly foodborne, are monitored more carefully and play a more significant role in the epidemiology of the virus [33]. Combining this preliminary national study with recorded outbreak and geographical epidemiological spread data for NoV will generate a comprehensive picture of the contribution of NoV to infection throughout Ireland. Increased communication between reference laboratories, hospitals and public health teams is imperative in order to fully understand and reduce NoV infection.

ACKNOWLEDGEMENTS

The authors acknowledge the contributions of colleagues at the Health Protection Surveillance Centre and in public health and infection control agencies in Ireland and the contribution of samples from all regional hospitals. Thanks are also due to Dr Jim Gray of the ENVRL, Health Protection Agency, UK, and Dr Jeff Connell, Principal Clinical Scientist, NVRL, Republic of Ireland, for their advice on this manuscript.

DECLARATION OF INTEREST

None.

REFERENCES

Molecular epidemiology of Norovirus strains circulating in Ireland


