Available online at www.sciencedirect.com

SCIENCE DIRECT

Virology 346 (2006) 312 – 323

VIROLOGY

www.elsevier.com/locate/yviro

Norovirus classification and proposed strain nomenclature

Du-Ping Zheng [a,b,c,*], Tamie Ando [a], Rebecca L. Fankhauser [a,b], R. Suzanne Beard [a], Roger I. Glass [a], Stephan S. Monroe [a]

[a] *Respiratory and Enteric Viruses Branch, Division of Viral and Rickettsial Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, GA 30333, USA*

[b] *Atlanta Research and Education Foundation, Decatur, GA 30033, USA*

[c] *Department of Pediatrics, Emory University School of Medicine, Atlanta, GA 30322, USA*

Received 14 June 2005; returned to author for revision 21 September 2005; accepted 12 November 2005
Available online 15 December 2005

Abstract

Without a virus culture system, genetic analysis becomes the principal method to classify norovirus (NoV) strains. Currently, classification of NoV strains beneath the species level has been based on sequences from different regions of the viral genome. As a result, the phylogenetic insights of some virus were not appropriately interpreted, and no consensus has been reached to establish a uniform classification scheme. To provide a consistent and reliable scientific basis for classifying NoVs, we analyzed the amino acid sequences for the major capsid protein of 164 NoV strains by first using an alignment based on the predicted 3D structures. A Bayesian tree was generated, and the maximum likelihood pairwise distances of the aligned sequences were used to evaluate the results from the uncorrected pairwise distance method. Analyses of the pairwise distances demonstrated three clearly resolved peaks, suggesting that NoV strains beneath the species level can be classified at three levels: strain (S), cluster (C), and genogroup (G). The uncorrected pairwise distance ranges for S, C, and G were 0–14.1%, 14.3–43.8%, and 44.9–61.4%, respectively. A scheme with 29 genetic clusters [8 in genogroup 1 (G1), 17 in G2, 2 in G3, and 1 each in G4 and G5] was defined on the basis of the tree topology with the standards provided and was supported by the distance analysis. Of these, five clusters in G2 and one in G1 are newly described. This analysis can serve as the basis for a standardized nomenclature to genetically describe NoV strains.

Keywords: Noroviruses; Norwalk-like viruses; Classification; Nomenclature; Phylogenetic analysis

Gastroenteritis is a major public health problem worldwide. With the application of new and sensitive diagnostic techniques, noroviruses (NoVs) are now recognized as the leading cause of outbreaks and cases of nonbacterial, acute gastroenteritis in humans. Up to 93% of these outbreaks and 60–85% of all gastroenteritis outbreaks within the United States, Europe, and Japan are associated with NoVs (Fankhauser et al., 2002; Inouye et al., 2000; Lopman et al., 2003). From 1997 to 2000, 233 outbreaks of gastroenteritis were reported to the Centers for Disease Control and Prevention (CDC); 86% (N = 201) were laboratory-confirmed to be associated with NoV (Fankhauser et al., 2002). NoV outbreaks occur in many settings, such as nursing homes, restaurants, schools, hospitals, and cruise ships (Fankhauser et al., 2002; Glass et al., 2000; Green et al., 2002; Widdowson et al., 2004). The viruses are very contagious and infect persons of all ages who consume contaminated food or water or who have close contact with infected persons (Rockx et al., 2002). Volunteer studies indicate that immunity to NoVs seems to be short lived and that immunity to one strain does not provide good protection from infection with other heterogeneous strains (Johnson et al., 1990; Matsui and Greenberg, 2000; Parrino et al., 1977; Wyatt et al., 1974). Observations of repeat infections in adults also suggest that long-term immunity may be absent (Fankhauser et al., 2002; Rockx et al., 2002). Recent studies indicate that susceptibility to NoV infection is associated with ABH histo-blood types, gut-expressed carbohydrates, and strain preferences (Harrington et al., 2004; Hutson et al., 2004; Meyer et al., 2004; Tan et al., 2004). No vaccine is currently available to prevent NoV disease in humans (Estes et al., 2000).

* Corresponding author. Respiratory and Enteric Viruses Branch, Division of Rickettsial Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, 1600 Clifton Rd., MS G04, Atlanta, GA 30333, USA. Fax: +1 404 639 3645.

E-mail address: dzheng@cdc.gov (D.-P. Zheng).

0042-6822/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2005.11.015

Noroviruses (formerly called "Norwalk-like viruses"), which were discovered in 1972 (Dolin et al., 1972; Kapikian et al., 1972), belong to the genus *Norovirus* in the family Caliciviridae (Green et al., 2000a; van Regenmortel et al., 2000). The viral genome is a plus-sense, single-stranded RNA of ~7.5 kb that contains three open reading frames (ORFs). ORF1 encodes the nonstructural polyprotein that is cleaved by viral 3C-like protease into probably 6 proteins, including the deduced RNA-dependent RNA polymerase [RdRp] (Belliot et al., 2003). ORF2 and ORF3 encode the major (VP1) and minor (VP2) capsid proteins, respectively (Green et al., 2001; Jiang et al., 1990, 1993). The VP1 protein forms two domains: P (protruding, P1 and P2) and S (shell). Most of the cellular interactions and immune recognition features are thought to be located in the P2 sub-domain, which extends above the viral surface and has the most sequence divergence in the genome (Chakravarty et al., 2005; Nilsson et al., 2003; Prasad et al., 1999; Tan et al., 2004). Therefore, it is believed that the capsid protein not only provides shell structure for the virus but also contains cellular receptor binding site(s) and viral phenotype or serotype determinants. The function of VP2 associates with up-regulation of VP1 expression in *cis* and stabilization of VP1 in the virus structure (Bertolotti-Ciarlet et al., 2003).

NoVs are genetically and antigenically diverse (Ando et al., 2000; Green et al., 1995, 1997, 2000b; Katayama et al., 2002). Historically, classification of NoVs was based on cross-challenge studies in volunteers and cross-reactivity analysis by immune electron microscopy (Green et al., 1995; Lewis et al., 1995; Okada et al., 1990; Wyatt et al., 1974). These antigenic classification schemes had poor accuracy and reproducibility, which were attributed to the cross-reactivity of antibodies (Ando et al., 2000). Direct serotyping based on neutralization is not possible because no cell culture system has been established for growing this virus (Duizer et al., 2004). Consequently, reverse transcription-polymerase chain reaction (RT-PCR) and genomic sequencing have become the major means for characterizing the viruses and to understand the relatedness of different strains (Ando et al., 2000; Jiang et al., 1990, 1993; Katayama et al., 2002). Accumulated sequence information has been especially useful for viral diagnosis and genotyping, and most investigators have chosen primers from conserved regions, such as the RdRp gene, in order to detect the greatest number of these diverse strains. The targeted regions are called region A (the RdRp gene located in ORF1), region B (the 3′-end of ORF1), region C (a short stretch close to the 5′-end of ORF2), and region D (located at the 3′-end of ORF2) (Ando et al., 2000; Vinjé et al., 2004). The region D primers were found to work well for detection of GI and GII viruses and to be suitable for strain differentiation; however, no evaluation has yet been based on the complete capsid sequences, and no clear criteria have been created for NoV classification. Short conserved sequences have been used successfully for NoV detection, especially when applied to multiplex PCR or real-time PCR (Kageyama et al., 2003; Richards et al., 2004; Yan et al., 2003), but such sequences become problematic for phylogenetic analyses because, with limited sequence variations, some strains cannot be distinguished from each other or be accurately classified into their proper clusters (Fankhauser et al., 2002; Kageyama et al., 2004).

Five genogroups (G) of NoVs have been tentatively assigned from the molecular characterization of complete capsid gene sequences (Fankhauser et al., 2002; Green et al., 1995; Karst et al., 2003; Oliver et al., 2003; Vinjé and Koopmans, 2000). Strains of three genogroups, GI, GII, and GIV, are found in humans (GII/11 are porcine), and GIII and GV strains are found in cows and mice, respectively. However, no consensus has been reached concerning the classification of NoV stains within genogroups. Few studies have examined complete capsid sequences for phylogenetic studies, and none has included sequences from all five genogroups. An early study of 35 NoV capsid sequences identified 7 sub-genogroups within GI and 8 sub-genogroups within GII (Green et al., 2000b). A later study of 39 capsid sequences from human NoVs classified these strains into 23 different genetic clusters (7 in GI, 15 in GII, and 1 in GIV) (Vinjé et al., 2004). Nevertheless, 14 genotypes in GI and 17 genotypes in GII, based on partial capsid and RdRp sequences (Kageyama et al., 2004), and 7 genetic clusters in GI, 8 plus 3 unassigned clusters in GII, and 1 cluster each in GIII and GIV, based only on partial capsid sequences of region C (Fankhauser et al., 2002), have been described.

To avoid confusion and to provide clear criteria for classification of NoVs and a consistent reliable basis for NoV nomenclature, we analyzed 164 deduced amino acid (AA) sequences of the NoV major capsid protein, including all sequences from five genogroups available at the time of the study. A well-defined phylogenetic scheme of NoVs was established, and strain clustering was evaluated by multiple methods of distance calculation. From our results, we propose an updated classification scheme for standardization of NoV nomenclature, which should aid molecular characterization and description of outbreak strains and provide genetic insights for future studies with this virus.

Results

Sequences and structure-based alignment

Of the 273 sequences obtained from GenBank by use of a BLAST search, only 145 met our criteria for study inclusion; an additional 19 new sequences were added from the CDC collection (Table 1). From the preliminary phylogenetic analysis of the 164 sequences (data not shown), 43 were chosen (including 1 or 2 from each of the major branches) to represent the diversity of NoV strains in this collection and used to construct an initial 3D structural alignment. The smaller number of the 43 sequences was aligned spatially first to avoid complexity during the structural modeling caused by limitations of computer memory and software handling. This structure-based alignment then served as a frame for aligning all of the 164 sequences. Overall, we created three structure-based alignments: (1) the 164-sequence alignment, 164-Cap, which included all available complete capsid sequences; (2)

Table 1
NoV sequences used for this study

G_Cluster	Name	Source	G_Cluster	Name	Source	G_Cluster	Name	Source
G1_1	**NV-USA93**	M87661	G2_3	SU18-JPN	AB039781	G2_6	SU3-JPN	AB039776
G1_1	Wtchest-USA	AY502016	G2_3	MG312-USA	AF414413	G2_6	SU4-JPN	AB039777
G1_1	KY89-JPN	L23828	G2_3	BB321-USA	AF414415	G2_6	SU17-JPN	AB039779
G1_1	Aich124-JPN	AB031013	G2_3	Mexico1-MEX	U22498	G2_6	BT274-USA	AF414408
G1_2	WR96-GBR	AJ277610	G2_3	OTH25-USA	L23830	G2_6	SU16-JPN	AB039778
G1_2	FB258-JPN	AB078335	G2_3	LV247-USA	AF414411	G2_6	UENO7k-JPN	AB078337
G1_2	**SOV-GBR93**	L07418	G2_4	**Bristol-GBR93**	X76716	G2_6	FL269-USA	AF414407
G1_2	C59-USA	AF435807	G2_4	Lsdale-GBR	X86557	G2_6	**Seacrof-GBR00**	AJ277620
G1_3	PD196-DEU	AF439267	G2_4	VA98387-USA	AY038600	G2_6	Miami292-USA	AF414410
G1_3	LR316-USA	AF414405	G2_4	CBW94-AUS	AF145896	G2_7	GN273-USA	AF414409
G1_3	**DSV-USA93**	U04469	G2_4	MD145-USA	AY032605	G2_7	**Leeds-GBR00**	AJ277608
G1_3	Stav-NOR99	AF145709	G2_4	DG259-DEU	AF425766	G2_8	SU25-JPN	AB039780
G1_3	HLL219-USA	AF414403	G2_4	NT104-JPN	AB078336	G2_8	**Amstdam-NLD99**	AF195848
G1_3	VA98115-USA	AY038598	G2_4	MD1347-USA	AY030098	G2_9	Idafall-USA	AY054299
G1_4	**Chiba-JPN00**	AB042808	G2_4	FL408-USA	AF080558	G2_9	**VABeach-USA01**	AY038599
G1_4	Koblenz-DEU	AF394960	G2_4	Mora97-SWE	AY081134	G2_10	**Erfurt-DEU01**	AF427118
G1_4	Valetta-MLT	AJ277616	G2_4	AZ379-USA	AF080556	G2_11	**SW918-JPN01**	AB074893
G1_4	BTM277-USA	AF414404	G2_4	FL384-USA	AF080557	G2_11	SWVA34-USA	AY077644
G1_4	Leed92-GBR	AJ313030	G2_4	ID366-USA	AF080554	G2_11	SW43-JPN	AB074892
G1_4	NO266-USA	AF414402	G2_4	DJ171-FRA	AF472623	G2_12	U1GII-JPN	AB067536
G1_5	SzUG1-JPN	AB039774	G2_4	OC112-DEU	AF427113	G2_12	SU1-JPN	AB039775
G1_5	**Musgrov-GBR00**	AJ277614	G2_4	95M14-AUS	AF080551	G2_12	GIFU96-JPN	AB045603
G1_5	AB318-USA	AF414406	G2_4	SC373-USA	AF080555	G2_12	SCH003-DEU	AF397905
G1_6	**Hesse-DEU98**	AF093797	G2_4	AZ364-USA	AF080553	G2_12	Chitta-JPN	AB032758
G1_6	WUG1-JPN	AB081723	G2_4	FL358-USA	AF080552	G2_12	HLL314-USA	AF414420
G1_6	Wiscon-USA	AY502008	G2_4	LA416-USA	AF080559	G2_12	Pirna110-DEU	AF427119
G1_6	CS841-USA	AY502007	G2_4	BL495-DEU	AF427123	G2_12	**Wortley-GBR00**	AJ277618
G1_6	VA497-USA	AF538678	G2_4	SC345-USA	AF080549	G2_13	**Faytvil-USA02**	AY113106
G1_7	**Wnchest-GBR00**	AJ277609	G2_4	AK140-DEU	AF425765	G2_13	KSW47-JPN	AB078334
G1_8	**Boxer-USA02**	AF538679	G2_4	BK124-DEU	AF427120	G2_14	**M7-USA03**	AY130761
G2_1	WO302-USA	AF414418	G2_4	BL159-DEU	AF425763	G2_15	Mex7076-USA	AF542090
G2_1	**Hawaii-USA94**	U07611	G2_4	LS218-DEU	AF427115	G2_15	**J23-USA02**	AY130762
G2_1	Miami81-USA	AF414416	G2_4	DG4770-AUS	AF406793	G2_16	Tonto-USA	AY502005
G2_1	DG391-DEU	AF425767	G2_4	EF007-DEU	AF427117	G2_16	Fayett-USA	AY502014
G2_1	RM283-USA	AF414419	G2_4	BL491-DEU	AF427122	G2_16	Fairfd-USA	AY502013
G2_1	PC301-USA	AF414421	G2_4	Oder170-DEU	AF427114	G2_16	Sandsk-USA	AY502012
G2_1	WD294-DEU	AF425769	G2_4	SG95-GBR	AJ277619	G2_16	Canton-USA	AY502011
G2_2	**Msham-GBR95**	X81879	G2_4	SC3452-USA	AF080550	G2_16	**Tiffin-USA03**	AY502010
G2_2	CF434-USA	AY054300	G2_4	BLD331-USA	AF414425	G2_16	Hiram-USA	AY502006
G2_2	SMV1-USA	AY134748	G2_4	LS221-DEU	AF427116	G2_16	Bradhe-USA	AY502015
G2_2	SMV2-USA	U70059	G2_4	MB326-USA	AF414424	G2_17	**CSE1-USA03**	AY502009
G2_3	NLV1157-SWE	AY247439	G2_4	BL238-DEU	AF425764	G3_1	**BoJena-DEU98**	AJ011099
G2_3	NLV2004-SWE	AY247432	G2_4	UK317-GBR	AF414417	G3_2	BoNA2-GBR	AF097917
G2_3	HB385-DEU	AF539439	G2_4	KO130-DEU	AF427121	G3_2	BoCV95-USA	AF542083
G2_3	OH455-DEU	AF539440	G2_4	FMHill-USA	AY502023	G3_2	BoCV186-USA	AF542084
G2_3	MD102-USA	AY030312	G2_4	CSG4-USA	AY502022	G3_2	BoPR55-GBR	AY126476
G2_3	MD1341-USA	AY030313	G2_4	CSG2-USA	AY502021	G3_2	BoDF94-GBR	AY126474
G2_3	AL96-NZL	U46039	G2_4	CSG1-USA	AY502020	G3_2	BoCH131-NLD	AF320113
G2_3	BB289-DEU	AF427112	G2_4	Anchor-USA	AY502019	G3_2	**BoCH126-NLD00**	AF320625
G2_3	**Toronto-CAN93**	U02030	G2_4	CSD1-CAN	AY502018	G3_2	BoAS24-GBR	AY126475
G2_3	ARG320-USA	AF190817	G2_4	Germtn-USA	AY502017	G4_1	**Alphatn-NLD99**	AF195847
G2_3	NO279-USA	AF414412	G2_5	MOH99-HUN	AF397156	G4_1	FLD560-USA	AF414426
G2_3	SU201-JPN	AB039782	G2_5	NO306-USA	AF414422	G4_1	SCD624-USA	AF414427
G2_3	TS313-USA	AF414414	G2_5	**Hilingd-GBR00**	AJ277607	G4_1	**Murine1-USA03**	AY228235
G2_3	RBH93-GBR	AJ277617	G2_5	WR290-USA	AF414423			

the 141-sequence alignment, 141-Cap(d), in which any columns with gaps were removed and 23 sequences in the 164-Cap alignment were no longer unique and were excluded; and (3) the 68-sequence alignment, 68 Cap(d), which was derived from the 141-Cap(d) in which only a prototype and 1–2 reference strain sequences (if available) from each cluster were selected.

Phylogenetic relationships and classification of NoVs

To avoid a bias by presenting results with incomplete data and to reflect the diversity of currently available NoV sequences, we constructed a consensus Bayesian tree, using the 141-Cap(d) alignment that contained all available unique complete capsid sequences (Fig. 1). The tree consists of five

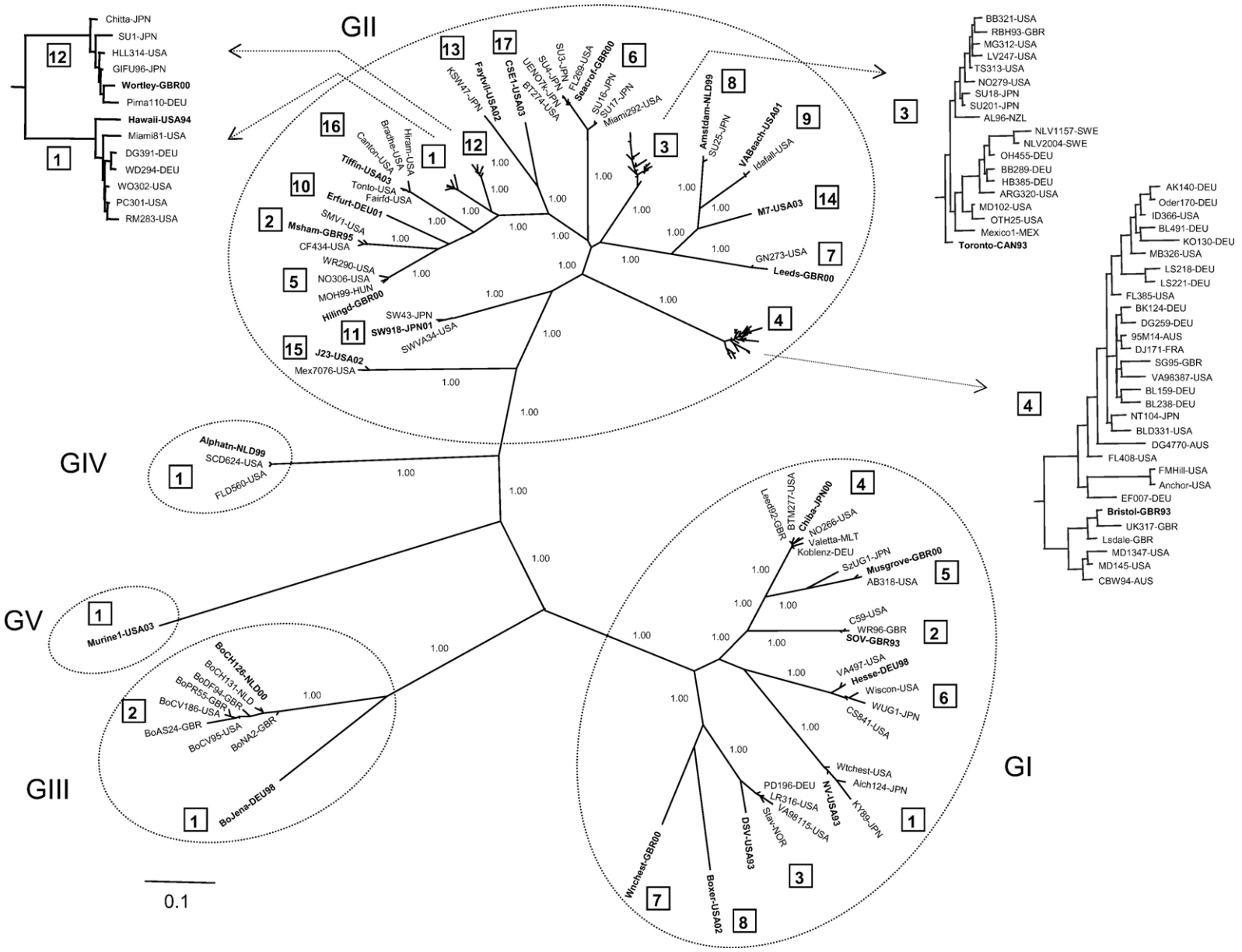

Fig. 1. Phylogenetic analysis of complete capsid AA sequences of 141 NoV strains. The tree was constructed with the structural alignment of 141 NoV sequences (column gaps removed) by using MrBayes program (Huelsenbeck et al., 2001). Tree topology was evaluated on the base of 1,000,000 generations (posterior possibility, 1.00 equals 100% out of 1,000,000). Formats of sequence name: Name–Country code (reference strains); Name–Country code Year (two digits) (prototype strains [in bold]). Number in square box is the cluster number within the genogroup (G). Country codes: AUS, Australia; CAN, Canada; DEU, Germany; FRA, France; HUN, Hungary; JPN, Japan; MLT, Malta; MEX, Mexico; NLD, Netherlands; NZL, New Zealand; NOR, Norway; SWE, Sweden; GBR, the United Kingdom; USA, the United States.

well-separated clades, each containing prototype strain(s) of the genogroups defined in this or earlier studies (Fankhauser et al., 2002; Green et al., 1995; Karst et al., 2003; Oliver et al., 2003; Vinjé and Koopmans, 2000). On the basis of the tree's topology and supported by the posterior probability of 1.00 for all branch partitions, these five clades were designated as genogroups GI–GV (Fig. 1). The two largest clades, GI and GII, include most of the diverse and common NoVs and are segregated into 8 clusters for GI and 17 clusters for GII. Consecutive numbers were assigned to each cluster, using method described previously (Ando et al., 2000) and on the basis of clusters already defined (Fankhauser et al., 2002; Vinjé et al., 2004). Among them, 1 cluster (GI/8) in GI genogroup and 5 clusters (GII/13–17) in GII genogroup were newly described or defined. Three previously unassigned clusters, based upon analysis of partial capsid sequence information, GII/j, GII/m, and GII/n (Fankhauser et al., 2002), were resolved and assigned as GII/13, GII/16, and GII/10, respectively. All sequences in GI and GII were from human viruses, with an exception of those in GII/11 that were from porcine viruses. Another genogroup of human strains, GIV, currently has only a single cluster. Two genogroups of NoVs of bovine (GIII) and murine (GV) origin have 2 and 1 clusters, respectively. All clusters in this classification scheme were well resolved with no intermediate clusters or genogroups found, and they were supported by the robust posterior probability (1.00, similar to bootstrapping 100%). In addition, we observed that tree topologies obtained from the analyses based on structural and nonstructural alignments remained the same (data not shown), and the classification scheme was consistent with results from our earlier work (Ando et al., 2000; Fankhauser et al., 2002).

Pairwise distance distribution of NoV sequences

Pairwise distances between sequences were calculated using either uncorrected or ML distances. To access the results obtained from the analysis using the uncorrected distance method (a common and conservative method) and to evaluate the classification scheme defined above (Fig. 1), we used the more reliable ML distance method for calculation of pairwise distances for 141 unique AA sequences. Numbers in the distance matrix obtained from ML distance analysis (data not shown) were plotted into a histogram that yielded three major peaks (indicated as S, C, and G) corresponding to the distance ranges of strains, clusters, and genogroups, respectively (Fig. 2A). These

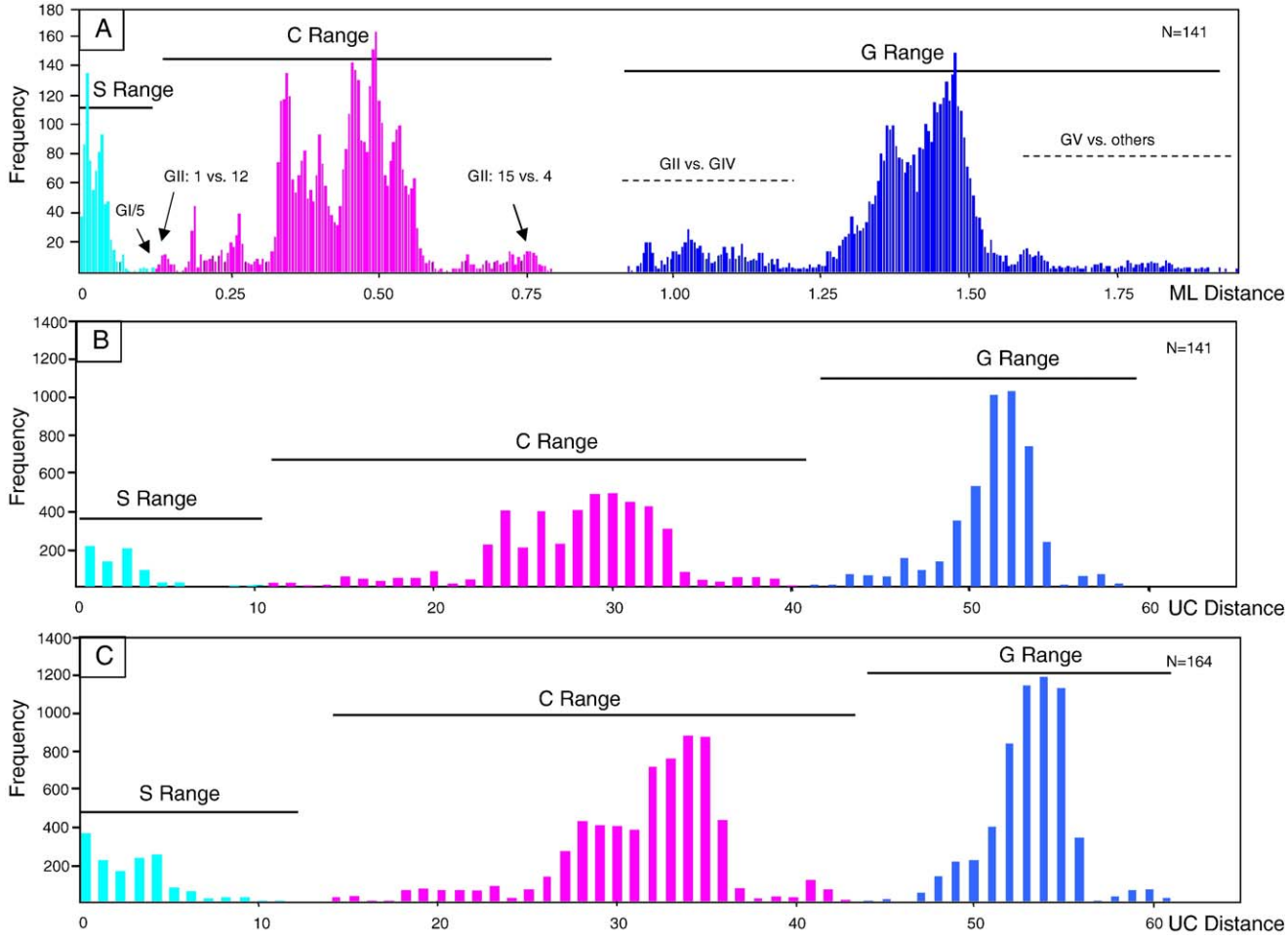

Fig. 2. Pairwise distance distribution of NoV strains. Pairwise distances of strains (S), clusters (C), and genogroups (G) were calculated by using (A) maximum likelihood (ML) method with 141-Cap (d) sequence alignment; (B) uncorrected distance (UC) method with 141-Cap (d) sequence alignment; and (C) uncorrected distance method with 164-Cap alignment. *x* axis: A, ML distances; B and C, percentages; *y* axis: frequency.

three distance ranges did not overlap, suggesting that NoVs can be unambiguously classified into these three distinct categories: strain (S), cluster (C), and genogroup (G). Some minor peaks occurred in the C and G ranges, indicating the diversity of NoV strains. The extension of the major peak of G to the right represents the distances between the GV genogroup (murine NoV) and the other genogroups, and the minor peaks to the left represent the distances between the GII and GIV genogroups. The highest diversity in the C range was between strains in clusters GII/4 and GII/15; the lowest was between strains in clusters GII/1 and GII/12, which was close to the highest distance cutoff for S. In the S range, the most diverse strains were among strains within GI/5 cluster. Distances were then calculated with the uncorrected method and plotted in the same way into the histograms (Fig. 2B, using the same alignment as in Fig. 2A; Fig. 2C, using the 164-sequence alignment with gaps). Both demonstrated three clearly resolved and non-overlapping peaks with very similar distribution patterns as indicated in Fig. 2A. These consistent results further support the classification scheme defined with the Bayesian phylogenetic analysis.

Defining distance ranges of NoV strains within and between clusters and between genogroups

In our earlier study, we suggested definitions for genetic cluster, genogroup and prototype, and reference strains (Ando et al., 2000). Since limited numbers of complete capsid sequences were available at that time, only 39 sequences of NoV strains were analyzed. In this study, 164 capsid sequences of NoV strains representing five genogroups were chosen to investigate the diversity of NoVs and to establish the criteria for classification. On the basis of the NoV distance matrixes (data not shown) and the histogram analyses above, we calculated the pairwise distance ranges of NoV strains within clusters, between clusters, and between genogroups (Tables 2 and 3). The maximum number in any group of pairwise distance ranges did not overlap with the minimum number in the next higher distance range group by any of the methods used. This observation indicates statistically that no intermediate strains between clusters or between genogroups were found in the window of the gene encoding the major structural capsid protein. The results produced by the uncorrected and the ML distance methods with different types of sequence alignments were consistent, suggesting that the pairwise distance ranges defined in Table 2 can be proposed as standards for classification of NoV strains. Differences among genogroups were further characterized (Table 3). The human genogroups had the closest distances: GIV and GII (41.88% and 44.91% in 141 and 164 alignments, respectively); while the murine genogroup, GV, had the greatest distances from the other genogroups, GI–III (58–61% in both alignments). These data all strongly support the classification scheme defined using Bayesian phylogenetic analysis.

Table 2
Pairwise distance ranges among noroviruses compared with 3 different sequence alignments

Sequence alignment	Distance range between		
	Strains (S)	Clusters (C)	Genogroups (G)
141 Cap (d) ML	0–0.1235	0.1299–0.7941	0.9207–1.9474
141 Cap (d) UC	0–10.62	11.04–40.42	41.88–58.54
164 Cap UC	0–14.07	14.26–43.78	44.91–61.41

Cap (d): gaps in the AA sequence alignment were deleted.
ML: maximum likelihood; UC: uncorrected.

Table 3
Pairwise difference ranges between NoV genogroups (%) calculated with 2 different sequence alignments

	G1	G2	G3	G4	G5
G1		48.54–55.21	45.62–52.29	48.75–53.75	55.62–58.33
G2	50.20–57.06		50.21–54.79	**41.88**–48.33	55.21–**58.54**
G3	47.57–53.94	52.07–56.89		52.50–54.38	56.25–58.12
G4	52.11–56.11	**44.91**–51.61	54.88–56.58		56.04
G5	58.14–60.74	58.17–**61.41**	57.74–59.68	58.56–58.75	

Differences were calculated by using the uncorrected distance method with the 141 Capsid (d) AA sequence alignment (top right portion) and with 164 Capsid AA sequence alignment (bottom left portion).

Proposed standards for NoV nomenclature

Based on the phylogenetic relationships and the classification scheme in the tree (Fig. 1), a maximum of 3 strain sequences from each cluster (i.e., 1 prototype strain and up to 2 reference strains) were selected to construct a standard tree for NoV nomenclature (Fig. 3). This figure, including 68 sequences, maintained the topology of the original tree (Fig. 1). All cluster branches in the tree were distinct from each other and were supported with robust posterior probability values. On the basis of the consistency of the topologies in both trees, we propose the classification scheme described in this study as standard nomenclature for NoVs: 29 genetic clusters are classified in the 5 genogroups, 8 in GI, 17 in GII, 2 in G3, and 1 each in G4 and G5. Among these, 5 clusters in GII (GII/13–17) and 1 cluster in GI (GI/8) are newly defined or described (Table 4). The standard for classification of a new cluster would be a 15–45% pairwise distance difference based on the complete capsid AA sequences (VP1) analyzed by the uncorrected distance method. Strains with distances below this range would be included with strains in the same cluster. Strains with distance above this range might represent different or new genogroups. Such outliers would require additional molecular and possibly biological data to support any new classification.

Characteristics of segmental capsid sequences of NoV in terms of diversity and pairwise distance range

Because of the difficulty amplifying the entire capsid gene of many strains due to their sequence diversities, various segmental regions in ORF1 and ORF2 have been used as shortcuts for diagnosis and genotyping (Ando et al., 2000; Vinjé et al., 2004). To explore statistically the molecular details

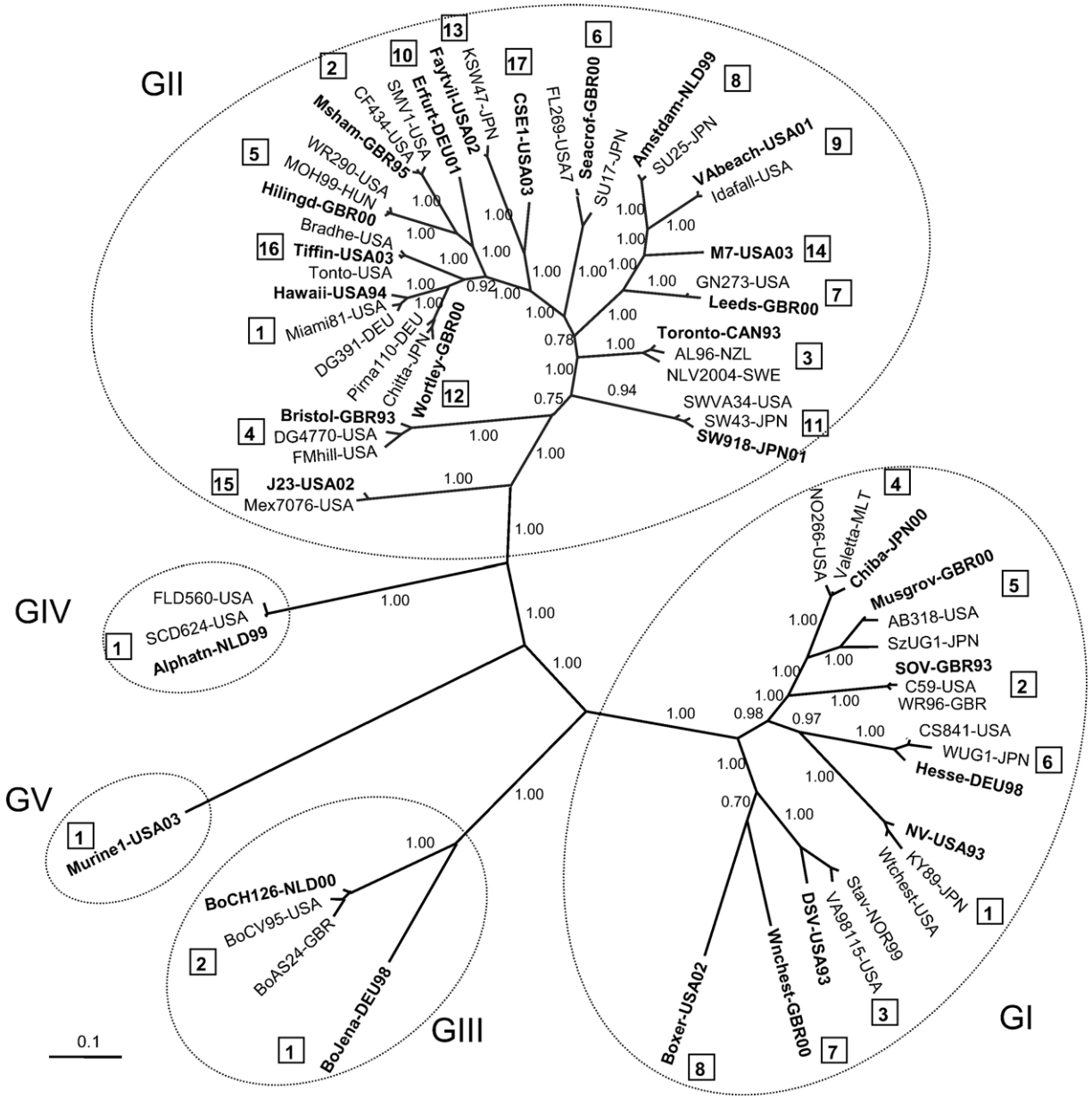

Fig. 3. Standard NoV classification tree proposed for nomenclature. Tree was made with the structural alignment of 68 NoV sequences (column gaps removed) by using MrBayes program (Huelsenbeck et al., 2001). Tree topology was evaluated on the base of 1,000,000 generations. See Fig. 1 legend for format of sequence names and country codes.

of ORF2, we evaluated 5 overlapping (~115 AA) segments 200 AA in length (each spanning from the N- to the C-terminal) and three structural domains of S, P1, and P2 (Prasad et al., 1999) for strain diversity and distance range as described above (Table 5). The least diversity in the 3 strain categories (S, C, and G) was located in segment 1 and the S domain, both corresponding to the 5′-end of ORF2; the greatest diversity was distributed in segment 4 and the P2 domain, another corresponding region within ORF2. This observation is consistent with the results of other investigations (Katayama et al., 2002; Prasad et al., 1999). However, the pairwise distance ranges of NoVs in the S, C, and G strain categories for all segments and domains overlapped, except in segment 5, where they were similar to those in the complete capsid protein: 0–18.18%, 19.23–52.20%, and 53.59–70.93% versus 0–14.07%, 14.26–43.78%, and 44.91–61.41%, respectively. These results suggest that, with

Table 4
Genogroups and clusters of NoV strains

Genogroup	Host	No. of sequences	Clusters	New clusters
G1	Human	30	8	1
G2	Human/Porcine	121	17	5
G3	Bovine	9	2	
G4	Human	3	1	
G5	Murine	1	1	
Total	5	164	29	6

Table 5
Pairwise distance ranges of segments and structural domains

Window	Location (AA)	Strains	Clusters	Genogroups
Capsid	3–546	0–14.07	14.26–43.78	44.91–61.41
Segment				
1	3–202	0–5.10	4.08–32.65	25.13–48.47
2	89–288	0–8.46	7.00–32.50	34.50–54.70
3	175–374	0–20.39	19.91–53.18	53.27–72.43
4	261–460	0–28.17	26.39–62.15	58.88–76.32
5	347–546	0–18.18	19.23–52.20	53.59–70.93
Domain				
S	50–225	0–6.85	3.65–30.14	25.69–48.86
P1	226–278	0–14.12	13.02–42.60	44.58–65.88
	406–520			
P2	279–405	0–33.09	30.40–74.64	66.94–84.21

Pairwise distance ranges were calculated by using the uncorrected method with 164 NoV sequences based on sequence window setting. Segments 1–5: 200 AA each, overlapped by ~115 AA, starting from N to C-terminal of capsid protein (based on NV strain, M87661); structural domains (S, P1, P2) (Prasad et al., 1999).

the exception of segment 5, NoV strains may be difficult to differentiate on the basis of analysis of segment or domain sequences alone.

Discussion

The International Committee on Taxonomy of Viruses has no guidelines for the classification of viruses below the species level, and, to date, no consensus for the classification of NoVs within genogroups has been reached (Green et al., 2000a; van Regenmortel et al., 2000). Furthermore, no simple immunological or biological methods are available to characterize human NoVs (e.g., serotypes); thus, sequence analysis has become the method of choice. To establish a standard classification system, we analyzed the phylogenetic relationships of 164 entire capsid sequences of NoV, using four unique approaches: (1) we selected only those strains with complete ORF2 sequences because ORF2 encodes the major structural protein that contains cell receptors and immunological determinants that might correlate with viral serotypes or phenotypes, and it is sufficiently long to allow proper differentiation between strains; (2) we included all possible NoV sequences from various hosts to present the true diversity of strains; (3) we used structure-based AA sequence alignment because phylogenetic outcomes are affected by the quality of sequence alignments, and alignment based on capsid structures also reflects more true homological relationships of the strains; and, (4) we used a Bayesian phylogenetic method that can efficiently handle a large number of sequences and produces the best trees based on likelihood. With these criteria, we defined a classification of NoVs with 29 genetic clusters, 8 in G1, 17 in G2, 2 in G3, and 1 each in G4 and G5. This phylogenetic scheme was statistically supported by pairwise distance analysis and was used to standardize the classification of NoVs. We also defined the ranges of pairwise distances for strains within a cluster, between clusters, and between genogroups, which provided clear criteria and a practical guide for characterization and description of outbreak strains.

Five well-segregated clades were observed in the tree generated from analysis of alignment of 141 sequences. These represent the current five NoV genogroups (GI–GV) as identified in other studies (Fankhauser et al., 2002; Green et al., 1995; Karst et al., 2003; Oliver et al., 2003; Vinjé and Koopmans, 2000). Among them, clades GI and II are the two largest genogroups and contain most of the diverse sequences in the collection (8 and 17 clusters, respectively). These include 6 newly defined or described clusters: GI/8 in GI and GII/13–17 in GII (Vinjé et al., 2004). Three unresolved clusters of j, m, and n in an earlier study were resolved and assigned clusters of GII/13, GII/16, and GII/10 (Fankhauser et al., 2002). The GII/4 is the most active cluster and contains the strains that have caused the most outbreaks worldwide since 1993, including the most recent US outbreaks (on land and aboard cruise ships) and national outbreaks in United Kingdom (Noel et al., 1999; Vipond et al., 2004; Widdowson et al., 2004). Except for GII/11, all sequences in GI and GII were from human viruses. Another human genogroup, the newly defined GIV, to date contains only 1 cluster (Fankhauser et al., 2002; Vinjé et al., 2004). Two other genogroups, GIII and GV, include strains from cows and mice and have 2 and 1 clusters, respectively. The murine strain, MNV-1, is the only NoV that grew successfully in cell culture and in a small animal model (Karst et al., 2003; Wobus et al., 2004). All together, 29 clusters in 5 genogroups were defined based on the tree topology shown in Fig. 1.

To evaluate the robustness of this classification scheme, we investigated the distributions of the pairwise distances between strains, using the ML or uncorrected method with different sequence alignments with or without gaps. All gave the same results, with three clearly resolved peaks that did not overlap. The 3-peak distribution pattern provides insight into the nature of the evolution of NoVs and evidence that NoVs can be further classified into 3 levels: strain (S), cluster (C), and genogroup (G). The same three levels also exist in “Sapporo-like viruses,” another genus within the Caliciviridae. However, the diversity of “Sapporo-like viruses” was substantially less than that for NoVs, with differences of 1–5%, 19–25%, and 49–55% for intra-cluster, inter-cluster, and inter-genogroup, respectively (Schuffenecker et al., 2001).

The existence of the minor peaks around large peaks (Fig. 2A) further demonstrates the greater diversity of NoVs. None of the distance distributions in major or minor peaks ranged across the classification line within cluster S, between clusters C, or between genogroups G. Therefore, the distance ranges of S, C, and G were used as criteria to classify new strains, which also extended our definitions for strains, clusters, and genogroups (Ando et al., 2000). The results (Fig. 2) obtained by using the three approaches provided parallel, consistent, and confirmatory results, demonstrating that the phylogenetic classification scheme of NoVs is well established and suited for making a standard tree for nomenclature (Fig. 3). We simplified the tree to include only one prototype (in bold) and up to two reference sequences selected from each cluster of the comprehensive tree and assigned consecutive numbers to the clusters that corresponded to the order of submission of the

prototype sequence to GenBank (Ando et al., 2000). Clearly, the tree topologies were maintained, and all cluster branches were supported by robust posterior probability values. We propose the phylogenetic classification scheme of NoVs plus the statistical measurements as guidelines for new NoV strain classification and nomenclature. By using the uncorrected distance method with untreated AA complete capsid sequences, the difference ranges were 45.0–61.4% for genogroups, 14.3–43.8% for clusters, and 0–14.1% for strains within a cluster.

The use of short conserved sequences has been successful for establishing diagnoses of NoV infection, but it became problematic for the classification or phylogenetic analysis because the small fragments made some strains indistinguishable and generated inconsistent and confused classification results. Strains with identical fragments were later found to be distinct when the full capsid was sequenced. A study based on the partial capsid N/S domain region defined 31 genotypes in GI and GII (Kageyama et al., 2004). When compared with the results of the entire capsid sequence analysis in this study, excluding 7 genotypes (4 in GI: GI/11–14 and 3 in GII: GII/11, 15, 16) that could not be compared due to only partially determined capsid sequences, 23 genotypes were found belonging to 25 clusters (16 agreed with each other [5 in GI: GI/1–4, 7 and 11 in GII: GII/1–10, 12]; 4 in GI [GI/5, 9 and GI/6, 8] belonged to 2 clusters [GI/5 and GI/6]; 2 in GII [GII/13, 14] were swapped; 1 [GI/10] had a different clustering number [GI/8]), and 1, defined as GII/17 (Alphatron-related viruses), was characterized as GIV in this and other studies (Fankhauser et al., 2002; Karst et al., 2003; Vinjé and Koopmans, 2000; Vinjé et al., 2004). The inconsistent classification results from partial sequence analyses were attributed to the different conservation rates along the genomic sequences (Table 5). Recombination within the capsid (Rohayem et al., 2005) also affected the correct classification. Therefore, we do not recommend using a partial sequence to classify new NoV strains.

NoV sequences have great genetic diversity, and the hypervariable region is located in the capsid gene. Consistent with the findings of others (Green et al., 1995, 2000b; Katayama et al., 2002; Prasad et al., 1999; Vinjé et al., 2000), we found that the capsid sequences varied by up to 60% between the five genogroups and up to 57% in human NoVs, a level of diversity much higher than that seen for other plus single-stranded RNA viruses, such as enteroviruses (polioviruses: ~20% in nts; Zheng et al., 1993) or respiratory viruses (rubella: ~10% in nts; Zheng et al., 2003), suggesting that the genogroups of NoV might be individual species or serotypes. Because no neutralization test is available for NoV serotyping, data on correlation of phylogenetic and antigenic characterizations are needed to support this hypothesis. The greater genetic diversity also suggests that NoVs might evolve faster than other RNA viruses, but the dynamic evolution of the virus remains unknown. We recently observed that virus transmitted during three outbreaks on a cruise ship in 18 days showed almost no mutation changes in ORF2 (Zheng, D.P., unpublished data). Nevertheless, an individual with chronic diarrhea shed viruses that had 32 AA changes accumulated in ORF2 during 1 year of illness (Nilsson et al., 2003). In addition to a higher mutation rate, recombination also contributes to the diversity and the evolution of RNA viruses. Human and bovine NoVs are no exception, and recombination at the junction of ORF1/2 and in ORF2 of human NoVs and in ORF1 of bovine NoVs was observed (Han et al., 2004; Katayama et al., 2002; Oliver et al., 2004; Rohayem et al., 2005; Vinjé et al., 2000). Recombination events may create new genotypes or species and cause difficulty in strain classification because the recombinants could be intermediates between genotypes or species. So far, no intermediate strains have been identified by capsid sequences analysis.

A recent study using the evolutionary trace method split the phylogenetic tree of NoVs into 10 even partitions and then grouped strains together that were nested in the same partition. Using these settings, bovine GIII strains were suggested to group with human GI strains and human GIV strains group with those of GII (Chakravarty et al., 2005). Our data demonstrate that NoV strains fall into 5 well segregated clades consistent with the 5 distinct genogroups previously defined. These 5 clades are statistically supported by our analysis of the pairwise distance distribution. Furthermore, the statistical differences in the ranges of pairwise distances that we observed between strains, clusters, and genogroups were not continuous and overlapping but were clearly divided, suggesting that they might well have a biological basis. Other criteria (http://www.ncbi.nlm.nih.gov/ICTVdb/ICTVdB/00.012.0.03.001.htm) often used to classify viruses include host range, and, unlike the results of Charavarty, our classification scheme places the bovine GIII strains into its own distinct genogroup, separate from the human GI strains. We also believe that the evolutionary trace method which is good for identifying conserved residues of genotypes or classes missed important information on genetic diversity that is critical for understanding molecular evolution and identifying epidemiologic links. Setting cutoff values at partition P03 to classify norovirus genogroups is quite arbitrary and does not reflect the clear statistical and natural relationships of NoVs that we observed. An added advantage of our approach is that, given any capsid sequence, the investigators can immediately classify their strain by genogroup, cluster, and strain after compared with those of prototype viruses we defined in this study.

In summary, we used a unique approach to standardize a classification scheme and provide clear criteria for NoV nomenclature below the genus level. Supporting data from biological or immunological characterization are not yet available due to the lack of a cell culture system; however, our results from phylogenetic analysis for genetic classification are reliable because the characterization was carried out with entire gene sequences of the major capsid protein, which correlated with strain phenotypes or serotypes, and the genetic groupings of G, C, and S were based statistically on their pairwise distance ranges, which reflected the natural relationships of NoVs. With rapid expansion of the number of sequences of diverse NoVs, this system provides guidelines for classifying new NoV strains while avoiding false classifications and confusion such as that generated from short conserved sequences. It can also enable selection of represen-

tative strains for preparation of virus-like particles for future biological or immunological studies, for the evaluation of diagnostic methods and results, and possibly for the study of candidate strains for vaccine production.

Materials and methods

NoV capsid (VP1) AA sequences

For this study, we selected 164 sequences of NoV strains (Table 1) that met the following criteria: the capsid gene sequence was complete and unique and had no ambiguous residues, with the exception of some prototype sequences (shown in bold). Of these, 145 were obtained from GenBank (submitted prior to Dec. 2003) by a BLAST search for *Norovirus* (taxonomy ID: 142786) with a sequence length of 1400–8000 nucleotides (nts), and 19 were from strains that had been newly sequenced at CDC using previously described methods (Fankhauser et al., 2002), except that, for this study, the primers were designed individually on the basis of each strain's sequences (data not shown). All nt sequences were translated into AA sequences for analysis.

Structure-based alignments of NoV AA sequences

A structure-based alignment of 43 capsid protein sequences was created based on the X-ray crystallographic structure of Norwalk strain capsid protein (Protein Data Bank identification, 1IHM) (Prasad et al., 1999) and other NoV capsid protein structures that were predicted by use of the MODELER program (INSIGHT II, version 2000.1, Accelrys Inc., San Diego, CA), including repetitive refinement of spatially aligned 3D models until the root mean square deviation between the model and 1IHM became 1.2 Å or less. To create the structure-based alignment with all 164 sequences, the remaining 121 AA sequences were grouped according to their homologies as determined by a preliminary phylogenetic analysis. Each of these sequences was then added and manually adjusted to the spatially aligned 43 sequences based on their obvious similarities.

Phylogenetic analysis

A preliminary phylogenetic analysis of the 164 sequences was done using the Pileup program (Wisconsin Package, version 10.3, Accelrys Inc., San Diego, CA) with the uncorrected distance algorithm. The phylogenetic analyses of structure-based alignments of 141 and 68 capsid sequences were performed using the MrBayes program (version 3, http://www.morphbank.ebc.uu.se/mrbayes/), a program for the Bayesian inference of phylogeny based upon a quantity called the posterior probability distribution of trees searched by the Markov chain Monte Carlo algorithm (Huelsenbeck et al., 2001). The input of the program (NEXUS file) includes a character matrix (the structure-based sequence alignment) and analysis settings (ngen = 1,000,000 [clustering generation number], nchains = 4 [searching chain number], samplefreq = 500 [sample frequency], swapfreq = 2 [swapping frequency number]). The best tree was collected from the pool of 1,000,000 tree generations, according to the highest likelihood number, viewed with TreeView software (Page, 1996) and edited with CorelDraw 10 (Corel Corporation, Ottawa, Canada).

Pairwise distance calculation

Pairwise distances of AA sequences in the 164-sequence alignment (with gaps) and in the 141-sequence alignment (without gaps) were calculated using both uncorrected distance (Wisconsin Package, version 10.3, Accelrys Inc., San Diego, CA) and the maximum likelihood (ML) distance (Tree-Puzzle 5.0) (settings for quartet likelihood searching: puzzling step [50,000], parameter estimates [exact], model of substitution [auto, JTT], AA frequencies [estimate from data set], model of rate heterogeneity [gamma distributed rates], gamma distribution parameter alpha [1.00], and number of gamma rate categories [8]) (Schmidt et al., 2002). The distances were sorted into three categories: S, the distances between strains within clusters; C, the distances between strains across clusters; and G, the distances between strains across genogroups, according to the classification scheme obtained from the phylogenetic analysis detailed above. These data were used to generate histograms of the distributions of pairwise distances.

Acknowledgments

We thank Dr. Larry Anderson, for his critical comments and suggestions, Ms. Kathy Murray and Claudia Chesley, for their excellent editorial assistance, and other colleagues for their help to make the manuscript possible.

References

Ando, T., Noel, J.S., Fankhauser, R.L., 2000. Genetic classification of "Norwalk-like viruses". J. Infect. Dis. 181 (Suppl. 2), S336–S348.

Belliot, G., Sosnovtsev, S.V., Mitra, T., Hammer, C., Garfield, M., Green, K.Y., 2003. In vitro proteolytic processing of the MD145 norovirus ORF1 nonstructural polyprotein yields stable precursors and products similar to those detected in Calicivirus-infected cells. J. Virol. 77, 10957–10974.

Bertolotti-Ciarlet, A., Crawford, S.E., Hutson, A.M., Estes, M.K., 2003. The 3′ end of Norwalk virus mRNA contains determinants that regulate the expression and stability of the capsid protein VP1: a novel function for the VP2 protein. J. Virol. 77, 11603–11615.

Chakravarty, S., Hutson, A.M., Estes, M.K., Prasad, B.V.V., 2005. Evolutionary trace residues in Noroviruses: importance in receptor binding, antigenicity, virion assembly, and strain diversity. J. Virol. 79, 554–568.

Dolin, R., Blacklow, N.R., DuPont, H., Buscho, R.F., Wyatt, R.G., Kasel, J.A., Hornick, R., Chanock, R.M., 1972. Biological properties of Norwalk agent of acute infectious nonbacterial gastroenteritis. Proc. Soc. Exp. Biol. Med. 140 (2), 578–583.

Duizer, E., Schwab, K.J., Neill, F.H., Atmar, R.L., Koopmans, M.P.G., Estes, M.K., 2004. Laboratory efforts to cultivate Noroviruses. J. Gen. Virol. 85, 79–87.

Estes, M.K., Ball, J.M., Guerrero, R.A., Opekun, A.R., Gilger, M.A., Pacheco, S.S., Graham, D.Y., 2000. Norwalk virus vaccines: challenges and progress. J. Infect. Dis. 181 (Suppl. 2), S367–S373.

Fankhauser, R.L., Monroe, S.S., Noel, J.S., Humphrey, C.D., Bresee, J.S.,

Parashar, U.D., Ando, T., Glass, R.I., 2002. Epidemiologic and molecular trends of Norwalk-like viruses associated with outbreaks of gastroenteritis in the United States. J. Infect. Dis. 186, 1–7.

Glass, I.R., Noel, J., Ando, T., Fankhauser, R., Belliot, G., Mounts, A., Parashar, U.D., Bresee, J.S., Monroe, S.S., 2000. The epidemiology of enteric Caliciviruses from humans: a reassessment using new diagnostics. J. Infect. Dis. 181, S254–S261 (Suppl.).

Green, S.M., Lambden, P.R., Caul, E.O., Ashley, C.R., Clarke, I.N., 1995. Capsid diversity in small round-structured viruses: molecular characterization of an antigenically distinct human enteric calicivirus. Virus Res. 37, 271–283.

Green, S.M., Lambden, P.R., Caul, E.O., Clarke, I.N., 1997. Capsid sequence diversity in small round structured viruses from recent UK outbreaks of gastroenteritis. J. Med. Virol. 52, 14–19.

Green, J., Vinje, J., Gallimore, C.I., Koopmans, M., Hale, A., Brown, D.W.G., 2000a. Capsid protein diversity among Norwalk-like viruses. Virus Genes 20, 227–236.

Green, K.Y., Ando, T., Balayan, M.S., Berke, T., Clarke, I.N., Estes, M.K., Matson, D.O., Nakata, S., Neill, J.D., Studdert, M.J., Thiel, H.-J., 2000b. Taxonomy of the Caliciviruses. J. Infect. Dis. 181 (Suppl. 2), S322–S330.

Green, K.Y., Chanock, R.M., Kapikian, A.Z., 2001. In: Knipe, D.M., Howley, P.M., et al. (Eds.), Human Caliciviruses: Fields Virology, 4th ed., vol. 1. Lippincott Williams and Wilkins, Philadelphia, pp. 841–874.

Green, K.Y., Belliot, G., Taylor, J., Valdesuso, J., Lew, J., Kapikian, A., Lin, F.-Y., 2002. A predominant role for Norwalk-like viruses as agents of epidemic gastroenteritis in Maryland nursing homes for the elderly. J. Infect. Dis. 185, 133–146.

Han, M.G., Smiley, J.R., Thomas, C., Saif, L.J., 2004. Genetic recombination between two genogroup III bovine noroviruses (BoNVs) and capsid sequence diversity among BoNVs and Nebraska-like bovine enteric caliciviruses. J. Clin. Microbiol. 42, 5214–5224.

Harrington, P.R., Vinje, J., Moe, C.L., Baric, R.S., 2004. Norovirus capture with histo-blood group antigens reveals novel virus–ligand interactions. J. Virol. 78 (6), 3035–3045.

Huelsenbeck, J.P., Ronquist, F., Nielsen, R., Bollback, J.P., 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. Science 294, 2310–2314.

Hutson, A.M., Atmar, R.L., Estes, M.K., 2004. Norovirus disease: changing epidemiology and host susceptibility factors. Trends Microbiol. 12 (6), 279–287.

Inouye, S., Yamashita, K., Yamadera, S., Yoshikawa, M., Kato, N., Okabe, N., 2000. Surveillance of viral gastroenteritis in Japan: pediatric cases and outbreak incidents. J. Infect. Dis. 181 (Suppl. 2), S270–S274.

Jiang, X., Graham, D.Y., Wang, K., Estes, M.K., 1990. Norwalk virus genome cloning and characterization. Science 250, 1580–1583.

Jiang, X., Wang, M., Wang, K., Estes, M.K., 1993. Sequence and genomic organization of Norwalk virus. Virology 195, 51–61.

Johnson, P.C., Mathewson, J.J., DuPont, H.L., Greenberg, H.B., 1990. Multiple-challenge study of host susceptibility to Norwalk gastroenteritis in US adults. J. Infect. Dis. 161 (1), 18–21.

Kageyama, T., Kojima, S., Shinohara, M., Uchida, K., Fukushi, S., Hoshino, F.B., Takeda, N., Katayama, K., 2003. Broadly reactive and highly sensitive assay for Norwalk-like viruses Based on Real-time quantitative Reverse Transcription-PCR. J. Clin. Microbiol. 41, 1548–1557.

Kageyama, T., Shinohara, M., Uchida, K., Fukushi, S., Hoshino, F.B., Kojima, S., Takai, R., Oka, T., Takeda, N., Katayama, K., 2004. Coexistence of multiple genotypes, including newly identified genotypes, in outbreaks of gastroenteritis due to Norovirus in Japan. J. Clin. Microbiol. 42, 2988–2995.

Kapikian, A.Z., Wyatt, R.G., Dolin, R., Thornhill, T.S., Kalica, A.R., Chanock, R.M., 1972. Visualization by immune electron microscopy of a 27 nm particle associated with acute infectious nonbacterial gastroenteritis. J. Virol. 10, 1075–1081.

Katayama, K., Shirato-Horikoshi, H., Kojima, S., Kageyama, T., Oka, T., Hoshino, F.B., Fukushi, S., Shinohara, M., Uchida, K., Suzuki, Y., Gojobori, T., Takeda, N., 2002. Phylogenetic analysis of the complete genome of 18 Norwalk-like viruses. Virology 299, 225–239.

Karst, S.M., Wobus, C.E., Lay, M., Davidson, J., Virgin IV, H.W., 2003. STAT1-dependent innate immunity to a Norwalk-like virus. Science 299, 1575–1578.

Lewis, D., Ando, T., Humphrey, C.D., Monroe, S.S., Glass, R.I., 1995. Use of solid-phase immune electron microscopy for classification of Norwalk-like viruses into six antigenic groups from 10 outbreaks of gastroenteritis in the United States. J. Clin. Microbiol. 33, 501–504.

Lopman, B.A., Reacher, M.H., Duijnhoven, Y.V., Hanon, F., Brown, D., Koopmans, M., 2003. Viral gastroenteritis outbreaks in Europe, 1995–2000. Emerg. Infect. Dis. 9 (1), 90–96.

Matsui, S.M., Greenberg, H.B., 2000. Immunity to Calicivirus infection. J. Infect. Dis. 181 (Suppl. 2), S331–S335.

Meyer, E., Ebner, W., Scholz, R., Dettenkofer, M., Daschner, F.D., 2004. Nosocomial outbreak of norovirus gastroenteritis and investigation of ABO histo-blood group type in infected staff and patients. J. Hosp. Infect. 56 (1), 64–66.

Nilsson, M., Hedlund, K.-O., Thorhagen, M., Larson, G., Johansen, K., Ekspong, A., Svensson, L., 2003. Evolution of human Calicivirus RNA in vivo: accumulation of mutations in the protruding P2 domain of the capsid leads to structural changes and possibly a new phonotype. J. Virol. 77, 13117–13124.

Noel, J.S., Fankhauser, R.L., Ando, T., Monroe, S.S., Glass, R.I., 1999. Identification of a distinct common strain of "Norwalk-like viruses" having a global distribution. J. Infect. Dis. 179, 1334–1344.

Okada, S., Sekine, S., Ando, T., Hayashi, Y., Murao, M., Yabuuchi, K., Miki, T., Ohashi, M., 1990. Antigenic characterization of small, round-structured viruses by immune electron microscopy. J. Clin. Microbiol. 28, 1244–1248.

Oliver, S.L., Dastjerdi, A.M., Wong, S., El-Attar, L., Gallimore, C., Brown, D.W.G., Green, J., Bridger, J.C., 2003. Molecular characterization of bovine enteric caliciviruses: a distinct third genogroup of norovirus (Norwalk-like viruses) unlikely to be of risk to humans. J. Virol. 77, 2789–2798.

Oliver, S.L., Brown, D.W.G., Green, J., Bridger, J.C., 2004. A chimeric bovine enteric Calicivirus: evidence for genomic recombination in genogroup III of the *Norovirus* genus of the Caliciviridae. Virology 326, 231–239.

Page, R.D.M., 1996. TREEVIEW: an application to display phylogenetic trees on personal computers. Comput. Appl. Biosci. 12, 357–358.

Parrino, T.A., Schreiber, D.S., Trier, J.S., Kapikian, A.Z., Blacklow, N.R., 1977. Clinical immunity in acute gastroenteritis caused by Norwalk agent. N. Engl. J. Med. 297, 86–89.

Prasad, B.V.V., Hardy, M.E., Dokland, T., Bella, J., Rossmann, M.G., Estes, M.K., 1999. X-ray crystallographic structure of the Norwalk virus capsid. Science 286, 287–290.

Richards, G.P., Watson, M.A., Kingsley, D.H., 2004. A SYBR green, real-time RT-PCR method to detect and quantitate Norwalk virus in stools. J. Virol. Methods 116, 63–70.

Rohayem, J., Münch, J., Rethwilm, A., 2005. Evidence of recombination in the norovirus capsid gene. J. Virol. 79, 4977–4990.

Rockx, B., Wit, M. de, Vennema, H., Vinjé, J., Bruin, E. de, van Duynhoven, Y., Koopmans, M., 2002. Natural history of Human Calicivirus infection: a prospective cohort study. Clin. Infect. Dis. 35, 246–253.

Schmidt, H.A., Strimmer, K., Vingron, M., von Haeseler, A., 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics 18, 502–504.

Schuffenecker, I., Ando, T., Thouvenot, D., Lina, B., Aymard, M., 2001. Genetic classification of "Sapporo-like viruses". Arch. Virol. 146, 2115–2132.

Tan, M., Hegde, R.S., Jiang, X., 2004. The P domain of norovirus capsid protein forms dimer and binds to histo-blood group antigen receptors. J. Virol. 78 (12), 6233–6242.

van Regenmortel, M.H.V., Fauquet, C.M., Bishop, D.H.L., Carstens, E.B., Estes, M.K., Lemon, S.M., Maniloff, J., Mayo, M.A., McGeoch, D.J., Pringle, C.R., Wickner, R.B., 2000. Virus Taxonomy: Seventh Report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego, NY.

Vinjé, J., Koopmans, M.P.G., 2000. Simultaneous detection and genotyping of

"Norwalk-like viruses" by oligonucleotide array in a reverse line blot hybridization format. J. Clin. Microbiol. 38, 2595–2601.

Vinjé, J., Green, J., Lewis, D.C., Gallimore, C.I., Brown, D.W.G., Koopmans, M.P.G., 2000. Genetic polymorphism across regions of the three open reading frames of "Norwalk-like viruses". Arch. Virol. 145, 223–241.

Vinjé, J., Hamidjaja, R.A., Sobsey, M.D., 2004. Development and application of a capsid VP1 (region D) based reverse transcription PCR assay for genotyping of genogroup I and II Noroviruses. J. Virol. Methods 116, 109–117.

Vipond, I.B., Caul, E.O., Hirst, D., Carmen, B., Curry, A., Lopman, B.A., Pead, P., Pickett, M.A., Lambden, P.R., Clarke, I.N., 2004. National epidemic of Lordsdale Norovirus in the UK. J. Clin. Virol. 30, 243–247.

Widdowson, M.A., Cramer, E., Hadley, L., Bresee, J., Beard, R.S., Bulens, S., Charles, M., Chege, W., Isakbaeva, E., Wright, J., Mintz, E., Massey, J., Glass, R.I., Monroe, S.S., 2004. Outbreaks of acute gastroenteritis on cruise ships and on land: identification of a predominant strain of norovirus (NV) United States 2002. J. Infect. Dis. 190, 27–36.

Wobus, C.E., Karst, S.M., Thackray, L.B., Chang, K.O., Sosnovtsev, S.V., Belliot, G., Krug, A., Mackenzie, J.M., Green, K.Y., Virgin IV, H.W., 2004. Replication of Norovirus in cell culture reveals a tropism for dendritic cells and macrophages. Plos Biol. 2 (12, e432), 0001–0009.

Wyatt, R.G., Dolin, R., Blacklow, N.R., Dupont, H.L., Buscho, T.R.F., Thornhill, S., Kapikian, A.Z., Chanock, R.M., 1974. Comparison of three agents of acute infectious nonbacterial gastroenteritis by cross-challenge in volunteers. J. Infect. Dis. 129, 709–714.

Yan, H., Yagyu, F., Okitsu, S., Nishio, O., Ushijima, H., 2003. Detection of norovirus (GI, GII) Sapovirus and astrovirus in fecal samples using reverse transcription single-round multiplex PCR. J. Virol. Methods 114, 37–44.

Zheng, D.P., Zhang, L.B., Fang, Z.Y., Yang, C.F., Mulders, M., Pallansch, M.A., Kew, O.M., 1993. Distribution of wild type 1 poliovirus genotypes in China. J. Infect. Dis. 168, 1361–1367.

Zheng, D.P., Frey, T.K., Icenogle, J., Katow, S., Abernathy, E.S., Song, K.J., Xu, W.B., Yarulin, V., Desjatskova, R.G., Aboudy, Y., Enders, G., Croxson, M., 2003. Global distribution of rubella virus genotypes. Emerg. Infect. Dis. 9, 1523–1530.