

G8 Rotavirus Strains Isolated in the Democratic Republic of Congo Belong to the DS-1-Like Genogroup

Jelle Matthijnsens,¹ Mustafizur Rahman,^{1,2} Xuelei Yang,³ Thomas Delbeke,¹ Ingrid Arijs,¹ Jean-Pierre Kabue,⁴ Jean-Jacques Tamfum Muyembe,⁴ and Marc Van Ranst^{1*}

Laboratory of Clinical and Epidemiological Virology, Department of Microbiology and Immunology, Rega Institute for Medical Research, University of Leuven, Leuven, Belgium¹; Laboratory of Virology, ICDDR, B, Centre for Health and Population Research, Dhaka, Bangladesh²; Xinjiang Pediatric Institute, People's Hospital, Urumqi, People's Republic of China³; and Institute for National Bio-Medical Research, Kinshasa, Democratic Republic of Congo⁴

Received 9 December 2005/Returned for modification 24 January 2006/Accepted 4 February 2006

Several G8P[6] and G8P[8] rotavirus strains were isolated from hospitalized patients in the Democratic Republic of Congo in 2003. To investigate their overall genomic relatedness and to determine to which genogroup they belonged, the complete genomes of strains DRC88 (G8P[8]) and DRC86 (G8P[6]) were determined. Genomic comparison of these two African G8 strains revealed that 10 out of their 11 gene segments, except for VP4, were nearly identical (>98.9% identical at the nucleotide level), suggesting that this rare G8P[8] rotavirus strain originated recently from a reassortment between a common G8P[6] strain and a strain with a P[8] specificity. A very close evolutionary relationship between 9 out of the 11 gene segments of DRC88 and DRC86 and rotavirus strains belonging to the DS-1-like (G2P[4]) “genogroup” was found, and several possible reassortment events preceding the occurrence of G8P[8] and G8P[6] human rotaviruses were hypothesized. Since the genes of G2P[4] rotavirus strains are very well adapted to infect humans, the acquirement of a new VP7 (G8) gene, and especially the replacement of P[6] (believed to be of animal origin) by P[8] (most common in human rotaviruses), might make DRC88-like rotaviruses very well equipped to become a predominant human rotavirus strain and an important pathogen on the African continent and the rest of the world. These findings have important implications for rotavirus vaccine development and highlight that typing of new rotavirus strains by merely sequencing their VP7 and VP4 genes provides us with only the tip of the iceberg regarding rotavirus diversity.

Group A rotaviruses (RVs) are the most important cause of diarrheal diseases and cause significant morbidity and mortality in young children and animals worldwide, especially in developing countries (34). RVs belong to the family *Reoviridae* and possess a genome of 11 segments of double-stranded RNA. A dual classification system based on the two major outer capsid proteins defines several G and P genotypes of group A RVs. To date, 15 G genotypes have been identified based on the glycosylated VP7, and 26 P genotypes have been identified based on the protease-sensitive VP4 (16, 37a, 56, 58).

At least 11 G types have been isolated from humans, but only G1, G2, G3, G4, and G9 are currently of epidemiological importance (16, 55). G8 RVs are common in cattle, but they also constitute a significant proportion of human RV (HRV) strains in African countries (11, 62). On single occasions, G8 RVs have also been found in a pig, a horse, two guanacos, and a macaque (24, 27, 28, 52). In humans, G8 RV strains (57 M, 69 M, B37, and B38) were first isolated between 1978 and 1981 on the Indonesian islands Sumatra and Java (5, 25). These four strains possessed an unusual “supershort” electrophoretic pattern. Since then, isolated human cases of G8 RV infection have been found in Finland, Italy, the United Kingdom, Brazil, Australia, Japan, and India with long and short electrophoretic

patterns and with SG I and SG II specificities (8, 11, 21, 31, 33, 48, 60, 63). By far, the most G8 HRVs were found all over the African continent (Malawi, Nigeria, Egypt, Kenya, South Africa, Guinea-Bissau, and Ghana), and in a few epidemiological surveys, G8 was the predominant G genotype (3, 7, 12, 13, 18, 19, 26, 44, 63). G8 HRVs have been detected in combination with a large number of VP4 specificities: P[1], P[2], P[4], P[6], P[8], P[10], and P[14] (13, 19, 22, 31, 33, 53, 61). In addition to this list, P[5] and P[11] (typical bovine P genotypes) have been found in combination with G8 in cattle (17, 20, 49).

RV diversity is generated through the processes of genetic drift and genetic shift. Genetic drift is due to the accumulation of point mutations due to the error-prone nature of RNA-dependent RNA polymerases. Genetic shift describes sudden changes in the rotavirus genome. The most important source of these sudden changes is gene reassortments. After infection of a host cell with two different RVs, theoretically, 2¹¹ progeny viruses with novel characteristics can arise (57). Other less frequently occurring examples of genetic shifts found in the literature are partial gene duplications and crossover between rotavirus gene segments belonging to different circulating strains (15, 40, 51).

To determine the overall genomic relatedness of different HRVs, RNA-RNA hybridization assays have been used to define three “genogroups” represented by the prototype HRV strains Wa, DS-1, and AU-1 (43). Several studies have used this genogrouping system to prove intergenogroup reassortments between HRVs or to investigate reassortments between RVs belonging to human and animal genogroups (42). Al-

* Corresponding author. Mailing address: Laboratory of Clinical and Epidemiological Virology, Department of Microbiology and Immunology, Rega Institute for Medical Research, Minderbroederstraat 10, B-3000 Leuven, Belgium. Phone: 32-16-347908. Fax: 32-16-347900. E-mail: marc.vanranst@uz.kuleuven.ac.be.

though RNA-RNA hybridization assays are an outstanding tool to quickly investigate the genetic relatedness of RVs, faint hybridization bands and different stringencies between assays can make results difficult to interpret and compare (42). As a complementary technique, complete or partial sequencing of the 11 genome segments of RVs can contribute to a better understanding of the overall genomic and evolutionary relationships of RVs, as was previously demonstrated in Finland by Maunula and Von Bonsdorff (41).

In this study, stool specimens collected from RV-infected patients during an outbreak of acute diarrhea among young children under 5 years of age in the Democratic Republic of Congo (DRC) were analyzed. The entire genomes of a G8P[8] RV strain and a common G8P[6] RV strain isolated in the same area were sequenced and compared. Their relationship to other known RV strains circulating all over the world was also investigated.

MATERIALS AND METHODS

Rotavirus strains. RV strains DRC88 and DRC86 were collected from patients in the Democratic Republic of Congo in 2003, during an outbreak of acute diarrhea among young children under 5 years of age, at the Kalembe-Lembe Pediatric hospital in Kinshasa, DRC. Fecal samples containing these RV isolates were collected at the Institut National de Recherche Bio-Medicale and were transported to the Rega Institute, University of Leuven, Belgium, using chromatography paper strips (54).

RT-PCR. Viral RNA was extracted using the QIAamp Viral RNA mini kit (QIAGEN/Westburg, Leusden, The Netherlands) according to the manufacturer's instructions. The extracted RNA was denatured at 95°C for 3 min, and reverse transcription followed by PCR (RT-PCR) was carried out using the QIAGEN OneStep RT-PCR kit (QIAGEN/Westburg). Primer sequences used are available upon request. RT-PCR was carried out with an initial reverse transcription step at 45°C for 30 min followed by a PCR activation step at 95°C for 15 min, 40 cycles of amplification, and a final extension step for 10 min at 70°C in a GeneAmp PCR System 9700 thermal cycler (Applied Biosystems Group, Foster City, CA). For the smaller gene segments encoding VP6, NSP1, NSP2, NSP3, NSP4, and NSP5, the amplification cycle conditions were as follows: 45 s at 94°C, 45 s at 47°C, and 2.5 min at 70°C. For the larger segments encoding VP1, VP2, VP3, and VP4, the cycle conditions were 45 s at 94°C, 45 s at 47°C, and 6 min at 70°C.

Nucleotide sequencing. The PCR amplicons were purified with the QIAquick PCR purification kit (QIAGEN/Westburg) and sequenced using the dideoxy nucleotide chain termination method with the ABI PRISM BigDye Terminator cycle sequencing reaction kit (Applied Biosystems) on an ABI PRISM 3100 automated sequencer (Applied Biosystems). The sequencing was performed with the same forward and reverse primers as those used for RT-PCR. Primer-walking sequencing was performed to cover the complete genome on both strands.

Determination of the 5'- and 3'-terminal sequences. To obtain the complete nucleotide sequences of strains DRC86 and DRC88, the 5'- and 3'-terminal sequences of the 11 gene segments were determined using the single-primer amplification method (36). Briefly, after RNA extraction, a modified amino-linked oligonucleotide (TGP-Linker [5'-PO₄-TTCCTTATGCAGCTGATCACT CTGTGTCA-spacer-NH₂-3']) was ligated into the 3' end of both strands of the viral double-stranded RNA with T4 RNA ligase (Promega, Leiden, The Netherlands). RT-PCR with primers TGP-3Out (5'-TGACACAGAGTGATCAGC-3') (complementary to TGP-Linker) and appropriate gene-specific primers (based on the known internal sequences of each segment) was carried out. The following thermal cycling conditions were used: an initial reverse transcription step at 45°C for 30 min, a PCR activation step at 95°C for 15 min, a 45-min step during which the temperature was gradually lowered from 83°C to 60°C (to allow the newly transcribed cDNA strands to anneal), a 10-min step at 72°C (to allow the DNA polymerase to repair the partial duplexes), 40 cycles of amplification (45 s at 94°C, 45 s at 45°C, and 1 min at 70°C), and a final extension step for 10 min at 70°C. These amplified products were purified and sequenced as described above.

RNA and protein sequence analysis. The chromatogram sequencing files were analyzed using Chromas 2.23 (Technelysium, Queensland, Australia), and con-

tigs were generated using SeqMan II (DNASTAR, Madison, WI). Nucleotide and protein sequence identity searches were performed using the National Center for Biotechnology Information (National Institutes of Health, Bethesda, MD) BLAST (Basic Local Alignment Search Tool) server of the GenBank database, release 143.0 (6). Multiple-sequence alignments were calculated using ClustalX 1.81 (64). Sequences were manually edited using the GeneDoc version 2.6.002 alignment editor (45). Genetic distances at the amino acid level were calculated using the Poisson correction parameter with MEGA version 2.1 software (35).

Phylogenetic analysis. Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 2.1 software (35), based on the deduced amino acid sequences of the different RV gene segments available in GenBank. Dendrograms were constructed using the neighbor-joining method.

Nucleotide sequence accession numbers. The complete nucleotide sequence data of the 11 gene segments of strains DRC88 and DRC86 have been deposited in GenBank under the accession numbers DQ005104 to DQ005114 and DQ005115 to DQ005125.

RESULTS

The complete nucleotide and deduced amino acid sequences of the 11 gene segments of DRC88 (G8P[8]) and DRC86 (G8P[6]) were determined. Both complete genomes contained 18,615 nucleotides and 5,799 deduced amino acids. For each segment, the amounts of nucleotides and amino acids were identical when DRC88 and DRC86 were compared. The exact number of nucleotides (amino acids) were as follows: 3,302 (1,088) for VP1, 2,687 (880) for VP2, 2,591 (835) for VP3, 2,359 (775) for VP4, 1,356 (397) for VP6, 1,062 (326) for VP7, 1,566 (493) for NSP1, 1,059 (317) for NSP2, 1,066 (313) for NSP3, 751 (175) for NSP4, and 816 (200) for NSP5. Comparison of the nucleotide (amino acid) sequences of the different gene segments of DRC88 and DRC86 revealed very high similarities, ranging from 98.9% to 99.9% (98.7% to 100%) except for gene 4, encoding different VP4 proteins (P[6] and P[8], 70.2% [76.7%]). The overall genomic identity was 96.5% (96.9%) (Table 1 and data not shown).

G-typing analysis of the VP7 gene. The VP7 genes of RV strains DRC88 and DRC86 showed a 96.2 to 97.7% nucleotide sequence identity and a 96.8 to 98.1% amino acid sequence identity with the following G8 RV strains: MW333 and MW23 isolated in Malawi, R291 isolated in Brazil, and human and bovine strains HMG035, HMG89, and NGRBg8 isolated in Nigeria (1–3, 12). Phylogenetic analysis of VP7 clustered DRC88 and DRC86 with these five African strains and one Brazilian RV strain (Fig. 1A and B). A very high degree of P-type diversity (P[1], P[4], P[6], and P[8]) was noticed for these eight G8 RV strains (Fig. 1B).

P-typing analysis of the VP4 gene. A phylogenetic tree that includes the complete VP4 amino acid sequences of our DRC88 and DRC86 strains and all known P types (except for P[22], P[23], and P[25], because no complete sequence data were available) was constructed (Fig. 1C). A detailed dendrogram of the P[8] and P[6] genotypes showing the relationship of DRC88 and DRC86 with other complete VP4 amino acid sequences is illustrated in Fig. 1D. For many RV strains, only the partial VP4 gene sequences are known; therefore, a second detailed dendrogram constructed with amino acids 12 to 286 is shown (Fig. 1E). DRC86 clustered closely with P[6] RV strains isolated in Malawi (G8P[6] strains), South Africa (G9P[6] strains), and the United States (G9P[6] and G12P[6] strains), with 95.1 to 96.7% identity (Fig. 1E). In contrast, DRC88 clustered very closely with G9P[8] RV strains isolated in Hungary and Australia (98.6 to 98.9% identity) and G1P[8],

TABLE 1. Percentages of amino acid identity of the DRC88 gene segments with the (almost) completely sequenced RV strains KU, DS-1, NR1, S2, IS2, Wa, TB-Chen, SA11, B4106, 30/96, OSU, RF, Uk(tc), DRC86, I321, and YM^a

Strain ^b	% Identity ^c								
	VP1	VP2	VP3	VP6	NSP1	NSP2	NSP3	NSP4	NSP5
G1P[8]/Hu/KU	88.3	89.3	78.0	91.6	62.0	90.0	78.3	80.5	83.1
G1P[8]/Hu/Wa	83.5 ^d	90.1	80.3	91.6	62.6	90.0	80.0	85.4	81.2
G2P[4]/Hu/TB-Chen	97.3	99.1	96.9	98.2	96.8	94.8	98.7	94.0	98.4
G2P[4]/Hu/DS-1	94.6^d	NS	95.7	NS	92.1	94.1	NS	95.4	96.7
G2P[4]/Hu/NR1	NS	NS	NS	99.7	97.9	98.7	99.7	94.7	98.9
G2P[4]/Hu/S2	NS	NS	NS	98.0	NS	NS	98.0	94.0	NS
G2P[4]/Hu/IS2	NS	NS	NS	99.2	NS	97.4	98.0	NS	NS
G3P[4]/Si/SA11	96.0	93.7	84.1	97.4	>40	86.8	77.1	94.7	87.5
G3P[4]/Hu/B4106	97.2	97.0	82.7	98.2	>40	93.4	81.2	89.1	82.5
G3P[4]/La/30/96	96.7	96.0	82.5	98.2	>40	93.7	80.4	86.2	80.7
G5P[4]/Po/OSU	NS	NS	79.3	91.3	62.0	89.6	80.8	84.7	83.1
G6P[4]/Bo/RF	97.4	96.8	90.3	99.0	47.7	93.1	81.6	94.0	68.9
G6P[4]/Bo/Uk(tc)	97.3	96.5	88.8	99.5	47.7	96.4	78.8	93.3	84.8
G8P[4]/Hu/DRC86	99.6	99.9	99.8	99.7	100.0	99.4	100.0	98.7	100.0
G10P[4]/Hu/I321	NS	NS	NS	99.2	61.6	97.8	80.4	93.7	NS
G11P[4]/Po/YM	88.3	NS	80.6	90.5	62.3	NS	NS	83.2	85.0

^a GenBank accession numbers used in this comparison were as follows: for KU, BAA84962 (VP1), BAA84963 (VP2), BAA84964 (VP3), BAA84965 (VP6), BAA84966 (NSP1), BAA84967 (NSP2), BAA84968 (NSP3), BAA84969 (NSP4), and BAB83818 (NSP5); for Wa, AAF19585 (VP1), CAA33074 (VP2), AAQ02692 (VP3), P03530 (VP6), P35424 (NSP1), AAA47301 (NSP2), CAA57193 (NSP3), AAA47309 (NSP4), and BAC65999 (NSP5); for TB-Chen, AAV65743 (VP1), AAV65742 (VP2), AAV65744 (VP3), AAV65735 (VP6), AAV65737 (NSP1), AAV65738 (NSP2), AAV65739 (NSP3), AAV65740 (NSP4), and AAV65741 (NSP5); for DS-1, AAF19587 (VP1), AAQ21041 (VP3), P35423 (NSP1), AAA47296 (NSP2), AAG09190 (NSP4), and P23048 (NSP5); for NR1, AAM44188 (VP6), AAQ07952 (NSP1), AAQ07953 (NSP2), AAQ07954 (NSP3), AAQ07957 (NSP4), and AAQ07965 (NSP5); for S2, CAA68495 (VP6), CAA57187 (NSP3), and AAB81290 (NSP4); for IS2, CAA64323 (VP6), CAA64262 (NSP2), and CAA54093 (NSP3); for SA11, AAC58684 (VP1), CAA34733 (VP2), CAA34198 (VP3), AAO32085 (VP6), P15687 (NSP1), 1L9V_A (NSP2), P03536 (NSP3), AAC61867 (NSP4), and AAA66881 (NSP5); for B4106, AAU43799 (VP1), AAU43798 (VP2), AAU43797 (VP3), AAU43795 (VP6), AAU43793 (NSP1), AAU43792 (NSP2), AAU43791 (NSP3), AAU43790 (NSP4), and AAU43789 (NSP5); for 30/96, DQ205221 (VP1), DQ205222 (VP2), DQ205223 (VP3), DQ205226 (VP6), DQ205225 (NSP1), DQ205227 (NSP2), DQ205228 (NSP3), DQ205230 (NSP4), and DQ205231 (NSP5); for OSU, AAQ21048 (VP3), AAK60604 (VP6), CAA78092 (NSP1), P09366 (NSP2), CAA57190 (NSP3), BAA13728 (NSP4), and CAA33540 (NSP5); for RF, P17468 (VP1), P12472 (VP2), AAM80568 (VP3), P04509 (VP6), P12475 (NSP1), CAA79755 (NSP2), CAA79754 (NSP3), AAM80569 (NSP4), and AAG15311 (NSP5); for Uk(tc), CAA39085 (VP1), P17462 (VP2), AAQ74387 (VP3), CAA37708 (VP6), CAA78093 (NSP1), P03538 (NSP2), AAA47317 (NSP3), P04513 (NSP4), and AAA47289 (NSP5); for I321, CAA64324 (VP6), AAA50481 (NSP1), CAA88010 (NSP2), CAA57192 (NSP3), and AAD50791 (NSP4); for YM, CAA54024 (VP1), AAQ74386 (VP3), CAA49243 (VP6), BAA20545 (NSP1), CAA49240 (NSP4), and CAA49241 (NSP5).

^b Species of origin of these strains were indicated by Hu (human), Si (simian), La (lapine), Po (porcine), and Bo (bovine).

^c NS, not sequenced. High amino acid identities are in boldface type.

^d Percent identity based on partial available gene sequences.

G3P[8], and G9P[8] RV strains isolated in Malawi, Taiwan, People’s Republic of China, India, and Denmark (97.8 to 98.6%) (Fig. 1E).

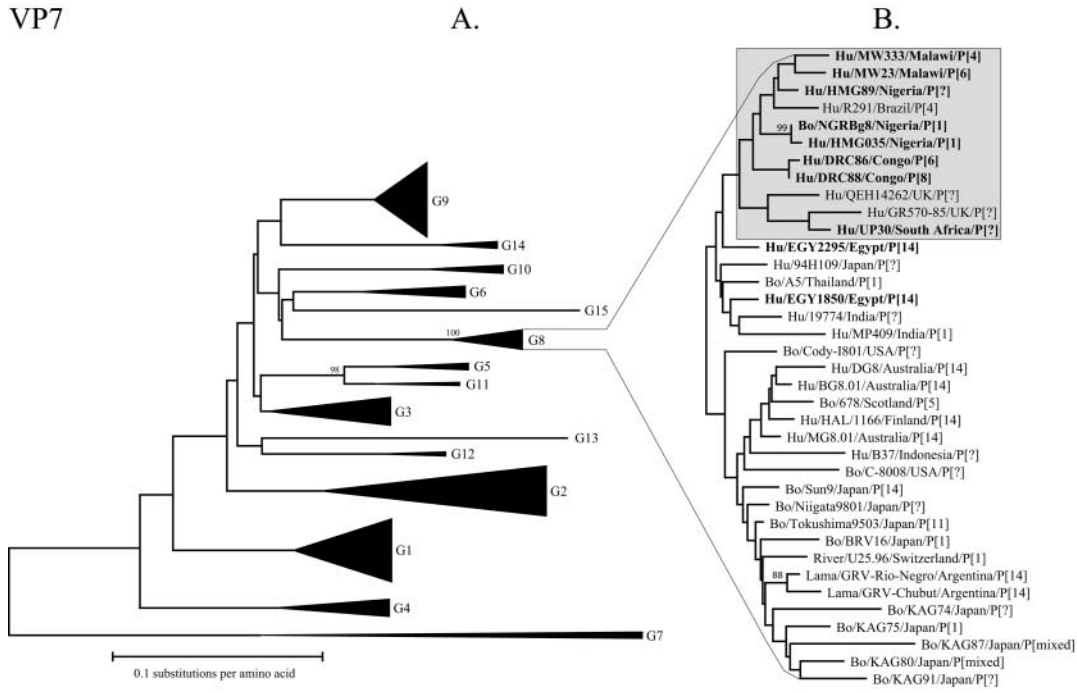
Genogrouping analysis of whole genomes. Initial comparisons of the other nine gene segments (encoding VP1, VP2, VP3, VP6, NSP1, NSP2, NSP3, NSP4, and NSP5) of strains DRC88 and DRC86 with the corresponding genes of other (nearly) completely sequenced RV strains found in GenBank revealed a very high degree of similarity between DRC88 and DRC86 and RV strains TB-Chen, DS-1, NR-1, S2, and IS2. Their VP4 and VP7 genes were only distantly related. Table 1 shows the comparison of these nine gene segments of the DRC88 and DRC86 strains with several RV reference strains (Ku, Wa, TB-Chen, SA11, B4106, 30/96, OSU, RF, UK, I321, and YM) and some of the partially sequenced, very closely related strains (DS-1, NR1, S2, and IS2). Note that all the strains that show a high degree of amino acid similarity with our DRC88 and DRC86 strains belong to the DS-1-like (G2P[4]) genogroup (TB-Chen, DS-1, NR-1, S2, and IS2) (Table 1). Phylogenetic dendrograms constructed with the complete amino acid sequences of these gene segments are shown in Fig. 2 and 3. The very close relationship between DRC88 and DRC86 and RV strains belonging to the DS-1-like genogroup is again reflected in the dendrograms of the different segments. A visual summary of these similarities is shown in Fig. 4.

NSP5 gene analysis. An unusually large NSP5 gene segment (816 nucleotides) was detected in both DRC88 and DRC86. This larger gene size was due to an AT-rich insertion into the 3’ end of the gene just behind the stop codon of the NSP5 protein. A very similar gene structure has been described previously for RV strains isolated in Argentina (v47, v51, v61, v115, v158, and v252), India (NR1), People’s Republic of China (TB-Chen), Japan (KUN), Australia (RV-5), and the United States (DS-1) (23, 39, 47). Percentages of nucleotide identity between the DRC88 and DRC86 NSP5 genes and the NSP5 gene of these strains ranged from 89.6% for the Argentinean, Australian, and American strains to 97.1% for the Japanese, Chinese, and Indian strains. These strains also clustered very closely together in the NSP5 phylogenetic dendrograms shown in Fig. 3.

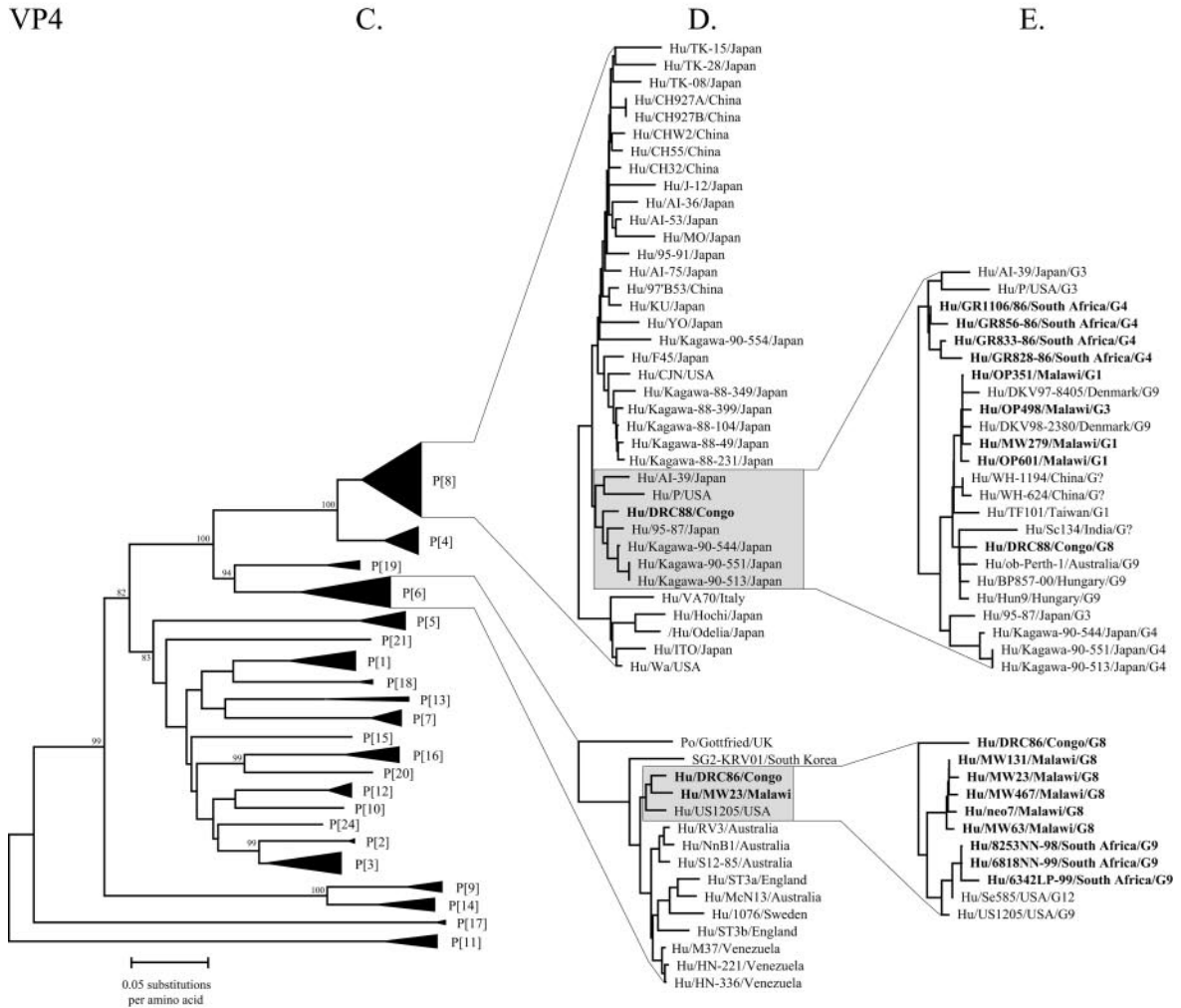
DISCUSSION

To our knowledge, this is the first study to report full genome sequence data for RVs from the Democratic Republic of Congo. Several G8P[6] and G8P[8] HRV strains were detected by partial sequencing of their VP7 and VP4 genes. In order to investigate their full genomic relatedness and their relationship to other known RVs, the complete genomes of DRC88 (G8P[8]) and DRC86 (G8P[6]) RV strains were determined. A very high degree of sequence identity (>98.9% at the nucleo-

VP7



VP4



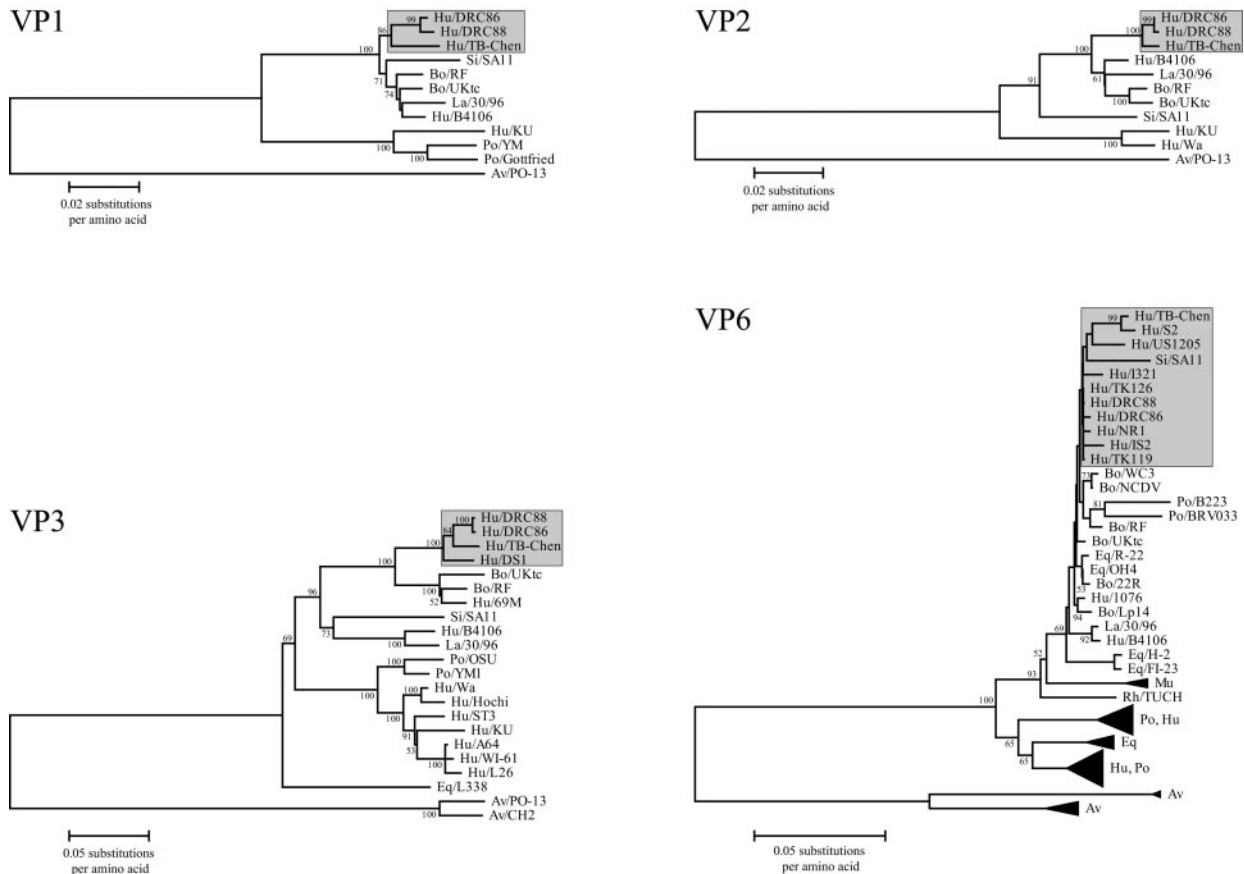


FIG. 2. Phylogenetic dendrograms based on the full-length amino acid sequences of RV VP1, VP2, VP3, and VP6 genes. Bootstrap values (2,000 replicates) above 50 are shown. Several clusters are replaced by a triangle for simplicity, and the host species are provided (Av, avian; Hu, human; Bo, bovine; Eq, equine; Po, porcine; Rh, rhesus; Si, simian; Mu, murine). Phylogenetic clusters containing DRC88 and DRC86 are boxed, and for each strain, the following data are given: species of origin/strain name.

tide level and >98.7% at the amino acid level) was found for 10 out of 11 gene segments. The VP4 genes of DRC88 (P[8]) and DRC86 (P[6]) were only distantly related (70.2% at the nucleotide level and 76.7% at the amino acid level). G8P[6] HRVs have recently emerged in Africa and have been isolated in Malawi, South Africa, Ghana, and Guinea-Bissau (9, 14, 18, 46, 63). On the other hand, G8P[8] HRVs were mentioned only very briefly in epidemiological studies in the United Kingdom, India, and Ghana between 1997 and 1999 and in Guinea-Bissau in 2002 (9, 30, 33, 46). Our findings suggests that a reassortment event has taken place recently between a common African G8P[6] HRV strain (a recent ancestor of DRC86) (7, 12, 13) and an RV strain with P[8] specificity, resulting in a novel strain with the G8P[8] gene combination. The VP4 phylogenetic tree (Fig. 1B) clustered DRC88 with G9P[8] RV

strains isolated in Australia and Hungary, suggesting that a G9P[8] RV strain might have been the donor of the P[8] VP4 gene segment of DRC88.

RNA-RNA hybridization assays are used to classify RVs into genogroups based on their 11 genome segments. Most HRVs can be assigned to one of two major genogroups, Wa or DS-1, or to one minor genogroup, AU-1, on the basis of overall genomic RNA homology (43). The same technique has been applied to define genogroups for several animal RVs (42). When the full amino acid sequences of the DRC88 and DRC86 gene fragments encoding VP1, VP2, VP3, VP6, NSP1, NSP2, NSP3, NSP4, and NSP5 are compared with those of several other (almost) completely sequenced RV strains (Table 1), a very high degree of amino acid identity among DRC88, DRC86, and DS-1-like RV strains TB-Chen, DS-1,

FIG. 1. (A) Phylogenetic dendrogram based on the full-length amino acid sequences of RV VP7 genes. Each G type is depicted as a triangle, and its size is an indication of the number of sequences it represents. (B) Detail of the phylogenetic relationship between G8 genotype RV strains. (C) Phylogenetic dendrogram based on the full-length amino acid sequences of the RV VP4 gene. (D) Detailed phylogenetic relationship between DRC88 and DRC86 and other RV strains with the P[8] and P[6] genotypes. (E) Second phylogenetic detail based on the partial amino acid sequence (amino acids 12 to 286) of the rotaviral VP4 gene. Bootstrap values (2,000 replicates) above 50 are shown. For each strain, the following data are given: species of origin/strain name/place of isolation/G or P type (Hu, human; Po, porcine; Bo, bovine). RV strains isolated in Africa are in boldface type.

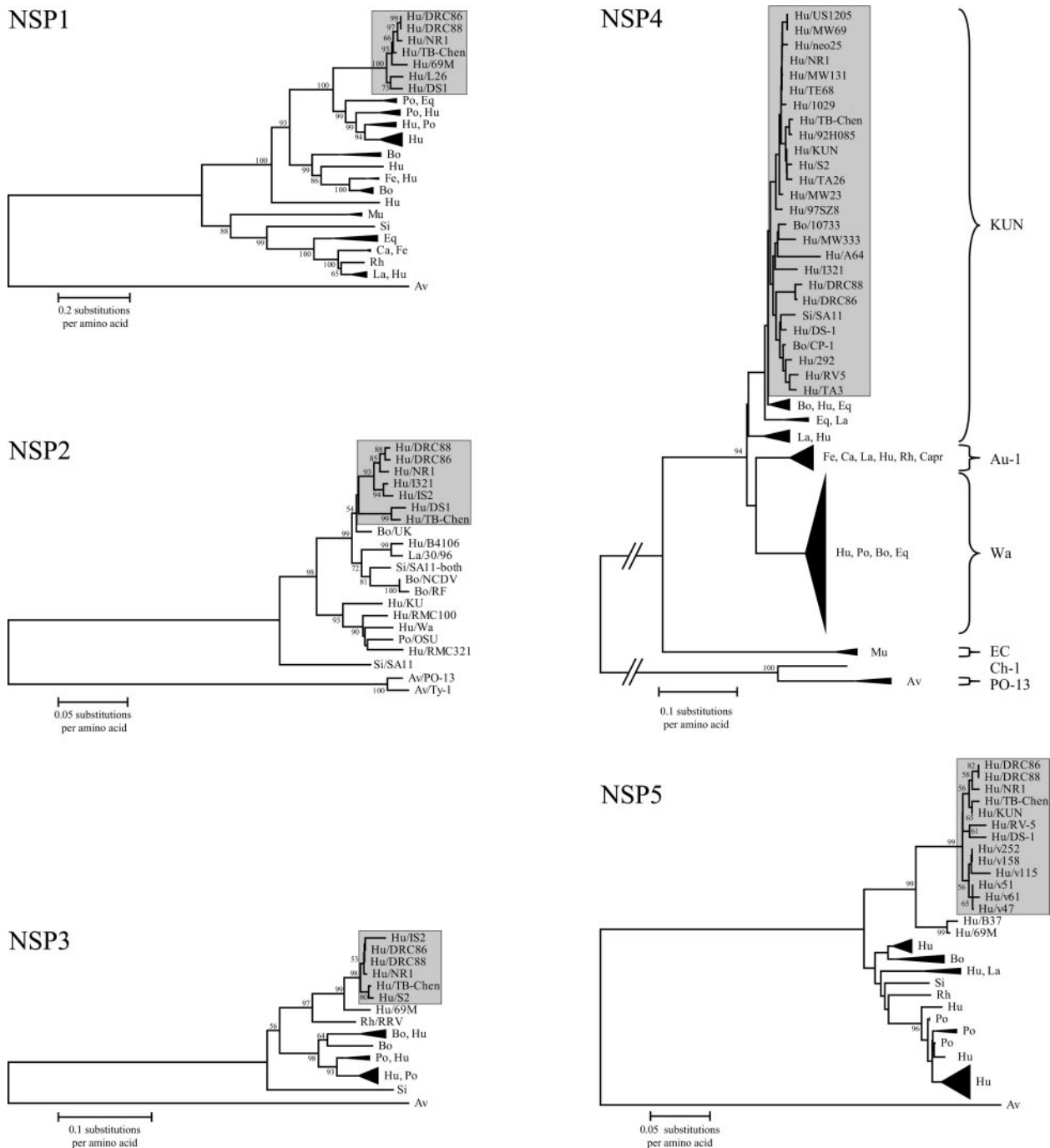


FIG. 3. Phylogenetic dendrograms based on the full-length amino acid sequences of RV NSP1, NSP2, NSP3, NSP4, and NSP5 genes. Bootstrap values (2,000 replicates) above 50 are shown. Several clusters are replaced by a triangle for simplicity, and the host species are provided (Av, avian; Hu, human; Bo, bovine; Eq, equine; Po, porcine; Rh, rhesus; Si, simian; Mu, murine; Fe, feline; Ca, canine; La, lapine; Capr, caprine). Phylogenetic clusters containing DRC88 and DRC86 are boxed, and for each strain, the following data are given: species of origin/strain name. For NSP4, the six established genogroups are also shown.

NR-1, S2, and IS2 was seen (Table 1). Therefore, DRC88 and DRC86 are designated DS-1-like HRVs. The close relationship between the G8P[8] and G8P[6] strains described in this study and the common G2P[4] HRVs is also reflected in the different phylogenetic dendrograms (Fig. 2 and 3). These findings indicate that strains DRC88 (G8P[8]) and DRC86

(G8P[6]), which at first sight did not seem to be related to the G2P[4] HRV strains at all, are actually very closely related to these strains.

Figure 4 shows an example of possible events that could have led to the genetic relationships between different HRVs and their gene segments and shows how the G8P[6] and G8P[8]

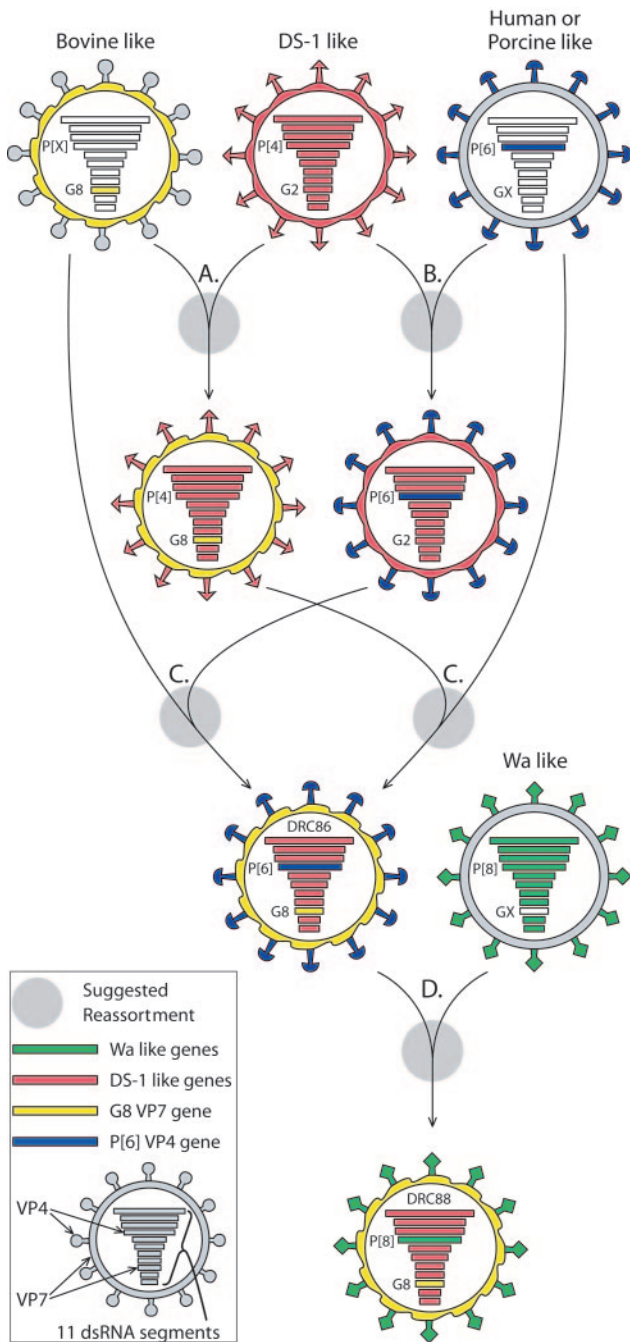


FIG. 4. Schematic representation of the several segmented rotavirus genomes. Suggested RV genomic reassortment events responsible for the occurrence of G2P[6], G8P[4], G8P[6], and G8P[8] HRV strains are depicted. Segments belonging to the DS-1-like genogroup based on sequence data, RNA-RNA hybridizations, or electropherotype are depicted in red, segments belonging to the Wa-like genogroup are shown in green, G8 is shown in yellow, and P[6] is shown in blue.

HRV strains might have occurred after several gene reassortments between RVs belonging to different genogroups. Wa-like and DS-1-like HRV strains are depicted in red and green, respectively. (G8P[4] and G2P[6] HRV strains were designated DS-1 like by RNA-RNA hybridization or electropherotype.) The G8 specificity (shown in yellow) is believed to be of

bovine origin, and it has been suggested that G8 was introduced into HRV strains during several separate reassortment events (1, 10, 11, 31, 44, 63). The P[6] specificity (shown in blue) is found in humans and pigs (38) and is present in more than a quarter of the HRV isolates found on the African continent (59). DS-1-like HRV isolates with G8P[4] specificity have been found on several occasions on the African continent (12, 13, 18, 44, 46, 63) and most likely acquired the G8 specificity through a reassortment event between a DS-1-like HRV and an RV with the G8 specificity (Fig. 4A). HRVs with DS-1-like features and the G2P[6] specificity were detected in 1997 in Brazil and South Africa (37, 50) and since then were found on several occasions on the African continent (in Ghana, Nigeria, Burkina Faso, and Guinea-Bissau) (3, 7, 46, 50). The P[6] specificity of these DS-1 like HRVs presumably arose after one or more reassortment events involving DS-1-like HRVs and RVs with the P[6] specificity (Fig. 4B). Both these HRVs with the G8P[4] and G2P[6] specificities are good candidates for being the possible ancestors of the DS-1-like G8P[6] RVs found in this study and in previous studies in Malawi, South Africa, Ghana, and Guinea-Bissau (Fig. 4C) (7, 9, 12–14, 18, 46, 63). This hypothesis is strengthened by the observation that in several studies, G6P[8] HRVs were shown to cocirculate with G8P[4] or G2P[6] HRV strains in the same area (7, 12, 13, 18, 46). The G8P[8] (DRC88) HRV described in this study has been shown to be highly similar in 10 out of the 11 gene segments compared to G8P[6] (DRC86), which was isolated in the same area. These data suggest that the reassortment that introduced the P[8] specificity (typical for Wa-like HRVs) into a DS-1-like G8P[6] HRV background, generating a G8P[8] HRV (Fig. 4D), is most likely a very recent event, different from the reassortment(s) that generated previously described G8P[8] HRV strains.

After the isolation of RV strains with an unusually large NSP5 gene, due to the insertion of a long noncoding AT-rich region in their 3' end, in Asia (China, Japan, India, and Australia), South America (Argentina), and North America (United States) (23, 39, 47), DRC88 and DRC86 are the first such HRV strains isolated and sequenced on the African continent. The first HRV strains with this rearranged NSP5 gene were isolated in 1977 (4, 32), and they have been isolated all over the world since then. Until now, they have always shown the G2 specificity. This again strengthens the hypothesis that DRC88 and DRC86 and G2P[4] RV strains belonging to the DS-1 genogroup have a common ancestor.

The fact that RV genotypes, which are normally uncommon in humans (G8 is of bovine origin, and P[6] is of porcine origin), are frequently detected in developing countries is most likely due to the close proximity in which humans and animals often live in these regions. Several studies also have indicated that mixed infections are more common in developing countries than in developed countries (46, 59). These are two important factors that favor the occurrence of unusual RV strains through reassortment in developing countries.

Although RNA-RNA hybridization is a fast way of comparing RV genomes and has been used successfully for the characterization of many human and animal (reassortant) RV strains, it is likely that with the increasing amounts of full RV genomes available in databases and the increasing sequencing possibilities in laboratories around the world, RNA-RNA hy-

bridizations will be gradually replaced by sequencing of RV genomes. As sequencing will become the new standard technique, RV genome analyses, as performed in this study, will be a powerful tool to investigate RVs and their genetic and evolutionary relationships.

The replacement of the P[6] specificity in DRC86 (G8P[6]), which is believed to be of animal origin, with the P[8] specificity, very well adapted to infect humans, might make this G8P[8] RV strain a successful human pathogen in a way similar to that described for the G9P[8] HRV strains all over the world (29, 55). This G8P[8] combination should be taken into account by vaccine companies, although the P[8] moiety included in several vaccines currently under development might convey some level of protection against severe gastroenteritis due to G8P[8] RV infection. The findings of this study also indicate that the typing of new RV strains by sequencing of their VP7 and VP4 genes alone will reveal only a small portion of the story. The G and P types of an RV strain are only the tip of the iceberg regarding RV diversity. Therefore, more attention should be paid to the full genomes of circulating RVs, and RV surveillance should not be restricted to humans.

ACKNOWLEDGMENTS

We thank all the colleagues of the Laboratory of Clinical and Epidemiological Virology, Department of Microbiology and Immunology, Rega Institute for Medical Research, University of Leuven, Belgium, for their helpful comments and discussion.

J.M. was supported by the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT Vlaanderen).

REFERENCES

- Adah, M. I., S. Nagashima, M. Wakuda, and K. Taniguchi. 2003. Close relationship between G8-serotype bovine and human rotaviruses isolated in Nigeria. *J. Clin. Microbiol.* **41**:3945–3950.
- Adah, M. I., A. Rohwedder, O. D. Olaleye, and H. Werchau. 1997. Nigerian rotavirus serotype G8 could not be typed by PCR due to nucleotide mutation at the 3' end of the primer binding site. *Arch. Virol.* **142**:1881–1887.
- Adah, M. I., A. Wade, and K. Taniguchi. 2001. Molecular epidemiology of rotaviruses in Nigeria: detection of unusual strains with G2P[6] and G8P[1] specificities. *J. Clin. Microbiol.* **39**:3969–3975.
- Albert, M. J., and R. F. Bishop. 1984. Cultivation of human rotaviruses in cell culture. *J. Med. Virol.* **13**:377–383.
- Albert, M. J., L. E. Unicomb, and R. F. Bishop. 1987. Cultivation and characterization of human rotaviruses with "super short" RNA patterns. *J. Clin. Microbiol.* **25**:183–185.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403–410.
- Armah, G. E., C. T. Payer, R. H. Asmah, F. R. Anto, A. R. Oduro, F. Binka, and D. Steele. 2001. Prevalence of unusual human rotavirus strains in Ghanaian children. *J. Med. Virol.* **63**:67–71.
- Beards, G., and C. Graham. 1995. Temporal distribution of rotavirus G-serotypes in the West Midlands region of the United Kingdom, 1983–1994. *J. Diarrhoeal Dis. Res.* **13**:235–237.
- Binka, F. N., F. K. Anto, A. R. Oduro, E. A. Awini, A. K. Nazzar, G. E. Armah, R. H. Asmah, A. J. Hall, F. Cutts, N. Alexander, D. Brown, J. Green, J. Gray, and M. Iturriza-Gomara. 2003. Incidence and risk factors of paediatric rotavirus diarrhoea in northern Ghana. *Trop. Med. Int. Health* **8**:840–846.
- Browning, G. F., D. R. Snodgrass, O. Nakagomi, E. Kaga, A. Sarasini, and G. Gerna. 1992. Human and bovine serotype G8 rotaviruses may be derived by reassortment. *Arch. Virol.* **125**:121–128.
- Cooney, M. A., R. J. Gorrell, and E. A. Palombo. 2001. Characterisation and phylogenetic analysis of the VP7 proteins of serotype G6 and G8 human rotaviruses. *J. Med. Microbiol.* **50**:462–467.
- Cunliffe, N. A., J. R. Gentsch, C. D. Kirkwood, J. S. Gondwe, W. Dove, O. Nakagomi, T. Nakagomi, Y. Hoshino, J. S. Bresee, R. I. Glass, M. E. Molyneux, and C. A. Hart. 2000. Molecular and serologic characterization of novel serotype G8 human rotavirus strains detected in Blantyre, Malawi. *Virology* **274**:309–320.
- Cunliffe, N. A., J. S. Gondwe, R. L. Broadhead, M. E. Molyneux, P. A. Woods, J. S. Bresee, R. I. Glass, J. R. Gentsch, and C. A. Hart. 1999. Rotavirus G and P types in children with acute diarrhea in Blantyre, Malawi, from 1997 to 1998: predominance of novel P[6]G8 strains. *J. Med. Virol.* **57**:308–312.
- Cunliffe, N. A., S. Rogerson, W. Dove, B. D. Thindwa, J. Greensill, C. D. Kirkwood, R. L. Broadhead, and C. A. Hart. 2002. Detection and characterization of rotaviruses in hospitalized neonates in Blantyre, Malawi. *J. Clin. Microbiol.* **40**:1534–1537.
- Desselberger, U. 1996. Genome rearrangements of rotaviruses. *Arch. Virol. Suppl.* **12**:37–51.
- Estes, M. K. 2004. Rotaviruses and their replication, p. 1747–1786. *In* D. M. Knipe, P. M. Howley, D. E. Griffin, R. A. Lamb, M. A. Martin, B. Roizman, and S. E. Straus (ed.), *Fields virology*, 4th ed. Lippincott Williams & Wilkins, Philadelphia, Pa.
- Falcone, E., M. Tarantino, L. Di Trani, P. Cordioli, A. Lavazza, and M. Tollis. 1999. Determination of bovine rotavirus G and P serotypes in Italy by PCR. *J. Clin. Microbiol.* **37**:3879–3882.
- Fischer, T. K., N. A. Page, D. D. Griffin, J. Eugen-Olsen, A. G. Pedersen, P. Valentiner-Branth, K. Molbak, H. Sommerfelt, and N. M. Nielsen. 2003. Characterization of incompletely typed rotavirus strains from Guinea-Bissau: identification of G8 and G9 types and a high frequency of mixed infections. *Virology* **311**:125–133.
- Fischer, T. K., H. Steinsland, K. Molbak, R. Ca, J. R. Gentsch, P. Valentiner-Branth, P. Aaby, and H. Sommerfelt. 2000. Genotype profiles of rotavirus strains from children in a suburban community in Guinea-Bissau, Western Africa. *J. Clin. Microbiol.* **38**:264–267.
- Fukai, K., T. Sakai, M. Hirose, and T. Itou. 1999. Prevalence of calf diarrhea caused by bovine group A rotavirus carrying G serotype 8 specificity. *Vet. Microbiol.* **66**:301–311.
- Gerna, G., A. Sarasini, L. Zentilin, A. Di Matteo, P. Miranda, M. Parea, M. Battaglia, and G. Milanesi. 1990. Isolation in Europe of 69 M-like (serotype 8) human rotavirus strains with either subgroup I or II specificity and a long RNA electrophoretotype. *Arch. Virol.* **112**:27–40.
- Gerna, G., J. Sears, Y. Hoshino, A. D. Steele, O. Nakagomi, A. Sarasini, and J. Flores. 1994. Identification of a new VP4 serotype of human rotaviruses. *Virology* **200**:66–71.
- Giambiagi, S., I. Gonzalez Rodriguez, J. Gomez, and O. Burrone. 1994. A rearranged genomic segment 11 is common to different human rotaviruses. *Arch. Virol.* **136**:415–421.
- Gouvea, V., N. Santos, and M. C. Timenetsky. 1994. Identification of bovine and porcine rotavirus G types by PCR. *J. Clin. Microbiol.* **32**:1338–1340.
- Hasegawa, A., S. Inouye, S. Matsuno, K. Yamaoka, R. Eko, and W. Suhayono. 1984. Isolation of human rotaviruses with a distinct RNA electrophoretic pattern from Indonesia. *Microbiol. Immunol.* **28**:719–722.
- Holmes, J. L., C. D. Kirkwood, G. Gerna, J. D. Clemens, M. R. Rao, A. B. Naficy, R. Abu-Elyazed, S. J. Savarino, R. I. Glass, and J. R. Gentsch. 1999. Characterization of unusual G8 rotavirus strains isolated from Egyptian children. *Arch. Virol.* **144**:1381–1396.
- Hoshino, Y., S. Honma, R. W. Jones, N. Santos, O. Nakagomi, T. Nakagomi, A. Z. Kapikian, and M. E. Thoulless. 2005. A rotavirus strain isolated from pig-tailed macaque (*Macaca nemestrina*) with diarrhea bears a P6[1]:G8 specificity. *Virology* **345**:1–12.
- Isa, P., A. R. Wood, T. Netherwood, M. Ciarlet, H. Imagawa, and D. R. Snodgrass. 1996. Survey of equine rotaviruses shows conservation of one P genotype in background of two G genotypes. *Arch. Virol.* **141**:1601–1612.
- Iturriza-Gomara, M., D. Cubitt, D. Steele, J. Green, D. Brown, G. Kang, U. Desselberger, and J. Gray. 2000. Characterisation of rotavirus G9 strains isolated in the UK between 1995 and 1998. *J. Med. Virol.* **61**:510–517.
- Iturriza-Gomara, M., B. Isherwood, U. Desselberger, and J. Gray. 2001. Reassortment in vivo: driving force for diversity of human rotavirus strains isolated in the United Kingdom between 1995 and 1999. *J. Virol.* **75**:3696–3705.
- Jagannath, M. R., R. V. Vethanayagam, B. S. Reddy, S. Raman, and C. D. Rao. 2000. Characterization of human symptomatic rotavirus isolates MP409 and MP480 having 'long' RNA electrophoretotype and subgroup I specificity, highly related to the P6[1]:G8 type bovine rotavirus A5, from Mysore, India. *Arch. Virol.* **145**:1339–1357.
- Kalica, A. R., M. M. Sereno, R. G. Wyatt, C. A. Mebus, R. M. Chanock, and A. Z. Kapikian. 1978. Comparison of human and animal rotavirus strains by gel electrophoresis of viral RNA. *Virology* **87**:247–255.
- Kang, G., J. Green, C. I. Gallimore, and D. W. Brown. 2002. Molecular epidemiology of rotaviral infection in South Indian children with acute diarrhea from 1995–1996 to 1998–1999. *J. Med. Virol.* **67**:101–105.
- Kapikian, A. Z., Y. Hoshino, and R. M. Chanock. 2001. Rotaviruses, p. 1787–1834. *In* D. M. Knipe, P. M. Howley, D. E. Griffin, R. A. Lamb, M. A. Martin, B. Roizman, and S. E. Straus (ed.), *Fields virology*, 4th ed. Lippincott Williams & Wilkins, Philadelphia, Pa.
- Kumar, S., K. Tamura, I. B. Jakobsen, and M. Nei. 2001. MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics* **17**:1244–1245.
- Lambden, P. R., S. J. Cooke, E. O. Caul, and I. N. Clarke. 1992. Cloning of noncultivable human rotavirus by single primer amplification. *J. Virol.* **66**:1817–1822.
- Linhares, A. C., J. D. Mascarenhas, R. H. Gusmao, Y. B. Gabbay, A. M. Fialho, and J. P. Leite. 2002. Neonatal rotavirus infection in Belem, northern Brazil: nosocomial transmission of a P[6] G2 strain. *J. Med. Virol.* **67**:418–426.

- 37a. **Martella, V., M. Ciarlet, K. Banyai, E. Lorusso, A. Cavalli, and M. Corrente.** 2006. Identification of a novel VP4 genotype carried by a G5 porcine RV. *Virology* **346**:301–311.
38. **Martella, V., K. Banyai, M. Ciarlet, M. Iturriza-Gomara, E. Lorusso, S. De Grazia, S. Arista, N. Decaro, G. Elia, A. Cavalli, M. Corrente, A. Lavazza, R. Baselga, and C. Buonavoglia.** 2005. Relationships among porcine and human P[6] rotaviruses: evidence that the different human P[6] lineages have originated from multiple interspecies transmission events. *Virology* **344**:509–519.
39. **Matsui, S. M., E. R. Mackow, S. Matsuno, P. S. Paul, and H. B. Greenberg.** 1990. Sequence analysis of gene 11 equivalents from “short” and “super short” strains of rotavirus. *J. Virol.* **64**:120–124.
40. **Matthijssens, J., M. Rahman, and M. Van Ranst.** 2006. Loop model: mechanism to explain partial gene duplications in segmented dsRNA viruses. *Biochem. Biophys. Res. Commun.* **340**:140–144.
41. **Maunula, L., and C. H. von Bonsdorff.** 2002. Frequent reassortments may explain the genetic heterogeneity of rotaviruses: analysis of Finnish rotavirus strains. *J. Virol.* **76**:11793–11800.
42. **Nakagomi, O., and T. Nakagomi.** 2002. Genomic relationships among rotaviruses recovered from various animal species as revealed by RNA-RNA hybridization assays. *Res. Vet. Sci.* **73**:207–214.
43. **Nakagomi, O., T. Nakagomi, K. Akatani, and N. Ikegami.** 1989. Identification of rotavirus genogroups by RNA-RNA hybridization. *Mol. Cell. Probes* **3**:251–261.
44. **Nakata, S., Z. Gatheru, S. Ukae, N. Adachi, N. Kobayashi, S. Honma, J. Muli, P. Ogaja, J. Nyangao, E. Kiplagat, P. M. Tukei, and S. Chiba.** 1999. Epidemiological study of the G serotype distribution of group A rotaviruses in Kenya from 1991 to 1994. *J. Med. Virol.* **58**:296–303.
45. **Nicholas, K. B., H. B. Nicholas, and D. W. Deerfield.** 1997. GeneDoc: analysis and visualization of genetic variation. *Embnet News* **4**:14.
46. **Nielsen, N. M., J. Eugen-Olsen, P. Aaby, K. Molbak, A. Rodrigues, and T. K. Fischer.** 2005. Characterisation of rotavirus strains among hospitalised and non-hospitalised children in Guinea-Bissau, 2002. A high frequency of mixed infections with serotype G8. *J. Clin. Virol.* **34**:13–21.
47. **Nuttall, S. D., C. P. Hum, I. H. Holmes, and M. L. Dyall-Smith.** 1989. Sequences of VP9 genes from short and supershort rotavirus strains. *Virology* **171**:453–457.
48. **Oka, T., T. Nakagomi, and O. Nakagomi.** 2000. Apparent re-emergence of serotype G9 in 1995 among rotaviruses recovered from Japanese children hospitalized with acute gastroenteritis. *Microbiol. Immunol.* **44**:957–961.
49. **Okada, N., and Y. Matsumoto.** 2002. Bovine rotavirus G and P types and sequence analysis of the VP7 gene of two G8 bovine rotaviruses from Japan. *Vet. Microbiol.* **84**:297–305.
50. **Page, N. A., and A. D. Steele.** 2004. Antigenic and genetic characterization of serotype G2 human rotavirus strains from the African continent. *J. Clin. Microbiol.* **42**:595–600.
51. **Parra, G. I., K. Bok, M. Martinez, and J. A. Gomez.** 2004. Evidence of rotavirus intragenic recombination between two sublineages of the same genotype. *J. Gen. Virol.* **85**:1713–1716.
52. **Parreno, V., K. Bok, F. Fernandez, and J. Gomez.** 2004. Molecular characterization of the first isolation of rotavirus in guanacos (*Lama guanicoe*). *Arch. Virol.* **149**:2465–2471.
53. **Qian, Y., and K. Y. Green.** 1991. Human rotavirus strain 69M has a unique VP4 as determined by amino acid sequence analysis. *Virology* **182**:407–412.
54. **Rahman, M., T. Goegebuier, K. De Leener, P. Maes, J. Matthijssens, G. Podder, T. Azim, and M. Van Ranst.** 2004. Chromatography paper strip method for collection, transportation, and storage of rotavirus RNA in stool samples. *J. Clin. Microbiol.* **42**:1605–1608.
55. **Rahman, M., J. Matthijssens, T. Goegebuier, K. De Leener, L. Vanderwegen, I. van der Donck, L. Van Hoovels, S. De Vos, T. Azim, and M. Van Ranst.** 2005. Predominance of rotavirus G9 genotype in children hospitalized for rotavirus gastroenteritis in Belgium during 1999–2003. *J. Clin. Virol.* **33**:1–6.
56. **Rahman, M., J. Matthijssens, S. Nahar, G. Podder, D. A. Sack, T. Azim, and M. Van Ranst.** 2005. Characterization of a novel P[25]G11 human group A rotavirus. *J. Clin. Microbiol.* **43**:3208–3212.
57. **Ramig, R. F.** 1997. Genetics of the rotaviruses. *Annu. Rev. Microbiol.* **51**:225–255.
58. **Rao, C. D., K. Gowda, and B. S. Reddy.** 2000. Sequence analysis of VP4 and VP7 genes of nontypeable strains identifies a new pair of outer capsid proteins representing novel P and G genotypes in bovine rotaviruses. *Virology* **276**:104–113.
59. **Santos, N., and Y. Hoshino.** 2005. Global distribution of rotavirus serotypes/genotypes and its implication for the development and implementation of an effective rotavirus vaccine. *Rev. Med. Virol.* **15**:29–56.
60. **Santos, N., R. C. Lima, C. F. Pereira, and V. Gouvea.** 1998. Detection of rotavirus types G8 and G10 among Brazilian children with diarrhea. *J. Clin. Microbiol.* **36**:2727–2729.
61. **Sato, M., T. Nakagomi, K. Tajima, K. Ezura, H. Akashi, and O. Nakagomi.** 1997. Isolation of serotype G8, P[6] bovine rotavirus from adult cattle with diarrhea. *J. Clin. Microbiol.* **35**:1266–1268.
62. **Steele, A. D., and B. Ivanoff.** 2003. Rotavirus strains circulating in Africa during 1996–1999: emergence of G9 strains and P[6] strains. *Vaccine* **21**:361–367.
63. **Steele, A. D., S. P. Parker, I. Peenze, C. T. Pager, M. B. Taylor, and W. D. Cubitt.** 1999. Comparative studies of human rotavirus serotype G8 strains recovered in South Africa and the United Kingdom. *J. Gen. Virol.* **80**:3029–3034.
64. **Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins.** 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**:4876–4882.