High Prevalence of Human Enterovirus A Infections in Natural Circulation of Human Enteroviruses

Elisabet Witsø,1* Gustavo Palacios,2 Ondrej Cinek,3 Lars C. Stene,1 Bjørn Grinde,1 Diana Janowitz,2 W. Ian Lipkin,2 and Kjersti S. Rønningen1

Norwegian Institute of Public Health, Oslo, Norway; Jerome L. and Dawn Greene Infectious Disease Laboratory, Mailman School of Public Health, Columbia University, New York, New York2; and Motol University Hospital, Second Medical School, Charles University in Prague, Prague, Czech Republic3

Received 27 March 2006/Returned for modification 22 June 2006/Accepted 21 August 2006

Human enterovirus (HEV) infections can be asymptomatic or cause only mild illness; recent evidence may implicate HEV infection in type 1 diabetes mellitus and myocarditis. Here, we report the molecular characterization of HEV obtained in serial monthly collections from healthy Norwegian infants. A total of 1,255 fecal samples were collected from 113 healthy infants beginning at age 3 months and continuing to 28 months. The samples were analyzed for HEV nucleic acid by real-time PCR. Fifty-eight children (51.3%) had HEV infections. One hundred forty-five positive samples were typed directly by nucleotide sequencing of the VP1 region. HEV-A was detected most frequently, with an overall prevalence of 6.8%. HEV-B was present in 4.8% of the samples and HEV-C in only 0.2% of the samples. No poliovirus or HEV-D group viruses were detected. Twenty-two different serotypes were detected in the study period: the most common were EV71 (14.5%), CAV6 (10.5%), CAV4 (8.9%), E18 (8.9%), and CBV3 (7.3%). These findings suggest that the prevalence of HEV infections in general, and HEV-A infections in particular, has been underestimated in epidemiological studies based on virus culture.

The genus Enterovirus belongs to the family Picomaviridae. Although the majority of Human enterovirus (HEV) infections are asymptomatic, enterovirus infections can cause upper respiratory illness, febrile rash, aseptic meningitis, pleuropneumonia, encephalitis, acute flaccid paralysis, and neonatal sepsis-like disease (35). HEV may also be implicated in the pathogenesis of severe chronic diseases, including type 1 diabetes mellitus (33), myocarditis and congestive cardiomyopathy (20), and neuromuscular diseases (8).

HEVs contain a linear (7.4-kb) single-stranded RNA genome comprising a 5’ and a 3’ noncoding region and a single, long open reading frame coding for a polyprotein of about 2,200 amino acids. HEVs are classified, based on molecular and biological properties, into five species (19, 36): (i) Poliovirus (types 1 to 3), (ii) Human enterovirus A (HEV-A) (CAV2 to CAV8, CAV10, CAV12, CAV14, CAV16, and EV71), (iii) Human enterovirus B (HEV-B) (CAV9, CBV1 to CBV6, E17, E9, E11 to E21, E24 to E27, E29 to E33, and EV69), (iv) Human enterovirus C (HEV-C) (CAV1, CAV11, CAV13, CAV15, CAV17 to CAV22, and CAV24), and (v) Human enterovirus D (HEV-D) (EV68 and EV70). Several new serotypes (EV73 to EV78 and EV89 to EV91) were recently described (28, 29, 31, 32).

Neutralization tests of cultured virus may not be sufficient to identify some serotypes (26). Many coxsackievirus A strains can be isolated and propagated only in suckling mice (24).

We address the need for a robust and universal typing system, many investigators have turned to the use of molecular methods (4, 30, 41). The introduction of molecular sequencing has extended the potential of virus surveillance and epidemiological studies by facilitating the identification of genotypes (14).

Most previous data on HEV circulation have been obtained from analysis of specimens from individuals with disease. Studies of enterovirus circulation in healthy populations antedate the advent of molecular technologies (12, 13, 16–18, 21, 23, 38, 39).

We have reported that 11.3% of fecal samples obtained prospectively from 113 healthy Norwegian infants were positive for enterovirus RNA (11). Here, we present the results of a comprehensive molecular characterization of these samples.

MATERIALS AND METHODS

Study design and subjects. Participating children were enrolled in a prospective study (MIDIA) to investigate potential environmental triggers of type 1 diabetes mellitus. The MIDIA study recruits newborn infants at their first visit to health care centers. Candidates are screened for the HLA genotype conferring the highest risk of type 1 diabetes (HLA-DRB1*0401-DQA1*03-DQB1*0201) (35) and born at 2.5 months of age or later. Over the period from September 2001 to October 2003, 113 of 116 infants identified as genetically susceptible were enrolled (60 males; 53 females). No more than one child in any nuclear family was enrolled. Beginning at 3 months of age and followed up to 28 months (for the first children recruited), monthly stool samples were obtained. Specimens were collected from September 2001 through November 2003; 96% (1,257/1,036) of the scheduled stool samples were received. Two samples were accidentally destroyed during handling. Eighty-nine percent of the stool samples were collected from children residing in the counties of Akershus (southwest Norway) and Hordaland (west coast), two counties separated by more than 400 km.

Processing of stool samples. Total nucleic acids were extracted from 140 μl of 1,255 stool suspensions using the QiAamp Viral RNA Mini Kit (QiAGEN). HEV nucleic acid was detected and quantitated using a one-step real-time reverse transcription (RT)-PCR targeting the 5’ nontranslated region (11).
Molecular typing. HEV serotypes of samples positive in real-time PCR were determined by VP1 sequencing of RNA (4, 44). In instances where samples could not be serotyped by this approach (39 of 145 samples), the RNA polymerase region was amplified and sequenced (4). Based on the RNA polymerase sequence, serotype- or species-specific VP1 primers were designed (Table 1) (4, 44). In total, 14 additional nesting primer sets (localized inside the sequence product of the first amplification) were designed, facilitating serotype assignment of 91.7% of the enterovirus-positive samples.

First-round VP1 amplification products obtained by the method of Casas et al. (4) or separate random-hexamer-primed RT uncoupled from the first round of PCR were used as templates with the new sense and antisense primers. PCR was carried out in a total volume of 20 μl with 10 pmol of each primer (Table 1), 1 to 2 μM of the RT-PCR products or cDNA, 2 μl 10× PCR buffer II (Applied Biosystems), 2.5 mmol MgCl2, 0.5 mmol deoxynucleoside triphosphates, and 2.5 U of Taq polymerase (Perkin-Elmer). The thermal-cycling program consisted of 5 min of denaturation at 94°C, followed by 40 cycles of 30 s of denaturation at 94°C, 30 s of annealing at 50 to 55°C, and 1 min of extension at 72°C.

Nucleotide sequence determination and analysis. Direct product sequencing was performed using the amplification primers. Both strands were sequenced by automatic methods (BigDye version 3.1 and ABI PRISM 310 Genetic Analyzer; Applied Biosystems, Foster City, CA). In instances where sequences could not be obtained directly, the products were cloned into a pGEM-T-Easy vector (Promega) and five clones from each sample were sequenced on both strands.

Raw sequence data were analyzed with Sequencher (version 4.2; Gene Codes Corporation, Ann Arbor, MI) to obtain the final consensus sequence. Ambiguous nucleotides were resolved by resequencing them.

Phylogenetic analysis. The VP1 capsid coding sequences determined in this study were included in a phylogenetic analysis with reference strains of all enterovirus serotypes. Multiple sequence alignments were performed with ClustalX version 1.82 (42). Phylogenetic analysis was performed using the Kimura two-parameter model as a model of nucleotide substitution and the neighbor-joining method to reconstruct the phylogenetic tree (MEGA program version 3.0 [22]). The statistical significance of the phylogenies constructed was estimated by bootstrap analysis with 1,000 pseudoreplicate data sets.

Infection episodes and coinfections. A new infection episode was defined as detection of a new genotype after a monthly sample was negative for that genotype. Coinfection was defined as the simultaneous presence of two or more different genotypes in the same stool sample.

Statistical analysis. The software packages used were Stata version 9.0 (Stata Corporation, College Station, TX) and SPSS version 12.0.1 (SPSS, Chicago, IL). The overall prevalence for each species over the whole study period was computed with 95% confidence intervals (CI). The confidence intervals for prevalence were computed based on standard errors estimated by bootstrapping, using each child as a unit to provide a simple adjustment for dependence between repeated infections (15). Relationships of sex, year, season, and geographic location with prevalence for the main HEV species were assessed.

Associations of viral load and duration of infection, sex, age, and enterovirus species were assessed in linear-regression models, using the viral load of the first positive sample in an infection period. The potential dependence in data from repeated measurements for each child was handled in the regression models by generalized estimating equations using the xtgee procedure in STATA (Stata cross-sectional time-series reference manual, release 8, StataCorp, Stata Press, College Station, TX).

The cumulative incidence of first infection during follow-up for individual children at different ages was estimated using Kaplan-Meier survival analysis. Differences between groups were tested using the log rank test. A P value of 0.05 or less was considered to be statistically significant.

Nucleotide sequence accession numbers. The 3′-end VP1 sequences reported here were deposited in the GenBank sequence database under accession no. DQ317159 to DQ317293. The RNA polymerase sequences were reported under accession no. DQ315510 to DQ315562.

RESULTS

Molecular typing and phylogenetic analysis. The HEV type was identified in 92% of the samples (133/145) by VP1 sequencing and in 6.2% (9 cases) by RNA polymerase sequenc-
ing. In total, 97.9% (142/145) of the real-time PCR-positive samples could be assigned to the species level. The VP1 sequencing and phylogenetic analysis allowed further subtyping. The three samples that were not identified had low viral loads. The positive samples represented 124 discrete infection episodes. A total of 22 different serotypes were detected (Table 2). The most prevalent serotypes were EV71 (14.5%), CAV6 (10.5%), CAV4 (8.9%), E18 (8.9%), and CBV3 (7.3%). Whereas CBV3 and E18 were detected throughout the study period, CAV4, CAV5, CAV6, CAV10, and CBV1 were detected only during short intervals. CAV16, E5, E11, E25, E30, E3, E13, and a CAV19-like strain were represented in only one or two infection episodes. The circulation of specific serotypes was not restricted to a single county or municipality.

Individual serotypes of the VP1 sequences were compared phylogenetically with all corresponding sequences of the respective serotypes in GenBank. Three serotypes (of the 22 analyzed) segregated in different clusters (CAV4, CAV9, and E18 [Fig. 1A to C, respectively], indicating that there have been at least two introductions of genotypes into the population for each of these serotypes. Two distinct clusters were observed for the CAV4-positive samples (Fig. 1A). Seven CAV4 sequences clustered together with sequences from Spain (1999) or Japan (2000 to 2004); six sequences segregated with a strain that circulated in Spain in 1998 (Fig. 1A). Some CAV9 sequences (Fig. 1B) were clustered with isolates from Cyprus (2002) and Spain (1999); one clustered with an unrelated cluster closer to an isolate from the United Kingdom (1962 to 1989). E18 sequences clustered with sequences isolated in Japan (2001 to 2004) or Spain (1995 to 1997), the United States (1996), and Sweden (1999) (Fig. 1C). Although more than one strain was found to be cocirculating in certain periods, as seen for CAV4 (Fig. 1A), the two clusters of CAV9 and E18 seem to have been introduced to the cohort at different times (with one exception) (Fig. 1B and C). The circulation of subgroups for each of the serotypes did not appear to be geographically restricted (Fig. 1A to C).

**Prevalence and epidemiology of HEV-A and HEV-B infections.** HEV was detected in 145 of 1,255 samples (11.6% of samples; 95% CI, 9.8% to 13.5%). HEV-A was detected most frequently, with an overall prevalence of 6.8% (95% CI, 5.0% to 8.6%); HEV-B was present in 4.8% of the samples (95% CI, 3.2% to 6.4%). HEV-C was present in only three samples (0.2%). No poliovirus or HEV-D group viruses were detected. Of 124 infection episodes, 70 (57%) were classified as HEV-A; 49 (40%) were classified as HEV-B (Table 2). HEV-B infections were more prevalent than HEV-A infections during the interval from March through August 2002 (2.9% to 4.4% versus 1.4% to 2.6%); however, during the interval from December 2002 through November 2003, HEV-A infections were more prevalent than HEV-B infections (3.9% to 10.4% versus 1.6% to 6.4%). From September through November 2002, HEV-A and HEV-B prevalence rates were similar (7.8%). The highest prevalence of HEV-A was observed from June through August 2003 (10.4%); the highest prevalence of HEV-B was observed from September through November 2002 (7.8%). The characteristic seasonal variation of HEV infections, with peak incidence in late summer and autumn, was seen for both HEV-A and HEV-B (data not shown).

Figure 2 shows the age-specific prevalences of HEV-A and HEV-B infections for children aged 3 to 23 months. The 3-month moving averages of the age-specific prevalences of positive samples are plotted against the ages of the children in months. In both species, the highest prevalence was in the second year of life.

The probability of having at least one HEV infection by the age of 12 months was approximately 40%; 90% of the children had been infected by the age of 2 years (see Fig. S1 in the supplemental material). There were no significant differences in prevalence of HEV-A or HEV-B between boys and girls or between the two main counties included in the study (data not shown).

**Prolonged excretion, viral load, and coinfections.** In 28 of the 124 infection episodes, the same viral serotype was observed in two or more consecutive samples: 23 cases of two consecutive months, 4 cases of three consecutive months, and 1 case of four consecutive months. Prolonged duration of infection was associated with a higher viral load in the first sample, but not with sex, viral species, or season. The initial viral load was on average eightfold higher in infections that persisted for at least 2 months than in single-month episodes ($P = 0.001$). A trend toward shorter duration of infection with increasing age was not statistically significant ($P = 0.078$).

Stool samples from 10 children showed evidence of coinfection with more than one HEV. Eight samples contained two serotypes; two samples had three serotypes. In six samples, the coinfections were limited to HEV-A; in three cases, there was coinfection with HEV-A and HEV-B; in one case, there was...
coinfection with HEV-A and HEV-C. The viral loads in the coinfected samples were not significantly different than in samples containing only a single serotype.

**DISCUSSION**

HEV infections were prospectively assessed in stools of healthy children by using a molecular typing strategy. HEVs were detected in 11.6% of the stool samples; 98% of the HEV-positive samples were identified to the species level. HEV-A species were detected in 6.8% of the samples; HEV-B was detected in 4.8% of the samples. Although these findings contrast with recent studies of HEV epidemiology based on culture and serology that reported only infrequent HEV-A infections in symptomatic (2, 5–7, 9, 25, 27, 37, 43) or healthy (13, 16, 21, 23, 38) children, they are consistent with earlier studies in which suckling mice were employed for in vivo assays and HEV-A infections were observed to be relatively frequent in healthy children (12, 17, 39). In a 1958 and 1959 study of stool samples obtained from 2,084 healthy children less than 5 years of age in London, the prevalence rates for HEV-A (CAV2, -4, -5, -6, -8, -10, and -12) and HEV-B (CBV3 and -4) isolates were 5.4% and 1.2%, respectively (17). In a study of clinical samples (primarily stools) obtained in 1966 and 1967 from 625 healthy children in Australia, the prevalence rates for HEV-A (CAV2, -3, -4, -8, and -10) and HEV-B (CBV2, -4, and -5 and CAV9) isolates were 12.1% and 4.3%, respectively (12). The Australian samples were collected in a single nursery; thus, the extent to which they represent the prevalence in the general population is uncertain. Finally, a Malaysian study of stool samples collected in 1971 and 1972 from healthy children less than 7 years of age found prevalences for HEV-A of 5.3%

**FIG. 1.** Rooted phylogenetic trees of CAV4 (A), CAV9 (B), and E18 (C) 3'-end VP1 sequences and available reference sequences and prototypes from GenBank for each of the serotypes. The evolutionary distances were calculated using the Kimura two-parameter model as a model of nucleotide substitution and the neighbor-joining method to reconstruct the phylogenetic tree (MEGA version 3.0 software package). The bootstrap values (the percentage of 1,000 pseudoreplicate data sets) supporting each cluster are shown at the nodes; for clarity, only values of >60% are shown. The outgroup taxa are the prototype strains of the closest serotypes. Sequence names for field strains (in boldface) are constructed as follows: municipality number (starting with 1, in Hordaland county; starting with 0, in Akershus county), month and year of sampling, child identifier, sample number, and serotype. The reference sequences have GenBank accession numbers. The scale bar represents the genetic distance (the number of nucleotide substitutions per site).
and for HEV-B of 3.4% (39). No species identification was reported.

We speculate that differences in the prevalences of individual HEV serotypes between our study and others reflect methodological differences. Whereas in vitro culture methods appear to be more sensitive for HEV-B, in vivo culture methods are more sensitive for HEV-A. To our knowledge, there are no previous studies reporting the prevalence of HEV-A in clinical samples using molecular methods. Our data indicate similar or equal sensitivities for both HEV-A and HEV-B. Thus, the findings reported here may represent accurate levels of circulating HEV serotypes.

Our analysis of seasonality confirmed the established pattern for HEV infections, where the highest prevalence in the Northern Hemisphere occurs between July and November (11). This observation held for both HEV-A and HEV-B (data not shown).

Analysis of age-specific frequencies of HEV-A and HEV-B showed an apparent trough between the ages of 9 and 12 months, with a distinct increase after 12 months of age (Fig. 2). These findings are consistent with those of Gamble et al. (17), in which the frequencies of both HEV-A and HEV-B infections were lower in the first year of life.

Phylogenetic analysis of the C-terminal domain of the VP1 gene has facilitated the differentiation of circulating serotypes and the association of specific HEV genotypes with disease (1, 3, 34, 40). Some serotypes are associated with global epidemics (1). During the period from 2001 to 2003, E13 and E18 were reported as the more common serotypes causing outbreaks of aseptic meningitis in the United States and elsewhere (6, 37). Most of the infections were E18 (8.9% versus 1.6%). Our VP1 sequence analysis revealed the circulation of several strains of the same serotype (Fig. 1).

HEV coinfections were commonly observed in the Norwegian infants followed in the present study. Although coinfections have been described in the developing world (23), they are only infrequently reported in children in industrialized or temperate areas (13).

FIG. 2. Three-month moving average of the age-specific prevalences of HEV-A- and HEV-B-positive samples for ages 3 to 23 months (n = 1,237 person-months). The curve was truncated at 24 months, due to few observations (18 stool samples) from 24 to 28 months of age. A logistic regression model was applied to adjust for the potential mutual confounding of age and season of year. Adjusting for season or year indicated that confounding by these factors was minimal or did not change the effect of age in the logistic-regression model.

We found no association between sex and either the frequency or duration of infection. These findings are in contrast with those reported for HEV-associated diseases, where infections were more frequent in males than in females, particularly for severe diseases, such as meningitis and carditis (35).

Our results suggest that the prevalence of HEV-A infections has been underestimated in recent studies that have relied on cell culture for detection. The diagnosis of HEV-A infections could be considerably improved by the use of molecular methods that facilitate detection and genotypic analysis.

ACKNOWLEDGMENTS

This study was supported by grants from the Norwegian Research Council (155300/320), the Ministry of Education of the Czech Republic (grant no. MSM0021620814), and the National Institutes of Health (AI55466 and AI51202). The MIDIA study is supported by the Norwegian Research Council (135895/330 and 156477/730).

We are grateful to the health care nurses and laboratory technicians in the MIDIA project and the laboratory technicians at the Department of Virology, Norwegian Institute of Public Health; to Trond Rasmussen for excellent data management and IT solutions; to Håkon Gjessing and Anders Skrondal for help with advanced statistical methods; and to Marian Rewers for sharing insights and for support.

REFERENCES


