Classification of Cryptosporidium Species from Patients with Sporadic Cryptosporidiosis by Use of Sequence-Based Multilocus Analysis following Mutation Scanning

Aaron R. Jex,1 Aradhana Pangasa,1 Bronwyn E. Campbell,1 Margaret Whipp,2 Geoff Hogg,2 Martha I. Sinclair,3 Melita Stevens,4 and Robin B. Gasser1*

Department of Veterinary Science, The University of Melbourne, Werribee, Victoria 3030, Australia; Melbourne Diagnostic Unit, Victoria 3052, Australia; Department of Epidemiology and Preventive Medicine, Monash University, Central and Eastern Clinical School, Alfred Hospital, Melbourne, Victoria 3004, Australia; and Melbourne Water Corporation, East Melbourne, Victoria 3002, Australia

Received 20 January 2008/Returned for modification 14 April 2008/Accepted 19 April 2008

In the present study, we analyzed genetic variation in Cryptosporidium species from humans (n = 62) with clinical cryptosporidiosis in South Australia. Sequence variation was assessed in regions within the small subunit of nuclear rRNA (p-SSU), the 70-kDa heat shock protein (p-hsp70), and the 60-kDa glycoprotein (p-gp60) genes by employing single-strand conformation polymorphism analysis and sequencing. Based on the analyses of p-SSU and p-hsp70, Cryptosporidium hominis (n = 38) and Cryptosporidium parvum (n = 24) were identified. The analysis of p-gp60 revealed eight distinct subgenotypes, classified as C. hominis IaA17R1 (n = 3), IbA9G3R2 (n = 14), IbA10G2R2 (n = 20), and IIA12G1R1 (n = 1), as well as C. parvum IiaA18G3R1 (n = 15), IIA20G3R1 (n = 6), IIA22G4R1 (n = 2), and IIA5G3R2 (n = 1). Subgenotypes Iia17R1 and IIA22G4R1 are new. Of the six other subgenotypes, IIA10G2R2, IIA18G3R1, IIA20G3R1, and IIA5G3R2 were reported previously from the state of Victoria. This is the fourth record in Australia of C. parvum subgenotype IIA18G3R1 from humans, which, to date, has been isolated only from cattle in other countries. This subgenotype might be a significant contributor to sporadic human cryptosporidiosis and may indicate a greater zoonotic contribution to the infection of humans in the area of study. Comparative analyses revealed, for the first time, the differences in the genetic makeup of Cryptosporidium populations between two relatively close, major metropolitan cities.

Cryptosporidium species are parasitic protists (of the Apicomplexa) that are commonly found in surface water globally (45) and parasitize many vertebrate classes, including birds, fish, mammals, and reptiles (21, 81). These species are transmitted via a highly resistant oocyst and usually infect the gastrointestinal tract, causing cryptosporidiosis. Cryptosporidiosis is commonly associated with enteritis and is usually characterized by acute, watery, or steatorrheic diarrhea and colic (16, 39), although asymptomatic infection can occur (15, 32, 55, 59, 68, 69). In humans, the transmission of Cryptosporidium infection is hypothesized to be most commonly anthropoontic (human to human) and, to a lesser extent, zoonotic (animal to human) (12, 62, 82). Cryptosporidiosis is usually short lived (days or weeks) and, in immunocompetent persons, eventually eliminated by a combined cellular and humoral immune response (63). In high-risk patient groups, such as infants and the elderly or people suffering from severe malnutrition, advanced human immunodeficiency virus/AIDS, or having undergone immunosuppressive therapy, cryptosporidiosis can be chronic, leading to severe dehydration and/or malabsorption (34). Although there has been progress with the development of drugs, such as paromomycin (10, 44, 56) and nitazoxanide (8, 65), currently there is limited accessibility to efficacious treatment in many countries (9, 13, 29, 43, 86). As a result, chronic cryptosporidiosis can be fatal and substantially affects the health of human communities, particularly in developing countries (39, 45). Therefore, the prevention and control of cryptosporidiosis are heavily reliant on diagnosis and an understanding of the epidemiology and genetics of Cryptosporidium species.

In recent years, Cryptosporidium species have received significant attention as the cause of epidemic outbreaks of disease (21, 37), often transmitted via drinking and/or recreational waters in developed cities in Australia (11, 61), Canada (54, 66, 75), Europe (67), and the United States (19, 20, 41, 85). Although such outbreaks can be substantial in scale and impact (41), sporadic cases are likely to be more common than widespread outbreaks, particularly in the developed nations. Elucidating the specific sources of sporadic cryptosporidiosis is often challenging and can be impractical. However, investigating the genetic makeup of Cryptosporidium species that are linked to sporadic cryptosporidiosis in particular geographical areas and understanding the risk factors associated with the disease provide a means of establishing the potential contributors to epidemics, the risk of zoonotic versus anthropoontic transmission to municipal populations, and the persistence of endemicity versus the emergence of “exotic” genetic variants of Cryptosporidium in the local environment.

* Corresponding author. Mailing address: Department of Veterinary Science, The University of Melbourne, Werribee, Victoria 3030, Australia. Phone: 61 3 9731 2294. Fax: 61 3 9731 2366. E-mail for A. Jex: ajex@unimelb.edu.au. E-mail for R. Gasser: robinbg@unimelb.edu.au.

† Published ahead of print on 30 April 2008.
Human cryptosporidiosis, whether sporadic or epidemic, predominantly is caused by Cryptosporidium hominis (for which only anthroponotic transmission has been reported) or Cryptosporidium parvum (for which both anthroponotic and zoonotic transmission have been reported) (12, 50). The population genetic analysis of C. parvum and C. hominis has revealed intraspecific variability in host specificity and geographical distribution (6, 24, 33, 51, 82, 84). For C. parvum, some of these intraspecific variants have been associated with substantial variation in minimum infective dosage and/or virulence (52, 53). Such variability in specificity, distribution, or virulence likely confers an effect on the risk that different strains can pose to human populations, which in turn might vary depending on the environment (e.g., urban versus rural or developed versus developing country).

Recently, the transmission dynamics of sporadic cryptosporidiosis in municipal regions of Australia was examined through a case control study of human populations in two major cities: Adelaide, South Australia (population of ~1 million), and Melbourne, Victoria (population of ~3.5 million) (64). The major risk factors associated with sporadic cases were inferred on the basis of positive statistical correlation (i.e., potential infection source versus incidence) to the use of public swimming pools (particularly those catering to young children); direct human-to-human contact; exposure to farm animals (particularly cattle and sheep); exposure to some food groups, such as cheeses produced from unpasteurized milk; and the consumption of untreated river, lake, or dam water in rural areas. Although Robertson et al. (64) examined the risk factors associated with sporadic cryptosporidiosis in humans, this study did not investigate the species/genetic variants of Cryptosporidium causing these cases. It is not possible, based on statistical data alone, to determine the contribution of anthroponotic or zoonotic transmission to sporadic infection in these cities, particularly for sources with greater potential for transmission via either pathway, such as risks related to untreated rural watersheds or foods. Such determinations require extensive and reliable epidemiological data sets that can be supported through the use of advanced molecular tools (23, 35, 70).

Although extensive molecular epidemiological studies have been carried out in a range of countries, surprisingly, there has been very limited study of the genetics of Cryptosporidium species isolated from humans in Australia (14, 36, 51). In addition, to date no study has examined the variation in the genetic makeup of Cryptosporidium populations that contribute to human infection between or among neighboring cities within the same country across a relatively small geographical range. A recent study (36) classified, for the first time, genotypes and subgroupotypes of Cryptosporidium that cause sporadic human cryptosporidiosis in the city of Melbourne and surrounding areas (in the state of Victoria, Australia). In the present study, we extended recent work to investigate the genetic variability in multiple loci within and among Cryptosporidium species from humans with sporadic cryptosporidiosis from Adelaide and surrounding areas in the state of South Australia (~1,000 km from Melbourne). The resultant data were compared to those presently available for sporadic cryptosporidiosis cases in Australia (14, 36, 51) and provided a first glimpse of the difference in the genetic makeup of Cryptosporidium populations between these two metropolitan cities and the extent to which the (epidemiological) risk factors identified by Robertson et al. (64) may have contributed to this difference.

MATERIALS AND METHODS

Isolation and purification of genomic DNA from fecal samples. Fecal samples (n = 62) from human patients with clinical cryptosporidiosis were collected (from 2001 to 2005) in Adelaide and surrounding areas in the state of South Australia; 26 of these samples originated from a previous epidemiological study but had not been genetically characterized (64). The parasitological diagnosis was made based on the detection of oocysts consistent with Cryptosporidium by using the modified Ziehl-Neelsen staining procedure (31). Genomic DNA was extracted from fecal samples as described previously (49). In brief, samples were suspended in a 1:4 ratio of phosphate-buffered saline (PBS). The suspension was homogenized, and 20 μl of suspension was aspirated and added to 80 μl of 10% polyvinylpyrrolidone (PVPP) (Sigma, St. Louis, MO) in H2O. After the addition of PVPP, samples were heated to 100°C for 10 min and then centrifuged at 10,000 × g for 30 s. Following centrifugation, each supernatant was transferred to a fresh tube and DNA was isolated using a minicolumn (Coprokit; Qiagen, Germany) according to the manufacturer’s protocol.

PCR-based amplification. All PCRs were set up in a PCR series laminar-flow cabinet (Clyde-APAC, Australia) using filter tips. The PCR was carried out in a 50-μl volume of 10 mM Tris-HCl (pH 8.4), 50 mM KCl, 2 mM MgCl2, 200 μM of each deoxynucleoside triphosphate, 25 pmol of each primer, and 1 U of GoTag polymerase (Promega, Madison, WI) in a GeneAmp PCR 2400 (Perkin Elmer) thermal cycle. Approximately 10 to 20 ng of template was included in the (primary) PCR amplification from genomic DNA; samples without templates were included in each run (as negative controls).

To identify the species of Cryptosporidium present in each DNA sample, an ~300-bp region of the small subunit of the nuclear rRNA (p-SSU) gene was amplified by PCR from genomic DNA (~10 to 20 ng) using the oligonucleotide primers 18SIF (forward; AGTGAACAGAAATAACATAACAGG-3′) and 18SIR (reverse; 5′-CCTGTTTAAGACCTCTATAATTC-3′) (24, 49) and subjected to sequencing (described below). The relative specificity of these primers for the amplification of Cryptosporidium species by PCR has been demonstrated previously (24, 49). The cycling conditions were one cycle of 94°C for 5 min (denaturation), followed by 5 cycles of 94°C for 30 s (denaturation), 58°C for 20 s (annealing), and 72°C for 30 s (extension), and then a final extension of 72°C for 7 min. The sequences derived from p-SSU amplicons were compared with selected, publicly available sequences representing C. hominis (accession no. AF093492 [80]) and C. parvum (accession no. AF115377 [83]).

To independently identify the Cryptosporidium species in each DNA sample, a part of the heat shock protein 70 gene (p-hsp70; ~450 bp) also was PCR amplified from genomic DNA using oligonucleotide primers HSPr4-F (forward; 5′-GGTGCTGTGATCCTTGTGATGTAC-3′) and HSPR4-R (reverse; 5′-CCTGTCATTAAAGCCTCTTAATTC-3′). These primers were originally designed to target relatively conserved regions of the heat shock protein 70 (hsp70) gene (38, 48) and have been used in previous studies of Cryptosporidium (3, 26). Sequences derived from p-hsp70 amplicons were compared with selected publicly available sequences representing C. hominis (accession no. DQ886255 [27]) and C. parvum (accession no. XM_001388291 [2]).

In order to assess genetic variation within each sample, a so-called variable region (~300 to 450 bp) within the gp60 gene (p-gp60) was amplified from DNA samples using a nested PCR approach. In the first round, the gp60 locus (~950 to 1,000 bp) was amplified from genomic DNA using primers gp15-ATG (forward; 5′-ATGAAATTTGTGCTATTACAT-3′) and gp15-STOP (reverse; 5′-TATACACAGGAATTACGTCG-3′) (71); the specificity of these primer pairs in PCR for the amplification of C. hominis and C. parvum has been demonstrated previously (36, 71). In the second round, 1 μl of the primary amplicon was transferred to a fresh tube (containing reagents) and subjected to PCR using primers gp15-15A (forward; 5′-GGCTTTCACCTACAGGGACAC-3′) and gp15-15E (reverse; 5′-CCACATATTACAAAGTGC-3′) (42) to amplify p-gp60 using the same cycling protocol, except that the annealing time was modified to 45 s. Known C. parvum (positive) controls and samples without DNA template (negative controls) as well as appropriate carrierover controls were included in each PCR run.

Agarose gel electrophoretic analysis, sequencing of amplicons, and phylogenetic analysis of sequence data sets. The quality and intensity of individual amplicons were examined on ethidium bromide-stained 1.5% agarose gels, using TBE (65 mM Tris-HCl, 27 mM boric acid, 1 mM EDTA, pH 9; Bio-Rad) as the
buffer and FX174-HaeIII (Promega) as a size marker. Amplicons were purified over minicolumns (Wizard PCR Prep; Promega), eluted in 30 μl H₂O, and then subjected to direct, automated sequencing (BigDye chemistry; Applied Biosystems, Foster City, CA) from both strands using the same primers (individually) as those used in the PCR (gp15-15A/gp15-15E for p-gp60). The sequences determined were compared with those available in current gene databases and published in quality, peer-reviewed scientific journals. For each locus (p-SSU, p-hsp70, or p-gp60), sequences were trimmed and then aligned using the program Clustal X (77), and the alignments were adjusted by employing the program BioEdit (30). Pairwise comparisons of the sequence differences (D) were made using the formula \( D = 1 - (ML) / (L) \), where \( M \) is the number of alignment positions at which the two sequences have a base in common and \( L \) is the total number of alignment positions over which the two sequences are compared; the differences were compared with automated calculations using PAUP*4.0b10 (74).

In the present study, we adopted the genotypic nomenclature proposed by Strong et al. (71), wherein each gp60 genotype is given the species-specific prefix for the species I (C. hominis) or II (C. parvum), followed by an alphabetical letter (e.g., Ia) that is assigned in reference to the sequence of the locus. Also, the present study recognizes and adopts the subgenotypic nomenclature for the gp60 locus proposed by Sulaiman et al. (72), wherein the subgenotypic assignment is based on the trinucleotide repeats within the gp60 microsatellite region, which encodes a polynucleotide tract within the protein OP45 (71). These repeats are categorized according to the number of TCA (A), TCG (G), TCT (T), and/or additional rare (R) repeats. Rare repeats were considered to relate to codons within the microsatellite region with a single point mutation at the first nucleotide position of the TCA trinucleotide in C. hominis or C. parvum (36). For example, a sample with a sequence consisting of 10 TCA repeats, 3 TCG repeats, and 2 ACA repeats would be identified as the subgenotype A10G3R2. As there is currently no consensus opinion on the precise definition of a genotype (or the threshold of intragenotypic versus intergenotypic sequence variation), a sample was assigned to a particular genotype based on its relationship with strongly supported clades after the phylogenetic analysis of the p-gp60 sequence data for all samples studied, including published sequences representing all currently recognized gp60 genotypes of C. hominis (Ia to If and Ib2 [7, 36, 71]) and C. parvum (Iia to IIk [1, 4, 7, 47, 58, 72, 76]) (for accession numbers, see Fig. 3).

The phylogenetic analysis of the p-gp60 nucleotide sequence data was conducted (separately for each species) using Bayesian inference and by employing the software package MrBayes 3.1.2 (http://mrbayes.csit.fsu.edu/index.php). Posterior probabilities (pp) were calculated according to the number of TCA (A), TCG (G), TCT (T), and/or additional rare (R) repeats. Rare repeats were considered to relate to codons within the microsatellite region with a single point mutation at the first nucleotide position of the TCA trinucleotide in C. hominis or C. parvum (36). For example, a sample with a sequence consisting of 10 TCA repeats, 3 TCG repeats, and 2 ACA repeats would be identified as the subgenotype A10G3R2. As there is currently no consensus opinion on the precise definition of a genotype (or the threshold of intragenotypic versus intergenotypic sequence variation), a sample was assigned to a particular genotype based on its relationship with strongly supported clades after the phylogenetic analysis of the p-gp60 sequence data for all samples studied, including published sequences representing all currently recognized gp60 genotypes of C. hominis (Ia to If and Ib2 [7, 36, 71]) and C. parvum (Iia to IIk [1, 4, 7, 47, 58, 72, 76]) (for accession numbers, see Fig. 3).

The method of single-strand conformation polymorphism (SSCP) was carried out according to protocol B (25). In brief, 10 μl of each amplicon was mixed with 5 μl of DNA sequencing stop solution (Promega), heat denatured at 94°C for 30 min, snap-cooled on a freeze block (−20°C), and screened using precast GMA S-2x25 gels (Eichrom Scientific AG, Cham, Switzerland). Samples were subjected to electrophoresis in a SEA 2000 apparatus (Eichrom Scientific AG) containing TAE buffer (0.04 M Tris base, 0.04 M acetate, 0.001 M EDTA) and were run at 110 V and 7.4°C (constant) for 6 h. Samples representing all of the profile variations that were detectable were selected from an initial analysis and subjected, under the same conditions, to electrophoresis in a GMA S-2x13 gel (Eichrom Scientific AG), using 3 μl of product in 10 μl of DNA sequencing stop solution at 74 V and 7.4°C for 18 h. A control sample was included on each gel to ensure the reproducibility of profiles, thus allowing the direct comparison of gels. Nucleotide sequence accession numbers. The sequences determined in the course of this work have been deposited in GenBank under accession numbers EU379540 to EU379551.

### RESULTS

Sixty-two samples from sporadic cases of cryptosporidiosis in humans from Adelaide and surrounding areas in South Australia were subjected to molecular analyses. Genetic variation was assessed in multiple loci (p-SSU, p-hsp70, and p-gp60) by SSCP analysis and targeted sequencing. Prior to SSCP and sequence-based analysis, all amplicons were subjected to agarose gel electrophoresis (not shown) and demonstrated to represent single, discrete bands. Amplicon sizes were ~250 bp (p-SSU), ~400 bp (p-hsp70), and ~300 to 350 bp (p-gp60). For both p-SSU and p-hsp70, the SSCP analysis of all amplicons divided samples into two distinct groups; within each group, there was no evidence of variation in profiles among amplicons. The sequencing of the p-SSU and p-hsp70 amplicons identified C. hominis (n = 38) and C. parvum (n = 24) based on comparisons with selected reference sequences (see Materials and Methods for accession numbers) (Table 1). The p-SSU sequences determined for C. hominis and C. parvum all were 250 and 247 bp, respectively. The hsp70 sequences obtained for C. hominis and C. parvum all were 403 bp. For both p-SSU and hsp70, no intraspecific sequence variation or polymorphism was detected within either C. hominis or C. parvum. Between C. hominis and C. parvum, sequence differences were 2% (two substitutions and three indels) in p-SSU and 1.2% (five substitutions) in p-hsp70. Representative sequences have been deposited under accession numbers EU379540 (C. hominis) and EU379541 (C. parvum) for p-SSU and EU379542 (C. hominis) and EU379543 (C. parvum) for p-hsp70.

Having identified C. hominis and C. parvum within the set of 62 samples, genetic heterogeneity was investigated further using the gp60 locus. Amplicons representing all 62 samples were se-
FIG. 1. Alignment of the four p-gp60 sequences representing Cryptosporidium hominis from 38 human patients from Adelaide and surrounding areas. A dot indicates a nucleotide identical to that of the top sequence; a dash indicates an insertion/deletion event. Sequences representing subgenotypes IaA17R1, IbA9G3R2, IbA10G2R2, and IfA12G1R2 are from the present study and have been deposited in GenBank under accession numbers EU379544, EU379545, EU379546, and EU379547, respectively. All other sequences represent reference sequences (accession numbers are given in the alignment at the left). Cryptosporidium parvum (Cp) was used as the outgroup for the phylogenetic analysis, which is depicted in Fig. 3.
quenced, and distinct p-gp60 sequence types were defined for *C. hominis* (*n* = 4) and *C. parvum* (*n* = 4) (Table 1). The p-gp60 sequences (accession numbers EU379544 to EU379551) varied in length from 302 to 326 bp (*C. hominis*) and from 263 to 314 bp (*C. parvum*) and had a G+C content of 45.1 to 49.2% (Fig. 1 and 2). According to the nomenclature proposed previously for *Cryptosporidium* genotypes/subgenotypes (36, 71, 72) and based on comparisons with reference sequences (Fig. 1 [C. hominis] and 2 [C. parvum] list the accession numbers), the 38 *C. hominis* samples were first classified genotypically/subgenotypically as Ia (IaA17R1), Ib (IbA10G2R2 and IbA9G3R2), or If (IIfA12G1R2), and the *C. parvum* samples were classified as Iia (IiaA18G3R1, IiaA20G3R1, and IiaA22G4R1) or Iic (IicA5G3R2) (Table 1). Sequence variation in p-gp60 was 0.3 to 52.4% (1 to 200 of 428 positions) within *C. hominis* and 2.4 to 51.3% (6 to 157 of 330 positions) within *C. parvum* (Tables 2 and 3). All sequence variation among all samples could be readily displayed visually by SSCP analysis (not shown).

For each species, all distinct p-gp60 sequence types that were determined were aligned (Fig. 1 and 2) with reference sequences representing all currently recognized gp60 genotypes, including some reported recently (36) from Melbourne, Australia (GenBank accession numbers EU164805 to EU164812). The nature and extent of sequence variation were evaluated. Among the four p-gp60 sequences determined for *C. hominis* (Fig. 1), relatively high levels of nucleotide variability, external to the microsatellite region (positions 7 to 81), were detected in regions between alignment positions 121 to 167, 171 to 229, 241 to 301, and 302 to 428. In regions ‘121 to 167’ and ‘171 to 229’, variation related predominantly to substitutions (42 and 90 sites, respectively) and some insertion/deletion events (16 and 45 sites, respectively). In regions ‘241 to 301’ and ‘302 to 428’, substitutions also were common (51 and 53 sites, respectively), but the number of insertion/deletion events was greater (47 and 54 sites, respectively) than that for the former two regions. Of the four p-gp60 sequences determined for *C. parvum* (Fig. 2), most of the nucleotide variation external to the microsatellite region (positions 7 to 108) was between alignment positions 121 to 230, 231 to 325, and 326 to 450. The nucleotide variability in regions ‘121 to 230’ and ‘326 to 450’ related predominantly to substitutions (56 and 87 sites, respectively), although insertion/deletion events were recorded (49 and 12 sites, respectively). In region ‘231 to 325’, insertion/deletion events were more common (85 sites) than substitutions (54 sites) and related mainly to the insertions in the sequences representing genotypes Ib and IIb (Fig. 2).

Having characterized the nature and extent of nucleotide variation and aligned all gp60 sequences for each species to achieve maximal positional homology, phylogenetic analyses were conducted using Bayesian inference (Fig. 3). Given the substantial sequence variability in p-gp60 between *C. hominis* and *C. parvum* and the inability to reliably align all sequences to achieve positional homology, the sequence data set for each species was subjected to separate phylogenetic analyses. The analysis of gp60 data showed that each recognized genotype of *C. hominis* (Ia to If) (7, 36, 71) and of *C. parvum* (Iia to IIk (1, 4, 7, 47, 58, 72, 76) resolved as a separate, strongly supported clade (pp of 1.00 for individual genotypic clades). The detailed appraisal of the *C. hominis* tree (Fig. 1) revealed the following clades (all with maximum nodal support [pp = 1.00]): genotype If grouped with Ie, IIe/Ie with Ib, and IIe/Ie/Ib with Ia to the exclusion of Id. A similar analysis of the *C. parvum* tree revealed the following clades (all with maximum nodal support [pp = 1.00]): genotype Iic grouped with Iib and Iic/Iib with Ile; genotype IIe grouped with IIk, IIk/Ik with Iia, Iia/Ii/Ik with IIg, Iia/Ii/Ig/Ik with Iib, Iia/Iib/Ii/Ig/Ik with III, and Ia/Iib/Ii/Ig/Ii/Ik with IIIe/Iib/Ii/Ik to the exclusion of IIe and IIj. Although genotype IId was shown to form a clade with all other *C. parvum* genotypes to the exclusion of IIj, the posterior probabilities calculated for this large clade (Iia to IIk to the exclusion of IIj) was low (pp = 0.50) and is not strongly supported by the present analysis. Importantly, for both species of *Cryptosporidium*, all of the gp60 genotypes determined herein grouped according to their preliminary classification based on genetic similarity alone (e.g., IaA17R1 with Ia and IbaA10G2R2 with Ib).

**DISCUSSION**

Using genotypic and subgenotypic definitions adopted from previous studies (36, 71, 72), the present *C. hominis* samples were initially classified as IaA17R1, IbaA9G3R2, Iba10G2R2, and IIfA12G1R2, and *C. parvum* samples were classified as IiaA18G3R1, IiaA20G3R1, and IiaA22G4R1. As there currently is no consensus view on the genetic boundaries or thresholds (based on nucleotide divergence) for the definition of gp60 genotypes or subgenotypes, we have suggested a cautious approach based on the phylogenetic analysis of sequence data (36). Thus, the classification of any new genetic variant within a species (or operational taxonomic unit) has been based on its relationship with known representatives (i.e., sequences from published studies and deposited in publicly available databases) and strong support at crucial nodes. In the present study, all previously identified gp60 genotypes of *C. hominis* and *C. parvum* were shown, based on Bayesian analysis, to form strongly supported clades consistent with their specific identification, as has been shown previously using other phylogenetic algorithms (1, 36, 60). In addition, the *C. hominis* and *C. parvum* gp60 variants from humans in Adelaide and surrounding areas in South Australia were positioned within these clades, supporting the initial classification based on sequence similarity.

FIG. 2. Alignment of the four p-gp60 sequences representing *Cryptosporidium parvum* from 24 human patients from Adelaide and surrounding areas (South Australia). A dot indicates a nucleotide identical to that of the top sequence; a dash indicates an insertion/deletion event. Sequences representing subgenotypes IiaA18G3R1, IiaA20G3R1, IiaA22G4R1, and IicA5G3R2 are from the present study and have been deposited in GenBank under accession numbers EU379548, EU379549, EU379550, and EU379551, respectively. All other sequences represent reference sequences (accession numbers are given in the alignment at the left). *Cryptosporidium hominis* (Ch) was used as the outgroup for the phylogenetic analysis, which is depicted in Fig. 3.
Of the gp60 subgenotypes identified in the present study, four (C. hominis IbA9G3R2, C. hominis IbA10G2R2, C. parvum IIA18G3R1, and C. parvum IIA20G3R1) have been reported previously in Adelaide (14) from a distinct subset of samples from the study by Robertson et al. (64). In addition to the C. hominis and C. parvum gp60 subgenotypes reported in the present study, Chalmers et al. (14) also identified subgenotypes C. hominis IeA12G3T3R1, C. parvum IIA15G2R1, and C. parvum IIA19G9R1. Of the C. hominis and C. parvum gp60 subgenotypes detected in humans in Adelaide/South Australia, all but two (C. hominis IaA17R1 and C. parvum IIA22G4R1) have been detected elsewhere. Within the C. hominis gp60 subgenotypes detected in Adelaide/South Australia, C. hominis IbA9G3R2 and IbA10G2R2 appear to be exceptionally common in humans and broadly distributed geographically, having been reported previously from numerous countries (6, 7, 87). Interestingly, the only other reports of this subgenotype in humans are from Australia (14, 36, 51). Cryptosporidium hominis IIA12G1R2 has been reported previously from humans in Australia (51) but appears to be much less prevalent than either IbA9G3R2 or IbA10G2R2 and has not been detected elsewhere. Within the C. parvum gp60 subgenotypes detected in Adelaide/South Australia or Melbourne/Victoria but not both, it should be noted that all of the subgenotypes found either in Adelaide/South Australia or Melbourne/Victoria (36, 51) and is recognized as a common subgenotype in humans globally (4, 6, 7, 40, 57, 58, 72).

Recently, we studied gp60 genotypes and subgenotypes among samples that represented sporadic cases of human cryptosporidiosis in Melbourne and surrounding areas (Victoria, Australia) (36). Four of the subgenotypes detected in humans in Adelaide (South Australia) in the present study, namely, C. hominis IbA10G2R2, C. parvum IIA18G3R1, C. parvum IIA20G3R1, and C. parvum IIA5G3R2, were identical in sequence to those reported previously in Melbourne (36). Given the relatively close proximity of the cities Adelaide and Melbourne (~1,000 km apart) and the states South Australia and Victoria and the high level of human transit that occurs between them, it was not surprising to encounter the same subgenotypes in both studies. Also interesting were the C. hominis and C. parvum gp60 subgenotypes that have been recorded either in Adelaide/South Australia or Melbourne/Victoria (36) but not in both cities or states. Subgenotypes (based on gp60 data) found in Adelaide/South Australia and not reported previously from Melbourne/Victoria were C. hominis IaA17R1, C. hominis IbA9G3R2, C. hominis IIA12G1R2, and C. parvum IIA22G4R1. Subgenotypes found in the previous study in Melbourne/Victoria (36) but not yet reported from Adelaide/South Australia were C. hominis IIA18G3R1, C. hominis IIA18G1R4, C. parvum IIA15G1R2, C. parvum IIA22G3R1, and C. parvum IIA23G3R1. Interestingly, only 4 (28.6%) of the 14 subgenotypes presently described from either of these regions (14, 36, and the present study) were found in both geographical regions. However, it should be noted that all of the subgenotypes found either in Adelaide/South Australia or Melbourne/Victoria but not both, e.g., Ia17R1 (n = 3), IeA12G3T3R1 (n = 2) (14), and IbA15G2R1 (n = 1) (14) in Adelaide/South Australia or IaA15G1R2 (n = 1), IIA22G3R1 (n = 1), and IIA23G3R1 (n = 1) (36) in Melbourne/Victoria, are relatively rare in the region from which they have been reported.
FIG. 3. Phylogenetic analysis of p-gp60 sequence data representing Cryptosporidium hominis (A) and C. parvum (B) from 62 human patients from Adelaide and surrounding areas of South Australia using key published sequence data for comparative purposes (1, 18, 40, 71–73). From the present study, sequences representing C. hominis subgenotypes IaA17R1, IbA9G3R2, IbA10G2R2, and IfA12G1R2 have been deposited under accession numbers EU379544, EU379545, EU379546, and EU379547, respectively; those representing C. parvum subgenotypes IIA18G3R1, IIA20G3R1, IIA22G4R1, and IIC5G3R2 are available under accession numbers EU379548, EU379549, EU379550, and EU379551, respectively. The accession numbers of other (reference) sequences are indicated in the trees. Numbers beneath the node of each clade represent bootstrap support obtained using Bayesian inference; the bar indicates distance (0.1 substitutions per site). An asterisk indicates a pp of 1.00.
cryptosporidiosis in Adelaide or South Australia also showed that *C. parvum* genotype IIa represented approximately one-third of the isolates investigated (14). Previous studies have reported that *C. parvum* genotype IIa is a major contributor to human cryptosporidiosis in Portugal, the United Kingdom, Slovenia, and parts of North America (6, 14, 28, 78), and it has been proposed in these studies that the predominance of this genotype is indicative of a high level of zoonotic transmission. In contrast, previous studies (12, 82) found that ~80% of cryptosporidiosis infections in urban populations in developed countries were associated with *C. hominis* infection. This 4:1 ratio of *C. hominis/C. parvum* infections appears to be the case for samples from humans with sporadic cryptosporidiosis in Melbourne, as a recent study (36) revealed that 75.5% of 98 samples contained *C. hominis*. The high prevalence of *C. parvum* subgenotypes within genotype IIa in sporadic human crypto-
sposporidiosis in Adelaide/South Australia also might indicate a significant zoonotic contribution. Indeed, a previous study (64) showed a positive correlation between samples from this re-
gion and likely zoonotic risk factors, such as human contact with calves, the consumption of some cheeses and unpasteur-
milk, or the consumption of untreated lake, river, and dam water in rural areas. However, it is important that the available data (51, 84) do indicate that *C. parvum* can be transmitted both by anthropo-
ontic and zoonotic routes, so an increased zoonotic risk to humans in this region cannot be concluded solely based upon the data presented here.

Interestingly, some subgenotypes, such as *C. hominis* IbA10G2R2 (6, 7, 18, 22, 40, 57, 72, 87) and *C. parvum* IIA15G2R1 (1, 6, 18, 57, 72, 76, 84, 87), have been reported from numerous countries and appear to be commonly associated with cryptosporidiosis in humans and/or animals globally. However, rare genetic variants, such as *C. hominis* IaA17R1, *C. hominis* IIA12G1R2, and *C. parvum* IIA22G4R1 (present study), indicate that localized, potentially geographically unique, subgenotypes are present even in large, urban popu-
lations in cosmopolitan cities, emphasizing the need for the increased study of underrepresented countries (such as those in Africa, Asia, and South America, as well as most island nations). There have been some suggestions (24, 57) that strain variation within *Cryptosporidium species* is correlated with geo-
ographical origin. Also, various authors (1, 5, 7, 51, 84, 87) have reported gp60 genotypes or subgenotypes that infect both hu-
man and animals and have formulated hypotheses regarding the importance of zoonotic transmission in causing sporadic cases or outbreaks of cryptosporidiosis in humans. However, to date neither the global distribution of gp60 genotypes and subgenotypes nor the range in host specificities of these types has been assessed comprehensively.

The present study classified *Cryptosporidium* species geno-
types/subgenotypes based on phylogenetic analysis and the for-
mation of strongly supported monophyletic clades (36). This approach was used because, although many gp60 genotypes and subgenotypes have been reported, threshold values for sequence variation (within species) and sequence difference (between species) have not yet been defined for the classifica-
tion of genotypes/subgenotypes. As such, the decision as to whether a new sequence type constitutes a new genotype or subgenotype is largely subjective; while phylogenetic analyses conducted in the present and previous studies (1, 36, 60, 72) indicate that the currently recognized genotypes usually form strongly supported monophyletic clades and thus appear to be valid, the current lack of defined genetic thresholds does not provide a solid framework for the classification of new genetic variants of *Cryptosporidium*. A comprehensive assessment of the nature and extent of sequence variation within and among gp60 and other informative genetic loci is required to establish a more definitive classification system for genotypes and sub-
genotypes globally.

ACKNOWLEDGMENTS

This study was supported by the Melbourne Water Corporation (R.B.G.) and the Australian Research Council (LP0561862). We thank Ivan Bastian for providing some of the samples.

REFERENCES

6. Alves, M., L. Xiao, F. Antunes, and O. Matos. 2006. Distribution of *Cryptosporid-
sation of *Cryptosporidium parvum* and *Cryptosporidium hominis* species. Int. J. Parasitol. 35:397–410.
17. Cohen, S., F. Dalle, A. Gallay, M. Di Palma, A. Bonnin, and H. D. Ward. 2006. Identification of Cppg80/15 type Ib as the predominant allele in iso-


