A New Restriction Fragment Length Polymorphism from *Cryptosporidium parvum* Identifies Genetically Heterogeneous Parasite Populations and Genotypic Changes Following Transmission from Bovine to Human Hosts

MARGARETHA CARRAWAY,¹ SAUL TZIPORI,^{1,2} AND GIOVANNI WIDMER^{1*}

Tufts University School of Veterinary Medicine, Department of Comparative Medicine, Division of Infectious Diseases, North Grafton, Massachusetts 01536,¹ and Division of Geographic Medicine, New England Medical Center, Boston, Massachusetts 02111²

Received 10 January 1997/Returned for modification 11 April 1997/Accepted 6 June 1997

Length and restriction site polymorphism within a 2.8-kb threonine-rich open reading frame from *Cryptosporidium parvum* was identified and used to determine the genotypes of isolates from calves and humans. In agreement with observations of other genetic loci, all calf isolates were identical at this locus. In contrast, human isolates showed two profiles, one found exclusively in humans and one a superposition of both profiles, which were indicative of heterogeneous parasite populations. PCR fingerprints were consistent with a change in the genetic profile of *C. parvum* isolates following transmission from bovine to human hosts.

Cryptosporidium parvum, an enteric protozoan parasite commonly found in neonatal calves, is an opportunistic infection in AIDS patients, frequently causing severe chronic diarrhea (4, 5). To advance our understanding of the epidemiology of *C. parvum*, genetic polymorphisms suitable for detection by PCR are being identified. Isolation of an open reading frame from *C. parvum* containing multiple polythreonine motifs (1) suggested the possibility of length polymorphism which is frequently associated with microsatellites (6). Here, we describe the presence of length and restriction site polymorphism at this locus and the application of these markers to fingerprinting *C. parvum* isolates from human patients or from calves or transmitted between different host species.

Isolation of a *C. parvum* threonine-rich open reading frame. A DNA probe was constructed by multimerizing a (ACA/T)₄ primer and its complement by PCR and cloning size-selected products into the TA cloning vector (Invitrogen). A genomic clone (pGC91) was identified by screening a λ ZapII *C. parvum* (GCH1 isolate) (7) phage library with this probe. Sequencing of pGC91 showed the presence of an open reading frame with 195 threonines (20% of the amino acids), 117 of which were localized within two regions (Fig. 1).

Identification of sequence polymorphism and isolates with mixed genotype. Amplification of a PCR fragment spanning most of the polythreonine open reading frame with primers cry32 and cry33 (Fig. 1) from seven *C. parvum* isolates (five from calves and two from AIDS patients) revealed an approximately 200-bp difference in the lengths of the amplified fragment between human (GCH2 and GCH3 [3]) and calf (UCP [7], GCH1 [7], MD, ICP [3], and IA) isolates (not shown). This was confirmed by a nested-PCR amplification targeting a 781-bp fragment with primers cry40 and cry41 (Fig. 2). With these primers, isolate GCH3 showed a mixed genotype, with one band shared with the calf and one band shared with the human GCH2 profile.

In order to circumvent the difficulties that were encountered in amplifying large polythreonine repeats, a 515-bp region of

the open reading frame flanked by primers cry37 and cry44 (Fig. 1) was screened for restriction fragment length polymorphism (RFLP) by using restriction endonuclease *Rsa*I. The four restriction fragments released by *Rsa*I migrated as three electrophoretic bands in isolates GCH1 and MD due to the similar sizes of the two smaller pieces (54 and 63 bp) (Fig. 3A, lanes 2 and 3). One of the *Rsa*I sites was absent from PCR products from GCH2. This resulted in an altered RFLP profile in this isolate (lane 4). In agreement with the detection of a mixed genotype in GCH3 with primers cry40 and cry41 (Fig. 2), RFLP analysis of the cry37-cry44 fragment also revealed a mixed profile (Fig. 3A). The profile recovered from GCH3 (lane 5) was a composite of the bovine profile (lanes 2 and 3) and the human profile seen in GCH2. A comparison of the undigested PCR products showed no length polymorphism in this region (Fig. 3B).

In order to obtain a simple RFLP marker, primer cry37 was replaced with cry39 (Fig. 1), resulting in a PCR fragment of 318 bp spanning the polymorphic *Rsa*I site. Fifteen *C. parvum* isolates were typed with this marker (Fig. 4). In some samples, restriction with *Rsa*I resulted in a 45-bp fragment and a 273-bp fragment. Confirming and extending the observations shown in Fig. 2, genotypes from nine animal-derived isolates were identical, whereas human isolates fell into two groups. One lacked the *Rsa*I site including GCH2, GCH3, F12, and F16 (Fig. 4; human samples 1, 2, 5, and 6), and two isolates had a mixed profile consisting mostly of *Rsa*I-positive DNA with some uncut PCR product (human samples 3 and 4). Except for GCH3, which was expected based on previous results to generate a mixed *Rsa*I profile, profiles from all samples matched the length polymorphism identified above.

Two ICP samples were included in this analysis, one from oocysts maintained in calves and one from oocysts recovered after a single passage through neonatal BALB/C mice. The mouse-passaged sample had retained its original profile as shown in Fig. 4 (sample 4). Two isolates originating from human patients, 740 and LL, and propagated once through calves had a calf profile (animal samples 8 and 9 [Fig. 4]). The same was found for GCH1, which is of human origin and has been maintained in calves since 1992 (7), and TAMU, which

* Corresponding author. Phone: (508) 839-7944. Fax: (508) 839-7977. E-mail: gwidmer@opal.tufts.edu.

```

1  NSILEGSIAG  IRSESCIVSE  LNFTSTTTGT  TDTSMNWPVS  ITSGELKDEN
51  KQATISGSRS  CGWKQGYSID  SSTGFRVDSI  TGLPTDPYSN  CPFNPVTGNL
101 VSRSTGKTIP  NTYAGVYRSN  ETKTTESAN  TNLLVDPKI  NAPCNSSENSF
151 EQQQIFDMGS  KVIYPTKCV  GVKHTTTTT  TTTTTTTTT  TTTTTTTTT
201 TTTTTTTTT  TTTTTTTTT  TTTTTTKKP  TTTTATTTT  TTSETESVIK
251 PDEWCWLEKN  GECEAKGATY  VGVIGKDGR  ENGMAFTMIP  NDDTHVRFRF
301 KVKDVGNTIS  VRCRKGAGKL  EFPDRSLDFT  IPPVAGHNSC  SIIVGVSGDG
351 KIHVSPYGSK  DVSLISAPIQ  PSELFPNEVC  DTCTAKYGAI  HSGYQTSADF
401 VTTTAKPTT  TTTGAPGQPT  TTTTGSPSKP  TTTTTKATT  TTTLNPTIT
451 TTTQKPTTT  TTKVPGKPPI  ATTTTLKPI  VTTTTKATT  TTTTVPTT
501 TTTKRDEMTT  TTTPLPDIGD  IEITPIPIEK  MLDKYTRMIY  DYNSGLLLDS
551 NDEPIPGSQA  GQIADTSNLF  PVQTHKSTGL  PIDPMVGLPF  DPKSGNLVHP
601 YTNQTMSGLS  VSYLAAKNLT  VDTDETYGLP  IDTLTGYPID  PVSLIPFNPE
651 TGELFDPISD  EIMNGTIAGI  VSGISASESL  LSQKSAPIDP  ATNMVVGFEF
701 GLLNPATGVM  IPGSLGPSEQ  TPFSPEIEDG  GIIPPEVAAA  NADKFKLSIP
751 PSVPESIEPK  DQKIDSISEL  MYDIESGRLI  GQVSKRPIPG  SIAGDLNPIM
801 KTPTQTDSVT  GKPIDTTGL  PFNPPTGHLI  NPTNNNTMDS  SFAGAYKYAV
851 SNGIKTDNVY  GLPVDEITGL  PKDPVSDIPF  NSTTGELVDP  STGKPINNST
901 AGIVSGKPL  PPIEDENGNL  FDPSTNLPID  GNNQLVNPET  NSTVSGSTSG
951 TTKPKP

```

FIG. 1. Protein translation of the *C. parvum* polythreonine locus. Shown is the translation of the 2,869-bp insert of clone pGC91. Threonine repeats are highlighted in boldface. PCR primers used in this study are indicated by horizontal arrows pointing 5' to 3'. Vertical arrow, location of the polymorphic *RsaI* restriction site.

was isolated from an accidental human infection and was also propagated in calves.

The human GCH4 and GCH5 isolates from two presumed accidental infections with GCH1 oocysts showed mostly the *RsaI*-positive profile with some uncut DNA (Fig. 4). The presence of *RsaI*-negative DNA in these samples suggested a change in the parasite population compared to GCH1. No uncleaved DNA was observed in GCH1 from subsequent pas-

FIG. 2. Length polymorphism in the *C. parvum* polythreonine locus. DNAs from five bovine (UCP, GCH1, MD, ICP, and IA) and two human (GCH2 and GCH3) isolates were amplified by a nested-PCR protocol with primers cry32 (tggtgaactgaaggatcc) with cry33 (gggttcaagtcaccagc) in the first reaction and cry40 (ctactacaaccaagaagacc) with cry41 (gttgaatgattgattaagag) in the second reaction. The PCR products identified a length polymorphism between isolates of human and calf origin and a mixed profile in GCH3. DNA marker (M) is a 100-bp ladder. neg, negative.

FIG. 3. PCR-RFLP analysis of four *C. parvum* isolates reveals a polymorphic *RsaI* restriction site and a mixed human isolate. (A) PCR with primers cry37 (cagcaagatatgaataccg) and cry44 (ctcttaatccaatcattacaac) were digested with *RsaI*. Three RFLP profiles were found in bovine (GCH1 and MD) isolates and two human (GCH2 and GCH3) isolates; GCH3 displayed a composite RFLP profile. (B) Undigested PCR products. Lane 1, 100-bp ladder.

sages. Since primers cry44 and cry39 do not amplify DNA from *C. parvum*-negative human stool (data not shown), we conclude that both the *RsaI*-positive band and the *RsaI*-negative band are of *C. parvum* origin. This view is consistent with sequence analysis of cloned cry44-cry37 PCR products. Clones with point mutations within the *RsaI* site were identified in GCH2 and GCH3 (data not shown).

To further substantiate the possibility of genotypic changes in human isolates, GCH4 and GCH5 were PCR fingerprinted

FIG. 4. RFLP analysis of eight bovine, one murine, and six human *C. parvum* isolates. *RsaI* digests of PCR products obtained with primers cry39 (gagtctaataataaaccactg) and cry44 show occurrence of two profiles, one of which is found exclusively in human samples. Some uncut DNA is visible in human isolates 3 and 4 (GCH4 and GCH5), suggesting genetic heterogeneity. Animal samples were as follows: 1, IA; 2, bovine ICP; 3, murine ICP; 4, TAMU; 5, 58EF; 6, 63H; 7, GCH1; 8, LL; and 9, 740. Human samples were as follows: 1, GCH2; 2, GCH3; 3, GCH4; 4, GCH5; 5, F12; and 6, F16. M, pBR/BstNI marker (fragment sizes are 1,857, 1,058, 929, 383, and 121 bp); C, undigested cry39-cry44 fragment from IA isolate.

with ribosomal primers within or flanking the ITS1 from bovine isolates. The expected 623-bp band was obtained from GCH1 but not from GCH5 or GCH4 (not shown). However, control amplification of GCH4 and GCH5 samples with conserved SSU ribosomal primers gave the expected products. Together, the *RsaI* RFLP and the ITS1 fingerprints demonstrate genetic differences between GCH1 and GCH4-GCH5.

In summary, a new polymorphism located within a three-nine-rich open reading frame was used to fingerprint 15 *C. parvum* isolates from different host species and geographical origin. In agreement with previous reports (2, 3), the two RFLP genotypes segregated according to host species: a homogeneously *RsaI*-positive bovine group and a genotypically mixed group from human patients. Analysis of human *C. parvum* isolates detected genetically heterogeneous populations in some individuals. In GCH3, this conclusion is supported by RFLP-PCR analysis with two primer pairs, detection of length polymorphism, and sequence analysis (not shown).

No correlation between the *RsaI* genotype of the six human isolates and the human immunodeficiency virus (HIV) status of the individuals from which they were isolated was found. GCH2 and GCH3 (the former being an isolate predominantly of human genotype [*RsaI* negative] and the latter being a mixed genotype) were both from HIV-positive individuals, whereas GCH4, GCH5, F12, and F16 were from HIV-negative individuals and were either of mixed or human genotype.

GCH4 and GCH5 were isolated from two laboratory workers who had recently been hired and who were involved in the purification of GCH1 oocysts and care of calves infected with this isolate. Since none of these patients reported any household members with intestinal symptoms preceding the onset of diarrhea, we assume that these infections resulted from the accidental ingestion of GCH1 oocysts. Based on this assumption, the ribosomal and *RsaI* fingerprints document a change in genotype between GCH1 and GCH4-GCH5. Similarly, the fingerprints of GCH2 and GCH3, which were two isolates recovered from HIV-positive patients living in the same household, provide circumstantial evidence for changes in *C. parvum* genotype. The individual from whom GCH3 was isolated was diagnosed with cryptosporidiosis before the one carrying GCH2, suggesting that GCH2 originated from GCH3 by direct transmission within the household. Under this assumption, the profiles of GCH3 and GCH2 show a transition from a mixed bovine-human-type parasite population in GCH3 to a human-type profile in GCH2. Together, *C. parvum* fingerprints of human isolates suggest a selective mechanism favoring propagation of parasites with the *RsaI*-negative genotype in humans.

In conclusion, data consistent with genotypic changes in *C. parvum* populations following transmission between different host species were found. This observation could explain the origin of genotypic differences between *C. parvum* isolates recovered from human and bovine hosts. In order to clarify this aspect of the epidemiology of *C. parvum*, genetic changes following transition from ruminant to nonruminant hosts are currently being investigated with animal models.

Nucleotide sequence accession no. The GenBank accession no. of pGC91 is U83169.

Contribution of oocysts and samples by the following individuals is gratefully acknowledged: Jennifer Clark, Lawrence Memorial Hospital, New London, Conn.; Joe Crabb, ImmuCell Corporation, Portland, Maine; Jeffrey Griffiths, Tufts University School of Veterinary Medicine, James McLaughlin, Central Public Health Laboratory, London, United Kingdom; Ravi Menon, University of Minnesota, Minneapolis; Cynthia Chappell and Karen Snowden, University of Texas, Houston; Furio Spano, La Sapienza University, Rome, Italy. Our thanks to Melissa Paris and Keith Johnson for assistance with animal experiments and oocyst purification.

This study received financial support from the U.S. Department of Agriculture, NRI grant 9403187, and Public Health Service contract N01-AI-25143 from the National Institute of Allergy and Infectious Diseases, Division of AIDS.

REFERENCES

1. Barnes, D. A., P. Doyle, S. Lewis, and C. Petersen. 1994. Surface antigens as targets for protective antibodies in cryptosporidiosis, abstr. C45, p. 76. In Abstracts of the 47th Annual Meeting of the Society of Protozoologists 1994. Society of Protozoologists, Cleveland, Ohio.
2. Bonnin, A., M. N. Fourmaux, J. F. Dubremetz, R. G. Nelson, P. Gobet, G. Harly, M. Buisson, D. Puygauthier-Toubas, F. Gabriel-Pospisil, M. Naciri, and P. Camerlynck. 1996. Genotyping human and bovine isolates of *Cryptosporidium parvum* by polymerase chain reaction-restriction length polymorphism analysis of a repetitive DNA sequence. FEMS Microbiol. Lett. 137: 207-211.
3. Carraway, M., S. Tzipori, and G. Widmer. 1996. Identification of genetic heterogeneity in the *Cryptosporidium parvum* ribosomal repeat. Appl. Environ. Microbiol. 62:712-716.
4. Pedersen, C., S. Danner, A. Lazzarin, M. P. Glauser, R. Weber, C. Katalama, S. E. Barton, J. D. Lundgren, and AIDS in Europe Study Group. 1996. Epidemiology of cryptosporidiosis among European AIDS patients. Genitourin. Med. 72:128-131.
5. Petersen, C. 1992. Cryptosporidiosis in patients infected with the human immunodeficiency virus. Clin. Infect. Dis. 15:903-909.
6. Su, X., and T. Wellems. 1996. Towards a high resolution *Plasmodium falciparum* linkage map: polymorphic markers from hundreds of simple sequence repeats. Genomics 33:430-444.
7. Tzipori, S., W. Rand, J. Griffiths, G. Widmer, and J. Crabb. 1994. Evaluation of an animal model system for cryptosporidiosis: therapeutic efficacy of paromomycin and hyperimmune bovine colostrum-immunoglobulin. Clin. Diagn. Lab. Immunol. 1:450-463.