ELSEVIER

Molecular and Biochemical Parasitology 108 (2000) 187–197

MOLECULAR AND BIOCHEMICAL PARASITOLOGY

www.elsevier.com/locate/parasitology

# Animal propagation and genomic survey of a genotype 1 isolate of *Cryptosporidium parvum*

G. Widmer [a], D. Akiyoshi [a], M.A. Buckholt [a], X. Feng [a], S.M. Rich [a], K.M. Deary [a], C.A. Bowman [a], P. Xu [b], Y. Wang [b], X. Wang [b], G.A. Buck [b], S. Tzipori [a,*]

[a] *The Division of Infectious Diseases, Tufts University School of Veterinary Medicine, North Grafton, MA 01536, USA*
[b] *Microbiology and Immunology, Virginia Commonwealth University, Richmond, VA 23198, USA*

Received 15 November 1999; received in revised form 10 February 2000; accepted 14 February 2000

**Abstract**

Human cryptosporidiosis is attributed to two major *Cryptosporidium parvum* genotypes of which type 1 appears to be the predominant. Most laboratory investigations however are performed using genotype 2 isolates, the only type which readily infects laboratory animals. So far type 1 has only been identified in humans and primates. A type 1 isolate, obtained from an individual with HIV and cryptosporidiosis, was successfully adapted to propagate in gnotobiotic piglets. Genotypic characterization of oocyst DNA from this isolate using multiple restriction fragment length polymorphisms, a genotype-specific PCR marker, and direct sequence analysis of two polymorphic loci confirmed that this isolate, designated NEMC1, is indeed type 1. No changes in the genetic profile were identified during multiple passages in piglets. In contrast, the time period between infection and onset of fecal oocyst shedding, an indicator of adaptation, decreased with increasing number of passages. Consistent with other type 1 isolates, NEMC1 failed to infect mice. A preliminary survey of the NEMC1 genome covering approximately 2% of the genome and encompassing 200 kb of unique sequence showed an average similarity of approximately 95% between type 1 and 2 sequences. Twenty-four percent of the NEMC1 sequences were homologous to previously determined genotype 2 *C. parvum* sequences. To our knowledge, this is the first successful serial propagation of genotype 1 in animals, which should facilitate characterization of the unique features of this human pathogen. © 2000 Elsevier Science B.V. All rights reserved.

*Keywords:* *Cryptosporidium parvum*; Genotype; Gnotobiotic pig; Genome sequence survey; Restriction fragment length polymorphism

*Abbreviations:* GKO, interferon-γ knock-out; GSS, genome sequence survey; PI, post-inoculation; RFLP, restriction fragment length polymorphism.

* Corresponding author. Tel.: +1-508-8397955; fax: +1-508-8397977.
*E-mail address:* stzipori@infonet.tufts.edu (S. Tzipori)

0166-6851/00/$ - see front matter © 2000 Elsevier Science B.V. All rights reserved.
PII: S0166-6851(00)00211-5

## 1. Introduction

*Cryptosporidium parvum*, an enteric apicomplexa, is a serious cause of acute diarrhea in immunocompetent humans and domestic animals, and chronic cryptosporidiosis in immunocompromised individuals [1,2]. At least two distinct genotypes of *C. parvum*, 1 and 2, are currently linked to human cryptosporidiosis [3–5]. Transmission in humans is either direct by the fecal–oral route or indirect through contamination of drinking water with oocysts, the environmentally resistant infectious forms. The main source of oocysts in drinking water, and hence the major cause of waterborne outbreaks of human cryptosporidiosis, was traditionally attributed to contamination with agricultural water run-offs. Consequently, the majority of human cases were assumed to be associated with genotype 2, the type which perpetuates in domestic animals. This assumption has been challenged by studies showing that a majority of human infections shed genotype 1 oocysts [3,5,6] although McLauchlin et al. [7] recently reported similar frequencies of type 1 and 2 infections among sporadic cases from the UK. With the exception of one infection in a captive primate, to date genotype 1 has been found exclusively in humans.

Despite the apparent epidemiological significance of genotype 1, the predominant cause of human cryptosporidiosis, investigations, including studies in human volunteers [8], were confined to type 2 isolates propagated in calves. Because of the inability of maintaining type 1 *C. parvum* in the laboratory, these isolates are rarely investigated, except for PCR-based epidemiological studies which require small amounts of DNA. In the absence of phenotypic studies on type 1 *C. parvum*, properties of type 2 oocysts are extrapolated to the entire species, possibly resulting in erroneous conclusions.

In this communication, we describe the adaptation to gnotobiotic piglets as well as selected phenotypic properties of a genotype 1 *C. parvum* isolate designated NEMC1. The isolate was obtained from an individual with HIV and cryptosporidiosis. The assignment to type 1 was based on the analysis of multiple genetic markers [9,10]. This isolate was sequentially passaged several times in gnotobiotic neonatal piglets [11] for a period of 7 months. As part of a preliminary genomic survey of NEMC1, approximately 200 kb, or an estimated 2% of the genome were sequenced and compared to published sequences from *C. parvum* genotype 2.

## 2. Materials and methods

### 2.1. Propagation of the NEMC1 isolate in gnotobiotic pigs

Piglets derived by cesarean section were housed in microbiological isolators under gnotobiotic conditions [11]. They were orally inoculated within 24 h with either fecal or intestinal fluid suspensions containing $10^5$–$10^6$ oocysts. Beginning 3 days post-inoculation (PI), fecal samples were collected directly from each piglet and smears were stained with acid-fast for microscopic examination. Feces were collected twice daily from infected animals. Fecal oocyst scores ranging from 0 to 5 were used as previously defined [11], to estimate the extent of oocyst excretion. Piglets were euthanized when the level of oocysts in feces started to decline. The small and large intestinal contents were collected and the oocysts were isolated using ether extraction followed by an isopycnic step gradient consisting of a low (15% w/v) and a high-density (30%) Nycodenz (Sigma) solution [3]. The NEMC1 isolate was passaged in piglets every 4–8 weeks, and oocyst DNA of fecal or intestinal origin from each animal was genotyped separately. The propagation from animal to animal was carried out under very strict precautions to minimize the possibility of contamination of the NEMC1 isolate with exogenous oocysts. A separate facility was used to handle and store NEMC1, and materials containing oocysts for passage were only handled in a sterile environment.

### 2.2. Mouse infectivity

Infectivity of NEMC1 oocysts in interferon γ knockout (GKO) mice was tested as described

[12,13]. Groups of seven GKO mice challenged with NEMC1 oocysts purified from different pig passages were kept in micro-isolators and sampled for oocyst shedding in feces regularly for at least 20 days.

### 2.3. Oocyst decay

Oocysts of isolate NEMC1 purified from the second pig passage by ether extraction and isopycnic centrifugation were suspended at a concentration of $10^6$ ml$^{-1}$ in sterile distilled water. The concentration of intact oocysts was determined six times over a period of 18 days using a hemocytometer. Four replicate counts were taken at each time point. The stability of oocysts from other isolates included in this analysis was determined in a similar fashion. Initial concentrations ranged from $5 \times 10^5$ to $10^7$ oocysts ml$^{-1}$.

### 2.4. Genotyping

Restriction fragment length polymorphism (RFLP) analysis included three markers: COWP [14] TRAP-C1 [15] and β-tubulin [10]. Based on their locations on chromosome bands 2 and 5 [9,16], these markers represent at least two separate linkage groups.

Sequences obtained from randomly picked clones originating from a *C. parvum* genotype 1 phage library generated from human isolate 2458L [3], were compared to published *C. parvum* sequences using the BLAST algorithm. Clone LIB13 (GenBank accession number AF190627) was found to be homologous to accession number B78618 originating from the genotype 2 IOWA isolate. The alignment of the LIB13 and B78618 sequences revealed two four-basepair (bp) deletions, one in each sequence. PCR primers 618F2 (ttctttaatttaatctatcatttaatt) and LIB13F2 (ttttacgccgccacgaaga) were designed such that their sequences are identical to the IOWA sequence at the two polymorphic regions. In contrast, primers 618F1 (ttctttaatttaatctatcatttaata) and LIB13F1 (ttttacgccgccacgaagt) are specific to the 2458L sequence. The specificity of these primers for each genotype was based on the fact that their 3′ ends matched only one sequence. Two control primers, LIB13R (aggaatattaccggccttgg) and 618R (tttggttctggaacttccttg) were located within conserved sequences upstream and downstream of the polymorphisms, respectively. To test for the presence of type 1 DNA, PCR was carried out using primers LIB13/LIB13F1 and 618R/618F1, and type 2 DNA was detected using the corresponding F2 primers in conjunction with LIB13R or 618R.

### 2.5. DNA preparation and library construction

Purified oocysts obtained from pig passage 4 were surface sterilized with 0.5% hypochlorite to remove exogenous DNA. Oocyst DNA was isolated using several freeze–thaw cycles and SDS/proteinase K treatment. A Sau3AI partial digestion was performed and DNA fragments in the 2–5 kb range were isolated by gel electrophoresis. These fragments were ligated into ZAP Express Bam HI/CIAP vector (Stratagene), packaged and plated with *E. coli* XL-1-Blue MRF′ cells onto X-gal and IPTG containing plates. This library had greater than $10^6$ independent clones. The clones were excised using ExAssist helper phage, mixed with XLOLR cells and plated onto LB plates containing kanamycin, X-gal and IPTG. White colonies were then streaked onto gridded LB plates containing kanamycin.

### 2.6. Template preparation and sequence analysis

Colonies for the genome sequence survey (GSS) were picked into deep 96-well plates containing 1.0 ml of Terrific Broth [17], with 50 mg ml$^{-1}$ kanamycin, and incubated for 17–22 h at 37°C with vigorous agitation. Plasmid template was isolated by alkaline lysis as previously reported [18]. Sequence reactions were run in 96-well reaction blocks using $0.5 \times$ PEB BigDye™ Terminator Reaction mix with T7 or T3 primer and $\sim 0.3$ μg of DNA in a 5-μl reaction volume as described by the manufacturer and analyzed on 7-h medium high throughput gels on a PEB 377 Prism™ Automated DNA Sequencer.

### 2.7. Processing and analysis of the GSS sequences

Sequence files were imported into Sequencher™ software (GeneCodes). Vector and ambiguous sequences were trimmed such that the first and last 25 bases contained 0.2 ambiguous base calls. Processed sequences were exported as ASCII files to a SGI Origin 2000 for database searches. Each sequence was searched against the non-redundant GenEMBL, the EST and GSS databases, and the SwissPlus and PIR databases (GCG) using BLASTX or BLASTN. These results were sorted according to probability significance scores. Sequences are accessible at http://views.vcu.edu/dnalab/cparvum/cparvum.html and in GenBank under accession numbers AQ988770–AQ989104.

## 3. Results

### 3.1. Propagation of isolate NEMC1 in piglets

Propagation of the NEMC1 isolate was initiated with the infection of two gnotobiotic piglets with $10^6$ oocysts extracted from human stool. Oocysts recovered from one piglet were inoculated 2 months later into four piglets. This procedure was repeated four more times. Repeated passage in piglets resulted in a shortened latent period, defined as the period between infection and onset of oocyst shedding. Whereas the first passage produced oocysts in both piglets on day 9 PI, in subsequent passages the latent period was progressively shortened (Fig. 1). In the last two passages, oocysts were observed at day 3 PI in one animal on the 5th passage, and in all three animals of the 6th passage. Later passages also showed an increase in oocyst counts, with mean peak scores ranging from 3 to 5 as compared to 1.2 to 2.5 for early passages (data not shown). The gradual adaptation of isolate NEMC1 to this new host, was manifested by earlier onset and larger amount of oocyst excretion with increasing number of passages, although only mild or no diarrhea was associated with the infection.

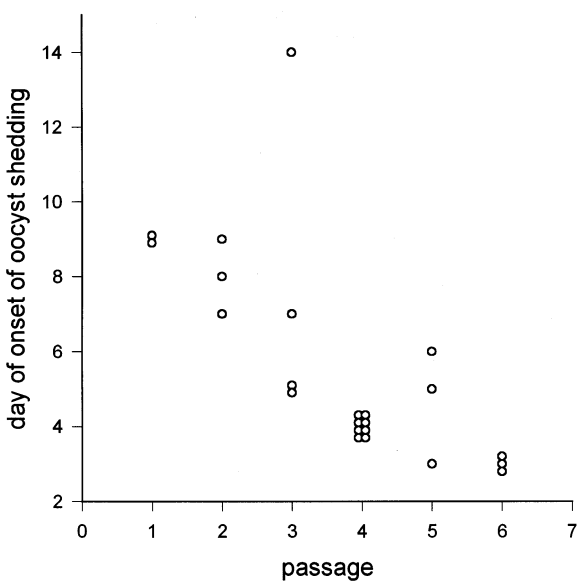


Fig. 1. First day of oocyst shedding for serial pig passage of *C. parvum* NEMC1 isolate. Each circle represents one animal. Overlapping symbols were artificially separated to show number of piglets used in each passage.

### 3.2. Mouse infectivity

The infectivity of NEMC1 oocysts for GKO mice was tested after each passage in piglets. Neither oocysts from the original human patient, nor oocysts isolated from piglets infected GKO mice.

### 3.3. Genotypic characterization of NEMC1

In order to verify that isolate NEMC1 was correctly classified as *C. parvum*, a 487-bp sequence from the SSU rRNA was determined. This sequence, which spans position 428–914 (GenBank accession #16996), includes several variable regions useful for species identification [19–21]. A multiple sequence alignment comparing 27 *Cryptosporidium* SSU rRNA sequences from GenBank with those from two PCR clones amplified from NEMC1 (pig passage 4) demonstrated that the sequences obtained from NEMC1 are diagnostic of *C. parvum* and *C. parvum* like genotypes (Fig. 2). Although not identical, clones NEMC1-2 (GenBank accession AF187984) and

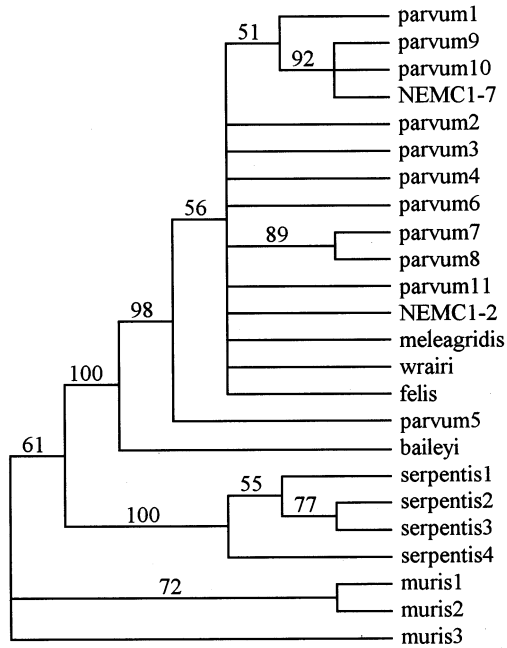


Fig. 2. Bootstrap consensus phylogeny of *Cryptosporidium* isolates from various sources, based on 100 resampling replicates of the neighbor-joining algorithm with distances evaluated by maximum likelihood. The monophyly of type-1 (NEMC1, parvum9, parvum10, and parvum11) and type-2 (parvum6, parvum7, parvum8) *C. parvum* isolates is strongly supported (98%). *C. muris* was used as outgroup based on its relatedness when this set of *Cryptosporidium* isolates was compared with *Theileria* and *Toxoplasma* as outgroup taxa (not shown). The following *C. parvum* GenBank accessions were used: parvum1, AF112569; parvum2, AF112570; parvum3, AF112571; parvum4, AF112572; parvum5, AF112573; parvum6, AF228682; parvum7, AF093490; parvum8, AF093493; parvum9, AF093489; parvum10, AF093491; parvum11, AF093492.

1-7 (AF187985) clustered with 10 *C. parvum* sequences of human and animal origin. GenBank entries with the highest similarity scores also included a *C. wrairi*, *C. meleagridis* and a *C. felis* sequence. The two NEMC1 sequences differed from each other at five nucleotide positions for an overall identity of 98%.

RFLP analysis was performed with the aim of identifying the *C. parvum* genotype of the original human isolate and subsequent pig passages. Restriction analysis of the COWP, TRAP-C1 and β-tubulin PCR products amplified from several pig passages, as well as from the original human NEMC1 oocysts, revealed identical RFLP patterns for all samples. The NEMC1 pattern was diagnostic for genotype 1 and was different from the genotype 2 profile displayed by isolate GCH1 (Fig. 3). Faintly stained DNA fragments co-migrating with the GCH1 control were seen with the COWP and TRAP-C1 markers (panels A and C). Although genotypically mixed isolates have been observed [22], we believe that these bands are not

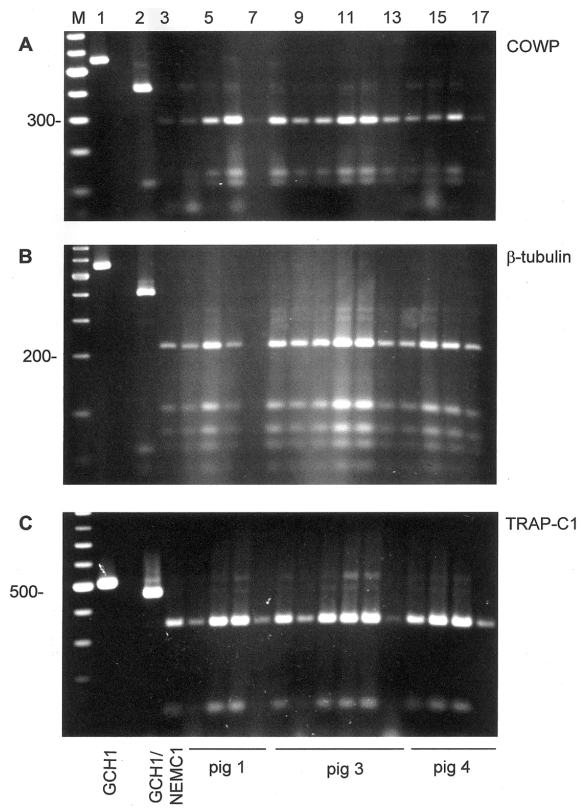


Fig. 3. RFLP analysis of NEMC1 human isolate and subsequent passages in piglets. Shown are RFLP profiles obtained with three genetically unlinked markers [9]. (A) COWP. (B) β-tubulin. (C) TRAP-C1. Lane 1, undigested PCR product from calf-propagated genotype 2 isolate GCH1; lane 2, restricted GCH1 PCR products; lane 3, restricted human NEMC1, lanes 4–17, pig passages as indicated below. Shown are multiple samples collected from individual pig passages. The restriction digests were performed with RsaI (COWP), Tsp509I (β-tubulin) and RsaI (TRAP-C1). The profiles displayed by the NEMC1 samples are diagnostic of genotype 1.

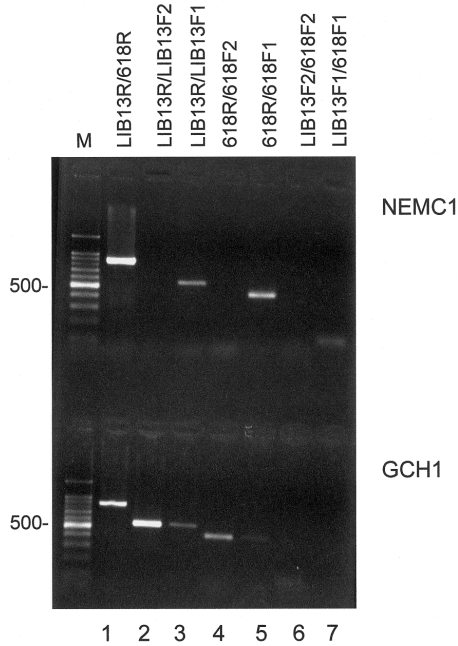


Fig. 4. Genotype-specific PCR analysis of NEMC1 oocysts originating from pig 4 and the genotype 2 GCH1 isolate using the LIB13 primer set. DNA from each isolate was amplified with the primers shown. F1 primers, designed based on the sequence of the human genotype 1 isolate 2458L, amplified NEMC1 DNA (upper row, lanes 3, 5, 7). No amplification product were obtained from NEMC1 DNA with F2 primers (lanes 2, 4, 6). Control PCR analysis with GCH1 DNA is shown below. Note faint PCR bands with F1 primers and GCH1 due to the presence of a subpopulation of genotype 1 DNA in this isolate. Amplification products with conserved R primers flanking the polymorphic region is shown in lane 1.

due to the presence of type 2 DNA because they only appeared in conjunction with full-length, uncut PCR product. This observation is indicative of incomplete digestion rather than heterogeneity of the DNA sample. The identity among RFLP profiles from different passages was consistent with NEMC1 having maintained the original type 1 profile during pig propagation.

To confirm the RFLP analysis, PCR primers specific for an anonymous *C. parvum* sequence (accession # B78618) were designed based on sequence heterogeneity between a type 1 and the homologous type 2 sequence. Type 1 and type 2 specific primers were used in combination with conserved primers to probe for the presence of type 1 and type 2 sequences in NEMC1 in parallel with the control type 2 GCH1 isolate. This analysis revealed only type 1 DNA from NEMC1 (Fig. 4). In contrast, the analogous PCR analysis with DNA from GCH1 identified predominantly type 2 DNA and traces of type 1 DNA. This observation is consistent with the RFLP data described above and with the previous evidence that GCH1 is genotypically mixed [10].

Included in the genotypic analysis of NEMC1 was also the sequence of the β-tubulin intron and adjacent exon 2 sequence. PCR products amplified from NEMC1 pig passage 4 with primers btub5 and btub2 [10] were cloned and three randomly picked clones were sequenced in both directions. Consistent with the RFLP data, all clones displayed the type 1-specific sequence, including a diagnostic Tsp509I restriction site at nucleotide position 399 (not shown). This restriction site distinguishes between type 1 and type 2 *C. parvum* isolates. We also noticed some sequence differences to other type 1 β-tubulin sequences, which is consistent with heterogeneity previously observed among type 1 isolates at this locus.

In summary, the genotypic analysis of multiple independent loci, indicate that isolate NEMC1 belongs to *C. parvum* and originated as a type 1 when recovered from a patient. This genotypic profile was maintained during six serial passages through piglets over a period of almost 7 months.

### 3.4. Stability of NEMC1 oocysts

The stability of NEMC1 oocysts stored in water at room temperature was determined. We previously found that the rate of decay of *C. parvum* oocysts fell into two groups; oocysts of genotype 2 and two genotype 1 samples displayed the stability typically associated with *C. parvum* oocysts [23], with a half-life at room temperature of 112 days. In contrast, a majority of genotype 1 oocysts decayed much faster, with a half-life of less than 7 days (Fig. 5). As illustrated, the NEMC1 decay curve was typical of the fast-decaying type and is consistent with the classification of this isolate as type 1. The observed rates of oocyst decay were not a function of the host as

the slow decaying groups consisted of oocysts originating from calves (GCH1), humans (MD-human, EC, PC1, 0676K) and mice (MD-mice). The rapid decaying groups included oocysts from humans (9897, 2066K, JHHIV) and the NEMC1 from pigs. Oocyst stability was also unrelated to their treatment prior to the experiment. The majority of the oocyst samples included in Fig. 5 were exposed to ether and floated on salt as part of the purification procedure. All samples were surface sterilized with bleach. Storage of stool samples in potassium dichromate prior to oocyst purification did not correlate with oocyst stability either.

### 3.5. Preliminary genome sequence survey (GSS) of NEMC1

A preliminary random GSS was performed using genomic DNA from the NEMC1 isolate of *C. parvum* cloned into the pBK-CMV plasmid (Stratagene). A total of 428 sequences were generated from 411 plasmid clones. Each of these sequences was stringently processed to remove vector and ambiguous 5′ and 3′ sequence using Sequencher™ software as described in Section 2. A total of approximately 200 kb of sequence was generated representing 340 unique contiguity groups and approximately 2.0% of the estimated 10.4 Mbp *C. parvum* genome [16]. Base composition analysis of these sequences suggested an overall 32% G + C content for this *C. parvum* isolate, consistent with what has recently been reported for coding (36%) and noncoding (32%) regions of a genotype 2 *C. parvum* isolate [24]. An analysis of di- and tri-nucleotide frequencies suggested no unusual features (data not shown), although the TA dinucleotide, which is generally under-represented in genomic DNA [25], was not under-represented in the NEMC1 isolate sequences. No complex repeats, but several dozen simple (mono-, di-, tri-, and tetra-nucleotide) microsatellites, all less than 40 bp in length, were identified in the GSS sequences (Table 1). No homologues of clones containing microsatellites were found among the genotype 2 sequences in GenBank.

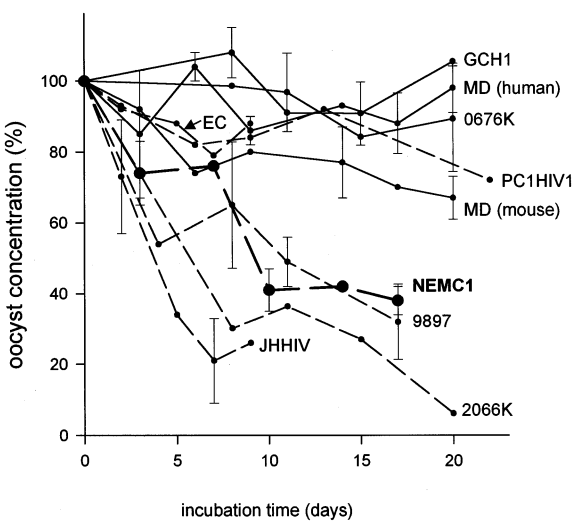


Fig. 5. Oocyst decay at room temperature. Oocysts of nine *C. parvum* isolates were surface-sterilized and incubated in sterile water. Counts of intact oocysts were determined over a period of up to 22 days. Mean counts and selected S.D. are shown. Dashed lines, genotype 1; continuous lines, genotype 2.

BLASTN searches of the GSS, EST and GenEMBL nucleotide databases (GCG) identified 67, 10 and four sequences, respectively, with significant ($P < e-5$) homologies. Each of the homologous sequences present in these databases was a previously sequenced gene from another *C. parvum* isolate. The 81 sequences with homology to *C. parvum* sequences in the databases represented 24% of the total nonredundant sequences, consistent with estimates that $>20\%$ of the *C. parvum* genotype 2 sequences are currently in publicly available databases [24]. These observations also validate the quality of the DNA used for generation of our library; i.e. no sequences homologous to *E. coli*, other possible contaminating microorganisms, or the experimental host (pigs) were detected. Similarity to the *C. parvum* sequences in the databases ranged from 85 to 100% with an average of $94.7 \pm 3.0\%$, clearly confirming that these sequences all derived from *C. parvum*. Moreover, since nearly all of the sequences in public databases originate from IOWA or other genotype 2 *C. parvum*, these results are consistent with the expected heterology between the NEMC1 genotype 1 sequence and genotype 2 sequences. However, the majority of the homologous *C. parvum* sequences identified were in

Table 1
Low complexity repeats

| Repeat[a] | Clone[b] |
|---|---|
| Mononucleotide | |
| A/T | 01a1a01ne.f1(9), 01a1a01ne.r1(9), 02c1a03ne.r1(6), 11d1c1.t3(6), 18c1f05ne.r1(6), 20c1g03ne.r1(6), 26a1a04ne.r1(7), 27 = T3.Seq(6), 36b1d06ne.r1(7), 41b1f04ne.f1(8) |
| Dinucleotide | |
| AC/GC | 02b1a03ne.f1(6)(8) |
| AG/CT | 16 = T3.Seq(13), 01a1a01ne.f1(8), 01a1a01ne.r1(8) |
| AT/AT | 02b1a09ne.f1(6)(8), 05c1b09ne.r1(7), 07e1c01ne.r1(16), 09a1c05ne.f1(6), 31d1c02ne.r1(12) |
| CC/GG | 10b1d01ne.f1(6) |
| CA/TG | 02b1a03ne.f1(8) |
| GA/TC | 16 = T3.Seq(13), 01a1a01ne.f1(8), 01a1a01ne.r1(8) |
| TA/TA | 02b1a09ne.f1(8), 05c1b09ne.r1(7), 07e1c01ne.r1(15), 31d1c02ne.r1(13) |
| Trinucleotide | |
| AAC/GTT | 08b1c03ne.r1(4), 45c1g06ne.r1(6) |
| AAT/ATT | 06c1b05ne.r1(4), 14d1f1.t3(4), 15d1e05ne.r1(4), 35b1d10ne.r1(4), 45c1g06ne.r1(5), 14a1e09ne.r1(6)(4), 14d1e03ne.r1(6), 38e1e04ne.r1(5), 45c1g12ne.r1(4) |
| ACA/TGT | 08b1c03ne.r1(4), 45c1g06ne.r1(7) |
| ACC/GGT | 04d1b01ne.r1(8)(5) |
| AGT/ACT | 11d1c1.t3(8) |
| ATA/TAT | 14d1f1.t3(4), 15d1e05ne.r1(4), 27b1a06ne.r1(4), 35b1d10ne.r1(4), 45c1g06ne.r1(5), 14d1e03ne.r1(6), 06c1b05ne.r1(4), 14a1e09ne.r1(6)(4), 38e1e04ne.r1(5) |
| ATC/GAT | 08b1c03ne.f1(10) |
| CAA/TTG | 08b1c03ne.r1(4), 45c1g06ne.r1(7) |
| CAC/GTG | 04d1b01ne.r1(4)(8)(5) |
| CAT/ATG | 08b1c03ne.f1(11) |
| CCA/TGG | 04d1b01ne.r1(4)(8)(5) |
| CTT/AAG | 31e1c02ne.r1(4) |
| GAA/TTC | 42c1f06ne.r1(4) |
| GTA/TAC | 11d1c1.t3(7) |
| TAA/TTA | 35b1d10ne.r1(4), 45c1g06ne.r1(6), 06c1b05ne.r1(4) |
| TAG/CTA | 11d1c1.t3(8) |
| TCA/TGA | 08b1c03ne.f1(10) |
| TCT/AGA | 31e1c02ne.r1(4) |
| TTA/TAA | 06c1b05ne.r1(5), 14a1e09ne.r1(5)(4), 14d1e03ne.r1(7), 38e1e04ne.r1(4), 45c1g12ne.r1(4) |

Table 1
Low complexity repeats

| Repeat[a] | Clone[b] |
|---|---|
| Tetranucleotide | |
| AATT/AATT | 19a1g07ne.r1(3) |
| ATTA/TAAT | 19a1g07ne.r1(3), 05a1b03ne.r1(3) |
| ATTT/AAAT | 35b1d04ne.r1(3) |
| TTAA/TTAA | 05a1b03ne.r1(3) |
| TTTA/TAAA | 06b1b05ne.r1(3) |

[a] Repeat sequence/complement.
[b] Number of repeat units in parentheses.

the GSS and EST databases, and therefore do not represent finished sequence. Therefore, the overall level of similarity between the NEMC1 sequences that we generated and genotype 2 sequences in the databases may be higher than that suggested by these data.

All of the nonredundant sequences were subjected to BLASTX searches against the PIR and SwissPlus (GCG) databases to identify clones with open reading frames homologous to known proteins. Of these sequences, 80 ( $\sim 24\%$) showed significant similarity ($P < e-5$) with known proteins, and these proteins were categorized into functional groups (Fig. 6) as recently described for the complete genome of *Saccharomyces cerevisiae* [26]. The results largely parallel those from a recent GSS of the Iowa isolate of *C. parvum* [24], which were categorized according to standard schemes for prokaryotes [27]. It is evident that genes involved in macromolecular and small molecule biosynthesis, cellular signaling, energy production, regulation, and transport are well represented in the *C. parvum* genome.

## 4. Discussion

This is the first report of a successful transmission and serial propagation of type 1 *C. parvum* in laboratory animals. Gnotobiotic piglets have proved to be a useful model to study enteric pathogens of human origin, which include among others diarrheogenic *E. coli*, microsporidia and others [11,28,29]. Piglets have also been used as a diarrhea model for the evaluation of chemothera-

peutic agents against *C. parvum*. Animals infected with type 2 isolates invariably develop profound diarrhea, anorexia, dehydration and wasting 3–4 days after challenge which leads to a serious weight loss [11,30]. Symptoms last 10–15 days and often result in death from severe dehydration and wasting. With the genotype 2 GCH1 isolate oocyst shedding begins 3–4 days after challenges and lasts 13–15 days. In contrast, infection with NEMC1 initially took 9 days before oocysts were detected in the feces. The onset of oocyst excretion on subsequent passages decreased to 3–4 days, as seen in type 2 isolates. However, little or no diarrhea, weight loss or other symptoms characteristic of type 2, were seen in piglets infected with NEMC1.

Propagation in piglets will facilitate research on *C. parvum* genotype 1 despite the added expense and the relatively low number of oocysts obtained from piglets as compared to calves. In addition, the half-life of the majority of type 1 oocysts investigated so far, including NEMC1, appears to be shorter than type 2. Consequently these isolates may require more frequent animal passage to maintain their viability. Despite these drawbacks, the propagation of isolate NEMC1 has led to a more detailed genotypic and phenotypic characterization of a type 1 isolate than had been possible before.

Prior to the isolation of NEMC1, we attempted to propagate in gnotobiotic piglets three type 1 isolates, two from humans and one from a captive macaque. In all cases, genotypic conversion from type 1 to type 2 was observed in the course of one or two passages (Tzipori, Widmer and Akiyoshi, unpublished data). We believe that a change from type 1 to type 2 in the course of these experiments indicates that some or even the majority of isolates may exist as a mixed genotypic population, and that a selective process favoring growth of type 2 over type 1 in this host had occurred. Although genotypically mixed infections in humans have been observed [22], it is unclear to what extent earlier type 1 isolates, which were not stably propagated in the pig, were genotypically mixed. An alternative possibility is that type 2 oocysts contaminated isolates from earlier passages resulting in overgrowth of type 2 parasites.

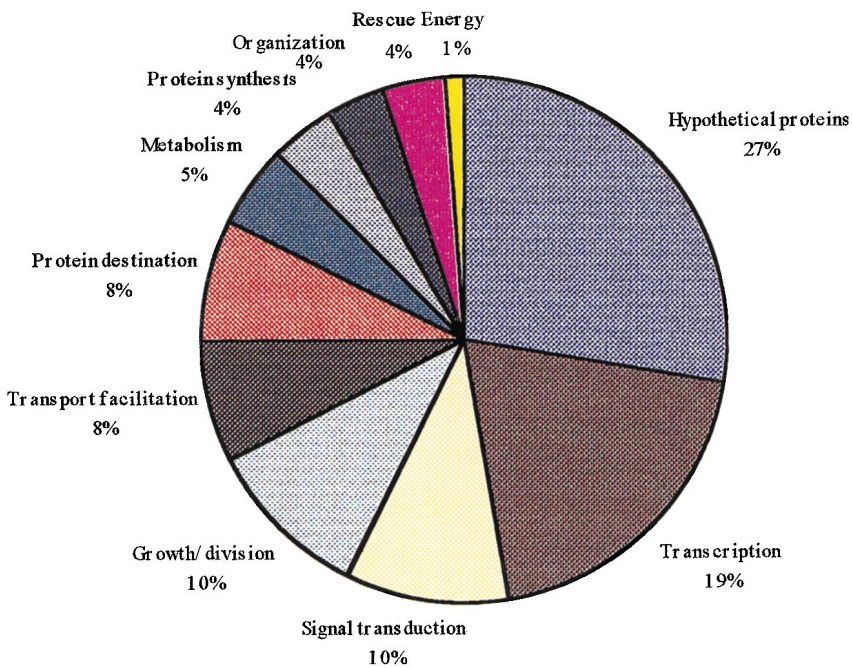


Fig. 6. Functional classification of proteins identified from the genome survey of *C. parvum* isolate NEMC1.

The observed genotypic stability of NEMC1 might have resulted from more stringent quality control measures adopted in the laboratory. These included physically separated facilities for storage and handling of NEMC1 samples and the use of dedicated personnel to avoid accidental carry-over.

Although some COWP and TRAP-C1 RFLP profiles revealed low levels of what appears as genotype 2 DNA, the following points argue against the presence of this genotype in NEMC1: (1) RsaI negative restriction fragments appear together with unrestricted PCR product, (2) the β-tubulin RFLP and LIB13 PCR profiles are devoid of type 2 bands, and (3) the potential for conversion of putatively mixed isolates to type 2 as discussed above suggests that NEMC1 might have lost its original type 1 profile had type 2 parasites been present in the original human oocyst sample.

As observed with other *C. parvum* isolates, sequence analysis of the SSU rRNA and β-tubulin (data not shown) indicated some heterogeneity within NEMC1. Consistent with our interpretation of the RFLP and LIB13 profiles, the β-tubulin sequences of all three PCR clones were diagnostic of type 1. The increase in virulence observed for NEMC1, manifested by shortening of the time period between infection and oocyst shedding over the course of multiple passages in piglets, is also indicative of heterogeneity in NEMC1 and argues that selection took place of phenotypes that propagate more readily in the pig. This process is consistent with a pre-existing genotypic heterogeneity defining a range of virulence.

Results from the genome survey are consistent with the genetic heterogeneity identified at selected loci [4,10,31]. Similarity between the GSS sequences from the NEMC1 genotype 1 isolate and sequences from the genotype 2 isolates in public databases ranged from 85 to 100% with an average of $\sim 95\%$. Since our sequences and most of the sequences in the public databases are random GSS sequences, and therefore are not confirmed finished sequence, these numbers may overestimate the differences between the two isolates. However, our sequences were carefully edited and we believe it is likely that the observed heterology largely reflects real genomic differences. If so, there is significant heterology between the genomic sequences of *C. parvum* genotypes 1 and 2.

It will now be possible to investigate to what extent genotypic differences correlate with clinical manifestations, as studies in human volunteers will no longer be restricted to genotype 2. Some phenotypic differences between NEMC1 and type 2 isolates were clearly apparent, including a significant difference in oocyst survival time and infectivity to GKO mice. The fact that GKO mice were refractory to infection with NEMC1 oocysts was unexpected considering the high susceptibility of this model to type 2 *C. parvum*. Experiments in GKO mice with four type 2 isolates demonstrated that a single oocyst is sufficient to generate an infection (data not shown).

## Acknowledgements

Our thanks to Arlen Avakian for isolating human NEMC1 oocysts, Sue Chapman and Jessica Brisben for technical support, and Cynthia Chappell for providing oocyst samples from infected individuals. Sequencing was performed in the Nucleic Acids Research Facility at VCU. Financial support from the NIH (U19 AI33384) and USDA (NRI grant # 9800919) is acknowledged. G. Widmer and D. Akiyoshi contributed equally to this study.

## References

[1] Tzipori S. Cryptosporidiosis in perspective. Adv Parasitol 1988;27:63–130.
[2] Guerrant RL. Cryptosporidiosis: An emerging, highly infectious threat. Emerg Infect Dis 1997;3:51–8.
[3] Widmer G, Tzipori S, Fichtenbaum CJ, Griffiths JK. Genotypic and phenotypic characterization of *Cryptosporidium parvum* isolates from people with AIDS. J Infect Dis 1998;178:834–40.
[4] Peng MM, Xiao L, Freeman AR, Arrowood MJ, Escalante AA, Weltman AC, et al. Genetic polymorphisms among *Cryptosporidium parvum* isolates: evidence of two distinct human transmission cycles. Emerg Infect Dis 1997;3:567–73.

[5] Patel S, Pedraza-Diaz S, McLauclin J, Casemore DP. Molecular characterization of *Cryptosporidium parvum* from two large suspected waterborne outbreaks, Commun Dis Public Health 1998;231–233.
[6] Sulaiman IM, Xiao L, Yang C, Escalante L, Moore A, Beard C, et al. Differentiating human from animal isolates of *Cryptosporidium parvum*. Emerg Infect Dis 1998;4:681–5.
[7] McLauchlin J, Pedraza-Diaz S, Amar-Hoetzeneder C, Nichols GL. Genetic characterization of *Cryptosporidium* strains from 218 patients with diarrhea diagnosed as having sporadic cryptosporidiosis. J Clin Microbiol 1999;37:3153–8.
[8] Okhuysen PC, Chappell CL, Crabb JH, Sterling CR, DuPont HL. Virulence of three distinct *Cryptosporidium parvum* isolates for healthy adults. J Infect Dis 1999;180:1275–81.
[9] Spano F, Putignani L, Crisanti A, et al. A multilocus genotypic analysis of *Cryptosporidium parvum* from different hosts and geographical origins. J Clin Microbiol 1998;36:3255–9.
[10] Widmer G, Tchack L, Chappell CL, Tzipori S. Sequence polymorphism in the β-tubulin gene reveals heterogeneous and variable population structure in *Cryptosporidium parvum*. Appl Environ Microbiol 1998;64:4477–81.
[11] Tzipori S, Rand W, Griffiths J, Widmer G, Crabb JH. Evaluation of an animal model system for cryptosporidiosis: therapeutic efficacy of paromomycin and hyperimmune bovine colostrum-immunoglobulin. Clin Diagn Lab Immunol 1994;1:450–63.
[12] Theodos C, Sullivan K, Griffiths J, Tzipori S. The profile of healing and non-healing *Cryptosporidium parvum* infection in mice with functional B and T lymphocytes: The extent of IFNg modulation determines infection outcome. Infect Immun 1997;65:4761–9.
[13] Griffiths J, Theodos C, Paris M, Tzipori S. The interferon-γ gene-disrupted mouse: a highly susceptible model for evaluation of therapeutic agents against *Cryptosporidium parvum*. J Clin Microbiol 1998;36:2503–8.
[14] Spano F, Putignani L, McLaughlin J, Casemore DP, Crisanti A. PCR-RFLP analysis of the *Cryptosporidium* oocyst wall protein (COWP) gene discriminates between *C. wrairi* and *C. parvum*, and between *C. parvum* isolates of human and animal origin. FEMS Microbiol Lett 1997;152:209–17.
[15] Spano F, Putignani L, Guida S, Crisanti A. *Cryptosporidium parvum*: PCR-RFLP analysis of the TRAP-C1 (thrombospondin-related adhesive protein of *Cryptosporidium*-1) gene discriminates between two alleles differentially associated with parasite isolates of animal and human origin. Exp Parasitol 1998;90:195–8.
[16] Blunt DS, Khramtsov NV, Upton SJ, Montelone BA. Molecular karyotype analysis of *Cryptosporidium parvum*: evidence for eight chromosomes and a low-molecular-size molecule. Clin Diagn Lab Immunol 1997;4:11–3.
[17] Sambrook J, Fritsch EF, Maniatis T. Molecular Cloning, A Laboratory Manual. New York: Cold Spring Harbor Laboratory Press, 1989.
[18] Roe BA, Crabtree JS, Khan AS. DNA Isolation and Sequencing. New York: Wiley, 1996.
[19] Leng X, Mosier DA, Oberst RD. Differentiation of *Cryptosporidium parvum*, *C. muris* and *C. baileyi* by PCR-RFLP analysis of the 18S rRNA gene. Vet Parasitol 1996;62:1–7.
[20] Morgan UM, Xiao L, Fayer R, Graczyk TK, Lal AA, Deplazes P, et al. Phylogenetic analysis of *Cryptosporidium* isolated from captive reptiles using 18S rDNA sequence data and RAPD analysis. J Parasitol 1999;85:525–30.
[21] Xiao L, Escalante L, Yang C, Sulaiman I, Escalante AA, Montali RJ, et al. Phylogenetic analysis of *Cryptosporidium* parasites based on the small-subunit rRNA gene locus. Appl Environ Microbiol 1999;65:1578–83.
[22] Carraway M, Tzipori S, Widmer G. New RFLP marker in *Cryptosporidium parvum* identifies mixed parasite populations and genotypic instability in response to host change. Infect Immun 1997;65:3958–60.
[23] Yang S, Healey MC, Du C. Infectivity of preserved *Cryptosporidium parvum* oocysts for immunosuppressed adult mice. FEMS Immunol Med Microbiol 1996;13:141–5.
[24] Liu C, Vigdorovich V, Kapur V, Abrahamsen MS. A random survey of the *Cryptosporidium parvum* genome. Infect Immun 1999;67:3960–9.
[25] Ellis J, Griffin H, Morrison D, Johnson AM. Analysis of dinucleotide frequency and codon usage in the phylum Apicomplexa. Gene 1993;126:163–70.
[26] Mewes HW, Albermann K, Bahr M, Frishman D, Gleissner A, Hani J, et al. Overview of the yeast genome [published erratum appears in Nature 1997, June 12;387(6634):737]. Nature 1997;387:7–65.
[27] Riley M. Functions of the gene products of *Escherichia coli*. Microbiol Rev 1993;57:862–952.
[28] Tzipori S, Gunzer F, de Montigny L, Donnenberg M, Kaper J, Donohue-Rolfe A. The role of the *eaeA* gene in diarrhea and neurological complications in a gnotobiotic piglet model of enterohemorrhagic *Escherichia coli* infection. Infect Immun 1995;63:3621–7.
[29] Kondova I, Mansfield K, Buckholt M, Stein B, Lackner A, Tzipori S. Transmission of *Enterocytozoon bieneusi* of human and primate origin in immunosuppressed gnotobiotic piglets. Infect Immun 1998;66:5515–9.
[30] Theodos CM, Griffiths JK, D'Onfro J, Fairfield A, Tzipori S. Efficacy of nitazoxanide against *Cryptosporidium parvum* in cell culture and animal models. Antimicrob Agents Chemother 1998;42:1959–65.
[31] Vasquez JR, Gooze L, Kim K, Gut J, Peterson C, Nelson RG. Potential antifolate resistance determinants and genotypic variation in the bifunctional dihydrofolate reductase-thymidilate synthase gene from human and bovine isolates of *Cryptosporidium parvum*. Mol Biochem Parasitol 1996;79:153–6.