High-throughput genotyping assay for the large-scale genetic characterization of *Cryptosporidium* parasites from human and bovine samples

J. L. ABAL-FABEIRO[1,2,3], X. MASIDE[1,2,3], J. LLOVO[3,4], X. BELLO[1,2,3], M. TORRES[5], M. TREVIÑO[4], L. MOLDES[4], A. MUÑOZ[6], A. CARRACEDO[1,7,8] *and* C. BARTOLOMÉ[1,2,3]*

[1] *Departamento de Anatomía Patolóxica e Ciencias Forenses, Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Spain*
[2] *Medicina Xenómica, CIMUS, Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Spain*
[3] *Xenómica Comparada de Parásitos Humanos, IDIS, Santiago de Compostela, Spain*
[4] *Servicio de Microbioloxía e Parasitoloxía, Complexo Hospitalario Universitario de Santiago de Compostela, 15706 Santiago de Compostela, Spain*
[5] *Centro Nacional de Genotipado, Universidade de Santiago de Compostela, Spain*
[6] *Departamento de Microbioloxía e Parasitoloxía, Facultade de Bioloxía, Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Spain*
[7] *Fundación Pública Galega de Medicina Xenómica (SERGAS)-CIBERER (USC), Complexo Hospitalario Universitario de Santiago de Compostela, 15706 Santiago de Compostela, Spain*
[8] *Center of Excellence in Genomic Medicine Research, King Abdulaziz University, 21589 Jeddah, Kingdom of Saudi Arabia*

(*Received 26 June 2013; revised 6 September 2013; accepted 8 September 2013; first published online 15 November 2013*)

SUMMARY

The epidemiological study of human cryptosporidiosis requires the characterization of species and subtypes involved in human disease in large sample collections. Molecular genotyping is costly and time-consuming, making the implementation of low-cost, highly efficient technologies increasingly necessary. Here, we designed a protocol based on MALDI-TOF mass spectrometry for the high-throughput genotyping of a panel of 55 single nucleotide variants (SNVs) selected as markers for the identification of common *gp60* subtypes of four *Cryptosporidium* species that infect humans. The method was applied to a panel of 608 human and 63 bovine isolates and the results were compared with control samples typed by Sanger sequencing. The method allowed the identification of species in 610 specimens (90·9%) and *gp60* subtype in 605 (90·2%). It displayed excellent performance, with sensitivity and specificity values of 87·3 and 98·0%, respectively. Up to nine genotypes from four different *Cryptosporidium* species (*C. hominis, C. parvum, C. meleagridis* and *C. felis*) were detected in humans; the most common ones were *C. hominis* subtype Ib, and *C. parvum* IIa (61·3 and 28·3%, respectively). 96·5% of the bovine samples were typed as IIa. The method performs as well as the widely used Sanger sequencing and is more cost-effective and less time consuming.

Key words: *Cryptosporidium* epidemiology, high-throughput SNV genotyping, MALDI-TOF mass spectrometry.

INTRODUCTION

Discrimination between *Cryptosporidium* species and subtypes is crucial for epidemiological studies and for the prevention and control of cryptosporidiosis. These require the screening of vast numbers of samples to determine the genotypes involved in human disease and the possible sources of contamination. Nowadays, Sanger DNA sequencing at a reduced number of genes is the most common tool used for molecular determination (Strong *et al*. 2000; Xiao *et al*. 2004). However, when large numbers of samples are analysed, this method is expensive and unpractical (Chalmers, 2008). Another limitation of the technique is that sometimes it is difficult to obtain high-quality sequences from stool DNA extractions containing inhibitors and/or DNA degrading substances (Chalmers, 2008; Jex *et al*. 2008).

There is a range of high-throughput cost-effective tools available to identify *Cryptosporidium* species and subtypes (Jex *et al*. 2008). The most widely used is the electrophoretic display of genetic variants using single-stranded conformation polymorphism (SSCP), based on changes in the mobility of single-stranded DNA in a non-denaturing polyacrylamide gel (Gasser *et al*. 2006; Jex *et al*. 2007*b*). This method is highly efficient, although very dependent on the genes used to discriminate polymorphisms (Chalmers *et al*. 2005; Jex *et al*. 2007*b*, 2008), so the selection of adequate loci is essential to obtain accurate results. A caveat to this technique is that it does not identify the nature of the variants detected and further sequencing is needed to determine it.

The aim of our study was to test and describe a new method for genotyping *Cryptosporidium* isolates

* Corresponding author: CIMUS, Despacho P2D2, Avda. Barcelona s/n, 15782 Santiago de Compostela, Galicia, Spain. E-mail: carolina.bartolome@usc.es

doi:10.1017/S0031182013001807

based on the combined use of single base extension (SBE) of a panel of selected single nucleotide variants (SNVs), that includes both polymorphic positions within a species (single nucleotide polymorphisms, SNPs) and fixed changes between species, followed by sequencing by mass spectrometry (Matrix-Assisted Laser Desorption/Ionization-Time Of Flight, MALDI-TOF). Its performance was tested in a total of 608 human isolates representative of community cases seeking medical assistance and 63 bovine samples. The results were compared with Sanger sequencing data from a panel of control samples.

MATERIALS AND METHODS

Cryptosporidium *samples*

The samples of human origin were routinely collected from diarrhoeal patients between years 2000 and 2008 at the Complexo Hospitalario Universitario de Santiago (CHUS, Santiago de Compostela, Spain) and diagnosed by staining oocysts with the auramine phenol method (Casemore *et al*. 1984). *Cryptosporidium* positive samples were stored at −20 °C in 1·5 mL plastic vials. The bovine stool samples were collected from calves (< 1 month of age) in 2007 and diagnosed as *Cryptosporidium*-positive at the Laboratorio de Sanidade e Producción Animal da Xunta de Galicia (Lugo, Spain) using immunochromatography.

DNA extraction

DNA was isolated from 180–200 mg of stool samples using the QIAamp DNA Stool Mini kit (QIAgen, Izasa, Barcelona, Spain) with slight modifications of the manufacturer's protocol to improve the disruption of *Cryptosporidium* oocysts. Three freeze-thawing cycles were performed after the addition of Buffer ASL by immersion in liquid nitrogen for 1 min and thawing in a water-bath at 95 °C for 2 min.

Genotyping

Samples were genotyped at single nucleotide variants from five different loci (*ACoAs, COWP1, gp60, gtub* and *ISWIr*), allowing us to distinguish between *Cryptosporidium* species and subtypes infecting humans using the MassARRAY system (Sequenom), which combines the iPLEX Gold genotyping technology with MALDI-TOF mass spectrometry and allows for the automated analysis of large numbers of samples. This assay is based on a primer extension aimed to detect sequence differences at the nucleotide level (Fig. 1). The protocol includes an initial PCR amplification of the region surrounding the sequence variation of interest, followed by the addition of a primer with mass-modified terminators that anneals

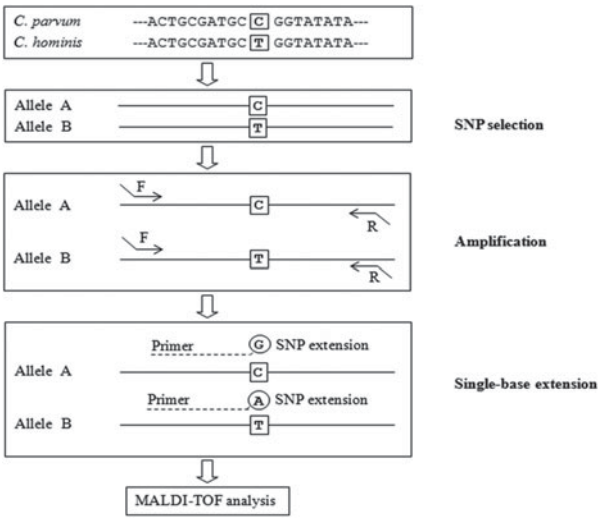

Fig. 1. Schematic diagram of the iPLEX™ Gold genotyping process. SNV markers are selected. The region around each SNV is PCR amplified in each sample. Specific oligonucleotides are added for a single-base extension. SNV alleles are identified by allele-specific differences in mass between alternative extension products.

immediately upstream of the polymorphic site and produces a specific single-base extension of the product complementary to the SNV (Gabriel *et al*. 2009). The mass difference of the single-base extension products enables allelic discrimination, which is performed by MALDI-TOF mass spectrometry (Sauer and Gut, 2002). Several SNVs of interest can be amplified simultaneously, a process known as 'multiplexing'.

The protocol was initially developed in two plexes of 20 and 15 SNVs. These were applied to 671 samples distributed into eight plates containing 90 samples and six appropriate controls each. However, given that some of the SNVs produced low-quality results (e.g. low call rates or dubious genotypes), 20 additional SNVs were also genotyped (Tables 1 and 2). These were arranged in three plexes of ten, six and four SNVs each that were applied to a single plate with 83 samples with dubious genotypes and the corresponding positive and negative controls. The designs of amplification and extension primers, as well as allele calling, were performed at the Universidade de Santiago de Compostela node of the Centro Nacional de Genotipado (CEGEN).

Diagnostic SNVs were chosen among those described in a multilocus variation study (Abal-Fabeiro *et al*. 2013) and from comparison of selected representative sequences retrieved from GenBank (Supplemental Table 1S – in Online version only). All SNVs were located in the coding sequences of five genes: *ACoAs, COWP1, gp60, gtub* and *ISWIr*, which have been mapped to four different chromosomes (1, 6, 6, 7 and 8 respectively; footnote 'a' of Supplemental Table S2 – in Online version only in

Table 1. Sensitivity and specificity assessment (%) of the 19 single nucleotide variant (SNV) markers used to identify *Cryptosporidium* species

SNV name	Allele	Identifies	*TP*	*FN*	*TN*	*FP*	*Sn*	*Sp*
ACOAS_235	C	*m*	1	0	81	0	100·0	100·0
	T	*p, h*	78	3	1	0	96·3	100·0
COWP1_111	G	*s*	0	0	82	0	–	100·0
COWP1_114	C	*p, m*	39	1	42	0	97·5	100·0
	T	*h, c, u, f*	40	2	40	0	95·2	100·0
COWP1_142	C	*f*	0	0	82	0	–	100·0
COWP1_189	G	*c*	0	0	82	0	–	100·0
COWP1_255	C	*u*	0	0	82	0	–	100·0
	A	*s*	0	0	82	0	–	100·0
COWP1_291	G	*f*	0	0	82	0	–	100·0
COWP1_333	C	*m*	1	0	81	0	100·0	100·0
	T	*p, h, f*	79	2	1	0	97·5	100·0
COWP1_336	G	*u*	0	0	82	0	–	100·0
	A	*s*	0	0	82	0	–	100·0
COWP1_399	T	*m*	0	1	81	0	0·0	100·0
	A	*p, h*	32	49	1	0	39·5	100·0
COWP1_435	C	*c*	0	0	82	0	–	100·0
GP60_73	G	*f*	0	0	82	0	–	100·0
	T	*p*	36	3	6	37	92·3	14·0
GTUB_ 96	A	*mu*	0	0	82	0	–	100·0
GTUB_360	G	*p*	36	3	42	1	92·3	97·7
	A	*h*	38	4	39	1	90·5	97·5
GTUB_735	T	*p*	34	5	43	0	87·2	100·0
	C	*h*	39	3	40	0	92·9	100·0
GTUB_804	G	*m*	1	0	81	0	100·0	100·0
	A	*p, h*	79	2	1	0	97·5	100·0
ISWIR_104	T	*mu*	0	0	82	0	–	100·0
ISWIR_120	C	*h*	39	3	40	0	92·9	100·0
	T	*p, m*	38	2	42	0	95·0	100·0
ISWIR_249	G	*m*	1	0	81	0	100·0	100·0
	A	*p, h*	72	9	1	0	88·9	100·0
Pooled			683	92	1728	39	88·1	97·8

m: *C. meleagridis*; *p*: *C. parvum*; *h*: *C. hominis*; *s*: *C. suis*; *f*: *C. felis*; *c*: *C. canis*; *u*: *C. ubiquitum*; *mu*: *C. muris*. *Sn*: *sensitivity*; *Sp*: specificity (both expressed as percentage). The number of true positive (*TP*), false negatives (*FN*), true negatives (*TN*) and false positives (*FP*) were estimated by comparing the iPLEX typing results with those obtained by Sanger sequencing (considered the gold standard, see Materials and Methods). Additional SNVs selected to reinforce the accuracy of base calling in a reduced number of samples with uncertain genotypes are typed in bold.

Abal-Fabeiro *et al*. 2013). Nineteen SNVs were selected for species identification and 36 for subtype assignment. The subtype of each sample was determined on the basis of the genotypes of various SNVs at *gp60*. Two other loci, *ACoAs* and *gtub*, were also used for verification purposes, taking advantage of the fact that the patterns of genetic variation in *Cryptosporidium* at these loci are structured according to the *gp60* haplotypes (Abal-Fabeiro *et al*. 2013).

Four types of internal quality-controls of the genotyping protocol were included: (i) 91 samples previously sent to the Center for Disease Control and Prevention (CDC, Atlanta, USA) for identification at the species level by Sanger sequencing of both the *COWP1* and *SSU rRNA* (82 of which were successfully characterized). (ii) A subset of 24 of these 82 samples typed by Sanger sequencing of *gp60* (Abal-Fabeiro *et al*. 2013). (iii) Repeated samples collected from the same individual at different stages of the infection, which were analysed blindly (i.e. 38 patients and 7 cows were sampled twice, and 5 other patients were sampled three times). (iv) Human blood DNA samples (*Cryptosporidium* free) and reagent blanks (negative controls).

Results were analysed in a sequential way; in the first round only the SNVs determining the species were evaluated, while in the second round the genotypes of a different group of SNVs were used to assign the subtype of isolates successfully characterized at the species level.

Given the high levels of within-species diversity described in *C. meleagridis* (Abal-Fabeiro *et al*. 2013) all the samples infected with this species were PCR amplified and sequenced at the *gp60* locus to check if the genotypes matched the Sanger sequencing results using primers and protocols previously described (Abal-Fabeiro *et al*. 2013). Other *Cryptosporidium* samples showing iPLEX inconclusive results were also sequenced at the *gp60* locus for species or subtype identification.

Table 2. Sensitivity and specificity assessment (%) of the 36 single nucleotide variant (SNV) markers used to identify *Cryptosporidium* gp60 subtypes

SNV name	Allele	Identifies	*TP*	*FN*	*TN*	*FP*	*Sn*	*Sp*
ACOAS_256	C	Ib	8	1	1	0	88·9	100·0
	T	Id	1	0	9	0	100·0	100·0
ACOAS_686	A	IIa	9	1	3	0	90·0	100·0
	G	IIn	3	0	10	0	100·0	100·0
GP60_33_1	C	Ia	0	0	10	0	–	100·0
GP60_33_2	A	Ig	0	0	10	0	–	100·0
GP60_45-1	G	Ia	0	0	10	0	–	100·0
GP60_45-2	T	Ig	0	0	9	1	–	90·0
GP60_110	G	IIIb	0	0	1	0	–	100·0
	A	IIIf, IIIi	0	1	0	0	0·0	–
GP60_111	T	IIIg	0	0	1	0	–	100·0
	C	IIIf, IIIh, IIIi	1	0	0	0	100·0	–
GP60_131	G	IIIh	0	0	1	0	–	100·0
GP60 – 144	C	IIb	0	0	13	0	–	100·0
	T	IIa, IId, IIf, IIg, IIj,IIn	0	10	3	0	0·0	–
GP60_162	T	If	0	0	10	0	–	100·0
GP60_171_1	T	IIf	0	0	13	0	–	100·0
GP60_171_2	A	IIIb	0	0	1	0	–	100·0
	C	IIIf, IIIg,IIIh, IIIi	1	0	0	0	100·0	–
GP60_181	G	IIj	0	0	13	0	–	100·0
GP60_183	C	IId	0	0	13	0	–	100·0
GP60_215	C	IIb	0	0	13	0	–	100·0
GP60_219	T	If	0	0	10	0	–	100·0
GP60_228	T	IId	0	0	13	0	–	100·0
	C	IIb, IIf	0	0	13	0	–	100·0
GP60_241	G	IIh	0	0	13	0	–	100·0
GP60_243	C	IIc	0	0	13	0	–	100·0
GP60_259	A	IIIf, g	1	0	0	0	100·0	–
	G	IIIb, IIIh, IIIi	0	0	1	0	–	100·0
GP60_260	C	IIj	0	0	13	0	–	100·0
GP60_261	C	Ia, Ig	0	0	10	0	–	100·0
GP60_270	C	IIg	0	0	13	0	–	100·0
GP60_274	A	IIc	0	0	13	0	–	100·0
GP60_276	A	IId	0	0	13	0	–	100·0
GP60_281	C	IIh	0	0	13	0	–	100·0
GP60_305	G	IIIb	0	0	0	0	–	–
GP60_318	T	Ie	0	0	10	0	–	100·0
GP60_335	C	IIIf	1	0	0	0	100·0	–
	G	IIIb, IIIg, IIIh, IIIi	0	0	1	0	–	100·0
GP60_349	A	Ie	0	0	10	0	–	100·0
GTUB_60	A	IIIj	0	0	1	0	–	100·0
	G	IIIf, IIIg, IIIh	1	0	0	0	100·0	–
GTUB_85	A	Ib	9	0	1	0	100·0	100·0
	C	Id	1	0	9	0	100·0	100·0
GTUB_87	C	IIIf, IIIh	1	0	0	0	100·0	–
	T	IIIg, IIIj	0	0	1	0	–	100·0
GTUB_417	T	Ib	8	1	1	0	88·9	100·0
	C	Id	1	0	9	0	100·0	100·0
GTUB_705	G	IIIh	0	0	1	0	–	100·0
	A	IIIf, IIIg, IIIj	1	0	0	0	100·0	–
Pooled			47	14	326	1	77·0	99·7

TP, FN, TN and *FP* were estimated by comparing the iPlex typing results with those obtained by Sanger sequencing (considered the gold standard) of *gp60* (Abal-Fabeiro *et al*. 2013). Additional SNVs selected to reinforce the accuracy of base calling in a reduced number of samples with uncertain genotypes are typed in bold.

To describe the diagnostic performance of each SNV allele, two widely used parameters were estimated: *sensitivity* (*Sn*), defined as the proportion of cases with the investigated attribute (i.e. a particular genotype) which are correctly identified and *specificity* (*Sp*), which measures the proportion of correct negative calls of the marker. $Sn = TP/(TP + FN)$, where TP is the number of true positive (TP) calls and FN the number of false negatives (FN). $Sp = TN/(TN + FP)$, where TN and FP stand for the

numbers of true negative and false positive calls, respectively. The number of TP, FN, true negatives (TN) and false positives (FP) were estimated by comparing the iPLEX typing results with those obtained by Sanger sequencing of diagnostic loci in a subset of samples (see above).

RESULTS

Identification of the species

Nineteen SNVs at five different genomic loci were selected to identify *Cryptosporidium* species (Table 1). Seven of the SNVs were monoallelic (i.e. the genotyping reaction is expected to work only if a particular allele is present) and 12 biallelic (two alternative results are expected). The calling performance of each SNV allele was determined by comparing their typing results with those obtained by successful Sanger sequencing of 82 out of 91 samples previously sent to the CDC for species identification. iPLEX genotyping permitted the identification of the *Cryptosporidium* species in 86 out of these 91 control isolates (94·5%), a slightly larger fraction than with the Sanger sequencing method, which was successful for 82 samples (90·1%; the difference is not statistically significant, $P = 0{\cdot}40$, in a two-tailed Fisher's exact test). The two methods produced coincident species determinations across the 91 control samples. The only discrepancies corresponded to two samples that were not typed by the iPLEX method, six by Sanger sequencing and three samples that were not typed by any of the two approaches.

Redundant SNVs were used to ensure the correct classification of the samples and to prevent against sporadic failures of the genotyping assays. This can be illustrated with the case of the two alleles of COWP1_399: the first one (nucleotide T), failed to detect the single *C. meleagridis* positive control (Table 1), as determined by Sanger sequencing and by other redundant SNVs such as ACOAS_235, COWP1_333, GT_804 or ISWIR_249, that correctly identified this species. The second allele (nucleotide A), only detected 32 out of the 81 *Cryptosporidium parvum* or *Cryptosporidium hominis* samples among the controls, which were also identified by alternative markers.

Despite the failure of some SNV callings, pooled sensitivity (*Sn*) and specificity (*Sp*) values across all markers (based on samples where Sanger sequencing produced a result) were high: 88·1 and 97·8%, respectively.

Identification of gp60 subtypes

Once the species was identified for each isolate, a new set of 36 SNVs selected for subtype determination were genotyped (22 monoallelic and 14 biallelic, Table 2). The results of the subtyping matched precisely those obtained by Sanger sequencing of the *gp60* gene in a subset of 24 human isolates that were used as positive controls for the subtyping process: Ib ($n = 9$), IIa ($n = 10$), IIn ($n = 3$), Id ($n = 1$) and IIIf ($n = 1$) (Abal-Fabeiro *et al*. 2013). Again, there was large redundancy between many of the markers and their performance varied significantly. Pooled *Sn* and *Sp* parameters values were high: 77·0 and 99·7, respectively. The lower sensitivity can be attributed to the performance of one marker (GP60_144; nucleotide T), which displayed a high proportion of false negative results.

Considering the two-step genotyping process as a whole (including the species and subtype determination), the overall *Sn* and *Sp* values for the combined set of markers were 87·3 and 98·0, respectively. These parameters would reach even higher values if only a selection of those SNVs with the best performance in terms of sensitivity and specificity was considered (pooled $Sn = 94{\cdot}5$ and $Sp = 99{\cdot}8$; Supplemental Table 2S – in Online version only).

Table 3. Subtyping results obtained by SNV-marker high-throughput genotyping in a panel of 671 human and bovine samples

Species	*gp60* subtype	Human	Cattle
C. hominis	Ia	3	0
	Ib	336	0
	Id	27	0
	Ie	4	0
	U	0	1
C. parvum	IIa	155	56
	IId	3	2
	IIn	16	0
	U	2	1
C. meleagridis	IIIb	1	0
	IIIf	2	0
C. felis		1	0
Not typed		58	3
Total		608	63

U: undetermined.

Analysis of a panel of 671 human and bovine samples

The analysis of the full panel of samples allowed species identification in 550 of the 608 human isolates (90·5%), 370 of which corresponded to *C. hominis* (60·8%), 176 to *C. parvum* (28·9%), three to *C. meleagridis* (0·5%) and one to *C. felis* (0·1%) (Table 3). Ambiguous results with genotype calls corresponding to *C. hominis* and *C. parvum* were obtained in two additional samples. Isolate 578 produced, among other base calls common to both species, a *C. parvum* specific G at GTUB_360 and a *C. hominis* specific C at GTUB_735 and ISWIR_120

(Supplemental Table 3S – in Online version only). The other one, sample 744, showed a heterozygous C/T at COWP1_114 compatible with an infection with several species (Supplemental Table 3S – in Online version only), although the remaining SNV calls were specific to *C. hominis*. To further investigate the existence of multiple infections, the *gp60* locus in these samples was sequenced by the Sanger method to determine not only the species but also the subtype/s causing the cryptosporidiosis. We found no evidence for variation in the length of the PCR amplicons and there were no double peaks in the sequencing electropherograms, as would be expected in the case of a mixed infection with two different *gp60* subtypes. The nucleotide sequences revealed that sample 578 contained *C. parvum* IIa and sample 744 *C. hominis* Ib.

Sixty (95·2%) of the 63 bovine isolates were successfully characterized and *C. parvum* was found in 59 cases (Table 3 and Supplemental Table 3S – in Online version only). One of the bovine samples (G27) showed a C call at ISWIR_120 compatible with the presence of *C. hominis*, something rarely reported to date (Smith *et al*. 2005). We failed to PCR amplify the *gp60* product for this allele.

The *gp60* subtype was determined in 605 out of the 671 isolates and the results obtained were consistent across samples collected from the same individual at different stages of the infection, except in three of them where one of the repeats failed to produce signal at any SNV (data not shown).

In humans the most frequent *C. hominis* subtype was Ib ($n = 336$; Table 3), followed by Id ($n = 27$), while the presence of Ia and Ie was uncommon (3 and 4 isolates, respectively). The predominant variant in *C. parvum* was IIa ($n = 155$), while other subtypes (IIn and IId) were found at much lower frequencies ($n = 16$ and 3, respectively). Two *C. parvum* isolates (10 and 572, Supplemental Table 4S – in Online version only) showed inconclusive results (i.e. lack of genotype call at the relevant SNVs) and only after sequencing their *gp60* gene they could be unambiguously assigned to IIa.

Two *C. meleagridis* isolates (95 and 451, Supplemental Table 4S – in Online version only) were also genotyped by Sanger sequencing of the *gp60* locus and subsequently assigned to subtype IIIf (Abal-Fabeiro *et al*. 2013) in good agreement with the iPLEX™ method. The third *C. meleagridis* isolate (161), typed as IIIb, could be not verified by Sanger sequencing.

Subtype was identified in 58 (92·1%) of the 63 bovine samples. The most prevalent subtype was IIa ($n = 56$), followed by IId ($n = 2$) (Table 3). One additional sample (G39) exhibited inconclusive results (Supplemental Table 4S – in Online version only), being classified as IIa only after its *gp60* nucleotide sequence was determined by Sanger sequencing.

DISCUSSION

The accurate diagnosis of cryptosporidiosis relies on the correct identification of *Cryptosporidium* species and subtypes, and it is essential for the study of the population structure, epidemiology and evolution of the parasite (Gasser *et al*. 2006; Jex *et al*. 2007*b*). The precise classification of the different genetic variants of this pathogen is central to detect potential sources of contamination and therefore to the prevention and control of these parasites (Jex *et al*. 2007*a*; Chalmers, 2008). To estimate the disease burden, identify outbreaks or understand population trends in surveillance data (Gasser, 2006; Chalmers, 2008; The ANOFEL Cryptosporidium National Network, 2010; Yoder *et al*. 2010) the genetic characterization of hundreds of samples is required, making increasingly necessary the application of efficient high-throughput technologies capable of genotyping large sample panels at low cost. Here we present one such method for the simultaneous genotype determination of a moderate number of SNVs at different loci and apply it to a large sample.

Marker selection

One major difficulty for the use of high-throughput technologies for molecular typing of non-model organisms is the lack of a panel of variants (preferably single nucleotide substitutions) that can be used as a reference. Here, we took advantage of the increasing number of nucleotide sequences of marker loci from different *Cryptosporidium* isolates that are available in public databases (GenBank, CryptoDB) and our own work (Abal-Fabeiro *et al*. 2013) to compile a panel of SNVs for the identification of *Cryptosporidium* genotypes commonly found in humans.

The ability of a technique to classify the different *Cryptosporidium* species depends on the choice of suitable genetic markers, which should combine low levels of polymorphism within species with a considerable degree of variation among them (Gasser, 2006). One of the most widely used genes for the classification of *Cryptosporidium* isolates is the small ribosomal subunit (*SSU rRNA*), which has been sequenced for all species (Jex *et al*. 2008). However, as found in the *C. parvum* and *C. hominis* genomes, this locus is present in several paralogous copies not necessarily identical (Strong and Nelson, 2000; Abrahamsen *et al*. 2004; Xu *et al*. 2004). In most *Cryptosporidium* species the number and sequences of the *SSU rRNA* copies are still ignored and the finding of any new genetic variant could be erroneously assigned to a novel species, when it could be a previously unidentified copy from a known species (Navarro-i-Martinez *et al*. 2003). The use of single copy genes as molecular markers is therefore highly recommended, since they facilitate the

sequence analyses and avoid misinterpretations of the data. Consequently, all marker SNVs selected for the multiplex study correspond to single copy nuclear loci (Abal-Fabeiro *et al*. 2013): *COWP1* and *gp60* had been extensively used for typing the main species involved in human and bovine infections (McLauchlin *et al*. 2000; Pedraza-Diaz *et al*. 2001; Kato *et al*. 2003; Leoni *et al*. 2006; Hunter *et al*. 2007; Geurden *et al*. 2009; Widmer and Lee, 2010) and the other three loci, (*ACoAs, gtub* and *ISWIr*) also present variants that are fixed between species (Abal-Fabeiro *et al*. 2013).

Furthermore, the use of a single locus to classify the isolates in different groups of alleles (subtyping) can be problematic (Widmer, 2009). The analysis of the microsatellite length polymorphism in *gp60* (Strong *et al*. 2000) as the main tool to identify *Cryptosporidium* variants exhibited inconsistencies in *C. hominis* and *C. parvum*, where sequences with the same microsatellite size carried different serine trinucleotides (Sulaiman *et al*. 2005; Gatei *et al*. 2007). As a consequence, many authors identified subtypes according to length polymorphism and nucleotide variation by counting the number of TCA, TCG or TCT repeats present in the *gp60* microsatellite region (Sulaiman *et al*. 2005; Jex and Gasser, 2010; Xiao, 2010). However, multilocus studies comparing the subtype classification obtained with *gp60* and other loci showed discrepancies; for example, samples sharing a *gp60* allele exhibited differences when other genetic markers, particularly micro and mini-satellites, were included. These observations were made after scoring isolates either according to amplicon lengths (Mallon *et al*. 2003; Tanriverdi and Widmer, 2006) or combining both amplicon lengths and SNVs (Gatei *et al*. 2007, 2008; Widmer and Lee, 2010). Considering that microsatellites are highly unstable and have some of the highest mutation rates observed at molecular loci (Goldstein and Pollock, 1997; Ellegren, 2004), polymorphisms either in length or in nucleotide composition are not excessively surprising. The finding of genetic differences at distinct microsatellite loci, is also predictable. Therefore, and to avoid such conflicts, we selected SNVs in *gp60* (Strong *et al*. 2000), in a region which did not include the microsatellite, and two additional single copy nuclear genes, *gtub* and *ACoAs*, which proved to be useful to classify samples into groups that can be associated with distinct *gp60* subtypes (Abal-Fabeiro *et al*. 2013).

Performance

We designed a two-step genotyping protocol based on the iPLEX methodology, consisting of an initial determination of the species, followed by the identification of the *gp60* subtype. It displayed very good performance at the two phases of the analysis, particularly at the identification of the species. However, two kinds of problems arose at this stage in some samples: the failure to produce an extension at the targeted SNVs and the presence of ambiguous genotype calls. Indeed, 49 out of the 61 samples that were not assigned to a particular species (Table 3) failed to produce an extension at all targeted SNVs (Supplemental Table 3S – in Online version only). This probably reflects lack of success in the single-base extension step due to below-requirements template DNA quantity or quality. In this respect it must be noted that the iPLEX methodology produced genotype data for six samples where Sanger sequencing failed. This performance is consistent with the observation that the former method has lower DNA quality restrictions and works comparatively better for samples with low DNA concentration or partially degraded DNA (Mendisco *et al*. 2011).

In samples with ambiguous calls corresponding to *C. hominis* and *C. parvum*, the lack of supporting evidence from the Sanger sequencing data does not rule out the possibility of mixed infections. Given the smaller amplicon sizes needed by the mass spectrometry method, the mixed results might reflect the presence of degraded DNA templates of a hypothetical second genotype, which would be missed by Sanger sequencing. A similar pattern would be observed if the parasite load of one of the species was much larger than the second one. To address the issue of mixed infections in which small amounts of DNA from a minor species could be involved, alternative methods like PCR-RFLP, SSCP or the use of species-specific primers have been proposed (Chalmers *et al*. 2005).

Genotyping results

MALDI-TOF genotyping performed as well as Sanger sequencing in our internal controls. In fact, the most informative SNVs (those selected because of their best performance in terms of sensitivity and specificity) showed mean sensitivity and specificity results comparable to those obtained with other molecular techniques such as real-time PCR (Stroup *et al*. 2006; Hadfield *et al*. 2011), RFLP analysis of *COWP1* and *SSU rRNA* (Chalmers *et al*. 2009), Sanger sequencing of *gp60* (Chalmers *et al*. 2008; Hijjawi *et al*. 2010) or fragment-size-analysis of three microsatellite markers (*gp60, ML1* and *ML2*) (Hunter *et al*. 2007).

We identified nine different *gp60* subtypes from four *Cryptosporidium* species in our panel of human-derived isolates (Table 3). The most common subtypes were Ib and IIa, which represent 61·3 and 28·3% of the typed human samples, respectively.

Subtypes Id and IIn were also abundant while the others were found in just a few isolates. The performance of probes designed to identify the less frequent *gp60* subtypes, like *C. hominis* If, Ig or *C. parvum* IIb, IIc, IIf, IIg, IIh, IIj could not be determined because none of them were detected in our sample panel despite the use of multiple SNVs.

Most of the markers that failed the primer extension were designed for the identification of *C. meleagridis* subtypes. Given the particularly high levels of within-species *gp60* diversity described in this species (Abal-Fabeiro *et al*. 2013), the possibility that these failures reflect undetected variation in the priming sites cannot be excluded.

The bovine sample group was not so diverse because most animals harboured *C. parvum*. The only exception was one calf putatively infected with *C. hominis*. Nevertheless, this result has to be taken with caution, since this isolate only produced base calls at three SNVs and all attempts to sequence the *gp60* gene in this sample were unsuccessful, probably due to the lack of good-quality DNA.

Although mass spectrometry has been widely used for the genotyping of SNVs for population surveys, often with medical purposes, as far as we know this is the first time this technology was used for the genotyping of *Cryptosporidium* samples. In our view, this method has the following advantages over Sanger sequencing: (i) it allows the simultaneous analysis of a large number of markers. This is particularly interesting because it allows the scrutiny of multiple genomic regions. (ii) The multiplexing design allows an easy update of the SNVs panel to be analysed. Further studies with the same purpose as this one could be limited to the sets of most informative SNVs (Supplemental Table 2S – in Online version only), greatly improving the method's performance indicators (sensitivity and specificity) without a noticeable loss of information. In addition, other SNVs of interest, for example to identify other species or *gp60* subtypes of interest, could also be included in multiplex designs. (iii) The method is particularly suitable for the genotyping of large numbers of samples, with high automatization and limited handling. (iv) It is much faster than other sequencing techniques; the full analysis of our panel of 55 SNVs in the 671 samples can be accomplished in less than 3 weeks, as compared with the several months necessary if other methods are used. (v) We also estimated that this technology is around 40 times cheaper than Sanger sequencing, which is now commonly used in *Cryptosporidium* studies, with 0·1 € per genotype as compared with over 4·0 € per Sanger sequencing reaction. The cost benefit increases with the sample size to be analysed. (vi) The small size of the amplicons means that the method can be used even with degraded template DNA. Finally (vii), the genotyping results are very robust, as demonstrated by the high repeatability in our replicated samples.

Overall, here we presented a method for the high-throughput genotyping of *Cryptosporidium* isolates in large sample panels. The method is based on the multiplex analysis of a collection of SNVs used as species and *gp60* subtype markers, by means of mass spectrometry. Compared to Sanger sequencing this method is practical, efficient, cheaper and less time consuming and it can be progressively adapted for the diagnosis of new species or subtypes as soon as their sequences are characterized.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit http://dx.doi.org/10.1017/S0031182013001807.

ACKNOWLEDGEMENTS

We are grateful to Alfredo Rúa and David Fernández-García (Laboratorio de Sanidade e Producción Animal, Lugo, Spain) who kindly provided *Cryptosporidium*-positive bovine samples. We thank Dr A. Campos (Servicio de Hematoloxía-CHUS, Santiago de Compostela, Spain) for lending us laboratory equipment, as well as the Fundación Pública Galega de Medicina Xenómica (SERGAS, Spain) for the use of their laboratory facilities, and Santiago Méndez-Torres (Servicio de Microbioloxía e Parasitoloxía-CHUS, Santiago de Compostela, Spain) for technical assistance, and Luis Navarro-i-Martinez, Fernando Bornay-Llinares (Universidad Miguel Hernández, Alicante, Spain) and Norman J. Pieniazek (CDC, Atlanta, USA) for the Sanger sequencing of samples used as positive controls. We also thank the anonymous reviewers whose careful review and suggestions helped to improve an earlier version of this manuscript.

FINANCIAL SUPPORT

This work was financed by a grant from Xunta de Galicia, Spain to CB (PGIDIT06PXIB228073PR). JLA-F was supported by a graduate student fellowship co-financed by Xunta de Galicia, Spain and the European Social Fund; XM was supported by a Programa Ramón y Cajal contract (Ministerio de Ciencia e Innovación, Spain); XB was supported by grant PGIDIT06PXIB228073PR (Xunta de Galicia. Spain) and CB was supported by a Programa Isidro Parga Pondal contract (Xunta de Galicia, Spain).

REFERENCES

Abal-Fabeiro, J. L., Maside, X., Bello, X., Llovo, J. and Bartolomé, C. (2013). Multilocus patterns of genetic variation across *Cryptosporidium* species suggest balancing selection at the *gp60* locus. *Molecular Ecology* **22**, 4723–4732. doi: 10.1111/mec.12425.

Abrahamsen, M. S., Templeton, T. J., Enomoto, S., Abrahante, J. E., Zhu, G., Lancto, C. A., Deng, M., Liu, C., Widmer, G., Tzipori, S., Buck, G. A., Xu, P., Bankier, A. T., Dear, P. H., Konfortov, B. A., Spriggs, H. F., Iyer, L., Anantharaman, V., Aravind, L. and Kapur, V. (2004). Complete genome sequence of the apicomplexan, *Cryptosporidium parvum*. *Science* **304**, 441–445. doi: 10.1126/science.1094786.

Casemore, D. P., Amstrong, M., Jackson, B., Nichols, G. and Thorn, B. T. (1984). Screening for *Cryptosporidium* in stools. *Lancet* **323**, 734–735. doi: 10.1016/S0140-6736(84)92245-1.

Chalmers, R. M. (2008). *Cryptosporidium*: from laboratory diagnosis to surveillance and outbreaks. *Parasite* **15**, 372–378. doi: 10.1051/parasite/2008153372.

Chalmers, R. M., Ferguson, C., Caccio, S., Gasser, R. B., Abs, E. L. O. Y. G., Heijnen, L., Xiao, L., Elwin, K., Hadfield, S., Sinclair, M. and Stevens, M. (2005). Direct comparison of selected methods for genetic categorisation of *Cryptosporidium parvum* and *Cryptosporidium hominis* species. *International Journal for Parasitology* **35**, 397–410. doi: 10.1016/j.ijpara.2005.01.001.

Chalmers, R. M., Hadfield, S. J., Jackson, C. J., Elwin, K., Xiao, L. and Hunter, P. (2008). Geographic linkage and variation in *Cryptosporidium hominis*. *Emerging Infectious Diseases* **14**, 496–498. doi: 10.3201/eid1403.071320.

Chalmers, R. M., Elwin, K., Thomas, A. L., Guy, E. C. and Mason, B. (2009). Long-term *Cryptosporidium* typing reveals the aetiology and species-specific epidemiology of human cryptosporidiosis in England and Wales, 2000 to 2003. *Eurosurveillance* **14**, pii: 19086.

Ellegren, H. (2004). Microsatellites: simple sequences with complex evolution. *Nature Reviews of Genetics* **5**, 435–445. doi: 10.1038/nrg1348.

Gabriel, S., Ziaugra, L. and Tabbaa, D. (2009). SNP genotyping using the Sequenom MassARRAY iPLEX platform. *Current Protocols in Human Genetics*, Chapter 2, Unit 2 12. doi: 10.1002/0471142905.hg0212s60.

Gasser, R. B. (2006). Molecular tools – advances, opportunities and prospects. *Veterinary Parasitology* **136**, 69–89. doi: 10.1016/j.vetpar.2005.12.002.

Gasser, R. B., Hu, M., Chilton, N. B., Campbell, B. E., Jex, A. J., Otranto, D., Cafarchia, C., Beveridge, I. and Zhu, X. (2006). Single-strand conformation polymorphism (SSCP) for the analysis of genetic variation. *Nature Protocols* **1**, 3121–3128. doi: 10.1038/nprot.2006.485.

Gatei, W., Das, P., Dutta, P., Sen, A., Cama, V., Lal, A. A. and Xiao, L. (2007). Multilocus sequence typing and genetic structure of *Cryptosporidium hominis* from children in Kolkata, India. *Infection, Genetics and Evolution* **7**, 197–205. doi: 10.1016/j.meegid.2006.08.006.

Gatei, W., Barrett, D., Lindo, J. F., Eldemire-Shearer, D., Cama, V. and Xiao, L. (2008). Unique *Cryptosporidium* population in HIV-infected persons, Jamaica. *Emerging Infectious Diseases* **14**, 841–843. doi: 10.3201/eid1405.071277.

Geurden, T., Levecke, B., Caccio, S. M., Visser, A., De Groote, G., Casaert, S., Vercruysse, J. and Claerebout, E. (2009). Multilocus genotyping of *Cryptosporidium* and *Giardia* in non-outbreak related cases of diarrhoea in human patients in Belgium. *Parasitology* **136**, 1161–1168. doi: 10.1017/S0031182009990436.

Goldstein, D. B. and Pollock, D. D. (1997). Launching microsatellites: a review of mutation processes and methods of phylogenetic interference. *Journal of Heredity* **88**, 335–342.

Hadfield, S. J., Robinson, G., Elwin, K. and Chalmers, R. M. (2011). Detection and differentiation of *Cryptosporidium* spp. in human clinical samples by use of real-time PCR. *Journal of Clinical Microbiology* **49**, 918–924. doi: 10.1128/JCM.01733-10.

Hijjawi, N., Ng, J., Yang, R., Atoum, M. F. and Ryan, U. (2010). Identification of rare and novel *Cryptosporidium GP60* subtypes in human isolates from Jordan. *Experimental Parasitology* **125**, 161–164. doi: 10.1016/j.exppara.2010.01.011.

Hunter, P. R., Hadfield, S. J., Wilkinson, D., Lake, I. R., Harrison, F. C. and Chalmers, R. M. (2007). Subtypes of *Cryptosporidium parvum* in humans and disease risk. *Emerging Infectious Diseases* **13**, 82–88. doi: 10.3201/eid1301.060481.

Jex, A. R. and Gasser, R. B. (2010). Genetic richness and diversity in *Cryptosporidium hominis* and *C. parvum* reveals major knowledge gaps and a need for the application of "next generation" technologies – research review. *Biotechnology Advances* **28**, 17–26. doi: 10.1016/j.biotechadv.2009.08.003.

Jex, A. R., Ryan, U. M., Ng, J., Campbell, B. E., Xiao, L., Stevens, M. and Gasser, R. B. (2007*a*). Specific and genotypic identification of *Cryptosporidium* from a broad range of host species by nonisotopic SSCP analysis of nuclear ribosomal DNA. *Electrophoresis* **28**, 2818–2825. doi: 10.1002/elps.200600772.

Jex, A. R., Whipp, M., Campbell, B. E., Caccio, S. M., Stevens, M., Hogg, G. and Gasser, R. B. (2007*b*). A practical and cost-effective mutation scanning-based approach for investigating genetic variation in *Cryptosporidium*. *Electrophoresis* **28**, 3875–3883. doi: 10.1002/elps.200700279.

Jex, A. R., Smith, H. V., Monis, P. T., Campbell, B. E. and Gasser, R. B. (2008). *Cryptosporidium* – biotechnological advances in the detection, diagnosis and analysis of genetic variation. *Biotechnology Advances* **26**, 304–317. doi: 10.1016/j.biotechadv.2008.02.003.

Kato, S., Lindergard, G. and Mohammed, H. O. (2003). Utility of the *Cryptosporidium* oocyst wall protein (COWP) gene in a nested PCR approach for detection infection in cattle. *Veterinary Parasitology* **111**, 153–159. doi: 10.1016/S0304-4017(02)00353-9.

Leoni, F., Amar, C., Nichols, G., Pedraza-Diaz, S. and McLauchlin, J. (2006). Genetic analysis of *Cryptosporidium* from 2414 humans with diarrhoea in England between 1985 and 2000. *Journal of Medical Microbiology* **55**, 703–707. doi: 10.1099/jmm.0.46251-0.

Mallon, M. E., MacLeod, A., Wastling, J. M., Smith, H. and Tait, A. (2003). Multilocus genotyping of *Cryptosporidium parvum* Type 2: population genetics and sub-structuring. *Infection, Genetics and Evolution* **3**, 207–218. doi: 10.1016/S1567-1348(03)00089-3.

McLauchlin, J., Amar, C., Pedraza-Diaz, S. and Nichols, G. L. (2000). Molecular epidemiological analysis of *Cryptosporidium* spp. in the United Kingdom: results of genotyping *Cryptosporidium* spp. in 1,705 fecal samples from humans and 105 fecal samples from livestock animals. *Journal of Clinical Microbiology* **38**, 3984–3990.

Mendisco, F., Keyser, C., Hollard, C., Seldes, V., Nielsen, A. E., Crubezy, E. and Ludes, B. (2011). Application of the iPLEX Gold SNP genotyping method for the analysis of Amerindian ancient DNA samples: benefits for ancient population studies. *Electrophoresis* **32**, 386–393. doi: 10.1002/elps.201000483.

Navarro-i-Martinez, L., Bornay-Llinares, F. J., Rueda, C., del Aguila, C., da Silva, A. J., Oleaga, A., Ramajo, V., Fenoy, S. and Pieniazek, N. J. (2003). Molecular characterization of *Cryptosporidium* sp. from animals in Spain. *Journal of Eukaryotic Microbiology* **50**(Suppl.), 553–554. doi: 10.1111/j.1550-7408.2003.tb00626.x.

Pedraza-Diaz, S., Amar, C. F., McLauchlin, J., Nichols, G. L., Cotton, K. M., Godwin, P., Iversen, A. M., Milne, L., Mulla, J. R., Nye, K., Panigrahl, H., Venn, S. R., Wiggins, R., Williams, M. and Youngs, E. R. (2001). *Cryptosporidium meleagridis* from humans: molecular analysis and description of affected patients. *Journal of Infection* **42**, 243–250. doi: 10.1053/jinf.2001.0839.

Sauer, S. and Gut, I. G. (2002). Genotyping single-nucleotide polymorphisms by matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry. *Journal of Chromatography B. Analytical Technologies in Biomedical and Life Sciences* **782**, 73–87. doi: 10.1016/S1570-0232(02)00692-X.

Smith, H. V., Nichols, R. A., Mallon, M., Macleod, A., Tait, A., Reilly, W. J., Browning, L. M., Gray, D., Reid, S. W. and Wastling, J. M. (2005). Natural *Cryptosporidium hominis* infections in Scottish cattle. *Veterinary Record* **156**, 710–711. doi: 10.1136/vr.156.22.710.

Strong, W. B. and Nelson, R. G. (2000). Preliminary profile of the *Cryptosporidium parvum* genome: an expressed sequence tag and genome survey sequence analysis. *Molecular and Biochemical Parasitology* **107**, 1–32. doi: 10.1016/S0166-6851(99)00225-X.

Strong, W. B., Gut, J. and Nelson, R. G. (2000). Cloning and sequence analysis of a highly polymorphic *Cryptosporidium parvum* gene encoding a 60-kilodalton glycoprotein and characterization of its 15- and 45-kilodalton zoite surface antigen products. *Infection and Immunity* **68**, 4117–4134. doi: 10.1128/IAI.68.7.4117-4134.2000.

Stroup, S. E., Roy, S., McHele, J., Maro, V., Ntabaguzi, S., Siddique, A., Kang, G., Guerrant, R. L., Kirkpatrick, B. D., Fayer, R., Herbein, J., Ward, H., Haque, R. and Houpt, E. R. (2006). Real-time PCR detection and speciation of *Cryptosporidium* infection using Scorpion probes. *Journal of Medical Microbiology* **55**, 1217–1222. doi: 10.1099/jmm.0.46678-0.

Sulaiman, I. M., Hira, P. R., Zhou, L., Al-Ali, F. M., Al-Shelahi, F. A., Shweiki, H. M., Iqbal, J., Khalid, N. and Xiao, L. (2005). Unique endemicity of cryptosporidiosis in children in Kuwait. *Journal of Clinical Microbiology* **43**, 2805–2809. doi: 10.1128/JCM.43.6.2805-2809.2005.

Tanriverdi, S. and Widmer, G. (2006). Differential evolution of repetitive sequences in *Cryptosporidium parvum* and *Cryptosporidium hominis*. *Infection, Genetics and Evolution* **6**, 113–122. doi: 10.1016/j.meegid.2005.02.002.

The ANOFEL Cryptosporidium National Network (2010). Laboratory-based surveillance for *Cryptosporidium* in France, 2006–2009. *Eurosurveillance* **15**, 19642.

Widmer, G. (2009). Meta-analysis of a polymorphic surface glycoprotein of the parasitic protozoa *Cryptosporidium parvum* and *Cryptosporidium hominis*. *Epidemiology and Infection* **137**, 1800–1808. doi: 10.1017/S0950268809990215.

Widmer, G. and Lee, Y. (2010). Comparison of single- and multilocus genetic diversity in the protozoan parasites *Cryptosporidium parvum* and *C. hominis*. *Applied and Environmental Microbiology* **76**, 6639–6644. doi: 10.1128/AEM.01268-10.

Xiao, L. (2010). Molecular epidemiology of cryptosporidiosis: an update. *Experimental Parasitology* **124**, 80–89. doi: 10.1016/j.exppara.2009.03.018.

Xiao, L., Fayer, R., Ryan, U. and Upton, S. J. (2004). *Cryptosporidium* taxonomy: recent advances and implications for public health. *Clinical Microbiology Reviews* **17**, 72–97. doi: 10.1128/CMR.17.1.72-97.2004.

Xu, P., Widmer, G., Wang, Y., Ozaki, L. S., Alves, J. M., Serrano, M. G., Puiu, D., Manque, P., Akiyoshi, D., Mackey, A. J., Pearson, W. R., Dear, P. H., Bankier, A. T., Peterson, D. L., Abrahamsen, M. S., Kapur, V., Tzipori, S. and Buck, G. A. (2004). The genome of *Cryptosporidium hominis*. *Nature* **431**, 1107–1112. doi: 10.1038/nature02977.

Yoder, J. S., Harral, C. and Beach, M. J. (2010). Cryptosporidiosis surveillance – United States, 2006–2008. *MMWR. Surveillance Summaries* **59**, 1–14.