Infection, Genetics and Evolution 31 (2015) 245–249

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Geographical segregation of *Cryptosporidium parvum* multilocus genotypes in Europe

Simone M. Cacciò [a,*], Valerie de Waele [b], Giovanni Widmer [c]

[a] Istituto Superiore di Sanità, Department of Infectious, Parasitic and Immunomediated Diseases, Viale Regina Elena 299, Rome 00161, Italy
[b] Institute of Tropical Medicine, Nationalestraat 155, 2000 Antwerpen, Belgium
[c] Cummings School of Veterinary Medicine at Tufts University, Department of Infectious Disease and Global Health, North Grafton, MA 01536, United States

ARTICLE INFO

Article history:
Received 14 October 2014
Received in revised form 22 December 2014
Accepted 6 February 2015
Available online 14 February 2015

Keywords:
Cryptosporidium parvum
Multilocus genotypes
Population structure
Principal coordinate analysis
Linkage disequilibrium

ABSTRACT

Cryptosporidium parvum is a common enteric protozoan pathogen of humans and livestock. Multilocus genotyping based on simple sequence repeat polymorphisms has been used extensively to identify transmission cycles and to investigate the structure of *C. parvum* populations and of the related pathogen *Cryptosporidium hominis*. Using such methods, the zoonotic transmission of *C. parvum* has been shown to be epidemiologically important. Because different genetic markers have been used in different surveys, the comparison of *Cryptosporidium* genotypes across different laboratories is often not feasible. Therefore, few comparisons of *Cryptosporidium* populations across wide geographical areas have been published and our understanding of the epidemiology of cryptosporidiosis is fragmented. Here we report on the genotypic analysis of a large collection of 692 *C. parvum* isolates originating primarily from cattle and other ruminants from Italy, Ireland and Scotland. Because the same genotypic markers were used in these surveys, it was possible to merge the data. We found significant geographical segregation and a correlation between genetic and geographic distance, consistent with a model of isolation by distance. The presence of strong LD and positive I_A^S values in the combined MLG dataset suggest departure from panmixia, with different population structures of the parasite prevailing in each country.

1. Introduction

Species within the genus *Cryptosporidium* are obligate intracellular parasites that multiply in the epithelial cells of the gastrointestinal tract of vertebrates. These pathogens have a complex life cycle that comprises two generations of asexual multiplication followed by a sexual cycle (Tzipori, 1988). In humans, *Cryptosporidium hominis* and *Cryptosporidium parvum* account for the vast majority of cases worldwide (Cacciò and Putignani, 2014; Widmer and Sullivan, 2012; Xiao, 2010). *C. hominis* is considered a human parasite, albeit experimental and natural infection of animals with this species have occasionally been reported (Connelly et al., 2013; Giles et al., 2009; Ryan et al., 2005; Smith et al., 2005; Tanriverdi et al., 2003). *C. parvum* is a common parasite of ruminants, particularly of young animals, and has an established zoonotic potential (Learmonth et al., 2004; Sopwith et al., 2005).

Infection is initiated by the ingestion of oocysts. Direct human-to-human or animal-to-human contact, or ingestion of contaminated water and food are common routes of transmission. The oocysts are remarkably resistant to environmental stress and can withstand chlorination of drinking water. These properties, coupled with a low infectious dose (Chappell et al., 2006; Chappell et al., 1996), explain the large number of waterborne outbreaks, including the largest ever recorded outbreak in Milwaukee in the United States (Mac Kenzie et al., 1994).

Understanding how natural parasite populations are structured and how population structures can vary in relation to ecological and epidemiological conditions has attracted a considerable interest, due to the implication such studies can have on the development of control measures. Indeed, as highly polymorphic genetic markers (mini- and micro-satellites) were identified (Cacciò et al., 2000; Cacciò et al., 2001; Feng et al., 2000; Mallon et al., 2003a), population genetics studies were initiated in different countries to evaluate genetic diversity among parasite isolates, to estimate the occurrence of mixed infections (Tanriverdi et al., 2003; Widmer et al., 2014) and to identify host-associated subpopulations (Drumo et al., 2012) and other population structures (Feng et al., 2014). These studies have shown the existence of human-adapted *C. parvum* multi-locus genotypes (MLGs) (Leav et al., 2002; Mallon et al., 2003b), or have demonstrated the influence of different husbandry practices on the structure of bovine

* Corresponding author. Tel.: +39 06 4990 3016; fax: +39 06 4990 3561.
E-mail address: simone.caccio@iss.it (S.M. Cacciò).

http://dx.doi.org/10.1016/j.meegid.2015.02.008
1567-1348/

C. parvum populations (Tanriverdi et al., 2006). Most studies have focused on isolates collected from single countries (Gatei et al., 2007; Hunter et al., 2007) or individual host populations (Cama et al., 2008). Progress on an informative comparison across studies has been hampered by the lack of a standardized genotyping scheme. Consequently, meta-analyses encompassing multiple surveys have been rare (Wang et al., 2014).

To investigate the structure of *C. parvum* on a large geographical scale, we combined MLGs from three published surveys of *C. parvum* isolates from Scotland, Italy and Ireland (Morrison et al., 2008; Drumo et al., 2012; De Waele et al., 2013). This approach enabled us to assemble a large and geographically diverse MLG dataset comprising 692 *C. parvum* isolates. By applying different statistical tests, we found evidence of geographical segregation and linkage disequilibrium, and observed a significant correlation between genetic and geographical distance.

2. Materials and methods

2.1. Source of data

MLG data were taken from three published studies (Morrison et al., 2008; Drumo et al., 2012; De Waele et al., 2013). The data from Scotland (Morrison et al., 2008) included isolates from cattle ($n = 212$), lambs ($n = 9$) and humans ($n = 64$). The data from Italy comprised isolates from cattle ($n = 122$), lambs ($n = 21$), goat kids ($n = 21$), and humans ($n = 9$). The data from Ireland comprised 234 isolates all from cattle (De Waele et al., 2013). The seven genetic markers included were MS1, GP15, MS9, TP14, MM5, MM18, and MM19. The MS1 marker contains a GGTGGTATGCCA repeat in the heat shock protein 70 gene (cgd2_20) located at positions 3136–5184 on chromosome 2. The GP15 marker contains a TCA repeated motif in a 975-bp gene (cgd6_1080) encoding a sporozoite surface protein located at positions 266,434–267,408 on chromosome 6. The MS9 marker contains a TGGACT repeat in a 2016-bp gene (cgd5_2850) encoding a hypothetical protein located at positions 640,137–642,152 on chromosome 5. The TP14 marker contains a CAA repeat in an 8421-bp gene (cgd8_1340) encoding a hypothetical protein located at positions 365,790–374,210 on chromosome 8. The MM5 marker contains a TCCTCCTCT repeat located in an 11,418-bp gene (cgd6_4290) located at positions 1002,285–1013,702 on chromosome 6. The MM18 marker contains a GGACCA repeat in the 5004-bp gene (cgd8_660) located at positions 165,295–170,298 on chromosome 8. The MM19 marker contains a GGAGCT repeat in the 7230-bp gene (cgd8_4840) located at position 1208,520–1215,749 on chromosome 8.

Since Morrison et al. (2008) used a different reverse PCR primer for the GP15 locus than that used by Drumo et al. (2012) and De Waele et al. (2013), GP15 alleles from Scotland were recoded by subtracting 66 nucleotides, based on the fact that the 5′ end of the two GP15 reverse primers are 66 nucleotides apart (see Supplementary Table S1 for the size of recoded GP15 alleles).

2.2. MS9 genotype of bovine isolates from Ireland

MS9 alleles were determined and added to the original dataset from Ireland. This marker was PCR amplified for this study from 234 of 245 bovine isolates from Ireland as described (Mallon et al., 2003a). The size of each PCR product was estimated by electrophoresis using a capillary apparatus (QiaXcel; Qiagen, Milan, Italy) by comparison to size standards. Samples representative of each allele were sequenced on both strands to confirm the estimated size. Each distinct allele was assigned a unique number indicating the estimated size in nucleotides.

2.3. MLG data analyses

The program LIAN 3.6 (Haubold and Hudson, 2000) (available at http://adenine.biz.fh-weihenstephan.de/cgi-bin/lian/lian.cgi.pl) was used to calculate the standardized index of association (I_A^S). This index, a derivation of the Maynard-Smith index of association (Maynard-Smith et al., 1993), measures the strength of the linkage disequilibrium (LD) and is independent of the number of loci analyzed. The null hypothesis of linkage equilibrium was tested by Monte Carlo simulation (10.000 iterations).

The MLG for each isolate was defined by combining the alleles from the seven genetic loci (Drumo et al., 2012). The eBURST software (http://eburst.mlst.net/default.asp) was used to visualize the structure of the combined *C. parvum* population from the three countries. Using this method, the clonal nature of related genotypes and putative "founder" genotypes were visualized (Feil et al., 2004). The most stringent setting was used, and only single-locus variants (SLVs) which differ at one locus only were assigned to the same cluster.

Principal Coordinate Analysis (PCoA) was applied to graphically display pairwise genetic distances between MLGs. A matrix of pairwise distances between MLGs was computed using the SSR distance metric. The SSR distance between two MLGs was calculated as the squared difference in estimated amplicon length summed over seven loci. This distance matrix was input into the PCoA calculator in GenAlex (Peakall and Smouse, 2012).

A Mantel test (Mantel, 1967) was used to assess whether genetic and geographical distances among isolates were correlated. Geographical distance was estimated using the latitude and longitude of the capital of the province or county from which the isolates originated. MLGs from 619 isolates obtained from all ruminants or isolated from calves only ($n = 568$) were included. The computations were performed with GenAlex.

3. Results

3.1. MLG analysis

To create a MLG dataset based on the same seven markers, the alleles at the MS9 locus were determined for 234 of the 245 originally studied Irish *C. parvum* isolates. The analysis of MS9 amplicons revealed that one allele of 450 bp in length predominated (231 of 234 isolates), whereas 3 other alleles (432 bp, 444 bp and 456 bp) were each found in a single isolate. Combining the MS9 locus data with six additional markers resulted in 75 different MLGs in Ireland, to which 102 unique MLGs from Italy and 89 unique MLGs from Scotland were added for the analyses. The complete list of alleles identified at each locus in each country is provided (Supplementary Table S1).

Most MLGs were found in one country only. Only 7 of 266 unique MLGs were found in *C. parvum* isolated from Italy and Ireland and were designated with code "EU" in Fig. 1. Specifically, MLG IT59 (found in 2 Italian isolates) was identical to IE02 (found in 10 Irish isolates, EU1), MLG IT60 (found in 1 Italian isolate) was identical to IE06 (found in 3 Irish isolates, EU2), MLG IT64 (found in 3 Italian isolates) was identical to IE07 (found in 6 Irish isolates, EU3), MLG IT72 (found in 5 Italian isolates) was identical to IE10 (found in 1 Irish isolate, EU4), MLG IT67 (found in 9 Italian isolates) was identical to IE11 (found in 7 Irish isolates, EU5), MLG IT88 (found in 1 Italian isolate) was identical to IE22 (found in 1 Irish isolate, EU6), and MLG IT97 (found in 1 Italian isolate) was identical to IE46 (found in 25 Irish isolates, EU7). Of note, the shared MLGs were all of bovine origin (the Italian panel also comprises 21 goats and 21 sheep samples). No MLGs were shared between Italy and Scotland, or between Ireland and Scotland (Fig. 1).

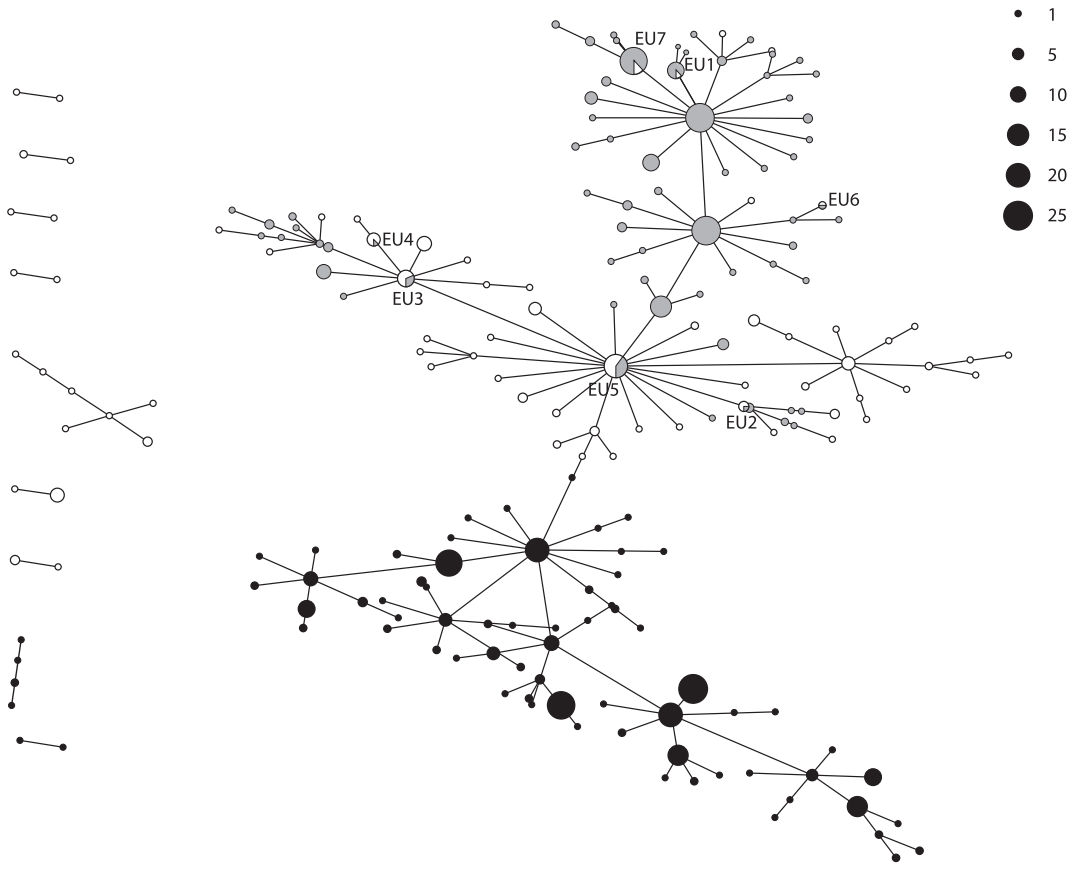

Fig. 1. eBURST analysis of *C. parvum* MLGs from three countries. Single-locus variants are connected. Black, Scotland; grey, Ireland; white Italy. Size of symbols is proportional to MLG abundance. Singletons ($n = 46$) are not shown.

3.2. Analysis of linkage disequilibrium

The level of LD was estimated using the Standardized Index of Association I_A^S, which is a measure of the association between alleles at all pairwise combinations of loci. This index takes a value of zero or becomes negative in randomly mating populations, but is positive for non-panmictic populations. LD analysis comprising all *C. parvum* MLGs from the 3 countries was consistent with strong LD, and this was observed also when only the bovine MLGs were considered. I_A^S values greater than zero indicative of LD were consistently obtained; the p value for Ireland was not statistically significant, albeit V_D was slightly larger than L, thus the null hypothesis of linkage equilibrium was not demonstrated (Table 1). Interestingly, the I_A^S value from all MLGs from 3 countries was not higher than that measured for the Italian population alone (Table 1).

Table 1
Analysis of linkage disequilibrium in *C. parvum* populations from three countries.

Population	N	I_A^S	p value	$V_D > L$	LD
All countries, all MLG	692	0.1295	<0.0001	Yes	Yes
All countries, cattle MLG	568	0.1179	<0.0001	Yes	Yes
All countries, no multiple MLG[a]	263	0.0668	<0.0001	Yes	Yes
Scotland, all MLG[b]	285	0.0677	<0.0001	Yes	Yes
Scotland, cattle MLG[b]	212	0.0346	<0.0001	Yes	Yes
Scotland, no multiple MLG[a]	89	0.0750	<0.0001	Yes	Yes
Ireland, all MLG[c]	234	0.0125	<0.0166*	Yes	Yes
Ireland, no multiple MLG[a]	75	−0.0152	9.6×10^{-01}*	No	No
Italy, all MLG	173	0.1317	<0.0001	Yes	Yes
Italy, cattle MLG	122	0.1187	<0.0001	Yes	Yes
Italy, no multiple MLG[a]	102	0.0991	<0.001	Yes	Yes

[a] Identical MLGs treated as a single data point.
[b] In addition to the seven markers included in our data, the original study (Morrison et al., 2008) included loci ML1, MS5 and MS12.
[c] The original study (De Waele et al., 2013) included locus MS5 but not locus MS9.
* Statistically not significant.

3.3. MLG networks

We used eBURST to generate a network connecting single-locus variants and visualize the relationship between *C. parvum* populations isolated from different countries. As illustrated in Fig. 1, separation of MLGs by country was evident. All MLGs from Scotland were separated from either Italian or Irish MLGs, being connected to the Italian-Irish cluster by only one MLG. Only 5 of 75 MLGs from Ireland did not have a SLV in the main cluster, whereas 15 of 89 MLGs from Scotland, and 26 of 102 MLGs from Italy were identified as singletons. To ensure that the observed MLG clusters were not affected by the recoding of GP15 alleles from Scotland, the eBURST analysis was repeated without GP15 data, i.e., based on six markers instead of seven. This analysis generated a similar eBURST topology, where MLGs from different countries formed separate clusters (not shown). The same observation was made after exclusion of the TP14 marker, instead of GP15, further supporting the country specific MLG clusters.

3.4. Principal coordinate analysis

PCoA of 568 livestock isolates showed two clusters, one enriched for isolates from Ireland, the other enriched for Italian isolates (Fig. 2). A similar topology was already noted in a previous analysis of the Italian MLGs (Drumo et al., 2012). In that survey, clusters

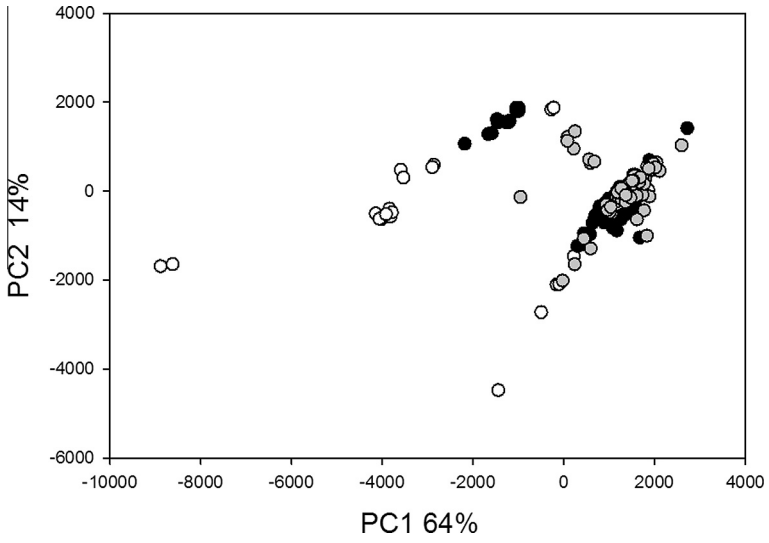

Fig. 2. Principal Coordinate Analysis of *C. parvum* isolates of bovine origin. The analysis is based on pairwise SSR distances as defined in Material and Methods. A total of 568 isolates are included. To improve the resolution of the plot, seven divergent isolates (1 from Ireland, 5 from Italy and 1 from Scotland) were excluded.

were found to be associated with host species and with year of collection. Visual inspection of Fig. 2 reveals a clear geographical partition among clusters, which is consistent with the eBURST analysis. Particularly apparent in the PCoA was the abundance of Irish isolates in the lower cluster and a paucity of isolates from this country in the upper cluster. There was nothing particular about the geographical origin of the 9 Irish isolates located between the cluster as they originated from 6 of the 19 counties represented in this survey. No other metadata than host species and geographical origin were available that could explain the non-random distribution of the isolates from Ireland and the presence of two clusters.

3.5. Analysis of geographical segregation

To further investigate the geographical segregation of MLGs, 619 MLGs from livestock were analyzed using a Mantel test. Consistent with the PCoA, the correlation between genetic and geographic distance exceeded the correlation obtained by random matrix permutation ($p < 0.01$). This result applies to the entire sample of 619 livestock isolates and to the bovine-only subsample of 568 isolates. Together, eBURST, PCoA and Mantel analyses showed that genetic distance among *C. parvum* MLGs was correlated with geographical distance. Further analysis of the relationship between geography and genotype for each marker did not show evidence consistent with a simple isolation-by-distance model caused by stepwise mutation of simple sequence repeats (Valdes et al., 1993). This model would predict that microsatellites alleles at northern- and southern-most locations would frequently be at the extreme of the observed range of allele length. Under this model, samples from Orkney (northernmost) and Italy (southernmost) would be expected to harbor extremely long or extremely short alleles, on average. In fact, the data showed no evidence of such a distribution, as for Orkney and Italy three markers had an average length that is extremely short or long, whereas intermediate locations (Thurso, Dumfries, Aberdeen, Ireland) had 1–4 loci where the average allele length was extreme.

4. Discussion

The present study demonstrates the power of meta-analysis of *C. parvum* MLGs to uncover population substructuring over large geographical areas. As several *Cryptosporidium* genotyping methods have been developed over the years (Robinson and Chalmers, 2012), data sets from different laboratories are typically not compatible and cannot be merged. Three MLG surveys conducted in Italy (Drumo et al., 2012), Scotland (Morrison et al., 2008) and Ireland (De Waele et al., 2013) used almost identical MLG markers, enabling the merging of MLG data without the need for extensive re-typing. The analysis of the merged dataset focused on elucidating the *C. parvum* population structure in space, revealing a geographical distribution limited to individual countries for most MLGs, resulting in a significant association between geographical and genetic distance. This observation extends and confirms a previous MLG analysis of *C. parvum* and *C. hominis* in various countries (Tanriverdi et al., 2008).

As pointed out in reference to bacterial populations by Maynard-Smith et al. (1993), the structure of microbial populations ranges from panmictic (unrestricted gene flow and linkage equilibrium), to clonal (largely restricted gene flow and in LD), to epidemic (underlying panmictic structure masked by an abundance of genetically identical clones). The main criterion distinguishing between these population structures is the presence of LD across loci, which arises as a consequence of reproductive isolation, limiting or excluding recombination among genetically dissimilar individuals.

Previous surveys of *Cryptosporidium* MLGs have provided evidence that all three population structures can be found. Panmixia has been found to characterize populations of bovine *C. parvum* in three Midwestern US states (Herges et al., 2012), in Ireland (De Waele et al., 2013) and in two regions of Scotland (Morrison et al., 2008). However, the Scottish study also found an epidemic population structure in two other regions, indicating that the presence of an underlying panmictic population was masked by the abundance of a few MLGs, and that the LD measured in these populations was no longer detectable after the exclusion of over-represented MLGs (Morrison et al., 2008). Finally, indication for a clonal population was obtained when analyzing MLGs from ruminants in Italy (Drumo et al., 2012). Here, we tested the effect of multiple hosts and of over-represented MLGs on I_A^S values calculated from the combined dataset and from individual countries (Table 1). In agreement with the conclusions reached in the original publications (Morrison et al., 2008; Drumo et al., 2012; de Waele et al., 2013), the *C. parvum* population from Ireland was panmictic, whereas statistically significant LD characterized the Italian and Scottish *C. parvum* populations at the different levels of analysis.

As observed with the malaria parasite *Plasmodium falciparum* (Anderson et al., 2000), it has been suggested that panmixia prevails in areas where transmission is high, as frequent transmission increases the chance of mixed infections and recombination. Likewise, animal trade is also expected to facilitate co-infection and recombination among genetically distinct genotypes found in different geographical regions (Tanriverdi et al., 2008). Indeed, cattle movement has been found to be significantly correlated with allele diversity in bovine populations from Scotland and Ireland (Morrison et al., 2008; De Waele et al., 2013). As movement involves a significant number of young animals (up to 60% of Irish cattle were likely to be moved between herds over a 4 year period; De Waele et al., 2013), different parasite genotypes can co-infect and recombine, leading to linkage equilibrium and lack of geographical segregation. The present observation that *C. parvum* MLGs are confined to individual countries is consistent with the wider geographical scale of the survey and indicates that trade of live cattle between countries did not influence significantly the population structure of *C. parvum*. This assumption could not be verified as we were unable to find data on cattle trade among European countries.

There are some limitations of this study that should be mentioned. First, the isolates have been collected in different time periods (Ireland: 2003–5; Italy; 1997–2010; and Scotland: prior to 2003) and this may have biased the analysis, if one assume that novel genotypes are introduced over time (Drumo et al., 2012). However, longitudinal studies addressing the stability of MLGs over time and space are difficult to conduct. A second potential limitation is that isolates were typed in different laboratories. The main concern arises with trinucleotide repeat markers (GP15 and TP14), as distinguishing alleles differing by 3 base pairs can be problematic. To ensure that the population structure we detected was not due to typing artefacts, we repeated the analyses without these two loci. Clustering of MLGs was substantially unaffected, indicating that MLG generated in different laboratories are comparable.

In conclusion, merging MLG datasets from three countries revealed significant geographical segregation of *Cryptosporidium* among countries. The correlation between genetic distance and geographic distance is in agreement with a model of isolation by distance. The presence of strong LD and positive I_A^S values in the combined MLG dataset from the three countries suggest departure from panmixia and indicate geographical segregation of MLGs, as also indicated by the eBURST network.

Acknowledgments

We thank Liam Morrison and Andy Tait for sharing MLG data from Scotland.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.meegid.2015.02.008.

References

Anderson, T.J. et al., 2000. Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*. Mol. Biol. Evol. 17, 1467–1482.
Cacciò, S. et al., 2000. A microsatellite marker reveals population heterogeneity within human and animal genotypes of *Cryptosporidium parvum*. Parasitology 120, 237–244.
Cacciò, S., Spano, F., Pozio, E., 2001. Large sequence variation at two microsatellite loci among zoonotic (genotype C) isolates of *Cryptosporidium parvum*. Int. J. Parasitol. 31, 1082–1086.
Cacciò, S.M., Putignani, L., 2014. Epidemiology of human cryptosporidiosis. In: Cacciò, S.M., Widmer, G. (Eds.), Cryptosporidium: Parasite and Disease. Springer, pp. 43–80.
Cama, V.A. et al., 2008. *Cryptosporidium* species and subtypes and clinical manifestations in children, Peru. Emerg. Infect. Dis. 14, 1567–1574.
Chappell, C.L. et al., 1996. *Cryptosporidium parvum*: intensity of infection and oocyst excretion patterns in healthy volunteers. J. Infect. Dis. 173, 232–236.
Chappell, C.L. et al., 2006. *Cryptosporidium hominis*: experimental challenge of healthy adults. Am. J. Trop. Med. Hyg. 75, 851–857.
Connelly, L. et al., 2013. Genetic diversity of *Cryptosporidium* spp. within a remote population of Soay sheep on St. Kilda Islands, Scotland. Appl. Environ. Microbiol. 79, 2240–2246.
De Waele, V. et al., 2013. Panmictic structure of the *Cryptosporidium parvum* population in Irish calves: influence of prevalence and host movement. Appl. Environ. Microbiol. 79, 2534–2541.
Drumo, R. et al., 2012. Evidence of host-associated populations of *Cryptosporidium parvum* in Italy. Appl. Environ. Microbiol. 78, 3523–3529.
Feil, E.J. et al., 2004. EBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. J. Bacteriol. 186, 1518–1530.
Feng, X. et al., 2000. Extensive polymorphism in *Cryptosporidium parvum* identified by multilocus microsatellite analysis. Appl. Environ. Microbiol. 66, 3344–3349.
Feng, Y. et al., 2014. Multilocus sequence typing of an emerging *Cryptosporidium hominis* subtype in the United States. J. Clin. Microbiol. 52, 524–530.
Gatei, W. et al., 2007. Multilocus sequence typing and genetic structure of *Cryptosporidium hominis* from children in Kolkata, India. Infect. Genet. Evol. 7, 197–205.
Giles, M. et al., 2009. *Cryptosporidium hominis* in a goat and a sheep in the UK. Vet. Rec. 164, 24–25.
Haubold, B., Hudson, R.R., 2000. LIAN 3.0: detecting linkage disequilibrium in multilocus data. Bioinformatics 16, 847–848.
Herges, G.R. et al., 2012. Evidence that *Cryptosporidium parvum* populations are panmictic and unstructured in the Upper Midwest of the United States. Appl. Environ. Microbiol. 78, 8096–8101.
Hunter, P.R. et al., 2007. Subtypes of *Cryptosporidium parvum* in humans and disease risk. Emerg. Infect. Dis. 13, 82–88.
Learmonth, J.J. et al., 2004. Genetic characterization and transmission cycles of *Cryptosporidium* species isolated from humans in New Zealand. Appl. Environ. Microbiol. 70, 3973–3978.
Leav, B.A. et al., 2002. Analysis of sequence diversity at the highly polymorphic Cpgp40/15 locus among *Cryptosporidium* isolates from human immunodeficiency virus-infected children in South Africa. Infect. Immun. 70, 3881–3890.
Mac Kenzie, W.R. et al., 1994. A massive outbreak in Milwaukee of *Cryptosporidium* infection transmitted through the public water supply. N. Engl. J. Med. 331, 161–167.
Mallon, M.E. et al., 2003a. Population structures and the role of genetic exchange in the zoonotic pathogen *Cryptosporidium parvum*. J. Mol. Evol. 56, 407–417.
Mallon, M.E. et al., 2003b. Multilocus genotyping of *Cryptosporidium parvum* type 2: population genetics and sub-structuring. Infect. Genet. Evol. 3, 207–218.
Mantel, N., 1967. The detection of disease clustering and a generalized regression approach. Cancer Res. 27, 209–220.
Maynard-Smith, J.M. et al., 1993. How clonal are bacteria? Proc. Natl. Acad. Sci. U.S.A. 90, 4384–4388.
Morrison, L.J. et al., 2008. The population structure of the *Cryptosporidium parvum* population in Scotland: a complex picture. Infect. Genet. Evol. 8, 121–129.
Peakall, R., Smouse, P.E., 2012. GenAlex 6.5: genetic analysis in Excel. Population genetic software for teaching and research–an update. Bioinformatics 28, 2537–2539.
Robinson, G., Chalmers, R.M., 2012. Assessment of polymorphic genetic markers for multi-locus typing of *Cryptosporidium parvum* and *Cryptosporidium hominis*. Exp. Parasitol. 13, 200–215.
Ryan, U.M. et al., 2005. Sheep may not be an important zoonotic reservoir for *Cryptosporidium* and *Giardia* parasites. Appl. Environ. Microbiol. 71, 4992–4997.
Smith, H.V. et al., 2005. Natural *Cryptosporidium hominis* infections in Scottish cattle. Vet. Rec. 156, 710–711.
Sopwith, W. et al., 2005. The changing epidemiology of cryptosporidiosis in North West England. Epidemiol. Infect. 133, 785–793.
Tanriverdi, S. et al., 2003. Identification of genotypically mixed *Cryptosporidium parvum* populations in humans and calves. Mol. Biochem. Parasitol. 130, 13–22.
Tanriverdi, S. et al., 2006. Emergence of distinct genotypes of *Cryptosporidium parvum* in structured host populations. Appl. Environ. Microbiol. 72, 2507–2513.
Tanriverdi, S. et al., 2008. Inferences about the global population structures of *Cryptosporidium parvum* and *Cryptosporidium hominis*. Appl. Environ. Microbiol. 74, 7227–7234.
Tzipori, S., 1988. Cryptosporidiosis in perspective. Adv. Parasitol. 27, 63–129.
Valdes, A.M., Slatkin, M., Freimer, N.B., 1993. Allele frequencies at microsatellite loci: the stepwise mutation model revisited. Genetics 133, 737–749.
Wang, R.L. et al., 2014. *Cryptosporidium parvum* IId family: clonal population and dispersal from Western Asia to other geographical regions. Sci. Rep. 4, 4208.
Widmer, G. et al., 2014. Population structure of natural and propagated isolates of *Cryptosporidium parvum*, *C. hominis* and *C. meleagridis*. Environ. Microbiol. http://dx.doi.org/10.1111/1462-2920.12447.
Widmer, G., Sullivan, S., 2012. Genomics and population biology of *Cryptosporidium* species. Parasite Immunol. 34, 61–71.
Xiao, L., 2010. Molecular epidemiology of cryptosporidiosis: an update. Exp. Parasitol. 124, 80–89.