Experimental Parasitology 184 (2018) 121–127

Contents lists available at ScienceDirect

Experimental Parasitology

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

# Population genetic characterization of *Cyclospora cayetanensis* from discrete geographical regions

Yaqiong Guo [a, b], Na Li [a], Ynes R. Ortega [c], Longxian Zhang [d], Dawn M. Roellig [b], Yaoyu Feng [a, *], Lihua Xiao [b]

[a] College of Veterinary Medicine, South China Agricultural University, Guangzhou, Guangdong, 510642, China
[b] Division of Foodborne, Waterborne and Environmental Diseases, National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, GA, 30329, USA
[c] Department of Food Science and Technology, University of Georgia, Griffin, GA, 30602, USA
[d] College of Animal Science and Veterinary Medicine, Henan Agricultural University, Zhengzhou, Henan, 450002, China

## HIGHLIGHTS

- Population genetic analysis of 64 *Cyclospora cayetanensis* specimens was conducted.
- The overall *C. cayetanensis* population has an epidemic structure.
- Clear geographical segregation was observed in *C. cayetanensis*.
- Frequent genetic exchange exist in C. cayetanensis in endemic area.

## ARTICLE INFO

*Article history:*
Received 24 May 2017
Accepted 18 December 2017
Available online 20 December 2017

*Keywords:*
*Cyclospora cayetanensis*
Multilocus sequence typing
Genetic diversity
Population genetics
Geographical segregation

## ABSTRACT

*Cyclospora cayetanensis* is an emerging pathogen that is endemic in developing countries and responsible for many large foodborne cyclosporiasis outbreaks in North America since 1990s. Because of the lack of typing targets, the genetic diversity and population genetics of *C. cayetanensis* have not been investigated. In this study, we undertook a population genetic analysis of multilocus sequence typing data we recently collected from 64 *C. cayetanensis* specimens. Despite the extensive genetic heterogeneity in the overall *C. cayetanensis* population, there were significant intra- and inter-genic linkage disequilibria (LD). A disappearance of LD was observed when only multilocus genotypes were included in the population genetic analysis, indicative of an epidemic nature of *C. cayetanensis*. Geographical segregation-associated sub-structuring was observed between specimens from China and those from Peru and the United States. The two subpopulations had reduced LD, indicating the likely occurrence of genetic exchange among isolates in endemic areas. Further analyses of specimens from other geographical regions are necessary to fully understand the population genetics of *C. cayetanensis*.



## 1. Introduction

*Cyclospora* spp. are important apicomplexan parasites in humans and various animals. Of the 20 known species, *Cyclospora cayetanensis* is the only one known to infect humans, causing acute diarrhea, fatigue, abdominal cramping, low-grade fever, and weight loss (Ortega and Sanchez, 2010; Li et al., 2015). Cyclosporiasis is endemic in some developing countries, and cases in developed countries are traditionally associated with travel to endemic areas (Sakakibara et al., 2010; Bednarska et al., 2015). However, since the mid-1990s, foodborne outbreaks have been reported in North America, primarily associated with imported fresh produce from cyclosporiasis-endemic areas such as Guatemala and Mexico (Hall et al., 2012; Milord et al., 2012; Buss et al., 2016). During the last four years, large multi-state outbreaks had occurred yearly in the United States (Centers for Disease Control and Prevention, 2016).

Although *C. cayetanensis* has become an important emerging pathogen, little is known of its genetic diversity and population structure. Current molecular characterizations of *C. cayetanensis* concentrate on ribosomal and heat shock protein genes and have

* Corresponding author.
*E-mail address:* yyfeng@scau.edu.cn (Y. Feng).

https://doi.org/10.1016/j.exppara.2017.12.006
0014-4894/

shown largely low genetic heterogeneity among isolates (Zhou et al., 2011; Sulaiman et al., 2013; Hussein et al., 2016). The low resolution of the molecular tools employed may result in an underestimate of genetic diversity. In the population genetic studies of other apicomplexan protozoa such as *Cryptosporidium parvum* and *Plasmodium falciparum*, high genetic heterogeneity and striking variations in population structure have been observed within specific pathogens, with population structure ranging from clonal, epidemic to panmictic and occurrence of host or geographic segregation (Annan et al., 2007; Herges et al., 2012; Feng et al., 2013).

Multiple factors contribute to the population genetic characteristics of pathogens, including host-specificity, transmission intensity, climate, and geographical barriers (Morrison et al., 2008; Feng et al., 2013). Although *C. cayetanensis* infects humans exclusively, it is widely distributed in developing countries and exhibits significant diversity in transmission intensity among geographic areas (Sherchand and Cross, 2001; Devera et al., 2005; Puente et al., 2006; Insulander et al., 2010). Thus, the population genetic structure of *C. cayetanensis* could be different among endemic areas.

To date, no population genetic characterization has been conducted for *C. cayetanensis*. This is mainly due to a lack of highly discriminatory markers. Recently, the sequencing of a *C. cayetanensis* genome has led to the development of a multilocus sequence typing (MLST) tool targeting five microsatellite loci, with which 64 *C. cayetanensis* specimens from several geographical areas were genotyped and noticeable geographic clustering was observed in an neighbor-joining analysis of the sequence data (Guo et al., 2016). In this study, we further analyzed the MLST data generated from the 64 *C. cayetanensis* specimens for the nature of genetic diversity and population structure. Even though only 34 of the 64 specimens have full MLST data from the five loci, we observed an overall epidemic population structure for *C. cayetanensis* and possible existence of population substructuring in the pathogen.

## 2. Materials and methods

### 2.1. Data source

Overall, MLST data from 64 *C. cayetanensis*-positive specimens obtained from a previous study (Guo et al., 2016) were analyzed in this study. The 64 specimens were collected from several geographical regions including China (n = 26), Nepal (n = 3), Indonesia (n = 1), Guatemala (n = 2), Peru (n = 8), Spain (n = 1), and the United States (n = 23). These sequence data were from five genetic loci, including CYC3, CYC13, CYC15, CYC21, and CYC22. Among the 64 specimens, 63, 61, 63, 62, and 64 generated the expected PCR products at CYC3, CYC13, CYC15, CYC21, and CYC22 loci, respectively, but due to the presence of mixed genotypes, 10, 6, 1, 4, and 11 specimens failed to produce readable sequences, respectively. Consequently, 56 specimens had sequence data from at least four loci, with 34 of them having sequence data from all five loci.

### 2.2. Analyses of genetic diversity and population structure

Initially, genetic diversity at each locus was measured by DnaSP (www.ub.es/dnasp/) analysis of sequence data from all 64 specimens. Haplotype numbers and diversity (Hd) were calculated based on the overall sequence polymorphism (including nucleotide substitutions and insertions or deletions), whereas intragenic linkage disequilibrium (LD) and recombination rates were calculated based on segregating sites. Pairwise intergenic LD among five loci was evaluated by using Arlequin analysis (http://cmpg.unibe.ch/software/arlequin35/) of allelic data from the 34 specimens with full MLST data at the five genetic loci. Nucleotide sequences from these 34 specimens were concatenated for analysis of genetic diversity among the multilocus genotypes (MLGs) by using DnaSP.

Population genetic analysis was conducted on the allelic data from the 34 specimens with full MLST data from the five loci. The standardized indexes of association ($\bar{I}^S_A$) of the overall population and subpopulations by sample source were calculated to assess multilocus LD using LIAN 3.7 (http://guanine.evolbio.mpg.de/cgi-bin/lian/lian.cgi.pl/query). In this analysis, negative or zero $\bar{I}^S_A$ values indicate that the population is in linkage equilibrium (LE). The variance of pairwise differences ($V_D$) and the 95% critical value ($L$) for $V_D$ were also calculated to confirm the population structure.

### 2.3. Subpopulation structure analysis

Subpopulation structure analyses were initially conducted on data from the 34 specimens with full MLST data. To determine the population substructure, allelic data from these specimens were analyzed using the Bayesian analysis tool Structure version 2.2 (http://pritch.bsd.uchicago.edu/structure.html) for K values (likely population numbers) of 2–5. In consideration of the small sample size of this analysis, we conducted an additional analysis of data from the 56 specimens with sequences from at least four genetic loci. The subpopulation structure and the ancestor-descendant relationship among different MLGs were further assessed using eBURST (http://eburst.mls.net). To confirm the population substructure, a median-joining algorithm analysis of the multilocus sequence data was performed by using Network version 4.6.1.0 (www.fluxus-ngineering.com/sharenet.htm). The validity of population substructures was assessed by calculation of Wright's index ($F_{ST}$) using Arlequin 3.5.

## 3. Results

### 3.1. Gene diversity at MLST loci

As reported previously, the sequence polymorphism at the five genetic loci included both copy number variations in the simple tandem repeats and single nucleotide substitutions, and CYC13 was the most polymorphic locus with 10 haplotypes identified (Guo et al., 2016). However, analyses conducted in this study indicated that CYC13 had gene diversity (Hd = 0.6855; 10 haplotypes) similar to CYC22 (Hd = 0.7025; 4 haplotypes), CYC21 (Hd = 0.6969; 8 haplotypes), and CYC3 (Hd = 0.5871; 4 haplotypes). In contrast, CYC15 had very low gene diversity (Hd = 0.0635; 2 haplotypes) (Table 1).

Intragenic LD analysis was feasible for only the four loci that had high gene diversity. Complete intragenic LD (|D'|, Y = 1.0000 ± 0.0000X) was observed at CYC3, CYC13, and CYC22, whereas a strong but incomplete intragenic LD (|D'|, Y = 0.9574–0.3904X) was detected at CYC21 (Table 1). The incomplete intragenic LD might have resulted from genetic recombination, as two potential recombination events (Rm) were identified at the locus.

### 3.2. Population structure

Concatenating sequences from the 34 specimens with full data at five genetic loci resulted in a 2317-bp multilocus alignment with 157 polymorphic sites. The 34 specimens yielded 25 MLGs with an Hd of 0.9733 (Table 1). Most of the MLGs only had one specimen, except for MS3, MS15, MS16, and MS17. They had MLG frequencies of 8.8% (3 specimens per MLG for MS3, MS16, and MS17) to 11.8% (4 specimens per MLG for MS15). LD analysis of the concatenated sequences indicated the occurrence of a strong but incomplete LD

**Table 1**
Pairwise intragenic linkage disequilibrium and recombination analysis of five genetic loci in *Cyclospora cayetanensis*.

| Locus | No. of sequence | No. of analyzed site | No. of polymorphic (segregating) sites | No. of haplotype | Haplotype diversity (Hd) | No. of polymorphic sites analyzed | No. of pairwise comparisons | ZnS | \|D'\| values | Estimate of R, per gene | Minimum number of recombination events, Rm |
|---|---|---|---|---|---|---|---|---|---|---|---|
| CYC3 | 53 | 486 | 46 | 4 | 0.5871 | 36 | 630 | 0.945 | Y = 1.0000 + 0.0000X | 0.001 | 0 |
| CYC13 | 54 | 556 | 38 | 10 | 0.6855 | 8 | 28 | 0.4346 | Y = 1.0000−0.0000X | 0.001 | 0 |
| CYC15 | 62 | 518 | 1 | 2 | 0.0635 | 1 | 0 | NA | NA | NA | NA |
| CYC21 | 58 | 383 | 51 | 8 | 0.6969 | 23 | 253 | 0.419 | Y = 0.9574−0.3904X | 0.001 | 2 |
| CYC22 | 53 | 379 | 41 | 4 | 0.7025 | 4 | 6 | 0.431 | Y = 1.0000−0.0000X | 7 | 0 |
| Concatenated sequence | 34 | 2317 | 157 | 25 | 0.9733 | 60 | 1770 | 0.4008 | Y = 0.9877−0.0609X | 0.5 | 2 |

(Y = 0.9877−0.0609X). Recombination event test results supported the occurrence of potential recombination (Rm = 2) (Table 1).

In the analysis of allelic data from the 34 specimens, the value of $I^S_A$ was positive and $V_D$ (1.2244) was more than $L$ (1.1779), indicating that the overall population was in LD (Table 2). Pairwise intergenic LD analysis across the five loci also identified 7/10 cases of significant LD (Fig. 1A). However, when each group of specimens with the same MLG was considered as one individual, the result of the analysis showed LE for the overall population, as the value of $I^S_A$ was negative and $V_D$ (0.9537) was less than $L$ (1.1343) (Table 2). Thus, the overall population had an epidemic structure.

### 3.3. Subpopulation structure

In an initial STRUCTURE analysis, we used K values of 2–5 to capture possible subpopulations in the 34 specimens with full multilocus sequence data, but no clear population subdivision was observed. This could be due to the small sample size in the analysis. Thus, we conducted another STRUCTURE analysis using data from the 56 specimens that had sequence data from at least four loci. Using K value of 2–5, the same population substructure pattern was obtained, which identified a clear separation between the specimens from China and other areas (Fig. 2). A similar geographic subdivision was observed in eBURST and Network analyses of the 34 specimens with full multilocus sequence data. In the eBURST analysis, all but two specimens from China were clearly separated from the specimens from other areas, and one MLG of from China was predicted to be the ancestor of MLGs in other areas (Fig. 3). In the Network analysis, although specimens from Peru, the United States, Spain, and Nepal were separated into two groups, they were distant from the group formed by specimens from China (Fig. 4). In addition, pairwise $F_{ST}$ analysis of the subpopulation formed by the specimens from China and the other one formed by the remaining specimens generated a relatively high $F_{ST}$ value ($F_{ST} = 0.403$, $P < .00001$), which further supported the significant genetic differentiation between the two subpopulations (Table 3).

In subpopulation-specific multilocus LD analysis, the two subpopulations which consisted of specimens from China and specimens from the other regions, were both in LE ($V_D < L$). LE was maintained for each subpopulation even if only MLGs were used in the analysis (Table 2). The lack of LD in these two subpopulations was confirmed by pair-wise intergenic LD analysis, which generated only 1/9 and 2/9 cases of significant LD, respectively (Fig. 1B and C). Thus, these two subpopulations were both panmictic.

## 4. Discussion

Data presented in this study have clearly shown an overall epidemic population structure for *C. cayetanensis*. Using a newly developed MLST tool, at least 25 MLGs were identified in the overall population, with an extremely high MLG diversity (Hd = 0.9733). Despite the extensive genetic heterogeneity in the overall *C. cayetanensis* population, significant intra- and inter-genic LD were observed, which suggests the population has a non-panmictic structure or contains genetically distinct subpopulations (Lymbery, 2002). The disappearance of LD in the analysis with only MLGs is indicative of the epidemic nature of the overall population. This was verified by the MLG frequency analysis, which has shown a common occurrence of four MLGs (MS3, MS15, MS16, and MS17) among the specimens analyzed.

Despite the epidemic nature of the overall population, there was clearly geographical segregation in *C. cayetanensis*. The STRUCTURE,

**Table 2**
Multilocus linkage disequilibrium (LD) analysis of allelic data at five genetic loci in *Cyclospora cayetanensis*.

| Population | No. of analyzed sequence | H (Genetic diversity) | $I^S_A$ | $P_{MC}$ | $V_D$ | $L$ | $V_D > L$ | LE or LD |
|---|---|---|---|---|---|---|---|---|
| All | 34 | 0.5383 ± 0.1079 | 0.0531 | 0.0250 | 1.2244 | 1.1779 | Y | LD |
| China | 16 | 0.4383 ± 0.1115 | −0.0562 | 0.9340 | 0.7613 | 1.2991 | N | LE |
| Peru | 6 | 0.3600 ± 0.1046 | 0.0638 | 0.1190 | 1.1714 | 1.7429 | N | LE |
| US | 9 | 0.3278 ± 0.1180 | −0.0391 | 0.6390 | 0.6944 | 1.3230 | N | LE |
| Nepal | 2 | 0.6000 ± 0.2449 | NA | <0.001 | NA | NA | NA | NA |
| Spain | 1 | NA | NA | NA | NA | NA | NA | NA |
| Peru & US | 15 | 0.3276 ± 0.1065 | 0.0091 | 0.4010 | 0.9062 | 1.1947 | N | LE |
| All countries excluding China | 18 | 0.3804 ± 0.1104 | 0.0273 | 0.2700 | 1.0364 | 1.2469 | N | LE |
| All [a] | 25 | 0.5893 ± 0.1127 | −0.0006 | 0.4920 | 0.9537 | 1.1343 | N | LE |
| China[a] | 14 | 0.4769 ± 0.1191 | −0.0899 | 0.9950 | 0.6171 | 1.2615 | N | LE |
| Peru [a] | 5 | 0.4000 ± 0.1095 | 0.0394 | 0.4960 | 1.1111 | 1.7778 | N | LE |
| US [a] | 6 | 0.4000 ± 0.1265 | −0.0877 | 0.9040 | 0.5714 | 1.7143 | N | LE |
| Nepal [a] | 2 | 0.6000 ± 0.2449 | NA | <0.001 | NA | NA | NA | NA |
| Spain [a] | 1 | NA | NA | NA | NA | NA | NA | NA |
| Peru & US [a] | 8 | 0.4429 ± 0.1182 | −0.0296 | 0.7210 | 0.8413 | 1.5079 | N | LE |
| All countries excluding China [a] | 11 | 0.4800 ± 0.1246 | −0.0169 | 0.6570 | 0.8741 | 1.3185 | N | LE |

[a] Considering each group of specimens with the same multilocus genotype as one individual.

| Locus | CYC3 | CYC13 | CYC15 | CYC21 |
|---|---|---|---|---|
| CYC22 | + | + | - | + |
| CYC21 | + | + | + | |
| CYC15 | - | - | | |
| CYC13 | + | | | |

A

| Locus | CYC3 | CYC13 | CYC15 | CYC21 |
|---|---|---|---|---|
| CYC22 | - | - | - | - |
| CYC21 | - | + | - | |
| CYC15 | - | - | | |
| CYC13 | - | | | |

B

| Locus | CYC3 | CYC13 | CYC15 | CYC21 |
|---|---|---|---|---|
| CYC22 | - | - | - | - |
| CYC21 | + | - | - | |
| CYC15 | - | - | | |
| CYC13 | + | | | |

C

**Fig. 1.** Pairwise intergenic linkage disequilibrium (LD) matrix among 5 polymorphic loci of *Cyclospora cayetanensis*. Significant LD between loci is marked yellow. The overall population has 7 cases of pairwise LD (A), while the Chinese subpopulation (B) and the other subpopulation (C) only have 1 and 2 cases of pairwise LD, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article).

eBURST, and Network analyses of allelic or sequence data all have revealed the existence of two geographically segregated subpopulations. Specimens from China formed one subpopulation genetically distant from the one formed by the remaining specimens (mostly from Peru and the United States). This is in concordance with the geographic clustering pattern in the neighbor-joining analysis in our previous report (Guo et al., 2016). The relative high $F_{ST}$ value (0.403) between Chinese and other subpopulations further supports the genetic differentiation between them (Table 3). As the both subpopulations have a panmictic population structure, there could be frequent genetic exchanges within each subpopulation. Therefore, although geographic segregation has reduced genetic exchanges between subpopulations of *C. cayetanensis*, genetic recombination during sexual reproduction of this parasite could play an important role in generating high genetic diversity within each subpopulation.

Results of this study improve our understanding of the transmission of *C. cayetanensis* in the United States. Since mid-1990s, the United States has experienced numerous foodborne outbreaks of cyclosporiasis. In epidemiological investigations, they have been mostly linked to imported fresh produce from cyclosporiasis-endemic areas from Mexico and Central America (Centers for Disease Control and Prevention, 2016). However, this has never been supported by genetic characterizations of the pathogen, which is largely due to the lack of genotyping tools. Results of the study indicate that in addition to possible case linkages, genotyping could potentially help with geographic tracking of pathogens in outbreaks associated with imported fresh produce or foreign travels. In this study, we have shown significant genetic differences in *C. cayetanensis* between China and Peru and a similar population genetic structure among parasites from outbreak cases in the United States and sporadic cases from Peru and Nepal. The genetic similarity between specimens from Peru and the United States is not surprising, as outbreaks of cyclosporiasis in the United States are mostly caused by consumption of fresh produce imported from Mexico and Central and South America (Ortega and Sanchez, 2010). As cyclosporiasis is endemic in these countries (Bern et al., 1999, 2002) and there are frequent population migrations between these countries (Brea, 2003), there could be significant genetic exchanges in *C. cayetanensis* in this area, as evident by the existence

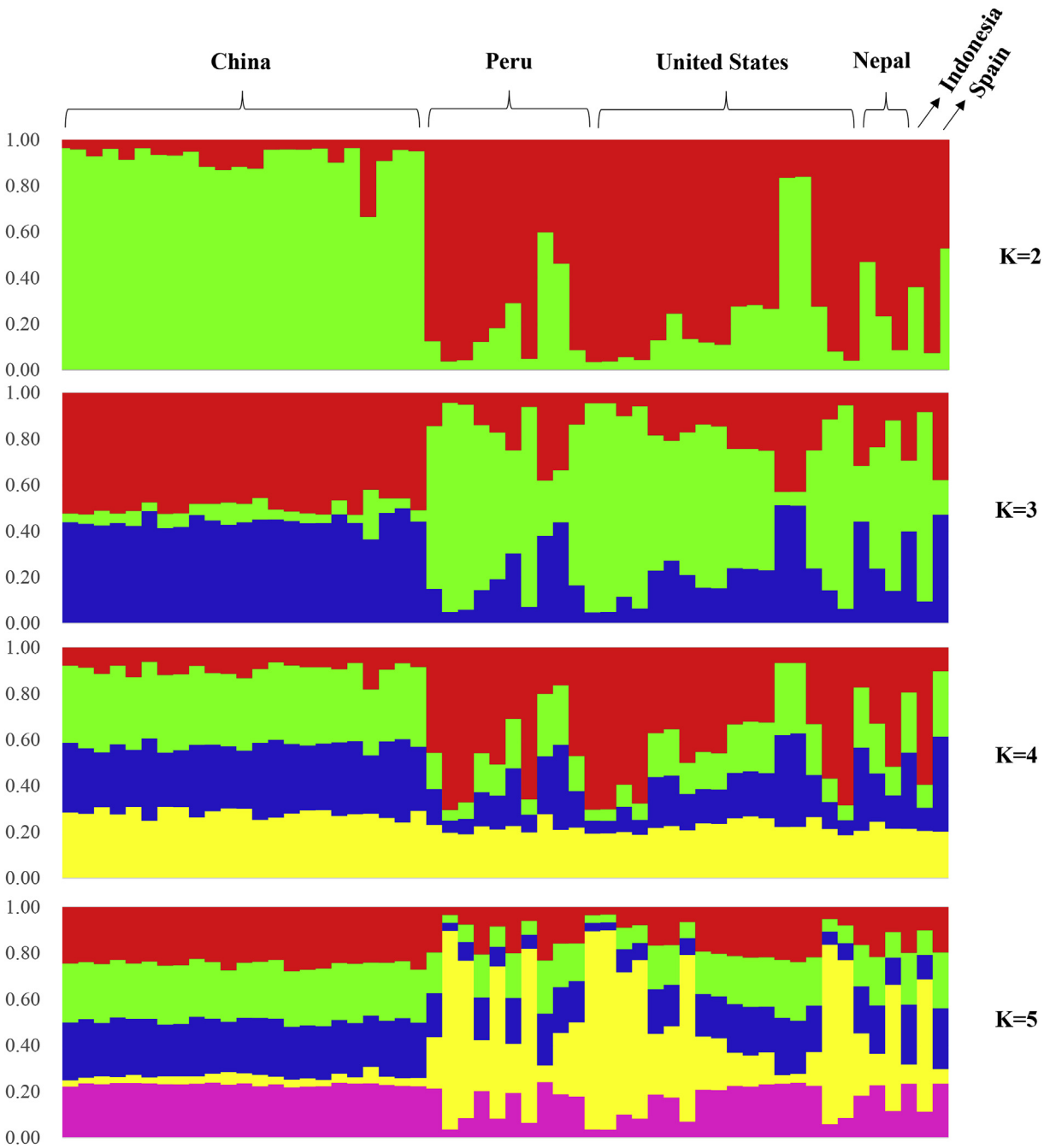


**Fig. 2.** Population substructure of 56 *Cyclospora cayetanensis* specimens with multilocus sequences at four or five loci by Bayesian analysis of allelic data. Sub-population patterns are shown for K = 2, 3, 4 and 5 used in the analysis.

of high genetic and haplotype diversity and panmictic nature of parasites from the United States and Peru. The high prevalence and foodborne nature of *C. cayetanensis* provide more opportunities for genetic exchange among different isolates and genotypes, leading to the occurrence of high genetic diversity and shared genotypes or subpopulations in geographically and commercially linked areas. In fact, the foodborne cyclosporiasis outbreak in the United States in 2013 was caused by multiple MLGs, including several ones identical to those in Peru (Guo et al., 2016). This suggests that the contamination of fresh produce by *C. cayetanensis* was likely not from a single source. A similar situation would be expected for other foodborne cyclosporiasis outbreaks in the United States. Interestingly, the two Nepalese specimens with full MLST data appear to be genetically more related to specimens from Peru and the United States than to those from China. This could be due to the small sample size of the parasite characterized.

In conclusion, we have demonstrated that the population genetics of *C. cayetanensis* can be studied with the data generated using the current MLST method, with patterns corresponding to an epidemic structure in the overall population, and the presence of geographically segregated sub-populations. The panmictic population structure of the two subpopulations identified in this study is indicative of a common occurrence of genetic exchange among isolates within endemic areas. Further analyses of specimens from

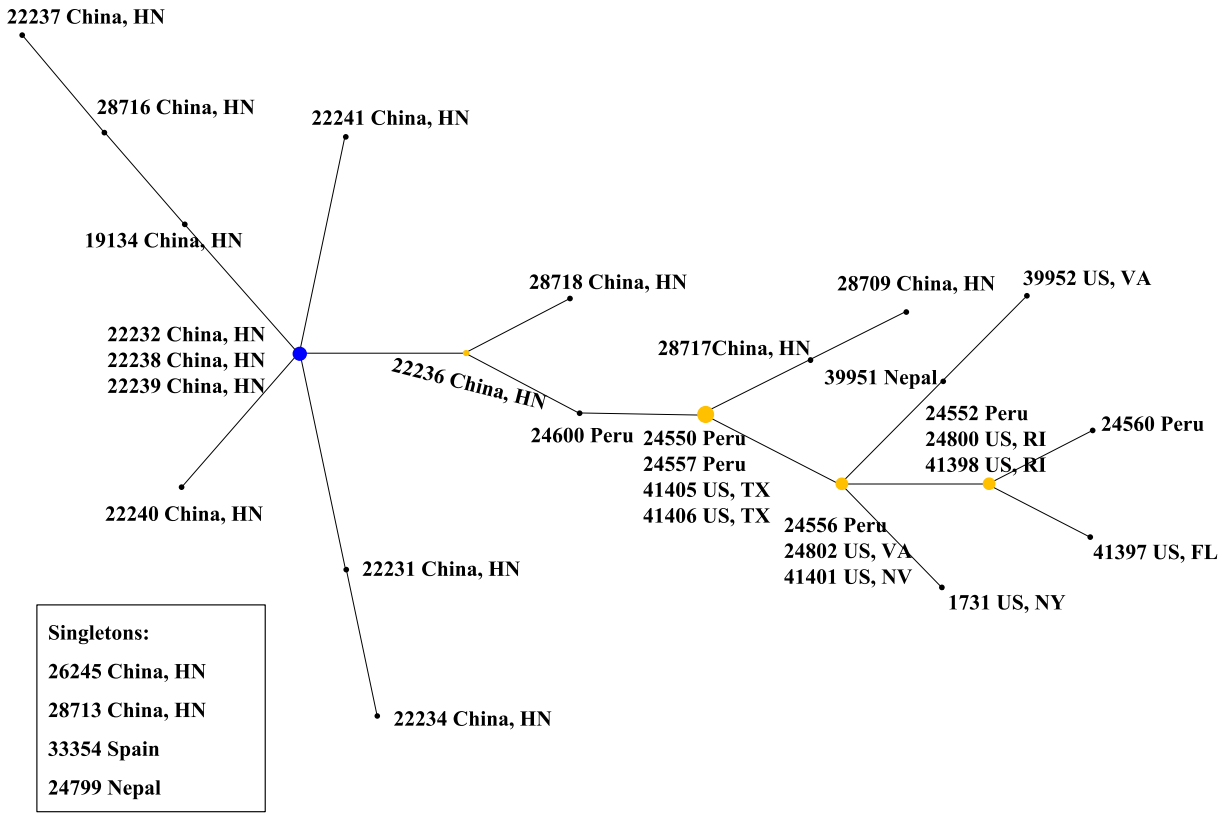


**Fig. 3.** Ancestor-descendant relationship among multilocus genotypes of 34 *Cyclospora cayetanensis* specimens as interpreted by eBURST analysis of allelic data. Blue dot is the predicted primary founder, while yellow dots are subgroup founders. The size of each dot is proportional to the number of specimens with the MLGs. Four specimens (26245, 28713, 33354 and 24799) and singletons are placed outside the diagram. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article).

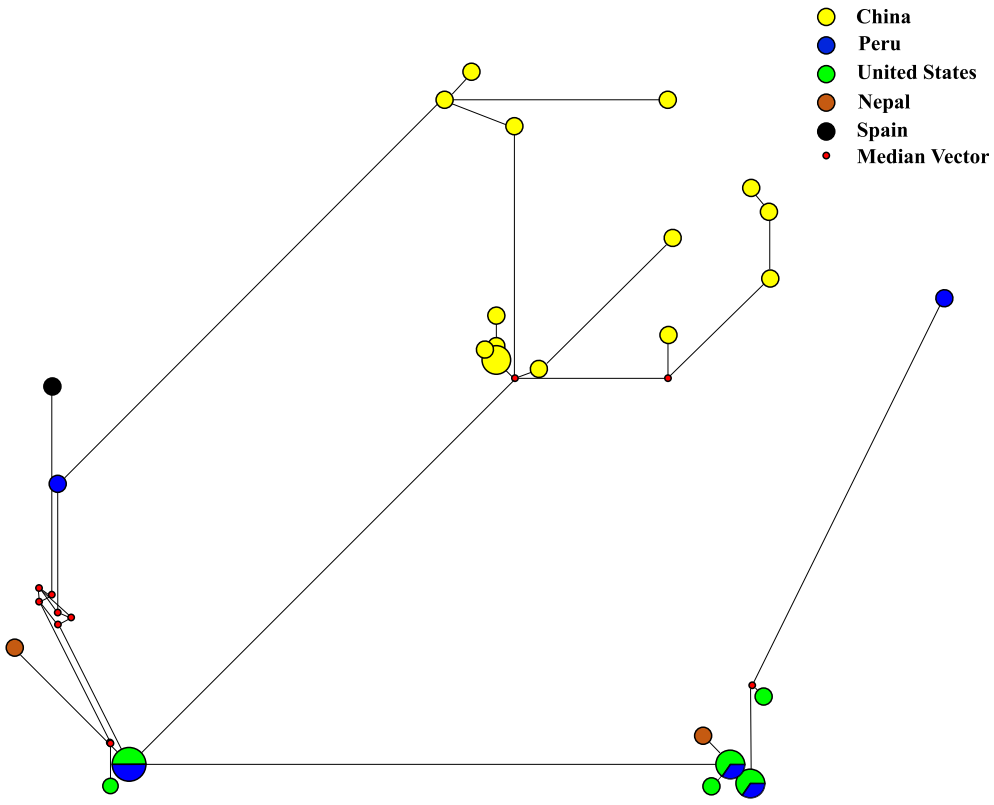


**Fig. 4.** Median-joining network of multilocus genotypes (MLGs) in 34 *Cyclospora cayetanensis* specimens. The size of each circle is proportional to the number of specimens with the sequence type. The colors in the circle represent geographical origins of specimens, except for the red dot (Median Vector). The length of lines connecting MLGs is proportional to the number of single-nucleotide polymorphisms. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article).

**Table 3**
Population differentiation among specimens from different geographic area in pairwise $F_{ST}$ analysis of 34 specimens with full MLST data.

| Subpopulation pairs (sample size) | $F_{ST}$ | P value |
|---|---|---|
| China (16)–Peru (6) | 0.3551 | .00000 ± .0000 |
| China (16)–United States (9) | 0.45267 | .00000 ± .0000 |
| China (16)–United States (9) & Peru (15) | 0.29583 | .00000 ± .0000 |
| China (16)–All others (18) | 0.40331 | .00000 ± .0000 |
| Peru (6)–United States (9) | 0.00285 | .95495 ± .0151 |

other geographical regions are required to fully understand the population genetics of *C. cayetanensis*.

## Acknowledgements

This work was supported by the National Natural Science Foundation of China (grant numbers 31425025, 31630078 and 31602042), the China Postdoctoral Science Foundation (grant number 2016M601531), and the U.S. Centers for Disease Control and Prevention.

The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the US Centers for Disease Control and Prevention.

## References

Annan, Z., Durand, P., Ayala, F.J., Arnathau, C., Awono-Ambene, P., Simard, F., Razakandrainibe, F.G., Koella, J.C., Fontenille, D., Renaud, F., 2007. Population genetic structure of *Plasmodium falciparum* in the two main African vectors, Anopheles gambiae and Anopheles funestus. Proc Natl Acad Sci USA 104, 7987–7992.

Bednarska, M., Bajer, A., Welc-Faleciak, R., Pawelas, A., 2015. *Cyclospora cayetanensis* infection in transplant traveller: a case report of outbreak. Parasites Vectors 8, 411.

Bern, C., Hernandez, B., Lopez, M.B., Arrowood, M.J., de Mejia, M.A., de Merida, A.M., Hightower, A.W., Venczel, L., Herwaldt, B.L., Klein, R.E., 1999. Epidemiologic studies of *Cyclospora cayetanensis* in Guatemala. Emerg. Infect. Dis. 5, 766–774.

Bern, C., Ortega, Y., Checkley, W., Roberts, J.M., Lescano, A.G., Cabrera, L., Verastegui, M., Black, R.E., Sterling, C., Gilman, R.H., 2002. Epidemiologic differences between cyclosporiasis and cryptosporidiosis in Peruvian children. Emerg. Infect. Dis. 8, 581–585.

Brea, J.A., 2003. Population dynamics in Latin America. Popul. Bull. 58, 3–36.

Buss, B.F., Joshi, M.V., O'Keefe, A.L., Allensworth, C.D., Garvey, A., Obbink, K., Mandernach, S., Safranek, T.J., 2016. Regional investigation of a cyclosporiasis outbreak linked to imported romaine lettuce - Nebraska and Iowa, June-August 2013. Epidemiol. Infect. 144, 1807–1817.

Centers for Disease Control and Prevention, 2016. Cyclosporiasis Outbreak Investigations and Updates for the United States. https://wwwcdcgov/parasites/cyclosporiasis/outbreaks/indexhtml. (Accessed 21 May 2017).

Devera, R., Blanco, Y., Cabello, E., 2005. High prevalence of *Cyclospora cayetanensis* among indigenous people in Bolivar State, Venezuela. Cad. Saúde Pública 21, 1778–1784.

Feng, Y., Torres, E., Li, N., Wang, L., Bowman, D., Xiao, L., 2013. Population genetic characterisation of dominant *Cryptosporidium parvum* subtype IIaA15G2R1. Int. J. Parasitol. 43, 1141–1147.

Guo, Y., Roellig, D.M., Li, N., Tang, K., Frace, M., Ortega, Y., Arrowood, M.J., Feng, Y., Qvarnstrom, Y., Wang, L., Moss, D.M., Zhang, L., Xiao, L., 2016. Multilocus sequence typing tool for *Cyclospora cayetanensis*. Emerg. Infect. Dis. 22, 1464–1467.

Hall, R.L., Jones, J.L., Hurd, S., Smith, G., Mahon, B.E., Herwaldt, B.L., 2012. Population-based active surveillance for *Cyclospora* infection–United States, foodborne diseases active surveillance network (FoodNet), 1997-2009. Clin. Infect. Dis. 54 (Suppl. 5), S411–S417.

Herges, G.R., Widmer, G., Clark, M.E., Khan, E., Giddings, C.W., Brewer, M., McEvoy, J.M., 2012. Evidence that *Cryptosporidium parvum* populations are panmictic and unstructured in the upper midwest of the United States. Appl. Environ. Microbiol. 78, 8096–8101.

Hussein, E.M., El-Moamly, A.A., Mahmoud, M.A., Ateek, N.S., 2016. Wide genetic variations at 18S ribosomal RNA locus of *Cyclospora cayetanensis* isolated from Egyptian patients using high resolution melting curve. Parasitol. Res. 115, 2797–2806.

Insulander, M., Svenungsson, B., Lebbad, M., Karlsson, L., de Jong, B., 2010. A foodborne outbreak of *Cyclospora* infection in Stockholm, Sweden. Foodb. Pathog. Dis. 7, 1585–1587.

Li, N., Ye, J., Arrowood, M.J., Ma, J., Wang, L., Xu, H., Feng, Y., Xiao, L., 2015. Identification and morphologic and molecular characterization of *Cyclospora macacae* n. sp. from rhesus monkeys in China. Parasitol. Res. 114, 1811–1816.

Lymbery, A., 2002. Clonal theory of parasitic protozoa. Trends Parasitol. 18, 149.

Milord, F., Lampron-Goulet, E., St-Amour, M., Levac, E., Ramsay, D., 2012. *Cyclospora cayetanensis*: a description of clinical aspects of an outbreak in Quebec, Canada. Epidemiol. Infect. 140, 626–632.

Morrison, L.J., Mallon, M.E., Smith, H.V., MacLeod, A., Xiao, L., Tait, A., 2008. The population structure of the *Cryptosporidium parvum* population in Scotland: a complex picture. Infect. Genet. Evol. 8, 121–129.

Ortega, Y.R., Sanchez, R., 2010. Update on *Cyclospora cayetanensis*, a food-borne and waterborne parasite. Clin. Microbiol. Rev. 23, 218–234.

Puente, S., Morente, A., Garcia-Benayas, T., Subirats, M., Gascon, J., Gonzalez-Lahoz, J.M., 2006. Cyclosporiasis: a point source outbreak acquired in Guatemala. J. Trav. Med. 13, 334–337.

Sakakibara, Y., Takigawa, A., Kawabata, Y., Hirotani, T., Mukai, K., Matsumoto, K., Nakanishi, F., Tanaka, Y., Masuda, E., Hijioka, T., 2010. An imported Japanese case of cyclosporiasis. Nihon Shokakibyo Gakkai Zasshi 107, 1290–1295.

Sherchand, J.B., Cross, J.H., 2001. Emerging pathogen *Cyclospora cayetanensis* infection in Nepal. Southeast Asian J. Trop. Med. Publ. Health 32 (Suppl. 2), 143–150.

Sulaiman, I.M., Torres, P., Simpson, S., Kerdahi, K., Ortega, Y., 2013. Sequence characterization of heat shock protein gene of *Cyclospora cayetanensis* isolates from Nepal, Mexico, and Peru. J. Parasitol. 99, 379–382.

Zhou, Y., Lv, B., Wang, Q., Wang, R., Jian, F., Zhang, L., Ning, C., Fu, K., Wang, Y., Qi, M., Yao, H., Zhao, J., Zhang, X., Sun, Y., Shi, K., Arrowood, M.J., Xiao, L., 2011. Prevalence and molecular characterization of *Cyclospora cayetanensis*, Henan, China. Emerg. Infect. Dis. 17, 1887–1890.