Infection, Genetics and Evolution 96 (2021) 105151

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Evolutionary and phylogenetic analyses of the barcoding region suggest geographical relationships among *Blastocystis* sp., ST3 in humans

Sara Nemati [a], Mohsen Falahati Anbaran [b,c], Hanieh Mohammad Rahimi [a], Monireh Sadat Hosseini [a], Sara Aghaei [a], Negar Khalili [a], Hamed Mirjalali [a,*], Mohammad Reza Zali [d]

[a] Foodborne and Waterborne Diseases Research Center, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[b] School of Biology, University of Tehran, 14155-6455 Tehran, Iran
[c] NTNU University Museum, Norwegian University of Science and Technology, Trondheim, Norway
[d] Gastroenterology and Liver Diseases Research Center, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran

ARTICLE INFO

Keywords:
Blastocystis
Subtype 3
Nucleotide diversity
Population analysis
Haplotype network

ABSTRACT

Blastocystis sp., has 21 distinct subtypes of which ST3 thought to be the most prevalent subtype. This study aims to analyze the global variations of ST3. In total, 496 sequences with more than 400 nucleotides from Asia, Europe, Africa, and America were included in this study. Results show that allele 34 was the most prevalent allele in all continents. The lowest and highest allele diversity were observed in Europe and Africa, respectively. The nucleotide diversity ranged from 0.0077 in Europe to 0.02 in Africa, and haplotype diversity ranged from 0.461 in America to 0.6 in Africa. The haplotype network and Bayesian structure showed at least two major clusters including Asia and Europe-Africa-America. Tajima's D values for all continents were negative and statistically significant, indicating an excess of rare nucleotide variants. Similarly, the Fu's FS test showed negative values for all regions, indicating an excess of rare haplotypes. Pairwise F_{ST} exhibited a high genetic differentiation between Asia and other continents. Mismatch analysis for all populations showed a unimodal distribution. Our findings indicate that there are two probable major clusters of *Blastocystis* sp. ST3, a cluster which is shared between Europe, Africa, and America, and a cluster which is restricted to Asia.

1. Introduction

Blastocystis sp., is a common intestinal anaerobic eukaryote, which colonizes the intestine of many animals including humans, worldwide (Mohammad Rahimi et al., 2021; Stechmann et al., 2008; Yoshikawa et al., 2007). Although this protist is reported in all continents, variations in the prevalence rate are related to geographical regions, studied populations, and methods of detection (Alfellani et al., 2013b; Javanmard et al., 2018). However, it's higher prevalence in developing countries is thought to be associated with poor hygiene standards (El Safadi et al., 2014). The main transmission routs of this protist are supposed to be the oral-fecal and close-contact to animals (Tan, 2008).

Based on the genetic features, several subtypes of *Blastocystis* sp., have been characterized (Hoevers et al., 2000; Mohammad Rahimi et al., 2019; Scicluna et al., 2006). The standard classification method is mainly based on sequencing of the barcoding region of the small subunit ribosomal RNA (SSU-rRNA) gene (Stensvold et al., 2007). Recently, 21 subtypes (ST1 - ST17 and ST23-ST26) have been characterized in mammalians and birds, but only half of them ($n = 10$) have been reported from humans (Stensvold and Clark, 2020; Stensvold et al., 2007).

Subtype 3 (ST3) is widely reported in epidemiological studies in both animals and humans, worldwide (Nemati et al., 2021; Rezaei Riabi et al., 2018; Segui et al., 2018; Stensvold et al., 2009). In addition, some clinical manifestations such as gastrointestinal disorders and urticarial have been suggested to be linked to *Blastocystis* sp., subtype 3 (Katsarou-Katsari et al., 2008; Mohamed et al., 2017). *Blastocystis* sp., ST3 also exhibits a wide geographic distribution and high genetic variation compared to other *Blastocystis* sp., subtypes (Meloni et al., 2012; Rezaei Riabi et al., 2018; Stensvold et al., 2012). However, to our knowledge, there is limited information on the genetic variations of this subtype and

* Corresponding author.
E-mail address: hamedmirjalali@sbmu.ac.ir (H. Mirjalali).

https://doi.org/10.1016/j.meegid.2021.105151
Received 27 February 2021; Received in revised form 11 November 2021; Accepted 16 November 2021
Available online 18 November 2021

the probable role of migration on the global distribution of *Blastocystis* sp. subtype 3. Regarding the wide distribution, high prevalence, and a probabile correlation between the presence of *Blastocystis* sp., subtype 3 and some clinical manifestations, a molecular analysis on available sequences of the barcoding region can provide interesting data about distribution and transmission in different communities, countries, and continents. Therefore, in this research, we aimed to determine the nucleotide variation and the geographic distribution of haplotypes of *Blastocystis* sp. ST3 isolated from humans across Asia, Africa, America, Europe continents.

2. Methods

2.1. Sequence data

We conducted a continental survey on the submitted nucleotide sequences of the barcoding region of *Blastocystis* sp. subtype 3 from four continents (Asia, Africa, America, Europe) in the GenBank database to the end of 2019. In this study, sequences with more than 400 nucleotides, which were submitted to the GenBank database, were selected. All sequences were limited to humans as host. To evaluate the reliability of sequences and exclude the probable wrong sequences (those which were wrongly downloaded), all sequences were compared again in the basic local alignment tools (BLAST; https://blast.ncbi.nlm.nih.gov/Blast.cgi). To analyze sequences with same length, the sequences were trimmed in BioEdit software (BioEdit V7.0), and aligned using CLUSTALW.

To characterize the alleles, all the SSU rRNA gene sequences corresponded for ST3, which were retrieved from the GenBank database, were submitted to https://pubmlst.org/blastocystis/.

2.2. Molecular diversity

At the first, grouping of the sequences was done based on the information related to the origin country of sequences as mentioned in the GenBank database. To compare sequences regarding the continents, various diversity parameters including haplotype (sequence-type) diversity (Hd), nucleotide diversity (π), number of segregating sites (S), and average number of pairwise nucleotide differences (K) were estimated. Tajima's D (Tajima, 1989) and Fu's Fs (Fu, 1997) neutrality indices were used to infer the demographic history of *Blastocystis* sp. subtype 3. Pairwise fixation index (*F*st) of the sequences of *Blastocystis* sp. subtype 3 regarding continents was employed to investigate the gene flow. DnaSP ver.5 (Librado and Rozas, 2009) was employed to analyze the indices. Accordingly, *F*st $\leq$ 0.1 represents low differentiation, populations with *F*st 0.15–0.2 are moderately differentiated, and *F*st $\geq$ 0.25 showed high differentiations. According to the references for the level of gene flow (Nm), Nm $>$ 1 and Nm $<$ 1 represent high gene flow and low gene flow, respectively (Govindaraju, 1989).

2.3. Networking and phylogenetic analysis

To investigate the global pattern of genetic relationship among different haplotypes of *Blastocystis* sp. subtype 3, haplotype TCS network (Clement et al., 2000), which is implemented in the PopART program ver. 1.7 (http://popart.otago.ac.nz/downloads.shtml), was employed. The demographic history and the distribution of pairwise differences (Rogers and Harpending, 1992; Slatkin and Hudson, 1991) between individual sequences were evaluated by mismatch distribution analysis using DnaSP 5.0 and Arlequin (Librado and Rozas, 2009).

The analysis of the Bayesian population structure was performed using STRUCTURE v2.3 http://web.stanford.edu/group/pritchardlab/structure.html) (Pritchard et al., 2000). An admixture model with correlated allele frequency was used for parameter estimation following 1×10^5 iterations after a burn-in of 2×10^5 iterations. The most likely number of cluster (K) was determined by evaluating the mean ln estimated probability of data with Structure Harvester V0.6.94 (http://taylor0.biology.ucla.edu/structureHarvester) (Earl and Vonholdt, 2012). The results for the proportion membership of individuals into estimated clusters were summarized with Clumpak. The global genetic structure of isolates was also investigated by plotting the ancestry coefficient over the geographic distribution. This analysis was performed with the tess3r package in R with K = 2 ancestral populations (Caye et al., 2016).

3. Results

3.1. Sequences analysis and allele discrimination

The allele analysis showed that allele 34 was the most prevalent allele in all continents. The lowest and highest number of alleles were found in America ($n = 4$) and Europe ($n = 19$), respectively. The analysis showed that three alleles 34, 36, and 37 were characterized in all continents, while other alleles were detected in three or less continents. Two alleles 23 and 56 were characterized only in Asia. Alleles 52, 101, and 136 were detected only in Africa, and alleles 44–47, 49–51, 53, 54, 127, and 128 were characterized only in Europe (Supplementary Table 1, Fig. 1).

3.2. Genetic diversity

The genetic diversity indices revealed a total of 195 S and 134 haplotypes, as well as a total Hd of 0.762 and π of 0.012 among all sequences. The genetic diversity analysis in each continent showed that the number of haplotypes ranged from 14 in America to 85 in Asia. As well, the number of segregating sites ranged 25 in America to 142 in Asia. The Hd was high, ranging from 0.461 $\pm$ 0.86 in America to 0.600 $\pm$ 0.76 in Africa, and nucleotide diversity ranged between 0.00771 in Europe to 0.02025 in Africa (Table 1). African sequence population exhibited the highest values of genetic diversity indices for ST3 in contrast to Europe population (Table 1).

3.3. Haplotype network and phylogenetic analysis

The number of haplotypes of each continent includes 85 in Asia, 19 in Africa, 14 in America, and 24 in Europe. The network showed a star topology radial distribution. The haplotype network of sequences showed at least two major clusters 1 and 2 (Supplementary Fig. 1, Figs. 2, 3). Cluster 1 was the dominant and almost all haplotypes in this cluster were from Asia, while sequences in cluster 2 were found in all 188 samples from Europe, Africa, America (Supplementary Fig. 1, Supplementary Table 1, Fig. 2).

3.4. Neutrality and mismatch distribution analysis

A negative value of Tajimas's D indicates past population growth. Neutrality tests revealed negative Tajima's D, which were significantly differed from zero for all populations, indicating excess of rare nucleotide variants compared to what would be expected under a neutrality (Table 1). Similarly, the results of Fu's FS test, which is based on the distribution of haplotypes, showed negative values for all populations, indicating an excess of rare haplotypes compared to what would be expected under neutrality (Table 1).

The pairwise genetic differentiations between Asia and other continents was relatively high (F_{ST} = 0.28–0.37). However, the pairwise F_{ST} was low for Africa-Europe, Africa-America, and America-Europe comparisons (F_{ST} = 0.015–0.44, Table 2).

The lowest F_{ST} values between the continental groups were seen between Europe vs. Africa (0.015), America vs. Europe (0.044), and Africa vs. America (0.0185) (Table 2). The result of gene flow showed high levels of Nm between America and Europe (201.52), followed by Africa vs Europe (133.79), and Africa vs America (50.68). In addition, the lowest levels of gene flow were observed between Asia and other

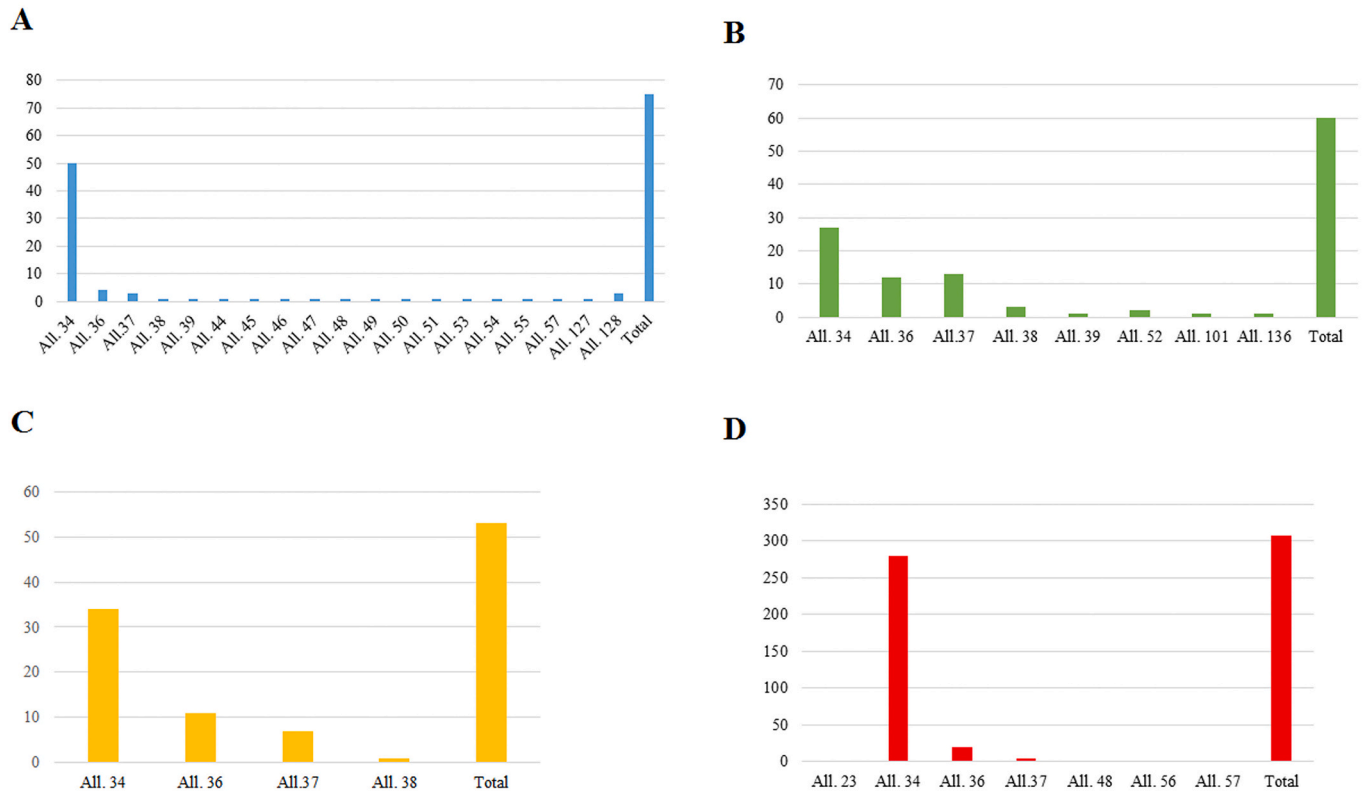

Fig. 1. Allelic frequency and diversity in different continents A) Europe, B) Africa, C) America, and D) Asia.

Table 1
Diversity statistics estimated for different populations of *Blastocystis* ST3 across the global distribution range.

Populations	N	S	h	Hd	π	Tajima's D	Fu's Fs
Asia	308	142	85	0.552 ± 0.36	0.00845	−2.85861 ***	−10.953**
Africa	60	77	19	0.600 ± 0.76	0.02025	−2.66667 ***	−13.01537**
America	53	25	14	0.461 ± 0.86	0.00966	−2.13315*	−13.01537**
Europe	75	42	24	0.558 ± 0.071	0.00771	−2.70963***	−13.01537**

N: number of sequences; *S*: number of segregating sites; *h*: number of haplotypes; Hd: haplotype diversity; π: nucleotide diversity. ** and *** indicate statistical significant at $P < 0.01$ and 0.001, respectively.

continents (Table 2). The mismatch distribution plot for all populations also showed unimodal graphs, demonstrating a recent population demographic and range expansion (Fig. 4).

3.5. Population structure

To determine the optimal number of genetic clusters, the model based analysis of population structure revealed the presence of two ancestral clusters (K = 2). These clusters were strongly differentiated from each other, where Asian sequences belonged to cluster 1, and sequences from other continents belonged to cluster 2 (Fig. 2). The result of tess3r represented the presence of two separated populations regarding geography distribution, confirming the findings of structure analysis (Supplementary Fig. 1, Fig. 5).

4. Discussion

Blastocystis sp., is a prevalent protist with worldwide distribution. Epidemiological studies have frequently reported the presence of three *Blastocystis* sp., subtypes 1–3 from humans with majority of ST3 across different geographic regions (Alfellani et al., 2013b; Stensvold and Clark, 2016). For example, in a study performed in the United Arab Emirates (UAE) ST3 was recognized among 58.9% of *Blastocystis* sp., carriers and was the most prevalent detected subtype (AbuOdeh et al., 2016). Abu-Madi et al. (2015) analyzed the subtype distribution of *Blastocystis* sp., among carriers in Qatar that the results showed high prevalence of ST3 (69.3%) among samples. Rezaei Riabi et al. (2018) investigated the prevalence of *Blastocystis* sp., and its subtypes in human subjects in Iran and claimed that ST3 was characterized among 45% of *Blastocystis* sp., carriers. Similar reports from other continents including European, African, and American countries have also reported ST3 as the most prevalent subtype. In a study in Brazil, *Blastocystis* sp. isolated from human samples in Rio de Janeiro were molecularly analyzed and subtype 3 was characterized among 49% of *Blastocystis* sp., carriers (Valenca Barbosa et al., 2017). In a review published by Jimenez et al. (2019), *Blastocystis* sp., ST3 was suggested to be the most prevalent subtype among human subjects in the South America (37.9%). In a large-scale study performed by Gabrielli et al. (2020) in Italy, 40% of *Blastocystis* sp., carriers were identified to harbor subtype 3. Forsell et al. (2017) analyzed the prevalence of *Blastocystis* sp., and its subtypes in Swedish subjects and identified ST3 as the most common subtype besides ST4. The predominance of ST3 in African countries is controversy (Abdulsalam et al., 2013; D'Alfonso et al., 2017; Poulsen et al., 2016), however, *Blastocystis* sp., ST3 was the major detected subtype among *Blastocystis* sp., carriers in western Angola (Dacal et al., 2018).

Our results represent that alleles 34, 36, and 37 are the most

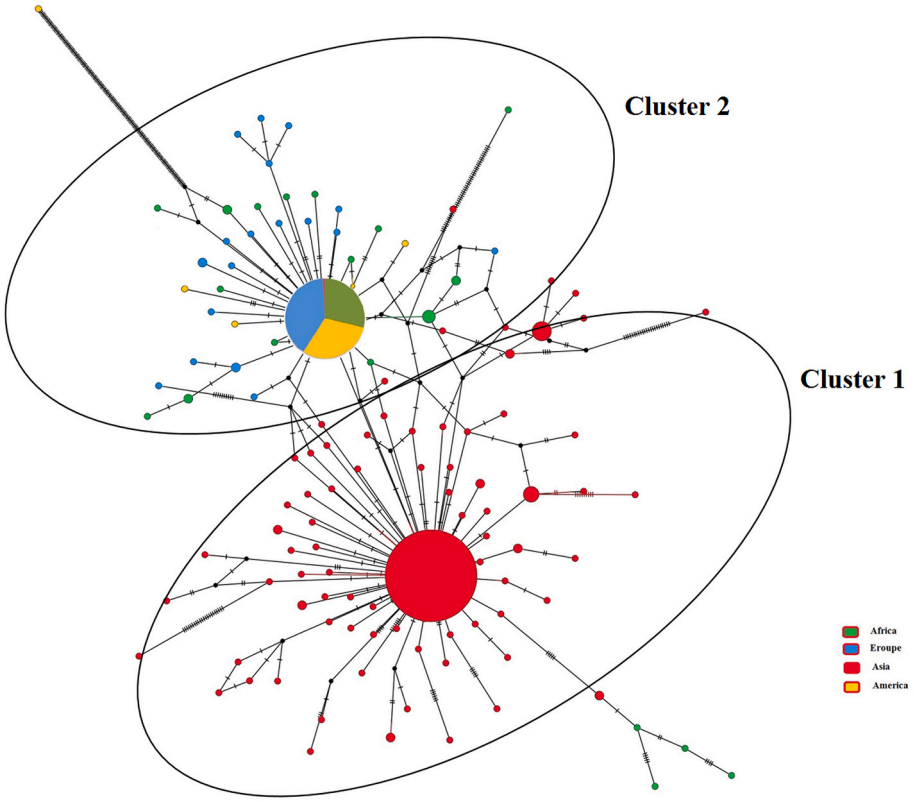

Fig. 2. Haplotype network of all studied sequences of *Blastocystis* sp., subtype 3 obtained from human populations. Each circle represents different haplotype, and each color represents different population. The size of each circle is proportional to the number of the samples in each population.

prevalent alleles and allele 34 is the predominant allele. Allele discrimination is usually performed based on the analysis of the barcoding region of SSU rRNA gene to identify alleles of subtypes ST1-ST10 (Alfellani et al., 2013a). In a study by Pandey et al. (2015), allele 34 was found as the major allele of ST3 among Indian populations. Rezaei Riabi et al. (2018) identified allele 34 as the most prevalent allele among both symptomatic and asymptomatic human subjects in Iran. This allele is suggested as the major strain colonizing the intestine of human population in South America, as well (Jimenez et al., 2019; Villamizar et al., 2019). Although there is no strong evidence to exhibit an association between certain allele and clinical manifestations, few studies have shown a linkage between allele 34 and urticarial (Casero et al., 2015). In addition, allele 34 is likely the main strain identified in animal samples (Alfellani et al., 2013a). However, there is no strong evidence explaining that ST3 allele 34 is a zoonotic strain.

The results show that the highest variable sites are found in African sequences. In addition, the high number of segregating sites, haplotype diversity (0.600 ± 0.76), and π (0.020) in African population indicate the high probability of cross transmission for *Blastocystis* sp. subtype 3 between humans and animals in Africa. This finding is supported by the results of previous studies that suggested correlation between high molecular diversity and high host/source multiplicity (Martinez-Hernandez et al., 2020; Mohammad Rahimi et al., 2020). In contrast, the lowest haplotype diversity and the number of segregating sites were found in America, proposing the lowest host/source multiplicity. Interestingly, the values of π and Hd in Africa was higher than other continents, indicating the probable population expansion in this continent. Importantly, the 5′ barcoding region of SSU rRNA gene, is a hypervariable fragment in ST3 (Alinaghizade et al., 2017); therefore, the lowest nucleotide diversity (π) in sequences from continents other than Africa suggests the clonally of *Blastocystis* sp. subtype 3 in Asia, Europe, and America compared to Africa (Ramirez et al., 2014; Stensvold, 2013; Stensvold et al., 2012).

The highest F_{ST} value (>0.25) was observed between Asia and other three continents (America, Africa and Europe), indicating high genetic differences or great separation. This pattern may be due to either the continental barriers and high geographical distance, or limited migration between Asia and other continents. On the other hand, negligible pairwise differentiation ($F_{ST} < 0.05$) observed between America - Europe, Africa - America, and Africa - Europe may indicate high genetic connectivity due to high gene flow among continents (Hamrick and Godt, 1996; Hastings and Harrison, 1994; Holsinger and Weir, 2009; Lang et al., 2012).

Gene flow represents the transfer of variations through the genomes of a population to others. Therefore, a high value of gene flow implies decreased genetic differentiations, which reflects the homogeneity, the integrity, and migration between two populations (Bolnick and Nosil, 2007; Seeburg et al., 1984; Su et al., 2003). The gene flow analysis showed the high level of Nm between America and Europe followed by Africa and Europe, and America and Africa. There are a number of factors that affect the rate of gene flow between different populations. However, this finding confirmed the results of *F*st and suggested that the high dispersal or mobility between these continents seems to be the main reason for this observation (Hamrick and Godt, 1996; Hastings and Harrison, 1994; Lang et al., 2012).

Haplotype network showed relatively high divergence and low number of common haplotypes in African sequences compared to other continents. This supports high nucleotide diversity and may suggest a link between high genetic diversity and hosts/source multiplicity in Africa (Martinez-Hernandez et al., 2020; Mohammad Rahimi et al., 2020).

Interestingly, the network analysis revealed a “star-like” topology with many singleton variations that suggests a recent rapid expansion (Avise et al., 1987; Slatkin and Hudson, 1991). The pattern also indicates the presence of at least two separate clusters of *Blastocystis* sp. subtype 3; cluster 1 belonging to Asia and cluster 2 belonging to the

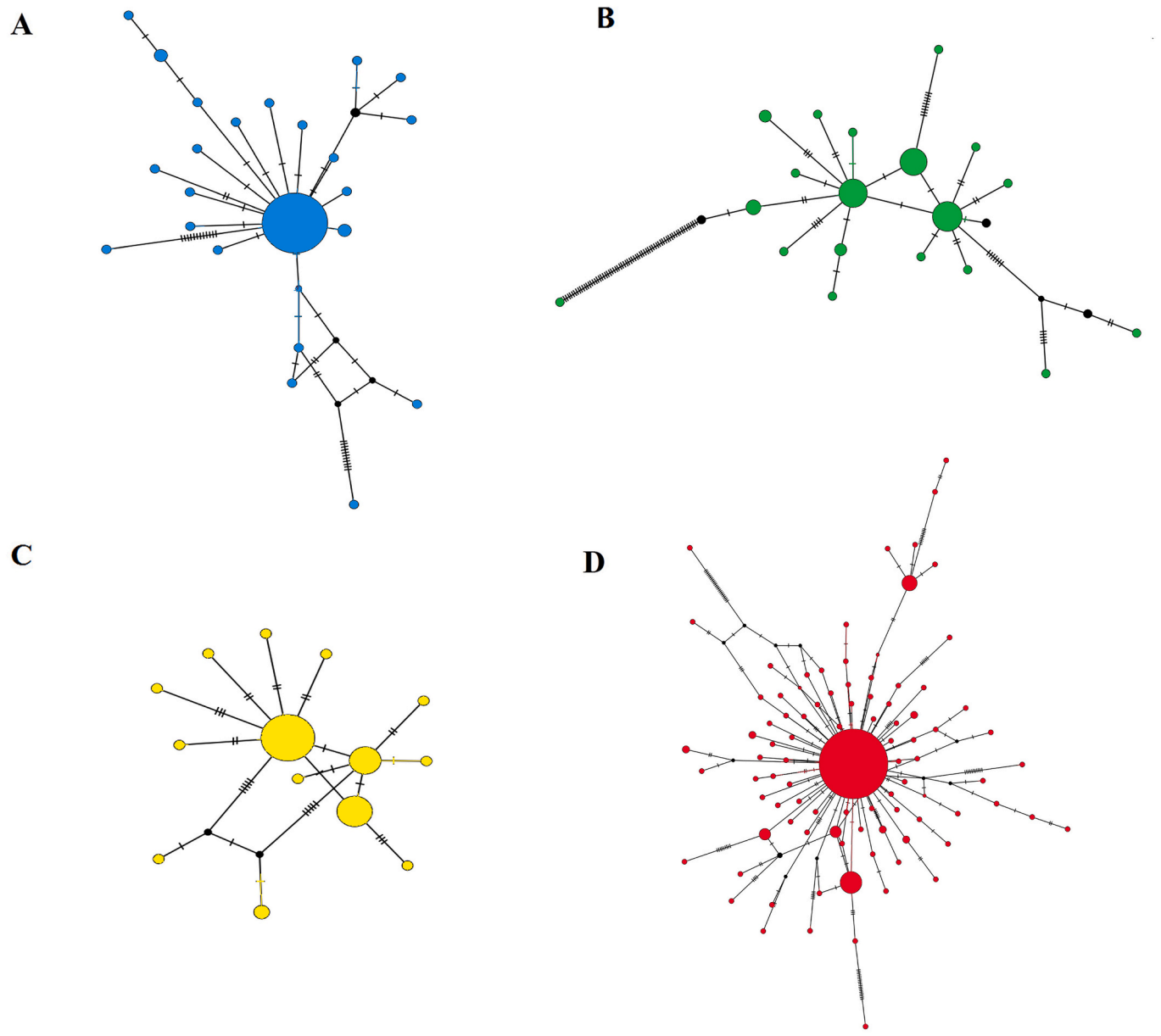

Fig. 3. Haplotype network of *Blastocystis* sp., subtype 3 populations in A) Europe, B) Africa, C) America, and D) Asia. The size of each circle is proportional to the frequency of the haplotype in each population.

Table 2
Estimates of pairwise F_{ST} based on the SSU-rRNA gene variation calculated for different populations of *Blastocystis* ST3.

Population 1	Population 2	Nm	F_{ST}	G_{ST}
Africa	America	50.68	0.0185	0.004
Africa	Europe	133.79	0.015	−0.0018
Africa	Asia	1.20	0.337	0.172
America	Europe	201.52	0.044	0.0001
America	Asia	1.09	0.282	0.186
Europe	Asia	0.99	0.37	0.186

other three continents. This finding supports the results of F_{ST} indicating the high similarity between *Blastocystis* sp. subtype 3 isolated from America, Europe, and Africa.

However, s "star-like" topology, which was observed in our study, is in contrast with the study published by Rojas-Velázquez et al. (2018). Actually, Rojas-Velázquez et al. (2018) claimed the presence of an ancestral clone for *Blastocystis* sp., and concluded that the haplotype 1, which was detected in their study, is probably originated in the Latin America and distributed to other regions of the world via human migration. However, our study, which was performed on higher number of sequences with wider distribution in the world suggested the presence of two probable major clusters. A cluster was restricted to Asian population and the other cluster contained sequences from America (mostly south America), Europe, Africa, and some samples from Asia. However, the main difference between our analysis and Rojas-Velázquez et al. (2018) seems to be due to the number of sequences and the number of nucleotides, which were included in analysis.

The demographic events with Tajima's D and Fu's FS tests, reveal excess of singleton segregating sites and rare haplotypes, respectively, that suggest recent demographic expansion across global distribution range. Here, a significant negative D and *Fs* values can be interpreted as signatures of population expansion (Tajima, 1989). The mismatch distribution results showed four unimodal graphs suggesting recent population demographic that supports the hypothesis of population

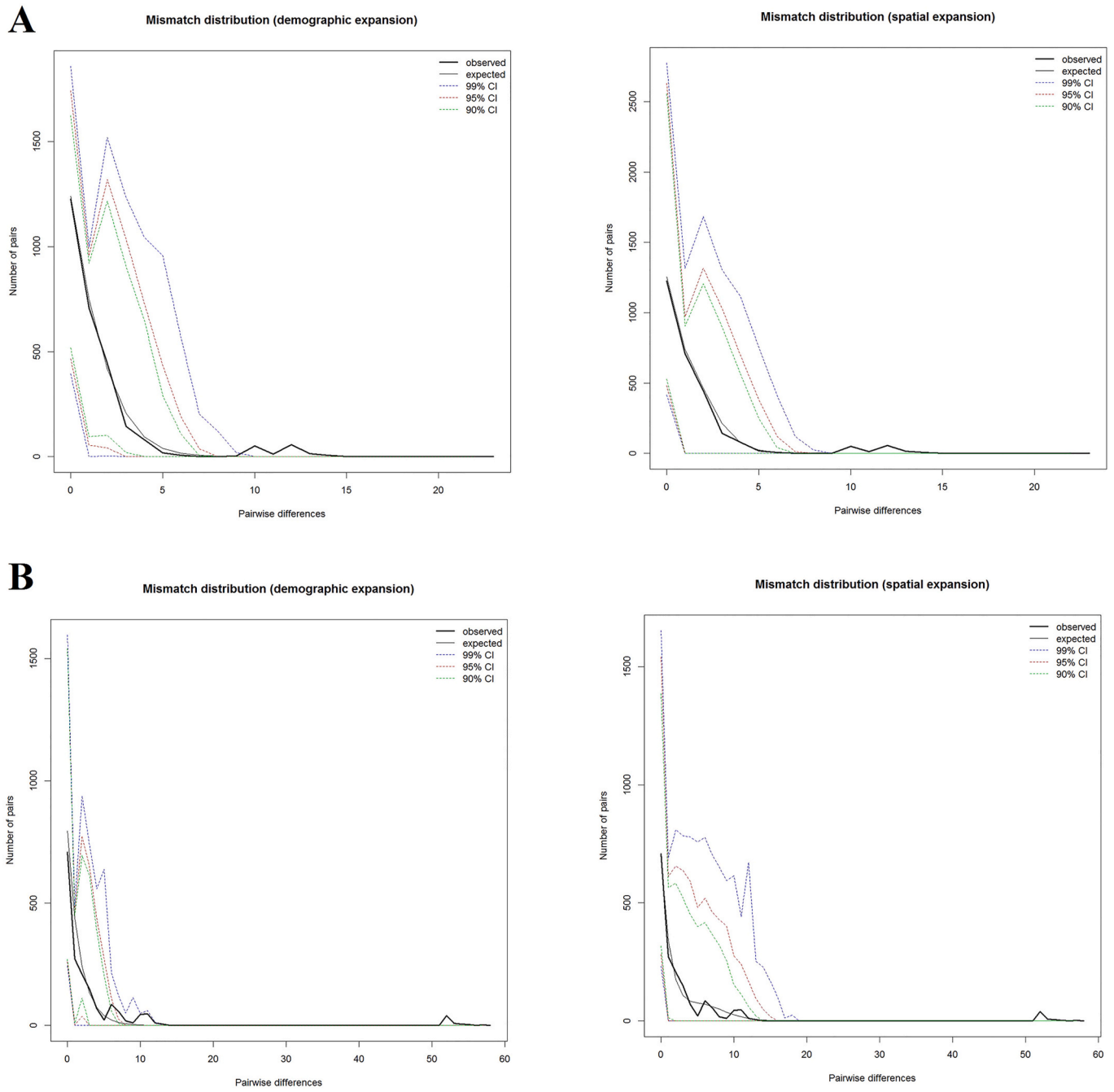

Fig. 4. The graphs obtained from mismatch distribution analysis for A) Europe, B) Africa, C) America, and D) Asia. The x axis shows the number of pairwise differences, the y axis shows the frequency of the pairwise comparisons. The observed frequencies are represented by red dotted line. The frequency expected under the hypothesis of population expansion model is depicted by continuous green line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

expansion (Slatkin and Hudson, 1991).

There are some limitations in this study. (1) this study was performed on about 500 sequences, which were previously submitted to the GenBank database, and some errors through the submitted genes by original authors may lead to misinterpretations. (2) most of molecular epidemiology studies on *Blastocystis* sp., have been performed on conserved non-coding genes and a few number of coding genes, like mitochondrial genes, are available in databases that limits the further genetic and ancestral analyses.

5. Conclusions

Our findings show that allele 34 is the most prevalent strain with variable frequency across the global distribution range. The lowest and highest allele frequency are seen in America and Europe, respectively. The highest genetic diversity was seen in Africa population, suggesting high probability of cross-transmission between different hosts in Africa compared to other continents. Haplotype network, Bayesian structure, and pairwise F_{ST} indicate that Asian population is clearly separated from European, African, and American populations, which could reveal the

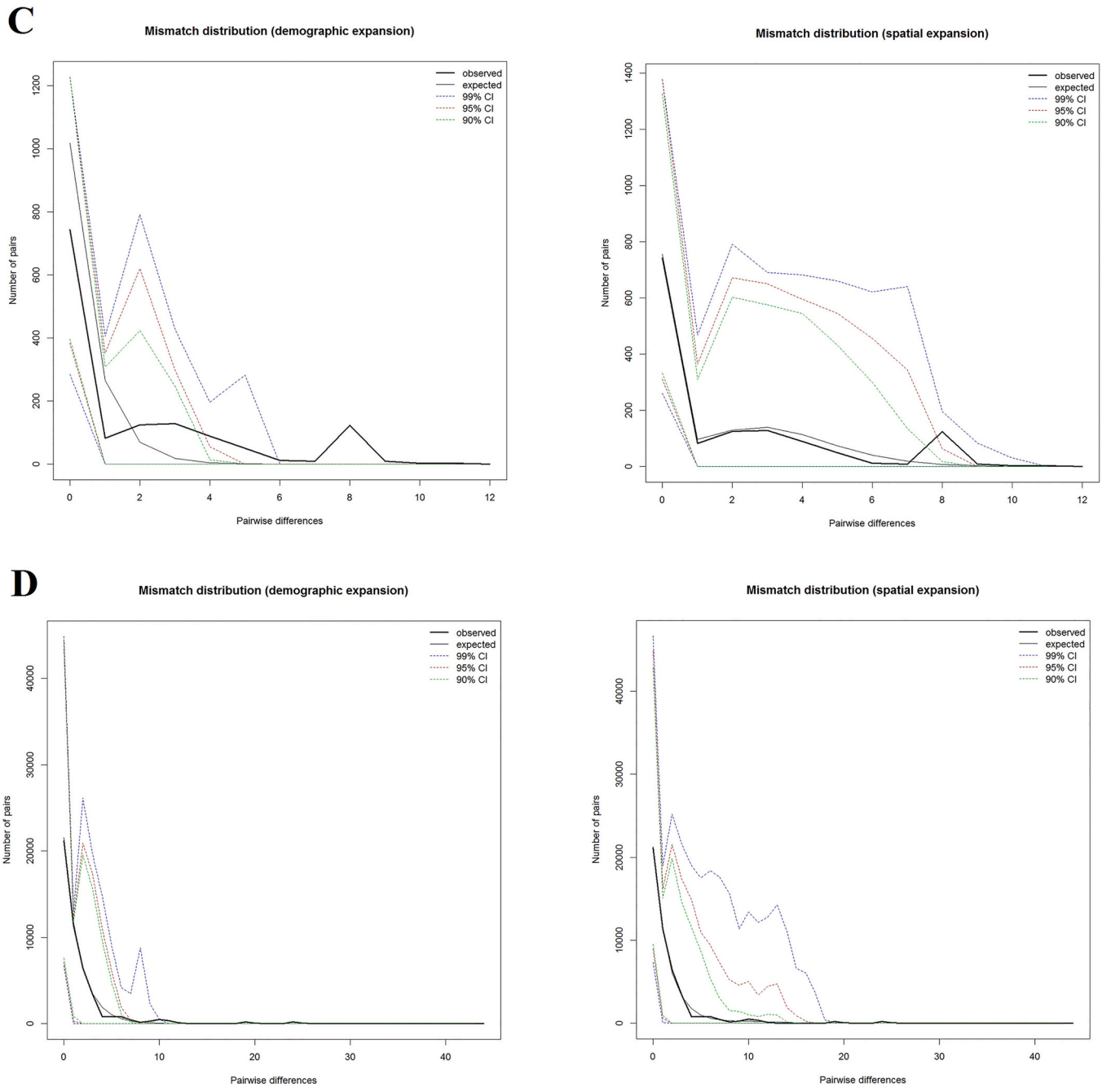

Fig. 4. (*continued*).

probable role of migration in distribution of *Blastocystis* sp., in the world. Taken together, unlike to Asia, high gene flow values between Africa, Europe, and America continents suggest a genetic uniformity among *Blastocystis* sp. subtype 3 in these continents. However, large molecular epidemiology on *Blastocystis* sp., subtype 3 isolated from humans from all continents would provide interesting data about the genetic variation of this protist across the world. Regarding the limited molecular studies on mitochondrial genes in *Blastocystis* sp., it seems that researches with large sample size on molecular analysis of conserved coding genes such as mitochondrial genes provide big datasets, which helps researchers to design population-based studies.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2021.105151.

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors declare that they have seen and approved the submitted version of this manuscript.

Availability of data and material

All generated data from the current study are included in the article or supplementary materials.

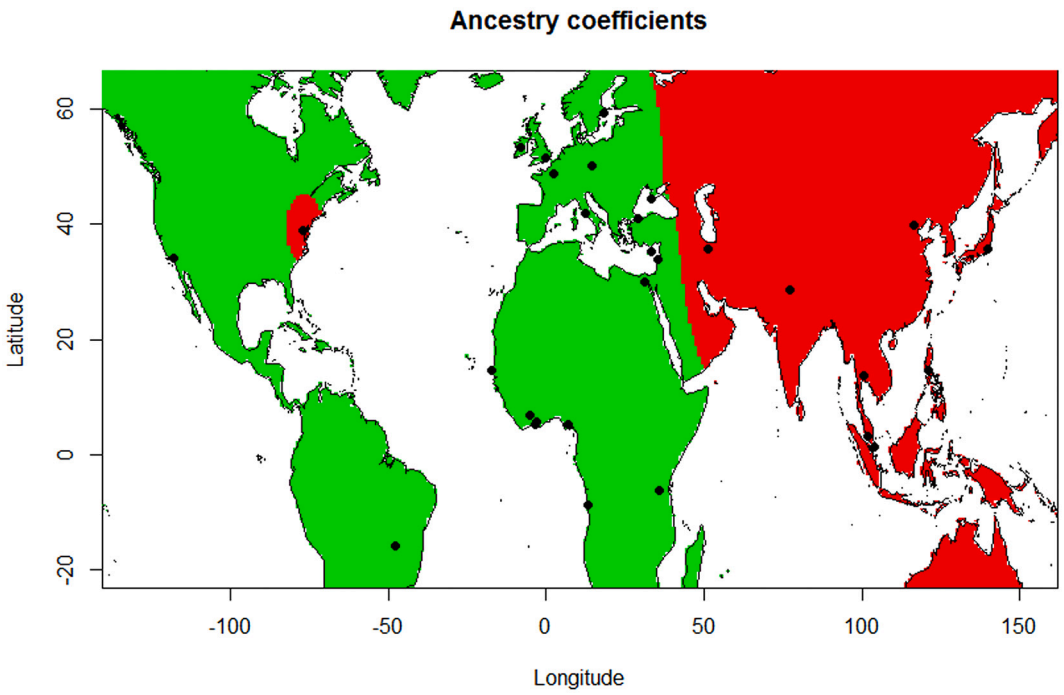

Fig. 5. The optimal number of clusters detected with Bayesian structure analysis. The map showed two potential ancestral separated clusters (Asia and other continents). Dots present haplotypes distribution based on the studied countries.

Competing interests

The authors declare that they have no conflict of interest.

Funding

None.

Author contribution

Conceived and designed the experiments: HM MRZ. Sequence gathering: HMR MSH, SA, NK. Analyzed the data: SN MFA. Wrote the paper: HM SN MFA. All authors read and approved the final version of the manuscript.

CRediT authorship contribution statement

Sara Nemati: Data curation, Formal analysis, Software, Validation, Visualization, Writing – review & editing. **Mohsen Falahati Anbaran:** Formal analysis, Software, Visualization, Writing – review & editing. **Hanieh Mohammad Rahimi:** Investigation. **Monireh Sadat Hosseini:** Investigation. **Sara Aghaei:** Investigation. **Negar Khalili:** Investigation. **Hamed Mirjalali:** Conceptualization, Data curation, Formal analysis, Project administration, Supervision, Validation, Writing – review & editing. **Mohammad Reza Zali:** Funding acquisition.

Acknowledgment

The authors thank all members of the Foodborne and Waterborne Diseases Research Center for their supports.

References

Abdulsalam, A.M., Ithoi, I., Al-Mekhlafi, H.M., Al-Mekhlafi, A.M., Ahmed, A., Surin, J., 2013. Subtype distribution of *Blastocystis* isolates in Sebha, Libya. PLoS One 8, e84372. https://doi.org/10.1371/journal.pone.0084372.

Abu-Madi, M., Aly, M., Behnke, J.M., Clark, C.G., Balkhy, H., 2015. The distribution of *Blastocystis* subtypes in isolates from Qatar. Parasit. Vectors 8, 465. https://doi.org/10.1186/s13071-015-1071-3.

AbuOdeh, R., Ezzedine, S., Samie, A., Stensvold, C.R., ElBakri, A., 2016. Prevalence and subtype distribution of *Blastocystis* in healthy individuals in Sharjah, United Arab Emirates. Infect. Genet. Evol. 37, 158–162. https://doi.org/10.1016/j.meegid.2015.11.021.

Alfellani, M.A., et al., 2013a. Diversity and distribution of *Blastocystis* sp. subtypes in non-human primates. Parasitology 140, 966–971. https://doi.org/10.1017/S0031182013000255.

Alfellani, M.A., Taner-Mulla, D., Jacob, A.S., Imeede, C.A., Yoshikawa, H., Stensvold, C. R., Clark, C.G., 2013b. Genetic diversity of *Blastocystis* in livestock and zoo animals. Protist 164, 497–509. https://doi.org/10.1016/j.protis.2013.05.003.

Alinaghizade, A., Mirjalali, H., Mohebali, M., Stensvold, C.R., Rezaeian, M., 2017. Inter- and intra-subtype variation of *Blastocystis* subtypes isolated from diarrheic and non-diarrheic patients in Iran. Infect. Genet. Evol. 50, 77–82. https://doi.org/10.1016/j.meegid.2017.02.016.

Avise, J.C., et al., 1987. Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. Annu. Rev. Ecol. Syst. 18, 489–522. https://doi.org/10.1146/annurev.es.18.110187.002421.

Bolnick, D.I., Nosil, P., 2007. Natural selection in populations subject to a migration load. Evolution 61, 2229–2243. https://doi.org/10.1111/j.1558-5646.2007.00179.x.

Casero, R.D., Mongi, F., Sanchez, A., Ramirez, J.D., 2015. *Blastocystis* and urticaria: examination of subtypes and morphotypes in an unusual clinical manifestation. Acta Trop. 148, 156–161. https://doi.org/10.1016/j.actatropica.2015.05.004.

Caye, K., Deist, T.M., Martins, H., Michel, O., François, O., 2016. TESS3: fast inference of spatial population structure and genome scans for selection. Mol. Ecol. Resour. 16, 540–548. https://doi.org/10.1111/1755-0998.12471.

Clement, M., Posada, D., Crandall, K.A., 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9, 1657–1659. https://doi.org/10.1046/j.1365-294x.2000.01020.x.

Dacal, E., et al., 2018. Prevalence and molecular characterization of *Strongyloides stercoralis*, *Giardia duodenalis*, *Cryptosporidium* spp., and *Blastocystis* spp. isolates in school children in Cubal, Western Angola. Parasit. Vectors 11, 67. https://doi.org/10.1186/s13071-018-2640-z.

D'Alfonso, R., et al., 2017. Blastocystis in cote d'Ivoire: molecular identification and epidemiological data. Eur. J. Clin. Microbiol. Infect. Dis. 36, 2243–2250. https://doi.org/10.1007/s10096-017-3053-1.

Earl, D.A., Vonholdt, B., 2012. Structure Harvester: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour. 4.

El Safadi, D., et al., 2014. Children of Senegal River basin show the highest prevalence of *Blastocystis* sp. ever observed worldwide. BMC Infect. Dis. 14, 164. https://doi.org/10.1186/1471-2334-14-164.

Forsell, J., Bengtsson-Palme, J., Angelin, M., Johansson, A., Evengard, B., Granlund, M., 2017. The relation between *Blastocystis* and the intestinal microbiota in Swedish travellers. BMC Microbiol. 17, 231. https://doi.org/10.1186/s12866-017-1139-7.

Fu, Y.X., 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147, 915–925.

Gabrielli, S., Furzi, F., Fontanelli Sulekova, L., Taliani, G., Mattiucci, S., 2020. Occurrence of Blastocystis-subtypes in patients from Italy revealed association of ST3 with a healthy gut microbiota. Parasite Epidemiol. Cont. 9 https://doi.org/10.1016/j.parepi.2020.e00134 e00134.

Govindaraju, D.R., 1989. Variation in gene flow levels among predominantly self-pollinated plants. J. Evol. Biol. 2, 173–181. https://doi.org/10.1046/j.1420-9101.1989.2030173.x.

Hamrick, J.L., Godt, M.J.W., 1996. Effects of life history traits on genetic diversity in plant species. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 351, 1291–1298. https://doi.org/10.1098/rstb.1996.0112.

Hastings, A., Harrison, S., 1994. Metapopulation dynamics and genetics. Annu. Rev. Ecol. Syst. 25, 167–188. https://doi.org/10.1146/annurev.es.25.110194.001123.

Hoevers, J., Holman, P., Logan, K., Hommel, M., Ashford, R., Snowden, K., 2000. Restriction-fragment-length polymorphism analysis of small-subunit rRNA genes of *Blastocystis hominis* isolates from geographically diverse human hosts. Parasitol. Res. 86, 57–61.

Holsinger, K.E., Weir, B.S., 2009. Genetics in geographically structured populations: defining, estimating and interpreting F(ST). Nat. Rev. Genet. 10, 639–650. https://doi.org/10.1038/nrg2611.

Javanmard, E., Niyyati, M., Ghasemi, E., Mirjalali, H., Asadzadeh Aghdaei, H., Zali, M.R., 2018. Impacts of human development index and climate conditions on prevalence of Blastocystis: a systematic review and meta-analysis. Acta Trop. 185, 193–203. https://doi.org/10.1016/j.actatropica.2018.05.014.

Jimenez, P.A., Jaimes, J.E., Ramirez, J.D., 2019. A summary of *Blastocystis* subtypes in north and South America. Parasit. Vectors 12, 376. https://doi.org/10.1186/s13071-019-3641-2.

Katsarou-Katsari, A., Vassalos, C.M., Tzanetou, K., Spanakos, G., Papadopoulou, C., Vakalis, N., 2008. Acute urticaria associated with amoeboid forms of *Blastocystis* sp. subtype 3. Acta Derm. Venereol. 88, 80–81. https://doi.org/10.2340/00015555-0338.

Lang, A.S., Zhaxybayeva, O., Beatty, J.T., 2012. Gene transfer agents: phage-like elements of genetic exchange. Nat. Rev. Microbiol. 10, 472–482. https://doi.org/10.1038/nrmicro2802.

Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics (Oxford, England) 25, 1451–1452. https://doi.org/10.1093/bioinformatics/btp187.

Martinez-Hernandez, F., et al., 2020. Molecular genotyping of *Blastocystis* spp. in wild mammals from Mexico. Parasitol. Res. 119, 97–104. https://doi.org/10.1007/s00436-019-06530-4.

Meloni, D., et al., 2012. Mixed human intra- and inter-subtype infections with the parasite *Blastocystis* sp. Parasitol. Int. 61, 719–722. https://doi.org/10.1016/j.parint.2012.05.012.

Mohamed, R.T., et al., 2017. Subtyping of Blastocystis sp. isolated from symptomatic and asymptomatic individuals in Makkah, Saudi Arabia. Parasit. Vectors 10, 174. https://doi.org/10.1186/s13071-017-2114-8.

Mohammad Rahimi, H., Mirjalali, H., Niyyati, M., Haghighi, A., Asadzadeh Aghdaei, H., Zali, M.R., 2019. Development and evaluation of high-resolution melting curve analysis for rapid detection and subtyping of *Blastocystis* and comparison the results with sequencing. Parasitol. Res. 118, 3469–3478. https://doi.org/10.1007/s00436-019-06486-5.

Mohammad Rahimi, H., Nemati, S., Mirjalali, H., Sharifdini, M., Zali, M.R., 2020. Molecular characterization and identification of *Blastocystis* and its subtypes from raccoon (*Procyon lotor*) in north of Iran. Parasitol. Res. 119, 2741–2745. https://doi.org/10.1007/s00436-020-06770-9.

Mohammad Rahimi, H., Mirjalali, H., Zali, M.R., 2021. Molecular epidemiology and genotype/subtype distribution of *Blastocystis* sp., *Enterocytozoon bieneusi*, and *Encephalitozoon* spp. in livestock: concern for emerging zoonotic infections. Sci. Rep. 11 (1), 17467. https://doi.org/10.1038/s41598-021-96960-x (PMID: 34471179; PMCID: PMC8410837).

Nemati, S., Zali, M.R., Johnson, P., Mirjalali, H., Karanis, P., 2021. Molecular prevalence and subtype distribution of *Blastocystis* sp. in Asia and in Australia. J. Water Health 19 (5), 687–704. https://doi.org/10.2166/wh.2021.011.

Pandey, P.K., et al., 2015. Prevalence and subtype analysis of *Blastocystis* in healthy Indian individuals. Infect. Genet. Evol. 31, 296–299. https://doi.org/10.1016/j.meegid.2015.02.012.

Poulsen, C.S., Efunshile, A.M., Nelson, J.A., Stensvold, C.R., 2016. Epidemiological aspects of *Blastocystis* colonization in children in Ilero, Nigeria. Am. J. Trop. Med. Hyg. 95, 175–179. https://doi.org/10.4269/ajtmh.16-0074.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959.

Ramirez, J.D., Sanchez, L.V., Bautista, D.C., Corredor, A.F., Florez, A.C., Stensvold, C.R., 2014. *Blastocystis* subtypes detected in humans and animals from Colombia. Infect. Genet. Evol. 22, 223–228. https://doi.org/10.1016/j.meegid.2013.07.020.

Rezaei Riabi, T., et al., 2018. Genetic diversity analysis of *Blastocystis* subtypes from both symptomatic and asymptomatic subjects using a barcoding region from the 18S rRNA gene. Infect. Genet. Evol. 61, 119–126. https://doi.org/10.1016/j.meegid.2018.03.026.

Rogers, A.R., Harpending, H., 1992. Population growth makes waves in the distribution of pairwise genetic differences. Mol. Biol. Evol. 9, 552–569. https://doi.org/10.1093/oxfordjournals.molbev.a040727.

Rojas-Velázquez, L., Morán, P., Serrano-Vázquez, A., Fernández, L.D., Pérez-Juárez, H., Poot-Hernández, A.C., Portillo, T., González, E., Hernández, E., Partida-Rodríguez, O., Nieves-Ramírez, M.E., Magaña, U., Torres, J., Eguiarte, L.E., Piñero, D., Ximénez, C., 2018. Genetic diversity and distribution of *Blastocystis* subtype 3 in human populations, with special reference to a rural population in central mexico. Biomed. Res. Int. 18 https://doi.org/10.1155/2018/3916263, 2018: 3916263.

Scicluna, S.M., Tawari, B., Clark, C.G., 2006. DNA barcoding of *Blastocystis*. Protist 157, 77–85. https://doi.org/10.1016/j.protis.2005.12.001.

Seeburg, P.H., Colby, W.W., Capon, D.J., Goeddel, D.V., Levinson, A.D., 1984. Biological properties of human c-Ha-ras1 genes mutated at codon 12. Nature 312, 71–75. https://doi.org/10.1038/312071a0.

Segui, R., et al., 2018. Prevalence of intestinal parasites, with emphasis on the molecular epidemiology of *Giardia duodenalis* and *Blastocystis* sp., in the Paranagua Bay, Brazil: a community survey. Parasit. Vectors 11, 490. https://doi.org/10.1186/s13071-018-3054-7.

Slatkin, M., Hudson, R.R., 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. Genetics 129, 555–562.

Stechmann, A., et al., 2008. Organelles in *Blastocystis* that blur the distinction between mitochondria and hydrogenosomes. Curr. Biol. 18, 580–585. https://doi.org/10.1016/j.cub.2008.03.037.

Stensvold, C.R., 2013. Comparison of sequencing (barcode region) and sequence-tagged-site PCR for *Blastocystis* subtyping. J. Clin. Microbiol. 51, 190–194. https://doi.org/10.1128/JCM.02541-12.

Stensvold, C.R., Clark, C.G., 2016. Current status of *Blastocystis*: a personal view. Parasitol. Int. 65, 763–771. https://doi.org/10.1016/j.parint.2016.05.015.

Stensvold, C.R., Clark, C.G., 2020. Pre-empting pandora's box: *Blastocystis* subtypes revisited. Trends Parasitol. https://doi.org/10.1016/j.pt.2019.12.009.

Stensvold, C.R., et al., 2007. Terminology for *Blastocystis* subtypes—a consensus. Trends Parasitol. 23, 93–96. https://doi.org/10.1016/j.pt.2007.01.004.

Stensvold, C.R., et al., 2009. Subtype distribution of *Blastocystis* isolates from synanthropic and zoo animals and identification of a new subtype. Int. J. Parasitol. 39, 473–479. https://doi.org/10.1016/j.ijpara.2008.07.006.

Stensvold, C.R., Alfellani, M., Clark, C.G., 2012. Levels of genetic diversity vary dramatically between *Blastocystis* subtypes. Infect. Genet. Evol. 12, 263–273. https://doi.org/10.1016/j.meegid.2011.11.002.

Su, H., Qu, L.J., He, K., Zhang, Z., Wang, J., Chen, Z., Gu, H., 2003. The great wall of China: a physical barrier to gene flow? Heredity 90, 212–219. https://doi.org/10.1038/sj.hdy.6800237.

Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123, 585–595.

Tan, K.S., 2008. New insights on classification, identification, and clinical relevance of *Blastocystis* spp. Clin. Microbiol. Rev. 21, 639–665. https://doi.org/10.1128/CMR.00022-08.

Valenca Barbosa, C., de Jesus, Batista R., Pereira Igreja, R., d'Avila Levy, C.M., Werneck de Macedo, H., Carneiro Santos, H.L., 2017. Distribution of *Blastocystis* subtypes isolated from humans from an urban community in Rio de Janeiro, Brazil. Parasit. Vectors 10, 518. https://doi.org/10.1186/s13071-017-2458-0.

Villamizar, X., et al., 2019. Molecular and descriptive epidemiology of intestinal protozoan parasites of children and their pets in Cauca, Colombia: a cross-sectional study. BMC Infect. Dis. 19, 190. https://doi.org/10.1186/s12879-019-3810-0.

Yoshikawa, H., Wu, Z., Howe, J., Hashimoto, T., Geok-Choo, N., Tan, K.S., 2007. Ultrastructural and phylogenetic studies on *Blastocystis* isolates from cockroaches. J. Eukaryot. Microbiol. 54, 33–37. https://doi.org/10.1111/j.1550-7408.2006.00141.x.