TITLE
Subtype analysis of Blastocystis isolates in Swedish patients.

AUTHORS
Joakim Forsell¹, Margareta Granlund¹*, C. Rune Stensvold², C. Graham Clark³, Birgitta Evengård⁴

¹Division of Clinical Bacteriology, Department of Clinical Microbiology, Umeå University, SE-901 87 Umeå, Sweden
²Department of Microbiological Diagnostics, Statens Serum Institut, Orestads Boulevard 5, DK-2300 Copenhagen S, Denmark
³Department of Pathogen Molecular Biology, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK.
⁴Division of Infectious diseases, Department of Clinical Microbiology, Umeå University, SE-901 87 Umeå, Sweden

*Corresponding author:
Margareta Granlund, Department of Clinical Microbiology, Umeå University, SE-901 87 Umeå, Sweden. Phone: +46 (0)90 785 17 72; Fax: +46 (0)90 785 35 91; Email: Margareta.Granlund@vll.se
ABSTRACT

*Blastocystis* is a genetically diverse and widespread intestinal parasite of animals and humans with controversial pathogenic potential. At least 9 subtypes of *Blastocystis* have been found in humans. Epidemiological studies of the relative distribution of the subtypes are emerging and may aid in the evaluation of the clinical significance of *Blastocystis*.

We here present the genetic diversity of *Blastocystis* examined in stool samples from 68 patients from the Stockholm area, Sweden. *Blastocystis* was identified by light microscopy, and subtyping was performed by sequencing the 5′-end of the small subunit ribosomal RNA gene. Five *Blastocystis* subtypes were identified in the 63 patients whose samples were successfully subtyped: ST1 (15.9%), ST2 (14.3%), ST3 (47.6%), ST4 (20.6%), and ST7 (1.6%). ST3 was more common in males compared to females (P= 0.049).

Comparative molecular analysis of *Blastocystis* sequences in this study revealed intra-subtype variations within the identified subtypes with the exception of ST4. Among ST4 sequences in this study, as well as in the majority of human GenBank sequences, a limited genetic diversity was found compared to the other common subtypes (ST1, ST2 and ST3). The relative prevalence of ST4 in this study was comparable to the overall distribution of ST4 in European cohorts (16.5%). This contrasts with the sparse reports of ST4 in studies from other continents, which may indicate that the distribution of this subtype is geographically heterogeneous.

KEYWORDS

*Blastocystis*, subtype, prevalence, Sweden, molecular epidemiology
INTRODUCTION

Blastocystis is an anaerobic unicellular eukaryote that can inhabit the intestinal tract of humans and many animals. It is the most commonly found non-fungal eukaryotic organism in human faecal samples [1-3]. The distribution of the parasite is worldwide and the prevalence in human faecal samples ranges from 7-20 % in developed countries to 30-60% in rural areas in developing countries [4-8]. Much is still unknown about Blastocystis. For instance, its life-cycle and pathogenic potential are still under debate [1,2,3,9,10]. Blastocystis is found in individuals both with and without symptoms. However, reports have been published of patients with gastrointestinal symptoms and Blastocystis as the only detected possible pathogen, and whose symptoms were relieved after successful treatment of Blastocystis [11-13]. Various gastrointestinal symptoms such as diarrhoea, abdominal pain, vomiting, constipation and flatulence have been linked to Blastocystis infection. As with other intestinal parasite infections, diarrhoeal episodes can alternate with normal defecation patterns or even constipation, conditions similar to chronic gastrointestinal illnesses such as irritable bowel syndrome (IBS). In fact, an increased prevalence of Blastocystis has been found in IBS patients [14-17]. Moreover, symptomatic Blastocystis infection has been reported to occur more often in immunocompromised adult patients than in controls [18] and has also been reported to be common in HIV-infected and immunocompromised children with gastrointestinal symptoms [11,19].

Molecular methods have shown extensive genetic variation among Blastocystis isolates [20,21]. A consensus approach has assigned Blastocystis isolates from humans, mammals and birds to one of nine subtypes, with sufficient genetic divergence between many of them to be classified as separate species [22].
Recently four additional subtypes have been identified but only in non-human hosts [23, 24]. This genetic diversity has supported the hypothesis that the variability in symptoms in patients positive for Blastocystis could be due to different pathogenic potential among the subtypes [25-31].

In a previous study investigating bacterial, viral and protozoan enteropathogens in Swedish patients with diarrhoea compared to healthy controls, Blastocystis was found in 4% of the patients and 9% of the controls [32] but no subtyping of Blastocystis was performed. We here present the first investigation of Blastocystis subtypes present in a Swedish collection of faecal samples from a diagnostic laboratory.

MATERIALS AND METHODS

Patient samples
Stool samples from patients in the Stockholm area were sent to the Laboratory of Diagnostic Parasitology of Karolinska University Hospital, Solna, Sweden, for routine ova and (oo)cyst examination. Notifications from clinicians on travel history and type of symptoms were only available in a minority of cases. Samples were sent fresh and Blastocystis was identified by light microscopy of iodine stained smears of faecal concentrates obtained by a modified formol-ether-concentration technique (FECT), utilising ethyl-acetate instead of ether [33]. Seventy-two Blastocystis positive samples were collected randomly during 2008 and the first few months of 2009. As a precautionary measure prior to DNA extraction all faecal samples were divided into two parts, one part was frozen...
without additives, and one part with approximately 1 ml faeces fixed in 4 ml 70% ethanol was stored at +4 °C, the latter according to a protocol for PCR examination for *Entamoeba* in faeces by Lebbad and Svärd [34].

**DNA extraction**

For each sample, DNA was extracted from approximately 200 mg stool using the QiaAMP DNA Stool Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer’s instructions, and eluted in 200 µl buffer AE (Qiagen). Frozen samples were not thawed before extraction. Samples fixed in ethanol were washed three times in phosphate-buffered saline (PBS) before DNA extraction.

**PCR**

The method described by Scicluna et al. [26] was used for subtype analysis. The method includes a conventional PCR with the primers RD5 (ATCGGTTTGGATCCTGCCAGT) and BhRD (GAGCTTTTTAACTGCAACAACG) amplifying an approximately 600 base pair (bp) fragment of the 1,800 bp small subunit ribosomal RNA gene (SSU-rDNA). This region of the SSU-rDNA has been shown to provide sufficient information for differentiating subtypes of *Blastocystis* [26]. One µl of extracted DNA was added to an amplification mixture of primers and BioMix™ (Bioline, London, UK) containing Taq DNA polymerase. The amplification profile consisted of 30 cycles of denaturation, annealing and extension at 94 °C, 59 °C and 72 °C (1 min for each step), with a final extension step of 2 min at 72 °C. Primarily, DNA extracted from frozen stool samples was used for the PCR. Where no PCR product was obtained, DNA was extracted from the ethanol fixed samples. Positive PCR products were purified using GeneJET™ Genomic DNA
Purification Kit (Fermentas, London, UK) according to the manufacturer’s recommendations.

**Sequencing and sequence analysis**

Purified PCR products were sequenced using the primer BhRDr, ABI BigDye sequencing kit version 3.1 (Applied Biosystems, Warrington, UK) and an ABI 3730 sequencer. The resulting chromatograms were analysed and edited in the computer software Chromas version 2.33 (Technelysium Pty. Ltd., Australia). The sequences obtained were compared to *Blastocystis* sequences in GenBank by BLAST searches at the National Center for Biotechnology Information (NCBI) [35]. Subtypes were identified by determining the exact match or closest similarity according to the classification by Stensvold et al. [22]. Sequences were aligned with SSU-rDNA sequences representing ST1-9 from GenBank using the software ClustalW [36] and MEGA version 4 [37] for comparative analysis of the sequence data.

Sequences obtained in the study are available as accession numbers JN003684-JN003686 and JN587541-JN587549 in the GenBank database.

**Phylogenetic analysis of Blastocystis**

To infer phylogenetic relationships among the sequences, phylogenetic trees were constructed with the neighbor-joining (NJ) method [38] using the software MEGA version 4 [37]. The flagellate *Proteromonas lacertae* (accession number U37108) was used as the outgroup as in previous phylogenetic studies [40,41]. Bootstrap confidence values for the branching reliability were calculated with 1,000 replicates.
Statistical analysis

For comparison of subtype prevalence of different groups the $\chi^2$ test (with collapsed categories when needed) was used. A P-value less than 0.05 was considered statistically significant. The software PASW version 18 (SPSS Inc, Chicago, IL, USA) was used for statistical analysis.

RESULTS

*Blastocystis* was identified by light microscopy in 72 faecal samples from 68 patients, 30 females and 38 males. The age of the patients ranged from 1 to 70 y, the median age was 30.5 y (interquartile range 16-46) in females and 38.5 y (interquartile range 26-49) in males. The age distribution is displayed in Fig.1. In 63 of the 68 patients the expected 600 bp fragment of the SSU-rDNA was amplified and successfully sequenced. In 4 of the 68 patient samples no PCR product was produced despite repeated DNA extractions and numerous attempts towards optimising PCR conditions. In one sample a PCR product of the correct size was obtained but the resulting sequence chromatogram showed double peaks, which may indicate a mixed infection of two or more *Blastocystis* subtypes.

Five *Blastocystis* subtypes were identified in the study, ST1, ST2, ST3, ST4, and ST7 (Fig. 2). ST3 was the most prevalent subtype, found in 30/63 patients (47.6%) followed by ST4 in 13/63 patients (20.6%). One patient harboured *Blastocystis* ST7, a subtype that has been found mostly in birds [42]. For comparison, *Blastocystis* subtype distributions found in previous European studies are displayed in Table 1.
There was a difference between men and women in the prevalence of ST3 and ST4 (Fig 3). Although supported by a marginal significance, ST3 was more common in males compared to females (P=0.049, collapsed categories). Among females ST4 was almost as common as ST3 but the difference in the relative frequency of ST4 among men and women was not statistically significant (P=0.13) (Fig. 3).

As previously reported [26] the sequence variation in the 5’-end of the SSU-rRNA gene of Blastocystis is extensive. The isolate BL051 (accession number JN003686), assigned to ST7, differs from its closest match in GenBank (DQ232821) by 7 nucleotides in the studied region of SSU-rDNA. However this level of divergence falls within the range of variation identified among previously published ST7 sequences (Fig 2). A similar level of diversity was seen among samples assigned to STs 1-3 identified in this study; ST1: up to 2 differences; ST2: up to 6 differences; ST3: up to 6 differences). In contrast, no variation was found among the thirteen identified ST4 sequences. A BLAST search revealed that the ST4 sequences of this study were identical to most published ST4 SSU-rRNA genes, among them 17 from human isolates deposited by Scicluna et al. [26], three human and four rat ST4 sequences from France [43,21] and three sequences from humans in Japan and Germany [44]. Single nucleotide polymorphisms (SNPs) differentiated these sequences from a group of five previously published ST4 sequences from rodents that all shared at least three of the SNPs [41,45]. The sequence diversity leads to the generation of two clades of ST4 in the phylogenetic tree (Fig. 2).
DISCUSSION

During the last few years several studies of the prevalence of *Blastocystis* subtypes have been published. This study is, to our knowledge, the first that describes the occurrence of different *Blastocystis* subtypes in Sweden. ST3 was the most common *Blastocystis* subtype found in the Stockholm area. The distribution of *Blastocystis* subtypes in human faecal samples varies from country to country. However, ST3 has dominated as the most prevalent subtype in most studies although the figures differ depending on populations studied [43].

ST4 was the second most prevalent subtype in our study, being found in 20.6% of the samples. This is in contrast to an average prevalence figure of 5.4% reported by Souppart et al. in 2009 in a summary of *Blastocystis* subtypes from different countries [43]. In most studies ST4 was the fourth most common subtype found in humans, after ST3, ST1 and ST2. However, when only studies from Europe are included in the comparison, and with additional data from the two last years included, the average prevalence for ST4 increases to 16.5% (Table 1). Using this figure, our findings are generally in concordance with the results from the rest of Europe. Since very few isolates of ST4 have been identified in studies from Asia, among them a large study of *Blastocystis* prevalence in China by Li et al. [8], the results may also indicate geographic differences in the prevalence of ST4. Interestingly, studies from Turkey, a country situated both in Europe and Asia, also reports very low numbers of ST4 (1/337 samples) (Table 1). It should be kept in mind, however, that differences in study populations and methodology used may influence the results.
Blastocystis carriers in the present study were diagnosed primarily by light microscopy. Routine examination using FECT is considered to have a low sensitivity of approximately 30-50% for the detection of Blastocystis [4,46]. The sensitivity has been shown to be lower for ST3 compared to the other subtypes [46]. Our selection method could therefore have underestimated the prevalence of ST3. The faecal samples in the present study were all from a diagnostic parasitology laboratory suggesting that the majority of the samples arise from patients with gastrointestinal symptoms. No PCR product could be amplified in the samples from four of the patients in this study. Possible reasons for this could be PCR inhibition, degradation of DNA due to long storage time or misidentification by light microscopy.

A difference in the occurrence of ST3 and ST4 between men and women was found. The higher relative frequency of ST3 in men was statistically significant, although the sample size in each subtype group was small. Differences in the relative proportions of subtypes between the sexes could indicate distinct pathways of transmission or exposure to the organism. Further studies are needed to confirm the gender difference in the distribution of subtypes of Blastocystis.

Overall the amplified fragment of Blastocystis SSU-rDNA exhibited extensive sequence variability. Although a limited number of Blastocystis were examined, variability was found within subtypes. This further supports the suggestion that the diversity in Blastocystis STs may qualify them to be classified as separate species, since the ribosomal RNA genes in most organisms are generally considered as too conserved to discriminate between subspecies variants. A4 The recent genome sequence of a Blastocystis ST7 strain [38] showed that some minor
sequence variation does exist among individual copies of the SSU-rRNA gene in the genome. However, intra-strain variation does not interfere with assignment of Blastocystis samples to subtypes as the sequence obtained from the PCR reaction represents only the dominant sequence in the sample. The assignment of Blastocystis isolates to subtypes generated by the SSU-rRNA gene may be insufficient to clarify questions about pathogenic potential and transmission patterns. A further subdivision of the subtypes with molecular techniques that have a higher discriminatory capacity, for example multilocus sequence typing (MLST), will possibly aid in investigations of the epidemiology of Blastocystis and possible associations between subtypes and pathogenicity. Aspects of zoonotic potential and transmission patterns may also be revealed by better discrimination among the subtypes. MLST has been used in many bacterial species but has recently also been used for subdivision of assemblages in the intestinal parasitic protozoan Giardia intestinalis [48,49].

An interesting observation was the low genetic variability found in the ST4 sequences compared to what was seen among the other common subtypes identified in the study. In a BLAST search these Blastocystis sequences were identical to human ST4 sequences from different parts of the world in addition to four Blastocystis sequences from rodent faeces with origins in France [21]. Another group of rodent isolates with at least three SNPs in common constituted another lineage among the ST4 sequences that was revealed in the phylogenetic tree (Fig. 2). A similar branching among ST4 sequences has been shown by Noël et al. [21] when examining mostly rodent isolates, and Stensvold et al. [31]. If the data indicating a low variability among human ST4 sequences can be confirmed by examination of other genes this may indicate that the ST4 lineage has
expanded more recently on an evolutionary scale. In epidemiological studies a low grade of variability within a subgroup of isolates can also be explained by a selection of a specific lineage caused by advantageous properties for colonization or infection [50,51]. Since the epidemiological context of the sequences compared in this study is unknown no such conclusions can be drawn. However, in relation to this hypothesis it is interesting that 76% of *Blastocystis* positive patients with acute diarrhoea harboured ST4 isolates in a recently published Danish study [31], indicating that ST4 may have a pathogenic potential.

In summary, *Blastocystis* of ST1, ST2, ST3, ST4, and ST7 were identified in 63 Swedish patient faecal samples with a dominance of ST3 (47.6%), followed by ST4 (20.6%). Intra-subtype sequence variations were identified in all the subtypes with the exception of ST4, which showed low variability in the 5’-end of SSU-rRNA. A comparison with recent studies in Europe indicates that the relative prevalence of ST4 in Europe may be higher than in other geographic regions.

**Acknowledgements**

The authors acknowledge the work involved in collecting the clinical samples and are thankful to Lillemor Karlsson, Christina Seeth Grönfors, Christine Stenström and Anna Dyrsmeds, all at Laboratory of Parasitology, Department of Clinical Microbiology, Karolinska University Hospital Solna, Stockholm, Sweden.
References


*Blastocystis* sp. subtype 4 is common in Danish Blastocystis-positive patients presenting with acute diarrhea. Am J Trop Med Hyg 84:883-885


Genome sequence of the stramenopile Blastocystis, a human anaerobic parasite. Genome Biol 12:R29


Table 1. *Blastocystis* subtype distribution in different European countries.

<table>
<thead>
<tr>
<th>Country [ref]</th>
<th>Method of subtyping b</th>
<th>No. of samples</th>
<th>Subtype distribution (%) c</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>ST1</td>
</tr>
<tr>
<td>Denmark [52]</td>
<td>Sequencing</td>
<td>29</td>
<td>3.4</td>
</tr>
<tr>
<td>Denmark [46]</td>
<td>Sequencing</td>
<td>28</td>
<td>17.9</td>
</tr>
<tr>
<td>Denmark [17]</td>
<td>Sequencing</td>
<td>116</td>
<td>26.7</td>
</tr>
<tr>
<td>Denmark [4]</td>
<td>Sequencing</td>
<td>99</td>
<td>20.2</td>
</tr>
<tr>
<td>Denmark [53]</td>
<td>Sequencing</td>
<td>22</td>
<td>50.0</td>
</tr>
<tr>
<td>Denmark [31]</td>
<td>Sequencing</td>
<td>25</td>
<td>4.0</td>
</tr>
<tr>
<td>France [43]</td>
<td>Sequencing</td>
<td>40</td>
<td>27.5</td>
</tr>
<tr>
<td>France [54]</td>
<td>Sequencing</td>
<td>27</td>
<td>3.7</td>
</tr>
<tr>
<td>Germany [55]</td>
<td>RFLP</td>
<td>166</td>
<td>24.1</td>
</tr>
<tr>
<td>Germany [44]</td>
<td>STS primers</td>
<td>12</td>
<td>25.0</td>
</tr>
<tr>
<td>Greece [56]</td>
<td>SSCP</td>
<td>45</td>
<td>20.0</td>
</tr>
<tr>
<td>Ireland [57]</td>
<td>Sequencing</td>
<td>14</td>
<td>7.1</td>
</tr>
<tr>
<td>Italy [58]</td>
<td>Sequencing</td>
<td>30</td>
<td>10.0</td>
</tr>
<tr>
<td>Spain [30]</td>
<td>RFLP</td>
<td>51</td>
<td>2.0</td>
</tr>
<tr>
<td>Sweden a</td>
<td>Sequencing</td>
<td>63</td>
<td>15.9</td>
</tr>
<tr>
<td>Turkey [28]</td>
<td>STS primers</td>
<td>92</td>
<td>20.6</td>
</tr>
<tr>
<td>Turkey [59]</td>
<td>Sequencing</td>
<td>87</td>
<td>9.2</td>
</tr>
<tr>
<td>Turkey [60]</td>
<td>STS primers</td>
<td>32</td>
<td>62.5</td>
</tr>
<tr>
<td>Turkey [29]</td>
<td>STS primers</td>
<td>35</td>
<td>8.6</td>
</tr>
<tr>
<td>Turkey [61]</td>
<td>STS primers</td>
<td>66</td>
<td>28.8</td>
</tr>
<tr>
<td>Turkey [62]</td>
<td>STS primers</td>
<td>25</td>
<td>36.0</td>
</tr>
<tr>
<td>UK [20]</td>
<td>RFLP</td>
<td>29</td>
<td>6.9</td>
</tr>
<tr>
<td>UK [26]</td>
<td>Sequencing</td>
<td>55</td>
<td>5.5</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>1188</td>
<td>19.4</td>
</tr>
</tbody>
</table>

a Present study  
b Abbreviations used: RFLP = restriction fragment length polymorphism of SSU-rRNA genes, STS primers = sequence-tagged site primers (specific for each subtype 1-7), SSCP = single strand conformational polymorphism  
c The percentages can add up to less or more than 100 in each study depending on if the *Blastocystis* subtypes involved in mixed infections were resolved
Fig 1.
Fig 2.
Fig. 3

<table>
<thead>
<tr>
<th></th>
<th>ST1</th>
<th>ST2</th>
<th>ST3</th>
<th>ST4</th>
<th>ST7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>13.9%</td>
<td>13.9%</td>
<td>56.3%</td>
<td>13.9%</td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>18.5%</td>
<td>14.8%</td>
<td>33.3%</td>
<td>29.6%</td>
<td>3.7%</td>
</tr>
</tbody>
</table>

p = 0.049

p = 0.13
Figure legends

**Fig 1.** The age distribution of patients positive for *Blastocystis* by light microscopy.

**Fig 2.** A phylogenetic tree comparing representative *Blastocystis* SSU-rDNA sequences generated in this study, designated ‘BL0--’, and reference sequences from GenBank, shown with accession numbers. The tree is inferred using the neighbor-joining method (with the maximum composite likelihood model) based on a hypervariable region at the 5’ end of the SSU-rRNA gene. *Proteromonas lacertae* (accession number U37108) is used as the outgroup. Bootstrap values (%) are indicated at the internal nodes (1,000 replicates). Bootstrap values of less than 50% are not shown. The subtype prevalence in the 63 samples is shown to the right as percentages. Italicised subtypes were not found in the present study.

**Fig 3.** Distribution of *Blastocystis* subtypes in males and females.