

## ORIGINAL ARTICLE

# Micro-eukaryotic diversity of the human distal gut microbiota: qualitative assessment using culture-dependent and -independent analysis of faeces

Pauline D Scanlan<sup>1</sup> and Julian R Marchesi<sup>1,2</sup>

<sup>1</sup>Alimentary Pharmabiotic Centre, Department of Microbiology, University College Cork, Cork, Ireland and

<sup>2</sup>School of Biosciences, Cardiff University, Cardiff, UK

Molecular ecological surveys of the human gut microbiota to date have focused on the prokaryotic fraction of the community and have revealed a remarkable degree of bacterial diversity and functionality. However, there is a dearth of information on the eukaryotic composition of the microbiota, and no culture-independent sequence-based surveys of human faeces are available. Culture-independent analyses based on DNA extraction and polymerase chain reaction targeting both the total eukaryotic 18S rRNA genes and fungal internal transcribed regions (ITS), together with culture-dependent analyses of fungi, were performed on a group of healthy volunteers. Temporal analysis was also included wherever possible. Collectively, the data presented in this study indicate that eukaryotic diversity of the human gut is low, largely temporally stable and predominated by different subtypes of *Blastocystis*. Specific analyses of the fungal populations indicate that a disparity exists between the cultivable fraction, which is dominated by *Candida* sp, and culture-independent analysis, where sequences identical to members of the genera *Gloeotinia/Paecilomyces* and *Galactomyces* were most frequently retrieved from both fungal ITS profiles and subsequent clone libraries. Collectively, these results highlight the presence of unprecedented intestinal eukaryotic inhabitants whose functional roles are as yet unknown in healthy individuals. Furthermore, differences between results obtained from traditionally employed culture-based methods and those obtained from culture-independent techniques highlight similar anomalies to that encountered when first analysing the bacterial diversity of the human faecal microbiota using culture-independent surveys.

The ISME Journal (2008) 2, 1183–1193; doi:10.1038/ismej.2008.76; published online 31 July 2008

**Subject Category:** microbe-microbe and microbe-host interactions

**Keywords:** gut microbiota; micro-eukaryotes; biodiversity; 18S rRNA analysis; DGGE; intergenic spacer analysis; *Blastocystis*

## Introduction

The micro-ecology of the human gastrointestinal (GI) tract is a burgeoning area of scientific research with significant commercial and clinical interest. The microbial community comprises a critical part of GI functioning and plays a key role in host health and disease (Blaut and Clavel, 2007; Marchesi and Shanahan, 2007) through their roles in the development and maintenance of host physiology, immu-

nology and also host nutrition (Cash *et al.*, 2006; Kelly *et al.*, 2007). It is now apparent from the application of culture-independent techniques, including 16S rRNA gene surveys and metagenomics, that the vast majority of bacterial species present are yet to be cultivated (Suau *et al.*, 1999; Eckburg *et al.*, 2005). Furthermore, in addressing the microbiota of the human gut as a total community, a relative wealth of both qualitative and quantitative information is available on diverse prokaryotic populations compared with that of the eukaryotic fraction of the community (Suau *et al.*, 1999; Eckburg *et al.*, 2005).

Micro-eukaryotic species are ubiquitous components of various natural ecosystems, including hydrothermal environments (Gadanhó *et al.*, 2003; Lopez-Garcia *et al.*, 2007), soil (van Elsas *et al.*, 2000; Moon-van der Staay *et al.*, 2006), the rumen

Correspondence: JR Marchesi, School of Biosciences, Cardiff University, Museum Avenue, Cardiff CF10 3US, UK.

E-mail: marchesijr@cardiff.ac.uk and

PD Scanlan, Alimentary Pharmabiotic Centre, Department of Microbiology, University College Cork, Cork, Ireland.

E-mail: paulinescanlan@yahoo.co.uk

Received 21 May 2008; revised 27 June 2008; accepted 1 July 2008; published online 31 July 2008

(Ranilla *et al.*, 2007) and food and beverage communities (Romano *et al.*, 2003; Fleet, 2007; Quiros *et al.*, 2008). In these environments, eukaryotes interact with other microbes to form complex communities that play primary roles in both ecosystem structure and function (Fleet, 1999, 2003; Lopez-Garcia *et al.*, 2007). The first culture-independent analysis of the fungal community populating a mammalian intestine revealed an unprecedented diversity and abundance of fungal species in the murine GI tract (Scupham *et al.*, 2006). 18S rRNA gene analysis identified the presence of fungal species in the murine gut belonging to nine different taxonomic clusters, including *Ascomycota*, *Basidiomycota*, *Mucorales* and *Scleroderma*, and quantitative analysis indicated that fungi account for approximately 2% of the total community present in caecal biofilms. The findings of this study are of great interest to human microbial ecology from both diversity and functional perspectives and provide further impetus for the work presented in this paper. In addition, mouse models are routinely applied to address questions pertaining to human nutrition, health and drug metabolism, and, therefore, any similarities or differences that exist between the human and murine gut are of significant interest to GI research.

Many culture-dependent studies on the eukaryotes inhabiting different human niches have readily isolated yeasts, such as *Candida* sp, from the mouth, fingernail, toenail and rectum of healthy hosts (Kam and Xu, 2002; Song *et al.*, 2005). However, culture-independent surveys of eukaryotic populations of the human GI tract are negligible and are perhaps limited to one publication with no sequence data reported (Kuhbacher *et al.*, 2006). As such, a requirement to characterize this community is apparent, and the primary aim of this study was to analyse the micro-eukaryotic inhabitants of the human intestinal tract to gain an insight into the diversity of eukaryotic species in the intestinal tracts of a cohort of healthy individuals.

## Materials and methods

### Sample collection

Faecal samples were collected from 17 healthy individuals and frozen immediately at  $-80^{\circ}\text{C}$ . Faecal samples are largely representative of the microbial community in the human distal colon. Eight individuals were sampled once and nine individuals were sampled over varying time points from 8 weeks to 3 years. An overview of the participants is provided in Table 1. A broad range of age groups was analysed, and individuals representing each decade of the age groups from 20 to 60 years were sampled along with a comparable female-to-male ratio (10:7). Fifteen individuals were born and raised in Ireland, and two individuals M2 and M3 were born and raised in the United Kingdom. No individuals analysed in this study had

travelled extensively or lived in a third world or developing country. In addition, for individuals who were sampled temporally, the samples obtained were evenly distributed over the time period studied.

### DNA extraction from human faecal samples

Three different DNA extraction methods were tested to ensure complete extraction of all representative DNAs and minimize bias (see Supplementary Table S1). DNA extractions were performed using Method A.

**Method A.** Total DNA was extracted from faecal samples using a combination of physical and chemical lyses, followed by consecutive precipitation and cleaning steps to remove any polymerase chain reaction (PCR) inhibitors for subsequent downstream application. A total of 250 mg of faecal sample was placed in a 2 ml tube containing a 200 mg mixture of 0.1, 0.5 and 2 mm zirconium beads, 500  $\mu\text{l}$  of 10 mM phosphate buffer (pH 7.4), 500  $\mu\text{l}$  (10% w/v) sodium dodecyl sulphate and  $\sim 400$   $\mu\text{l}$  phenol: chloroform (1:1). Alternate bead beating at 3200 r.p.m. per 60 s and cooling on ice for 10 min were carried out three times, with a final centrifugation of the lysed suspension for 10 min at 16 000 g. Three additional phenol:chloroform extractions were performed and DNA was then collected by centrifugation following precipitation on ice using 0.7 volumes of isopropanol. DNA was resuspended in 10 mM phosphate buffer and further cleaned using 3M sodium acetate and ethanol precipitation on ice. DNA was collected by centrifugation for 10 min at 16 000 g and the nucleic acid pellet was washed with 1 ml of ice-cold 70% ethanol and then air-dried for 5 min before resuspending it in 200  $\mu\text{l}$  of 10 mM phosphate buffer (pH 7.4). Prior to PCR and to increase the efficiency of amplification, 100  $\mu\text{l}$  of DNA was further cleaned using saturated ammonium acetate and chloroform followed by ethanol precipitation and subsequent centrifugation to collect the nucleic acid pellet. DNA was resuspended in 50  $\mu\text{l}$  of distilled water, added to 150  $\mu\text{l}$  of Chelex (10% w/v distilled water) and heat-treated for 5 min at  $37^{\circ}\text{C}$  and 5 min at  $95^{\circ}\text{C}$ , and then cooled to  $4^{\circ}\text{C}$ ; the supernatant was removed and stored at  $-20^{\circ}\text{C}$  until further use. DNA was visually assessed on 1% agarose containing ethidium bromide (0.25 mg ml $^{-1}$ ) and  $0.5 \times$  TAE buffer (Sambrook *et al.*, 2000) with an applied voltage of 5 V cm $^{-1}$ . DNA was visualized by UV illumination (302 nm), and DNA concentration was determined using the Nanodrop spectrophotometer.

### PCR of 18S rRNA genes and fungal internal transcribed regions

Polymerase chain reaction mixtures (50  $\mu\text{l}$ ) contained  $1 \times$  buffer (20 mM Tris (pH 8.4), 50 mM KCl), 3 mM MgCl $_2$ , 200  $\mu\text{M}$  of each dNTP, 1.25 U of *Taq*

**Table 1** Overview of participants

Subject code	Age	Sex	Medical history	No. of samples obtained and time period over which samples were collected
M1	54	M	Treated for diverticulitis	N=3, 12 weeks
M2	25	M	None	N=3, 12 weeks
M3	40	M	None, history of intermittent antibiotic usage	N=3, 12 weeks
M4	26	M	None	Once
M5	26	M	None	N=6, 2 years
M6	62	M	None	Once
M7	59	M	None	Once
F1	38	F	None	N=6, 3 years
F2	28	F	None	N=6, 2 years
F3	54	F	None	N=3, 12 weeks
F4	27	F	None	N=4, 3 years
F5	28	F	None	Once
F6	23	F	None	N=3, 6 months
F7	61	F	None	Once
F8	64	F	None	Once
F9	63	F	None	Once
F10	65	F	None	Once

Abbreviations: F, female; M, male.

polymerase (Invitrogen, Paisley, UK) and 10 pmol of each primer. Bovine serum albumin and appropriately diluted DNA (80–100 ng) was added to the final PCR. The PCR program for primers Euk1a and 516r and 516r-GC (Supplementary Table S1) consisted of an initial denaturation at 94 °C for 5 min, followed by 35 cycles of denaturation at 94 °C for 30 s, annealing at 55 °C for 40 s, extension at 72 °C for 60 s and a final extension of 5 min at 72 °C. 18S rRNA gene PCR products were obtained from all individuals in this study using the primers Euk1a and 516r. The PCR program for primers ITS1F and ITS4R (Supplementary Table S1) consisted of an initial denaturation at 95 °C for 5 min, followed by 35 cycles of denaturation at 94 °C for 30 s, annealing at 50 °C for 30 s, extension at 72 °C for 60 s and a final extension of 5 min at 72 °C. Fungal internal transcribed region (ITS) PCR products were generated from 14 of the 17 individuals in this study using the primer pair ITS1F and ITS4R.

#### Denaturing gradient gel electrophoretic analysis and band excision

18S rRNA gene PCR products were separated according to the specifications of Muyzer *et al.* (1993) using the DCODE system (Bio-Rad Laboratories, Hemel Hempstead, UK) with the following modifications: polyacrylamide gels (dimensions: 200 × 200 × 1 mm<sup>3</sup>) consisted of 8% v/v polyacrylamide (37.5:1 acrylamide/bisacrylamide) and 0.5 × TAE. The final denaturant gradient for primer pair Euk1A and 516r-GC was optimized to 5–35%. Prior to polymerization of the denaturing gel, a stacking gel without denaturants was added. Electrophoresis was performed for 16 h at 85 V in 0.5 × TAE buffer at a constant temperature of 60 °C. Gels were stained with SybrGold according to the manufacturer's

instructions. All bands present on 18S rRNA gene denaturing gradient gel electrophoretic profiles were excised using a gel spot extractor pipette and DNA was eluted in PCR grade water for 30 min at room temperature. A volume of 1 µl was used as template DNA in PCR, as previously described, using primers Euk1A and 516r-GC. The resulting product was resolved on a gel with the original sample to identify and ensure the correct band selection for sequencing. PCR products were cleaned using the Qiagen QIAquick PCR cleanup kit according to the manufacturer's instructions prior to sequencing.

#### Fungal internal transcribed spacer analysis

Fungal intergenic spacer PCR products were separated using 4% polyacrylamide gels (dimensions: 160 × 215 × 1 mm<sup>3</sup>) and electrophoresis was performed for 4 h at a constant voltage of 180 V. All gels were stained with SybrGold (Molecular Probes, Paisley, UK) according to the manufacturer's instructions. Bands of interest were excised and analysed as previously outlined for DGGE band excision.

#### Clone library construction and analysis

For the construction of clone libraries, PCR cycle numbers were reduced to 25 to minimize any bias imposed by the C<sub>0</sub>T effect (Mathieu-Daude *et al.*, 1996) or preferential DNA amplification. For each clone library, four PCR reactions on the same sample were pooled and cleaned using the Qiagen PCR cleanup kit prior to clone library construction. PCR products were transformed into *Escherichia coli* cells using TOPO XL PCR cloning kit (Invitrogen, Dublin, Ireland) according to the manufacturer's

instructions. 18S rRNA gene clone libraries were constructed from PCR products generated using the primer set Euk1a and 516r as per the protocol previously outlined for PCR of 18S rRNA genes. A total of 11 libraries were constructed and 36 randomly selected clones from each 18S rRNA gene library were selected for sequencing. Fungal intergenic spacer clone libraries were constructed using PCR products generated using the primer set ITS1F and ITS4R. A total of 11 libraries were generated and, owing to natural size and sequence variation of fungal ribosomal intergenic spacer regions, 48 clones from each library were picked, re-amplified by PCR and the products were size checked by PCR on 1.5% agarose gels prior to plasmid preparation (Qiagen mini prep kit) and sequencing.

#### *Bionumerix gel analysis*

To extrapolate meaningful data from the DGGE and fungal internal transcribed spacer analysis gel profiles, the Gel Compar function of the Bio Numerics software programme (Applied Maths, St. Martens-Latem, Belgium) was used to calculate similarities between samples from individuals over time (intra-individual similarity/individual temporal stability) and between samples from different individuals (inter-individual similarity). The similarities were determined by calculating similarity indices based on the DICE similarity coefficient and the unweighted pair-group method using arithmetic averages (UPGMA). The DICE coefficient is also referred to as Sorenson's pair-wise similarity coefficient ( $C_s$ ) and is commonly used to compare species composition of different ecosystems.

#### *Fungal culturing and identification*

Faecal samples were serially diluted and fivefold dilutions were spread-plated in triplicate on Sabouraud dextrose agar, potato dextrose agar and malt agar supplemented with chloramphenicol ( $40 \mu\text{g ml}^{-1}$ ) and kanamycin ( $50 \mu\text{g ml}^{-1}$ ). The plates were incubated both aerobically and anaerobically for 48 h at  $37^\circ\text{C}$  and counted. Random colonies and any colonies exhibiting different morphologies were restreaked to obtain pure cultures. DNA extraction from pure fungal cultures was performed as per the protocol described by Gadanho *et al.* (2003). In brief, 200 mg of overnight spread-plate cultures was suspended in a 2 ml tube containing 200  $\mu\text{l}$  volume of 500  $\mu\text{m}$  of glass beads and 500  $\mu\text{l}$  of lysing buffer (50 mM Tris, 250 mM NaCl, 50 mM EDTA, 0.3% w/v sodium dodecyl sulphate (pH 8)). After a 2 min bead beat step, samples were incubated at  $65^\circ\text{C}$  for 1 h. Suspensions were then centrifuged at 16 000 g for 10 min. The collected supernatant was diluted to 1:100 and 2  $\mu\text{l}$  used in a 50  $\mu\text{l}$  PCR, as outlined for both primer sets Euk1a and 516r and ITS1 and ITS4. The PCR products were cleaned as previously outlined prior to sequencing. Ninety-five PCR

products from pure culture isolates were generated for analysis.

#### *Bioinformatic sequence analysis and phylogenetic tree construction*

All sequences from pure culture PCR products, DGGE and ITS extracted bands and clone libraries were analysed using the tBLASTn (Altschul *et al.*, 1990) functions for nucleotide identification. Phylogenetic trees were constructed with unambiguously aligned nucleotide sequences using the MEGA4 software package (Tamura *et al.*, 2007). The evolutionary history was inferred using the neighbour-joining method (Saitou and Nei, 1987) from nonredundant 18S rRNA gene sequences obtained from DGGE band excision and clone libraries. The bootstrap consensus tree inferred from 500 replicates is taken to represent the evolutionary history of the taxa analysed (Archie and Felsenstein, 1993). The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) is shown next to the branches. The tree is drawn to scale, with branch lengths expressed in the same units as the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the maximum composite likelihood method (Tamura *et al.*, 2004) and are expressed as the number of base substitutions per site. All positions containing gaps and missing data were eliminated from the data set ('Complete deletion' option). Owing to partial sequence reads obtained from DGGE band excision sequencing, 394 nucleotides were included in the final data set.

#### *Accession numbers*

Accession numbers for 18S rRNA genes and fungal ITS sequences retrieved in this study are available from EMBL database numbers AM992460–AM992473 and AM998784–AM998792, respectively.

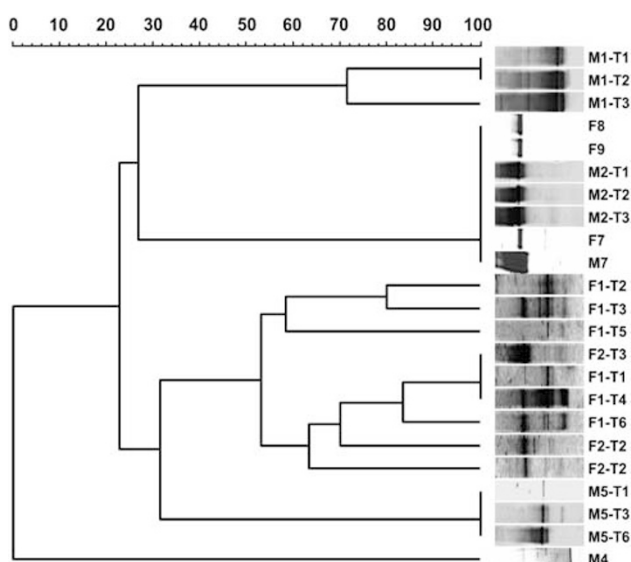
## **Results and discussion**

A paucity of information exists on the incidence and diversity of eukaryotes in the GI tract of humans, and the primary aim of this paper was to outline suitable culture-independent methodologies that address this concern. As such, appropriate DNA extraction and PCR methods were optimized (see Supplementary Table S1) and the composition of the micro-eukaryotic fraction of the human gut microbiota was determined using a three-tiered experimental design; DNA extraction followed by (1) PCR amplification and analyses of 18S rRNA genes to survey total eukaryotes, (2) PCR amplification and analyses of fungal ITS regions to assess the fungal community and finally (3) determination of the congruency of results between cultivation-independent and -dependent techniques through the comparative analysis of fungal 18S rRNA genes and ITS regions.

*Blastocystis* is a common and stable feature of the human healthy distal colon

Denaturing gradient gel electrophoretic analysis of 18S rRNA gene PCR products indicates that the diversity of eukaryotes in the human intestinal tract is low (between one and three bands per individual) and relatively stable over time (see Figure 1). Software analysis of generated DGGE profiles shows inter-individual variation and similarity, and variability with respect to intra-individual temporal stability (see Figure 1). Band excision and analysis, supported by clone library analysis, revealed that *Blastocystis* sp were the most common eukaryotes recovered (14/17 individuals analysed), with fungal sequences infrequently recovered as well. Although *Blastocystis* was the dominant eukaryote common to all individuals analysed, with three exceptions, comparative DGGE gel and sequence analysis, followed by phylogenetic tree construction, indicated that *Blastocystis* sp present between individuals were widely diversified (see Figure 2). Furthermore, temporal analysis indicated that when present in DGGE profiles, *Blastocystis* sp were the dominant and constant feature of all time points analysed.

*Blastocystis* sp are obligate anaerobes of controversial epidemiological status and are phylogenetically placed within the Stramenopiles (Silberman *et al.*, 1996), a heterogeneous taxonomic group that includes uni- and multicellular protists, such as brown algae and diatoms (Patterson, 1994). Although these are routinely cited as a common human intestinal parasite or pathogen and have been suggested as an aetiological agent in many GI illness, including diarrhoea, inflammatory bowel disease and irritable bowel syndrome (Giacometti *et al.*, 1999; Yakoob *et al.*, 2004; Stark *et al.*, 2007), it is believed that the potential pathogenicity of *Blastocystis* is subtype-dependent (Kaneda *et al.*, 2001; Tan *et al.*, 2006). Genetic and polymorphic variations of *Blastocystis* are well reported in the literature and classification schemes based on 18S rRNA genes and elongation factor genes have been used to infer phylogenetic relationships and discrete lineages (Ho *et al.*, 2000; Abe, 2004; Yoshikawa *et al.*, 2004), and the phylogenetic tree constructed from *Blastocystis* sequences obtained from different individuals analysed in this study falls into distinct phylogenetic clades or subtypes that have been previously defined (Arisue *et al.*, 2003; Stensvold *et al.*, 2007a). The majority of individuals harboured *Blastocystis* sp that cluster with subtypes II, III and IV, and interestingly no sequences representing *Blastocystis* sp from subtypes V, VI and VII were obtained. An interesting observation from both DGGE and phylogenetic tree analyses shows that four out of the six individuals analysed from the  $\geq 60$  years age group (F7, F8, F9 and M7) clustered together in DGGE profiles and with subtype III species, whereas the two other members from this age group clustered with subtype II (M6 and F10). In

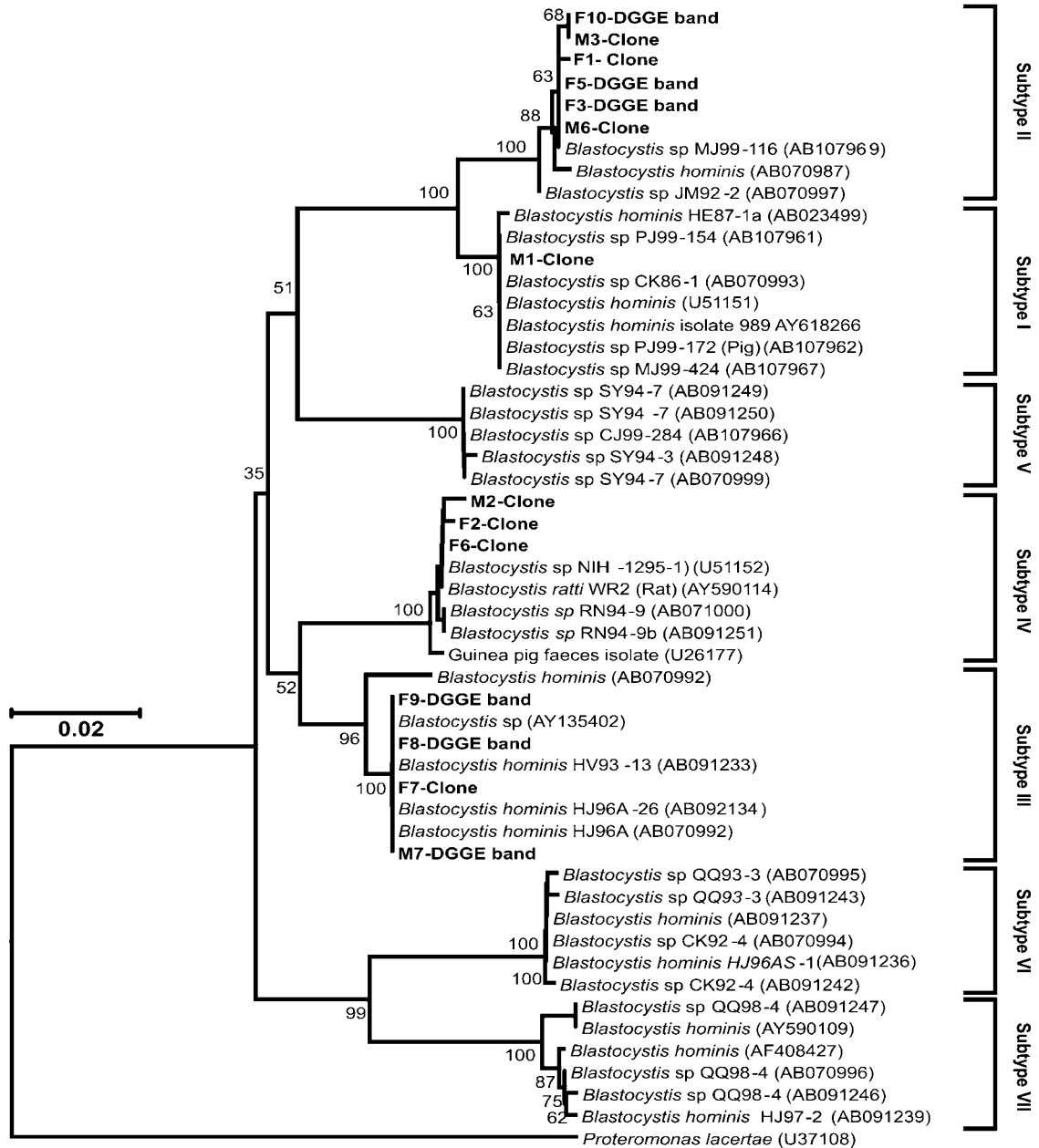


**Figure 1** Denaturing gradient gel electrophoretic profile analysis. DICE correlation coefficients yielded the calculated similarity indexes (expressed as percentages) of DGGE profiles generated from a selection of different individuals using the eukaryotic primer pair Euk1a and Euk516r-GC. The dendrogram is a statistically optimal representation of the similarities between DGGE profiles based on the matrix of DICE correlation coefficients and UPGMA. The code for each individual is outlined in Table 1 and the letter T followed by a number present after the code represents the time point analysed where applicable. This figure highlights both inter-individual variation and similarities and variations in intra-individual temporal stability are evident.

addition, the three individuals that clustered with group IV were all under the age of 28 years. Otherwise, no clear pattern could be determined with respect to *Blastocystis* incidence across individuals and the results presented here largely indicate that sex is not a factor in clustering, as both males and females are represented when present.

It is apparent from previous studies and the data presented here that many *Blastocystis* sp are commonly shared across human and animal species, which indicates low host specificity and a versatile genome, with a capacity for survival in the potentially harsh GI environments of wide range of hosts (Tan, 2004). The manner by which humans acquire *Blastocystis* sp is still uncertain, as many modes of transmission have been proposed, including human–human and animal–human (Yan *et al.*, 2006). M1 and F3, a couple married for 32 years, hosted different *Blastocystis* sp that clustered with subtypes I and II, respectively. In addition, their daughter (F4) did not have any *Blastocystis* sp present on analyses, which might indicate that human–human transmission is not important, but *Blastocystis* might be more likely acquired from an environmental source (Karanis *et al.*, 2007).

A high incidence of *Blastocystis* is usually associated with poor sanitation and surveys have typically focused on developing countries (Ostan



**Figure 2** Phylogenetic tree of partial 18S rRNA gene sequences of *Blastocystis* sp.

*et al.*, 2007). The considerable variation in *Blastocystis* incidence among broad and disparate human populations is likely owing to the specific diagnostic procedures employed (Stensvold *et al.*, 2007a, b). Previous analysis of individuals from developed countries has shown prevalence rates of 0.5% in Japan (Horiki *et al.*, 1997) and up to 23% in America (Amin, 2002); however, it is most probable that the high numbers of individuals positive for *Blastocystis* (82%) presented in this study may be as a result of the increased sensitivity of PCR-based assays over traditional methods (Stensvold *et al.*, 2007b). With respect to the data presented here, the incidence, dominance and temporal stability, together with inter-individual similarity and diversity, of *Blasto-*

*cystis* in this study of non-symptomatic healthy individuals from a developed Western country are of significant interest and require further investigation. As such, research should focus on the clinical relevance of this eukaryote and on establishing whether certain *Blastocystis* sp have functional roles in the gut of asymptomatic individuals.

Three individuals analysed using DGGE and clone library analysis of the 18S rRNA gene differed from the *Blastocystis* trend; in one individual (M4), *Entamoeba coli* was identified as the only species present. Individuals (F4 and M5) hosted a fungal species with 100% 18S rRNA gene sequence identity to both *Verticillium leptobactrum* (AB214657)/*Verticillium insectorum* (AB214657)

and *Paecilomyces fumosoroseus* (AB086629). This sequence was present at all time points for both individuals. In addition, individual F4 was positive for *Saccharomyces cerevisiae* over a 3-year period (these latter two results were supported by ITS profile and clone library sequence data where sequences 100% identical to both *Gloeotinia temulenta* (DQ235697) and *P. fumosoroseus* (AB086629) were also recovered (see Table 2). 18S rRNA gene sequences homologous to the fungi *Aspergillus* and *Rhizochaete*, and the insect genus *Mydas*, were retrieved from one individual (F2) who was analysed over a 2-year period. These sequences were only retrieved from single time points and not temporally, which could indicate that there is also a transient nature to the composition of the micro-eukaryotic GI community, a feature similar to the fungal ITS temporal analysis. As a method to characterize both similarities and differences of eukaryotic communities, the usefulness of the DGGE method is best suited to temporal analyses of individuals. Although DGGE was also able to

identify true similarities between different individuals, some individuals that clustered together on gel analysis had different band products upon sequence analysis (see Figure 2). Therefore, sequence analyses of PCR products and DGGE bands are required to resolve any potential ambiguities when applying the techniques outlined in this study.

*Fungal diversity of human intestinal tract is low, relatively stable and incongruent between culture and culture-independent analysis*

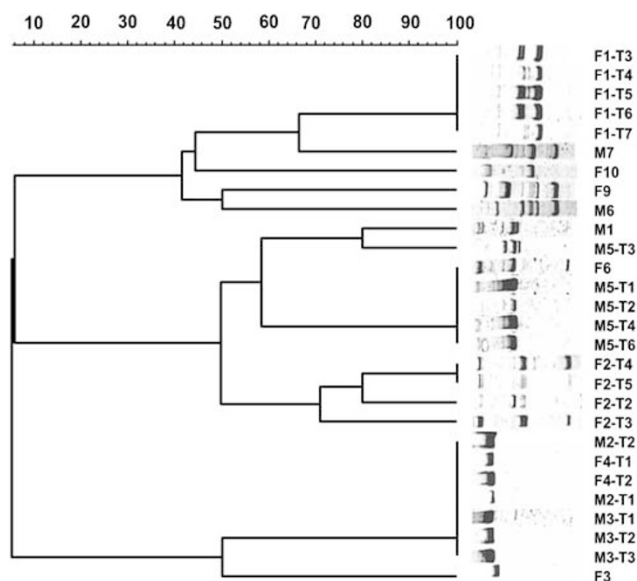
Fungal ITS PCR products were obtained from 14 out of 17 individuals analysed in this study. The diversity of fungi in the human intestinal tract is low (one to three bands per individual) and both fungal ITS profile and sequence analyses demonstrated both inter-individual similarity and variation across the group of individuals (see Figure 3 and Table 2). Temporal analysis of ITS profiles showed that each individual's profile was largely stable over time with

**Table 2** Results from fungal cultivation and cultivation-independents analysis of health subjects

Subject code	Mean fungal counts CFU g <sup>-1</sup> (37°C)	Pure-culture isolate classification (closest relative on database)	Fungal ITS sequence classification recovered from profiles and clone libraries (closest relative on database)
M1	6.0 × 10 <sup>2</sup>	<i>Candida parapsilosis</i> (EF152414)	<i>Gloeotinia temulenta</i> (DQ235697)/ <i>Paecilomyces fumosoroseus</i> (AB086629) and <i>Cephalosporium</i> sp JS2015 (AM176722) <sup>a</sup>
M2	1.9 × 10 <sup>4</sup>	<i>Candida parapsilosis</i> (EF152414)	<i>Gloeotinia temulenta</i> (DQ235697)/ <i>Paecilomyces fumosoroseus</i> (AB086629) and <i>Psathyrella candolleana</i> (DQ093736) <sup>a</sup>
M3	5.7 × 10 <sup>2</sup>	<i>Candida parapsilosis</i> (EF152414)	<i>Saccharomyces cerevisiae</i> (AB280540)
M4	ND	NA	No PCR signal
M5	1.1 × 10 <sup>3</sup>	<i>Candida parapsilosis</i> (EF152414)	<i>Gloeotinia temulenta</i> (DQ235697)/ <i>Paecilomyces fumosoroseus</i> (AB086629), <i>Ceratobasidium</i> sp (DQ278930) <sup>a</sup> and <i>Aspergillus versicolor</i> (AJ937752) <sup>a</sup>
M6	ND	NA	<i>Galactomyces geotrichum</i> (DQ668351) and <i>Psathyrella candolleana</i> (AF345810)
M7	ND	NA	<i>Galactomyces geotrichum</i> (DQ668351) and <i>Penicillium roquefortii</i> strain FRR 849 (AY373929)
F1	6.1 × 10 <sup>5</sup>	<i>Candida parapsilosis</i> (EF152414)	<i>Galactomyces</i> sp BPY-54 (DQ286062)
F2	1.1 × 10 <sup>4</sup>	<i>Candida parapsilosis</i> (EF152414)	<i>Gloeotinia temulenta</i> (DQ235697)/ <i>Paecilomyces fumosoroseus</i> (AB086629)
F3	1.0 × 10 <sup>3</sup>	<i>Candida albicans</i> ATCC 10231	<i>Gloeotinia temulenta</i> (DQ235697)/ <i>Paecilomyces fumosoroseus</i> (AB086629) and <i>Saccharomyces cerevisiae</i> (AB212259)
F4	1.0 × 10 <sup>2</sup>	<i>Candida parapsilosis</i> (AY520248)	<i>Gloeotinia temulenta</i> (DQ235697)/ <i>Paecilomyces fumosoroseus</i> (AB086629) and <i>Saccharomyces cerevisiae</i> (AM90040)
F5	2.0 × 10 <sup>3</sup>	<i>Candida albicans</i> ATCC 10231	<i>Saccharomyces cerevisiae</i> (AM90040) and <i>Candida albicans</i> ATCC 10231
F6	1.4 × 10 <sup>3</sup>	<i>Candida parapsilosis</i> (AY520248)	<i>Gloeotinia temulenta</i> (DQ235697)/ <i>Paecilomyces fumosoroseus</i> (AB086629)
F7	ND	NA	No PCR signal
F8	1.5 × 10 <sup>4</sup>	<i>Candida albicans</i> ATCC 10231 (EU266569)	No PCR signal
F9	1.6 × 10 <sup>6</sup>	<i>Candida albicans</i> ATCC:10231 (EU266569)	<i>Candida albicans</i> ATCC: 10231 (EU266569) and <i>Galactomyces geotrichum</i> (DQ668351)
F10	ND	NA	<i>Cf. Acremonium</i> sp FSU2858 (AY633563)

Abbreviations: NA, not applicable; ND, not detected.

<sup>a</sup>These fungal species were found at single time points for temporally analysed individuals.



**Figure 3** Fungal internal transcribed spacer analysis (FITSA) profile analysis. Example of FITSA gel cluster analysis highlighting inter-individual variation, and similarities with intra-individual temporal stability are also evident. DICE correlation coefficients yielded the calculated similarity indexes (expressed as percentages on horizontal bar) of gel profiles generated from different individuals using the eukaryotic primer pair ITS1F and ITS4R. The code outlined in Table 1 and the numbers present after the code represent the time point analysed, where applicable. The dendrogram is a statistically optimal representation of the similarities between gel profiles based on the matrix of DICE correlation coefficients and UPGMA analysis was performed using the Gel Compar software programme.

the occasional detection of transient fungal species (see Figure 3). Table 2 describes the different ITS sequence recovered from both fungal internal transcribed spacer analysis gel profiles and ITS libraries. The most commonly recovered species retrieved from seven individuals was 100% identical to *G. temulenta* (DQ235697)/*P. fumosoroseus* (AB086629). Temporal analysis of individuals indicated that, when present, this fungal species was a constant feature of all time points analysed. *Galactomyces* was also recovered from four individuals and, when present, was also a constant feature of all time points analysed. In addition, this species was also recovered from 18S rRNA gene DGGE analysis of one individual (F1). Many other species, including likely food borne fungi, such as *Penicillium roquefortii*, were retrieved at single time points using ITS clone library and gel analysis. *S. cerevisiae* was also recovered from cultivation-independent analysis, but was not recovered from plates, which may be owing to initial freezing of samples, the presence of temperature-sensitive strains or the application of suboptimal cultivation methods for the isolation of *S. cerevisiae*.

There was considerable variation in the total counts of viable fungi isolated at 37 °C from different individuals ranging from 0 up to  $1.45 \times 10^6$  CFU g<sup>-1</sup> of faeces from all individuals analysed. Uniform colony morphologies were evident on plates and

18S rRNA gene and ITS sequence analyses revealed that this uniform morphological type present across individuals was most closely related to *Candida*. Of particular interest is the lack of congruency between cultivation and cultivation-independent analysis. It is generally believed that *Candida* sp are the dominant fungal species in the human gut. However, the fact that *Candida* sp sequences were never detected in either 18S rRNA gene DGGE analysis and only recovered from the two ITS gel profiles indicate that they are not in fact the dominant eukaryotes in the distal colon of healthy individuals. One of these individuals (F9) had the highest *Candida* counts in this study ( $>10^6$  CFU g<sup>-1</sup> of faeces), which may be the threshold of detection for these species using the cultivation-independent approaches. As such, it is possible that *Candida* sp cell numbers were below the threshold levels of detection for the PCR protocols outlined in this study. However, in individual F5, *Candida albicans* counts did not exceed  $10^3$  CFU g<sup>-1</sup> of faeces and a signal was obtained by ITS-PCR; this disparity needs further work to clarify the limits of detection for the PCR. Previous studies using DGGE to characterize the micro-eukaryotic communities in food and beverages have also encountered this anomaly where yeasts isolated by culture were not represented in 18S rRNA gene DGGE profiles (Fleet, 2007). Media such as Sabouraud dextrose may provide a selective enrichment bias toward yeasts, such as *Candida*. Initial optimization and exploratory analysis using Sabouraud dextrose agar plates incubated aerobically at 30, 20 and 12 °C, isolated fungal species such as *Rhodotorula mucilaginosa* and *Penicillium* sp. These particular species were not recovered from plates incubated at 37 °C, which emphasizes the biases imposed when selecting *a priori* parameters required for cultivation-based analysis. In addition, the impact of freezing on faecal samples is unknown and may impact the recovery of cultivable fungi. Overall, the application of both cultivation-dependent and -independent analyses of fungal diversity within the human GI tract revealed a relatively limited range of different species across individuals compared with the diversity of fungal species present in the murine gut (Scupham *et al.*, 2006), and also in comparison with the diversity of bacterial species present in the distal colon of healthy human hosts (Eckburg *et al.*, 2005; Suau *et al.*, 1999). Although a wide range of primers were tested during optimization (see Supplementary Table S1), the impact of any bias, selectivity and coverage of the primers used to generate data in this study may limit the complete characterization of all species present.

The introduction and establishment of the GI microbiota can occur from numerous sources, including contact with the environment. Genes encoding pH, bile, anaerobic tolerance and the ability to compete for nutrients and space are necessary for competing species to survive. A ready



source of microbial inocula includes diet, where food and beverages represent complex microbial ecosystems comprising yeasts, filamentous fungi, bacteria and their viruses. Both *Candida* sp and, in particular, *Saccharomyces* sp are routinely used in the brewing and baking industry; *Galactomyces* sp are routinely found in cheese microbial consortia where they are used in starter cultures (Goerges *et al.*, 2008); and *Galactomyces geotrichum* has also been identified as an agent in soft drinks spoilage (Ancasi *et al.*, 2006). As such, when fermented foods and beverages or food containing eukaryotic species are consumed on a continuous basis, they provide ready inocula for the host. Furthermore, fungal species such as *P. fumosoroseus* are commonly used as bio-control agents against economically important agricultural pests, and this may be responsible for its incidence in humans (Madsen *et al.*, 2007). To the authors' knowledge, this is the first account of the *Galactomyces*, *Paecilomyces* and *Gloetinia* genera of fungal species addressed in the context of a healthy human GI tract, and as such no reference material is available. The presence of such species in the distal colon may be a reflection of the diet and/or the environment of the individual, but the presence of these species across different individuals and over time points conceivably indicates that they are common and stable inhabitants of the healthy human intestinal tract. With this novel data, we can now begin to further investigate the roles of these species in the human distal colon.

#### Concluding remarks

Molecular ecological analysis of microorganisms from complex communities is an extremely difficult undertaking considering that the vast majority of both bacterial and eukaryotic species in natural environments are refractory to cultivation (Amann *et al.*, 1990; Pace, 1997; Hugenholtz *et al.*, 1998; Suau *et al.*, 1999). To circumvent the requirement for laborious optimization and concurrent large-scale culturing that is necessary to facilitate the complex and distinct growth requirements of eukaryotes from natural environments, culture-independent approaches are requisite to complete community diversity surveys (Edgcomb *et al.*, 2002; Lepere *et al.*, 2006).

Prior to this investigation, the eukaryotic members of the human gut microbial community have been primarily assessed using cultivation-based methods, and this has led to the belief that they are not widely distributed or abundant in the human microbiota (Rajilic-Stojanovic *et al.*, 2007). This statement is indeed true for the cultivable fraction of the community, but through the application of cultivation-independent techniques, such as DNA extraction and PCR, this study provides evidence that stable eukaryotic communities are widespread in the healthy human gut. Furthermore, allochthonous species were routinely encountered during analysis

of temporally sampled individuals, indicating a transient characteristic of the micro-eukaryotic community of the GI tract. *Blastocystis* sp and fungi identical to *Gloetinia/Paecilomyces* and *Galactomyces* are common GI inhabitants, and, when present, are stable and constant features of the micro-eukaryotic community. This raises many interesting ecological questions pertaining to micro-eukaryotes and their impact on both health and disease of the host. For example, what factors are involved in niche colonization and competition, what are their substrate requirements and what is the impact of their metabolic output and extracellular metabolite production? Does the sheer size and volume of a eukaryotic cell compared to that of a bacterial cell influence competition and interaction with other components of the GI microbiota and also the host cell? Do some GI tracts possess a bacterial community that is suppressive toward pathogenic fungi, a phenomenon seen in the soil ecosystem (Borneman and Becker, 2007)? What are the potential roles eukaryotes may play in immune stimulation and modulation; for example, the role of yeasts and the presence of anti-*S-cerevisiae* antibodies in Crohn's disease (Schaffer *et al.*, 2007) and the use of certain helminths as a treatment for CD (Reddy and Fried, 2007)?

Results presented in this study highlight fungal culturing biases that are similar to the anomalies that scientists encountered when they began to investigate the bacterial fraction of the human GI microbiota (Suau *et al.*, 1999; Hayashi *et al.*, 2002), where readily cultivable species, such as *E. coli*, were overrepresented on plates. Therefore, as with bacterial culturing, it would appear that the application of traditionally used cultivation media such as Sabouraud dextrose agar for the cultivation of fungi from the human GI alone does not by itself provide an accurate account of the extant diversity. With this novel information, we can begin not only to apply different cultivation methods to the human GI tract but also design better cultivation techniques to isolate fungi, assess the functionality of *Blastocystis* and further investigate the long-neglected role of micro-eukaryotic species and their role in gut health.

#### Acknowledgements

We thank all the participants for providing samples for this study and making it possible. This research was supported by Science Foundation Ireland.

#### References

- Abe N. (2004). Molecular and phylogenetic analysis of *Blastocystis* isolates from various hosts. *Vet Parasitol* **120**: 235–242.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. (1990). Basic local alignment search tool. *J Mol Biol* **215**: 403–410.

- Amann RI, Binder BJ, Olson RJ, Chisholm SW, Devereux R, Stahl DA. (1990). Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Appl Environ Microbiol* **56**: 1919–1925.
- Amin OM. (2002). Seasonal prevalence of intestinal parasites in the United States during 2000. *Am J Trop Med Hyg* **66**: 799–803.
- Ancasi EG, Carrillo L, Benitez Ahrendts MR. (2006). Moulds and yeasts in bottled water and soft drinks. *Rev Argent Microbiol* **38**: 93–96.
- Archie JW, Felsenstein J. (1993). The number of evolutionary steps on random and minimum length trees for random evolutionary data. *Theor Popul Biol* **43**: 52–79.
- Arisue N, Hashimoto T, Yoshikawa H. (2003). Sequence heterogeneity of the small subunit ribosomal RNA genes among *Blastocystis* isolates. *Parasitology* **126**: 1–9.
- Blaut M, Clavel T. (2007). Metabolic diversity of the intestinal microbiota: implications for health and disease. *J Nut* **137**: 751S–755S.
- Borneman J, Becker JO. (2007). Identifying microorganisms involved in specific pathogen suppression in soil. *Annu Rev Phytopathol* **45**: 153–172.
- Cash HL, Whitham CV, Behrendt CL, Hooper LV. (2006). Symbiotic bacteria direct expression of an intestinal bactericidal lectin. *Science* **313**: 1126–1130.
- Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M *et al.* (2005). Diversity of the human intestinal microbial flora. *Science* **308**: 1635–1638.
- Edgcomb VP, Kysela DT, Teske A, de Vera Gomez A, Sogin ML. (2002). Benthic eukaryotic diversity in the Guaymas Basin hydrothermal vent environment. *Proc Natl Acad Sci USA* **99**: 7658–7662.
- Fleet GH. (1999). Microorganisms in food ecosystems. *Int J Food Microbiol* **50**: 101–117.
- Fleet GH. (2003). Yeast interactions and wine flavour. *Int J Food Microbiol* **86**: 11–22.
- Fleet GH. (2007). Yeasts in foods and beverages: impact on product quality and safety. *Curr Opin Biotechnol* **18**: 170–175.
- Gadanhó M, Almeida JM, Sampaio JP. (2003). Assessment of yeast diversity in a marine environment in the south of Portugal by microsatellite-primed PCR. *Ant Van Leeuwenhoek* **84**: 217–227.
- Giacometti A, Cirioni O, Fiorentini A, Fortuna M, Scalise G. (1999). Irritable bowel syndrome in patients with *Blastocystis hominis* infection. *Eur J Clin Microbiol Infect Dis* **18**: 436–439.
- Goerges S, Mounier J, Rea MC, Gelsomino R, Heise V, Beduhn R *et al.* (2008). Commercial ripening starter microorganisms inoculated into cheese milk do not successfully establish themselves in the resident microbial ripening consortia of a south German red smear cheese. *Appl Environ Microbiol* **74**: 2210–2217.
- Hayashi H, Sakamoto M, Benno Y. (2002). Phylogenetic analysis of the human gut microbiota using 16S rDNA clone libraries and strictly anaerobic culture-based methods. *Microbiol Immunol* **46**: 535–548.
- Ho LC, Armiugam A, Jeyaseelan K, Yap EH, Singh M. (2000). *Blastocystis* elongation factor-1 $\alpha$ : genomic organization, taxonomy and phylogenetic relationships. *Parasitology* **121**(Pt 2): 135–144.
- Horiki N, Maruyama M, Fujita Y, Yonekura T, Minato S, Kaneda Y. (1997). Epidemiologic survey of *Blastocystis hominis* infection in Japan. *Am J Trop Med Hyg* **56**: 370–374.
- Hughenoltz P, Goebel BM, Pace NR. (1998). Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *J Bacteriol* **180**: 4765–4774.
- Kam AP, Xu J. (2002). Diversity of commensal yeasts within and among healthy hosts. *Diagn Microbiol Infect Dis* **43**: 19–28.
- Kaneda Y, Horiki N, Cheng XJ, Fujita Y, Maruyama M, Tachibana H. (2001). Ribodemes of *Blastocystis hominis* isolated in Japan. *Am J Trop Med Hyg* **65**: 393–396.
- Karanis P, Kourenti C, Smith H. (2007). Waterborne transmission of protozoan parasites: a worldwide review of outbreaks and lessons learnt. *J Water Health* **5**: 1–38.
- Kelly D, King T, Aminov R. (2007). Importance of microbial colonization of the gut in early life to the development of immunity. *Mutat Res* **622**: 58–69.
- Kuhbacher T, Ott SJ, Helwig U, Mimura T, Rizzello F, Kleessen B *et al.* (2006). Bacterial and fungal microbiota in relation to probiotic therapy (VSL#3) in pouchitis. *Gut* **55**: 833–841.
- Lepere C, Boucher D, Jardillier L, Domaizon I, Debroas D. (2006). Succession and regulation factors of small eukaryote community composition in a lacustrine ecosystem (Lake Pavin). *Appl Environ Microbiol* **72**: 2971–2981.
- Lopez-Garcia P, Vereshchaka A, Moreira D. (2007). Eukaryotic diversity associated with carbonates and fluid-seawater interface in Lost City hydrothermal field. *Environ Microbiol* **9**: 546–554.
- Madsen AM, Hansen VM, Meyling NV, Eilenberg J. (2007). Human exposure to airborne fungi from genera used as biocontrol agents in plant production. *Ann Agric Environ Med* **14**: 5–24.
- Marchesi J, Shanahan F. (2007). The normal intestinal microbiota. *Curr Opin Infect Dis* **20**: 508–513.
- Mathieu-Daude F, Welsh J, Vogt T, McClelland M. (1996). DNA rehybridization during PCR: the 'C<sub>0</sub> effect' and its consequences. *Nucleic Acids Res* **24**: 2080–2086.
- Moon-van der Staay SY, Tzeneva VA, van der Staay GW, de Vos WM, Smidt H, Hackstein JH. (2006). Eukaryotic diversity in historical soil samples. *FEMS Microbiol Ecol* **57**: 420–428.
- Muyzer G, de Waal EC, Uitterlinden AG. (1993). Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol* **59**: 695–700.
- Ostan I, Kilimcioglu AA, Girginkardesler N, Ozyurt BC, Limoncu ME, Ok UZ. (2007). Health inequities: lower socio-economic conditions and higher incidences of intestinal parasites. *BMC Public Health* **7**: 342.
- Pace NR. (1997). A molecular view of microbial diversity and the biosphere. *Science* **276**: 734–740.
- Patterson D. (1994). Progress in Protistology. In: Hausmann H, Hulsmann N (eds). *Proceedings of the IX International Congress of Protozoology*. G. Fisher: Berlin, pp 1–14.
- Quiros M, Martorell P, Querol A, Barrio E, Peinado JM, de Sioniz MI. (2008). Four new *Candida cretensis* strains isolated from Spanish fermented sausages (chorizo): taxonomic and phylogenetic implications. *FEMS Yeast Res* **8**: 485–491.

- Rajilic-Stojanovic M, Smidt H, de Vos WM. (2007). Diversity of the human gastrointestinal tract microbiota revisited. *Environ Microbiol* **9**: 2125–2136.
- Ranilla MJ, Jouany JP, Morgavi DP. (2007). Methane production and substrate degradation by rumen microbial communities containing single protozoal species *in vitro*. *Lett Appl Microbiol* **45**: 675–680.
- Reddy A, Fried B. (2007). The use of *Trichuris suis* and other helminth therapies to treat Crohn's disease. *Parasitol Res* **100**: 921–927.
- Romano P, Fiore C, Paraggio M, Caruso M, Capece A. (2003). Function of yeast species and strains in wine flavour. *Int J Food Microbiol* **86**: 169–180.
- Saitou N, Nei M. (1987). The neighbor-joining method—a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**: 406–425.
- Sambrook J, Russell D, Sambrook J, Russell D. (2000). *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor Laboratory Press: New York.
- Schaffer T, Muller S, Flogerzi B, Seibold-Schmid B, Schoepfer AM, Seibold F. (2007). Anti-*Saccharomyces cerevisiae* mannan antibodies (ASCA) of Crohn's patients crossreact with mannan from other yeast strains, and murine ASCA IgM can be experimentally induced with *Candida albicans*. *Inflamm Bowel Dis* **13**: 1339–1346.
- Scupham AJ, Presley LL, Wei B, Bent E, Griffith N, McPherson M *et al.* (2006). Abundant and diverse fungal microbiota in the murine intestine. *Appl Environ Microbiol* **72**: 793–801.
- Silberman JD, Sogin ML, Leipe DD, Clark CG. (1996). Human parasite finds taxonomic home. *Nature* **380**: 398.
- Song X, Eribe ER, Sun J, Hansen BF, Olsen I. (2005). Genetic relatedness of oral yeasts within and between patients with marginal periodontitis and subjects with oral health. *J Periodontol Res* **40**: 446–452.
- Stark D, van Hal S, Marriott D, Ellis J, Harkness J. (2007). Irritable bowel syndrome: a review on the role of intestinal protozoa and the importance of their detection and diagnosis. *Int J Parasitol* **37**: 11–20.
- Stensvold CR, Arendrup MC, Jespersgaard C, Molbak K, Nielsen HV. (2007b). Detecting *Blastocystis* using parasitologic and DNA-based methods: a comparative study. *Diagn Microbiol Infect Dis* **59**: 303–307.
- Stensvold CR, Suresh GK, Tan KS, Thompson RC, Traub RJ, Viscogliosi E *et al.* (2007a). Terminology for *Blastocystis* subtypes—a consensus. *Trends Parasitol* **23**: 93–96.
- Suau A, Bonnet R, Sutren M, Godon JJ, Gibson GR, Collins MD *et al.* (1999). Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. *Appl Environ Microbiol* **65**: 4799–4807.
- Tamura K, Dudley J, Nei M, Kumar S. (2007). MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* **24**: 1596–1599.
- Tamura K, Nei M, Kumar S. (2004). Prospects for inferring very large phylogenies by using the neighbor-joining method. *PNAS* **101**: 11030–11035.
- Tan KS. (2004). *Blastocystis* in humans and animals: new insights using modern methodologies. *Vet Parasitol* **126**: 121–144.
- Tan TC, Suresh KG, Thong KL, Smith HV. (2006). PCR fingerprinting of *Blastocystis* isolated from symptomatic and asymptomatic human hosts. *Parasitol Res* **99**: 459–465.
- van Elsas JD, Duarte GF, Keijzer-Wolters A, Smit E. (2000). Analysis of the dynamics of fungal communities in soil via fungal-specific PCR of soil DNA followed by denaturing gradient gel electrophoresis. *J Microbiol Methods* **43**: 133–151.
- Yakoob J, Jafri W, Jafri N, Khan R, Islam M, Beg MA *et al.* (2004). Irritable bowel syndrome: in search of an etiology: role of *Blastocystis hominis*. *Am J Trop Med Hyg* **70**: 383–385.
- Yan Y, Su S, Lai R, Liao H, Ye J, Li X *et al.* (2006). Genetic variability of *Blastocystis hominis* isolates in China. *Parasitol Res* **99**: 597–601.
- Yoshikawa H, Abe N, Wu Z. (2004). PCR-based identification of zoonotic isolates of *Blastocystis* from mammals and birds. *Microbiology* **150**: 1147–1151.

Supplementary Information accompanies the paper on The ISME Journal website (<http://www.nature.com/ismej>)