# DETECTION OF A NEGATIVE BIOMARKER FOR AMEBIASIS (*ENTAMOEBA HISTOLYTICA*) IN THE HUMAN GUT MYCOBIOME

ABULJADAYEL, D.[1] – ATEF, A.[1] – AL-MATARY, M.[1] – EDRIS, S.[1,2,3] – AL-QUWAIE, D. A. H.[4] – ALSUBHI, N. H.[4] – AL-HINDI, R. R.[1] – SABIR, J. S. M.[1] – HALL, N.[1,5] – BAHIELDIN, A.[1,2*]

[1]*Department of Biological Sciences, Faculty of Science, King Abdulaziz University (KAU), P.O. Box 80141, Jeddah 21589, Saudi Arabia*

[2]*Department of Genetics, Faculty of Agriculture, Ain Shams University, Cairo, Egypt*

[3]*Princess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-HD), Faculty of Medicine, King Abdulaziz University (KAU), Jeddah, Saudi Arabia*

[4]*Department of Biological Sciences, Rabigh College of Science and Arts, King Abdulaziz University (KAU), Rabigh, Saudi Arabia*

[5]*The Genome Analysis Center, Norwich Research Park, Norwich NR4 7UH, UK*

*\*Corresponding author*
*e-mail: abmahmed@kau.edu.sa*

**Abstract.** Amebiasis caused by *Entamoeba histolytica* parasite is among the worst gut diseases worldwide. The study involved detection of gut mycobiome signature of the infected subjects in a trial to improve ability to diagnose and treat this disease. The results indicated that diversity of samples slightly changed due to infection. Dendrogram indicated that the most common phyla in the human gut are Ascomycota, Basidiomycota and Zygomycota. Species *Pichia kudriavzevii*, *Candida glabrata*, *Saccharomyces cerevisiae* of phylum Ascomycota and *Rhodotorula mucilaginosa* of phylum Basidiomycota were found highly OTU abundant. Weighted unifrac diversity distances resulted in complete separation of the two groups. Differential abundance analysis indicated that four phyla, six families, nine genera and ten species showed considerable changes in gut mycobiome signatures within and among groups. Unclassified species of genus *Malassezia* (phylum Basidiomycota) showed high abundance in infected subjects as compared to healthy subjects, while families Saccharomycetaceae and Trichocomaceae, genus *Saccharomyces* and species *S. cerevisiae* showed opposite results. Linear discriminant analysis effect size (LEFSE) indicated that family Saccharomycetaceae, of which *S. cerevisiae* is a downstream taxon, can be a negative biomarker for amebiasis. In conclusion, the study provides new insights into possible use of *Saccharomyces* as a probiotic against amebiasis.
**Keywords:** *encystation, yeast, Crabtree effect, oxidative stress, chitin formation*

## Introduction

*Entamoeba histolytica* parasite is among the most common causes of death worldwide (WHO, 1997). It is estimated that around 50-100 million people are infected with the disease each year. Infection can be symptomatic (dysentery or amoebic colitis) or asymptomatic, and can also be invasive (amoebic liver abscess) (Stanley, 2003; Ximénez et al., 2009; Lozano et al., 2010; Verkerke et al., 2012). Furthermore, amoebae can cause abscess in the brain, pericarditis in the heart or pleuropulmonary disease in the lung (Stanley, 2003). Infection is usually drug-treated (Ali and Nozaki, 2007), however, in severe cases surgical intervention might be necessary (Ishida et al., 2003;

Gupta et al., 2009). Amebiasis is proven to be mainly associated with water supplies contaminated with human feces resulting in 50,000-100,000 deaths reported annually, which make this protozoan parasite a significant cause of death worldwide (WHO, 1997; Soares et al., 2019).

*E. histolytica* parasite is transmitted via the ingestion of water contaminated with cysts. After excystation, the *Entamoeba* trophozoite divides and colonizes the colon, then, eventually produces new cysts that are excreted into the environment (Ali and Nozaki, 2007). Two major processes are crucial for the existence and cycling of these two forms of *Entamoeba*. They are encystation and trophozoite adhesion to human cells. If one of the two processes is impaired, then, existence or life cycling of the parasite will be in jeopardy. Encystation requires a number of processes including Gal-lectin production and chitin metabolism (Mi-ichi et al., 2016). Chitin in the cyst wall comprises of two chitin synthases (Chs), one of which is unique (EhCHS-1) to *Entamoeba*, while the other (EhCHS-2) resembles those in insects, fungi and nematodes (Van Dellen et al., 2006). Adhesion of trophozoites to human cells is the main event of subsequent cytopathogenic activities, contact-mediated cytolysis and phagocytosis, intestinal colonization and invasion (Eaton et al., 1970; Trissl et al., 1978; Ravdin et al., 1980; Mi-ichi et al., 2016). Phagocytosis is a mandatory process for survival of this organism, while blocking this process results in the inhibition of its proliferation and loss of pathogenicity (Hirata et al., 2007; Iyer et al., 2019). Other process for survival of *Entamoeba* is the modifications in the structure of transmembrane proteins that makes the parasite unidentified, thus, helps in escaping the host immune system (Khomkhum et al., 2019).

The human gut ecosystem contains a microbiome representing the three domains of life, with fungi—the mycobiome—predominates gut Eukarya (Rajilic-Stojanovic et al., 2007; Scanlan and Marchesi, 2008). The gut microbiome is presumed sterile at birth and the primary route for microbes to enter the gastrointestinal (GI) tract occurs via ingestion (Penders et al., 2006) of which Saccharomycetalean yeasts are the first fungi detected in the infant gut (Bliss et al., 2008). The exact relationship of host gut and its ecosystem, comprising prokaryotes, fungi, viruses and parasites, is extremely obscure (Hugon et al., 2017). The composition of the different microbiomes of the human body reflects health and disease parameters (Moyes and Naglik, 2012; Wang et al., 2014; Gouba and Drancourt, 2015; Mukherjee et al., 2015; Richard et al., 2015). Thus, studying the gut microbiota of *Entamoeba*-infected subjects might help in improving our understanding of these parameters that help in diagnosing and treating the disease more efficiently.

Although human mycobiomes have received little attention compared with bacteriomes, a number of recent articles have evidenced the importance of this organ in the whole body (Cui et al., 2013; Huffnagle and Noverr, 2013; Seed, 2014; Underhill and Iliev, 2014), especially in the gut (Ianiro et al., 2014; Kirschner et al., 2015; Suhr and Hallen-Adams, 2015). Therefore, the present study was planned to address the possible influence of human gut mycobiome on the occurrence and level of *Entamoeba* infection. The study also attempted to detect new biomarkers linked to amebiasis.

## Materials and methods

### *Recruitment of participants and sample collection*

Samples have been collected during summer 2019 from dysentery-infected subjects in four hospitals in Jeddah, KAU. An ethical approval (no. A00451) was issued by the Ministry of Health, Saudi Arabia following the regulation of the General Administration

for Research and Studies at the Ministry of Health (registration no. 1195437) and the National Committee for Medical and Biological Ethics (registration no. H-02-J-002). Consent forms were filled by healthy and infected subjects or their relatives at sampling time. Patient's fecal samples were diagnosed by microscopic examination and positive samples were kept in iSWAB-TM Microbiome©2016 (Mawi DNA Technologies LLC. USA/Canada) to ensure the integrity of DNA required for deep sequencing. Genomic DNA was extracted using the QIAamp® DNA Mini kit (Qiagen®51306; Hilden, North Rhine-Westphalia, Germany) according to the manufacturer's instructions. DNA purity was evaluated via A260/A280 ratio using NanoDrop 7000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA), and DNA integrity was checked by 1% agarose gel electrophoresis. PCR amplification of the V5-V7 region of 18S rRNA gene was performed using the universal primers SSU0817F (forward) 5′-TTAGCATGGAATAATRRAATAGGA -3′ and SSU1196R (reverse) 5′-TCTGGACCTGGTGAGTTTCC -3′. The PCR program was: initial denaturation at 95 °C for 5 min; 25 cycles of denaturation at 95 °C for 30 s, annealing at 56 °C for 30 s, and extension at 72 °C for 40 s; and final extension of 72 °C for 10 min. Amplicons were run on agarose gel (1.2%). Based on PCR results, original high-quality DNAs were shipped to Beijing Genome Institute (BGI) in China for library construction and deep sequencing on Illumina Miseq platform to recover ≥ 300 bp pair-ended reads of the V5-V7 region of 18S rRNA gene.

## 18S dataset processing and statistics

The ends of each read were overlapped to generate high-quality, full-length reads. The resulted sequencing data will eventually be deposited in the European Nucleotide Archive (ENA) (https://www.ebi.ac.uk/ena/submit/sra/#studies). The raw data were pre-processed to get clean data by in-house procedure as the following: (1) truncation of sequence reads not having an average quality of 20 over a 25 bp sliding window based on the Phred algorithm (Ewing and Green, 1998), (2) removal of trimmed reads having less than 75% of their original length, as well as its paired read; (3) removal of reads contaminated by adapter (default parameter: 15 bases overlapped by reads and adapter with maximal 3 bases mismatch allowed); (4) removal of reads with ambiguous base (N base), and its paired reads; and (5) removal of reads with low complexity (default: reads with 10 consecutive same base). The consensus sequence was generated by FLASH (Fast Length Adjustment of Short reads, v1.2.11; Magoc and Salzberg, 2011) in case the two paired-end reads overlapped with the following parameters: (1) minimal overlapping length: 15 bp; (2) mismatching ratio of overlapped region: < = 0.1; and (3) paired end reads without overlaps were removed.

Then, tags were clustered to OTU (Operational Taxonomic Unit) by scripts of software USEARCH (v7.0.1090; Edgar, 2010; http://drive5.com/usearch/manual/uparse_pipeline.html) as the following: (1) the tags were clustered into OTU with a 97% threshold by using UPARSE (Edgar, 2013), and the OTU unique representative sequences were obtained; (2) *de novo* chimeras were filtered out by using UCHIME (v4.2.40, Edgar et al., 2011); and (3) all tags were mapped to each OTU representative sequences using USEARCH GLOBAL. Then, the tags number of each OTU in each sample was summarized in OTU abundance table. OTU representative sequences were taxonomically classified using Ribosomal Database Project (RDP) Classifier (v.2.2; Cole et al., 2014), using 0.6 confidence values as cutoff. Silva (default) V119 database was used for species annotation (Quast et al., 2013).

OTUs were filtered as the following: (1) unassigned OTUs were removed; and (2) OTUs not assigned to the target species were removed. Then, filtered tags were clustered into OTUs to be used in downstream processing.

Alpha diversity was addressed by Shannon and Simpson indices. Both parameters were calculated by Mothur (v1.31.2; Schloss et al., 2009), and the corresponding boxplot of alpha diversity and rarefaction curve were drawn by software R (v3.1.1; https://cran.r-project.org/bin/windows/base/old/3.1.1/). Drawing rarefaction curve was based on calculating OTU numbers of the extracted tags (in multiples of 500), then, detecting the maximum depth (no. reads) permitted to retain all samples in the dataset. Sequences were extracted randomly according to the minimum sequence number for all samples, and the extracted sequences formed a new 'OTU table biom' file.

The representative sequences were aligned against the Silva core set (Silva_108_core_aligned_seqs) using PyNAST (Caporaso et al., 2010b) by 'align_seqs.py'. Then, a representative OTU phylogenetic tree was constructed using the QIIME (v1.80, Caporaso et al., 2010b) built-in scripts including the fast-tree method for tree construction. The tags with the highest abundance of each genus was chosen as the corresponding genus representative sequences, and genus level phylogenetic tree was obtained by the same way of OTU phylogenetic tree. Then, the phylogeny tree was imaged by software R (v3.1.1).

Venn diagram was drawn by Venn program of software R (v3.1.1), while differences in the relative abundances of taxa at the phylum, genus and species levels were analyzed using Metastats (Paulson et al., 2011). Benjamini–Hochberg false discovery rate (FDR) correction (Benjamini and Hochberg, 1995) was used to correct for multiple hypothesis testing where applicable.

Heat maps were generated using the package 'gplots' of software R (v3.1.1). The used distance algorithm is 'euclidean' and the clustering method is 'complete'. At phylum level, all species were used to draw the heat map and taxa of which abundance is less than 0.5% in all samples were classified as 'others'. To minimize the differences degree of the relative abundance value, the values were all log transformed.

To detect the beta diversity within and between groups, the weighted and unweighted UniFrac distances were calculated (Lozupone et al., 2011) and plotted *via* principal coordinate analysis (PCoA) using package 'ade4' of software R (v3.1.1). UniFrac uses the system evolution information to compare the composition of community species between samples. The results can be used as a measure of beta diversity. It takes into account the distance of evolution between species, and the higher the index values, the greater the differences between samples or groups. The UniFrac is divided into weighted UniFrac and unweighted UniFrac of which the weighted UniFrac considers the abundance of sequences, while unweighted UniFrac puts more emphasis on species presence or absence.

Linear discriminant analysis (or LDA) Effect Size (LEFSE) was done to determine the features (organisms, clades, operational taxonomic units, genes, or functions) explaining the differences between the two groups of subjects by coupling standard tests for statistical significance in addition to the tests encoding biological consistency and effect relevance as described (Abarenkov et al., 2010). In the LEFse tree, different colors indicate different groups. Colored node (red) indicates an important microbe biomarker in either group and the biomarker name lists in the upper right corner of the tree. The yellow nodes represent the biomarkers, which do not show any importance in either group.

## Results

### *Statistics of gut 18S rRNA sequence datasets*

Illumina MiSeq was used in analyzing a number of two healthy subjects as well as nine patients proved via microscopic examination to be infected with *Entamoeba*. The statistical analysis of gut mycobiome was done for the 11 gut mycobiome samples based on the 18S rRNA as shown in *Table 1*. The average sequence length per read was 296 bp across different samples ranging from 293 to 299 bp. Average percentages of N base and low-quality reads were as little as 0.097% and 0.00%, respectively. A total of 9,086,626 raw sequence reads were generated across subjects with an average of 1,514,437 reads per sample. While, a total of 9,050,377 clean sequence reads were generated across subjects with an average read number of 822,762 per sample. Average percentage of data utilization was as high as 99.61% (*Table 1*). The data in *Figure 1* indicates that the average number of clean reads is 791,462 across healthy subjects, while 829,717 across patients. The highest value of clean reads was given to patient 5 or P5 (1,127,017), while the lowest was given to P1 (692,848) (*Table 1*; *Fig. 1*).

***Table 1.*** *Statistics of data generated from deep sequencing for healthy and Entamoeba-infected individuals*

| Sample ID | Reads length (bp) | Raw data (Mbp) | N base (%) | Low quality (%) | Clean data (Mbp) | Raw reads | Clean reads | Read utilization (%) |
|---|---|---|---|---|---|---|---|---|
| H1 | 294:298 | 470.19 | 0.098 | 0 | 469.07 | 794239 | 792348 | 99.76 |
| H2 | 294:297 | 469.35 | 0.096 | 0 | 467.23 | 794155 | 790575 | 99.55 |
| P1 | 294:296 | 409.75 | 0.094 | 0 | 408.78 | 694500 | 692848 | 99.76 |
| P2 | 293:299 | 473.51 | 0.098 | 0 | 472.17 | 799844 | 797581 | 99.72 |
| P3 | 293:298 | 435.56 | 0.094 | 0 | 434.65 | 736986 | 735454 | 99.79 |
| P4 | 293:297 | 506.06 | 0.099 | 0 | 504.42 | 857721 | 854957 | 99.68 |
| P5 | 295:299 | 673.04 | 0.097 | 0 | 669.45 | 1133058 | 1127017 | 99.47 |
| P6 | 295:298 | 421.7 | 0.100 | 0 | 419.55 | 711138 | 707511 | 99.49 |
| P7 | 295:297 | 550.46 | 0.096 | 0 | 546.63 | 929832 | 923365 | 99.3 |
| P8 | 295:296 | 446.04 | 0.099 | 0 | 444.03 | 754715 | 751321 | 99.55 |
| P9 | 294:299 | 522.1 | 0.096 | 0 | 520.3 | 880435 | 877400 | 99.66 |

A total of 8,776,506 tag-linked sequences were generated across subjects with an average number of 783,091 sequences across healthy subjects and 801,147 sequences across patients (*Fig. 1*). The highest value of tag-linked sequences was given to patient P5 (1,061,387), while the lowest was given to P1 (685,566) (*Table 1*; *Fig. 1*), with an average connect percentage of 97.16%. These sequence tags were assigned to 146 eukaryotic OTUs (operational taxonomic units) across samples with an average of 60,113 sequence tags per OTU. The results for the number of observed species ($S_{obs}$) or number of OTUs indicated a summation of 698 OTUs for the 11 samples with an average of about 63 OTUs per sample ranging from 47 (P6) to 82 (P4) OTUs and averages of 60 and about 64 OTUs for healthy subjects and patients, respectively (*Fig. 1*). Number of OTUs in relation to relative abundance of OTUs for different healthy subjects (H) and *Entamoeba*-infected patients (P) is shown in *Figure A1* in the *Appendix*. The results in this figure complement those of *Figure 1* as P4 showed the highest number and abundance of OTUs, while P6 showed the lowest.
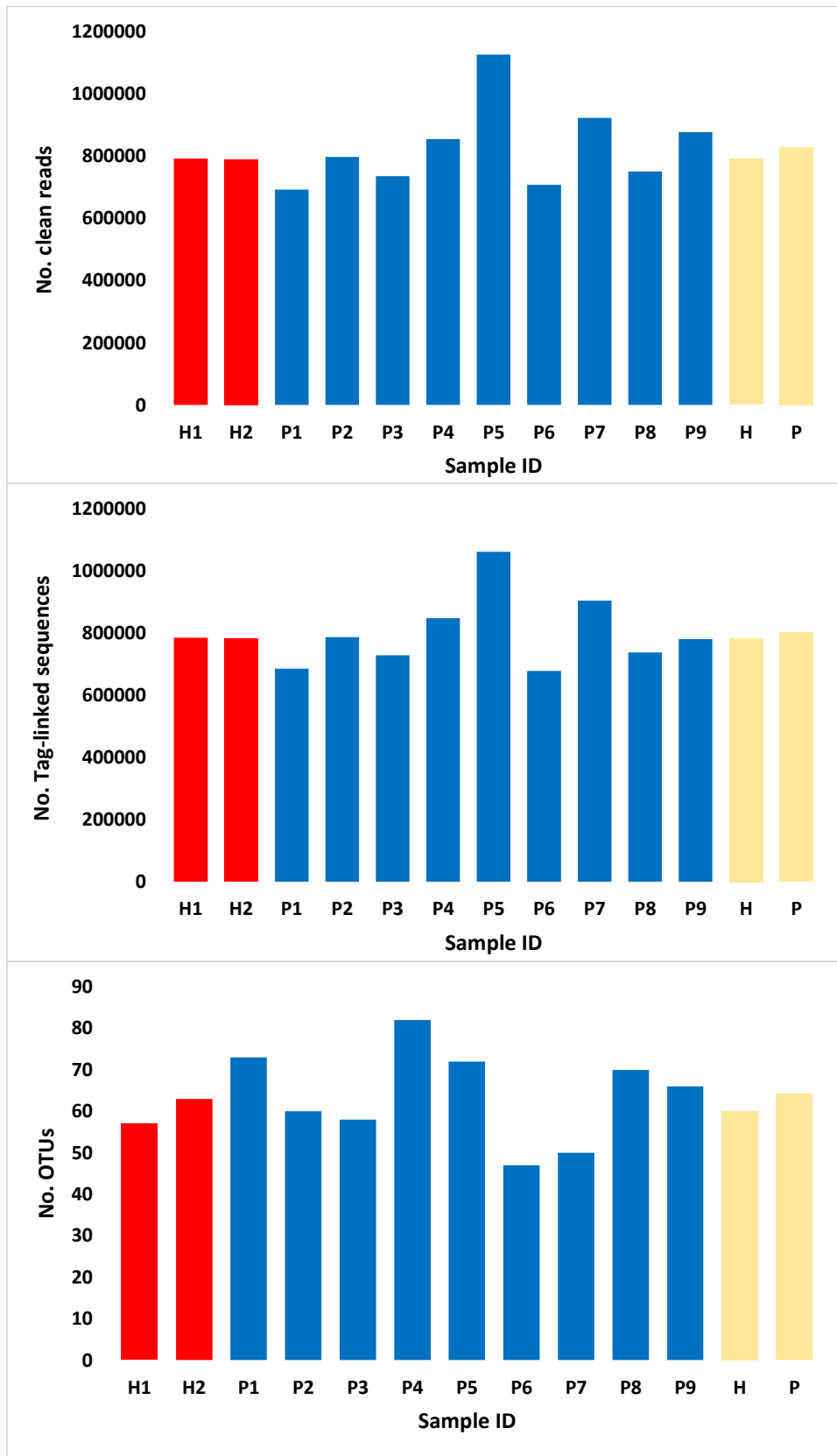
***Figure 1.*** *Comparison between numbers of clean reads, tag-linked sequences and no. OTUs among and across healthy individuals (H) and Entamoeba-infected patients (P). Healthy = red, Infected = blue, Mean = beige*

### Diversity and rarefaction curve analysis

Alpha diversity analysis, as Shannon and Simpson indices, was applied to detect complexity among species and the deviation due to the disease. Shannon and Simpson values reflect the species diversity of the community at both richness and evenness levels. But Shannon index poses more weight on sequence richness, while Simpson index poses more weight on evenness. With the same species richness, the greater the species evenness, the greater the community diversity. Trend of values for the two measures likely to be opposite. The results in *Figure 2* exactly aligned with the latter statement as subject with high richness shows low evenness and vice versa. Subjects with the highest richness and lowest evenness are P2 followed by P9, while those with the lowest richness and highest evenness are P3 followed by P7 (*Fig. 2*). Overall, the results indicated that alpha diversity slightly changed due to the infection indicating an influence of the parasite on the type and intensity of the other eukaryotic microbes hosted in the gut.
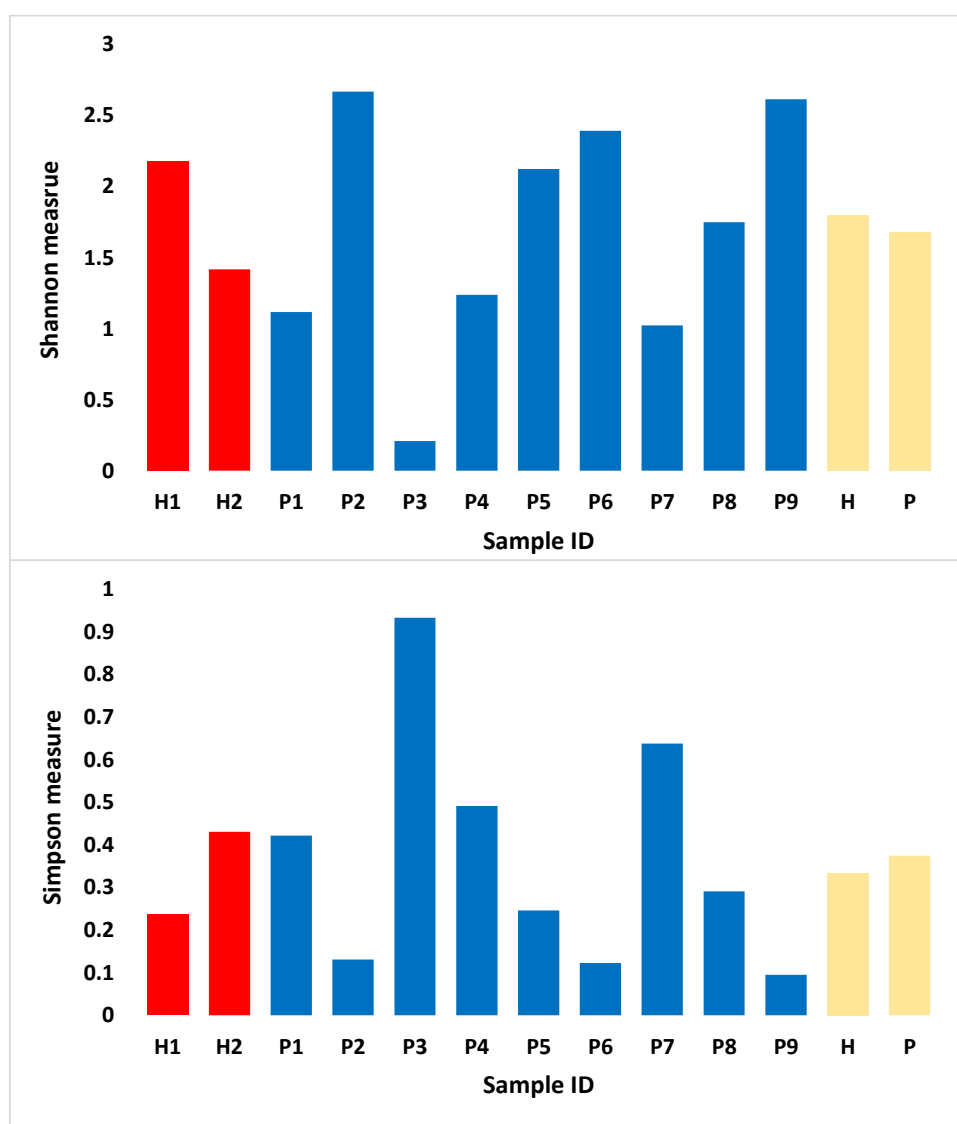


***Figure 2.*** *Alpha diversity measures among and across healthy (H) and Entamoeba-infected (P) individuals to describe sample or group richness (Shannon index) and evenness (Simpson index). Healthy = red, Infected = blue, Mean = beige*

Principal Coordinate Analysis (PCoA) was used in order to address the differences of OTU composition in different samples and groups. PCoA summarizes factors that are responsible for this difference. When similarity between two groups is high, then, the two groups are closely located. Based on the OTU abundance information, the relative abundance of each OTU in each sample and group was calculated, thus the PCoA of OTUs was plotted (*Fig. 3*). The PCoA plot completely showed a similar tendency in the distances within and between groups. The diversity of healthy subjects was localized in the positive directions of PCoA 1 (PC1) and PC2, while diversity of patients was higher towards the negative direction of PC2, and distributed evenly in the PC1 direction. These results indicated that the gut mycobiome signatures of the two groups differed due to amebiasis.
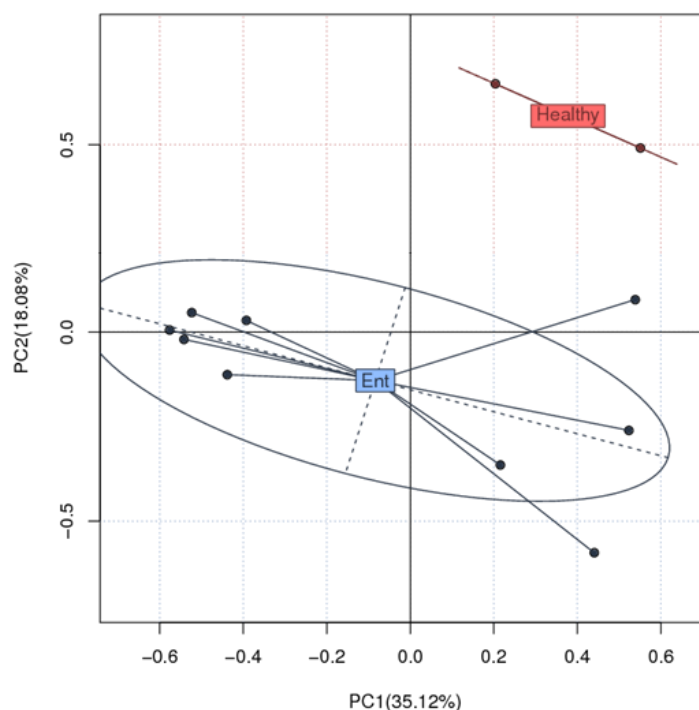


***Figure 3.*** *PCoA based on OTU abundance of different samples. Blue box indicates samples of Entamoeba-infected patients, while red box indicates samples of healthy individuals. X-axis is the first principal coordinate and Y-axis is the second. Number in brackets represents contributions of PCoAs to differences among samples. A dot represents each sample, and different colors represent different groups. Healthy = red, Infected = blue*

Rarefaction curves based on stacked number of OTUs and number of observed species were drawn (*Fig. 4*). Stacked number of OTUs were analyzed, as rarefaction measures, in order to describe the maximum depth permitted to retain all samples in the dataset for studying taxonomic relative abundance and to evaluate if produced data is enough to cover all species in the microbial community. The point where a stacked value of a given sample approaches dropping or moving downwards (indicated by arrow in the figure) indicates the sealing at which further analysis that holds homogeneity in the datasets can be done. The more the curve continues to climb with increasing sequencing reads, the higher the complexity in samples, i.e., there still be species uncovered by the sequencing data. The rarefaction measures indicated that the

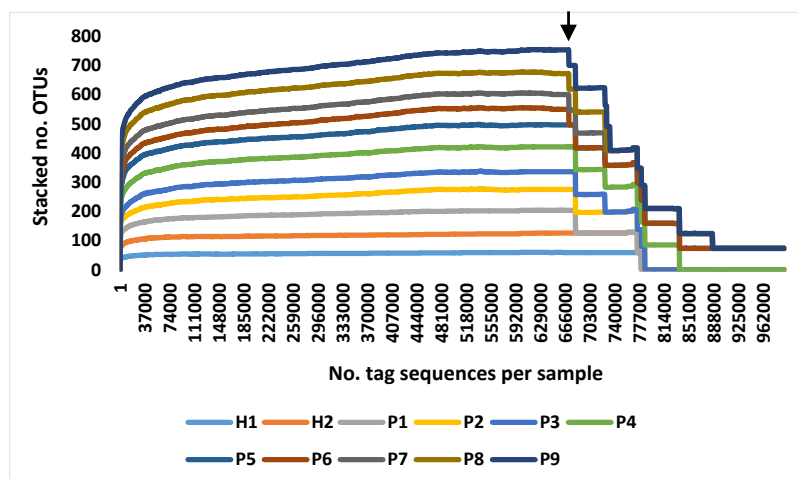maximum number of sequence reads to be used for further analysis of taxonomy abundance is as high as ~667,000 (*Fig. 4*).



***Figure 4.*** *Stacked number of OTU tags as rarefaction measures to describe the maximum depth permitted to retain all samples in the dataset for studying taxonomic relative abundance. The arrow indicates the suitable sample size for analyzing taxonomy abundance (~660,000 sequence reads) for healthy individuals (H) and Entamoeba-infected patients (P)*

### Description of gut microbiome signatures

Description of observed species detected from OTU annotation is shown in *Table A1* (see electronic appendix). Phylogenetic tree describing taxonomic groups of oral microbiomes at the phyla and genera levels are shown in *Figure 5* and organized in *Table 2*. A phylogenetic tree is a branching diagram showing the inferred evolutionary relationships among various biological taxa based upon similarities and differences in their physical or genetic characteristics. The evolution distance between taxa is closer if the branch length is shorter. Besides the taxa composition and abundance analysis, phylogenetic tree could further clarify the species evolution relationships. The results expectedly indicated that the most common phyla in the gut are Ascomycota (11 genera), Basidiomycota (nine genera) and Zygomycota (two genera).

Six genera of Ascomycota belong to family Saccharomycetaceae, while the rest belong to separate families. The nine genera of Basidiomycota belong to nine different families, while the two genera of Zygomycota belong to the family Mucoraceae (*Table 2*). Family information of one genus of Basidiomycota namely *Tilletiopsis* is not available. Abundance of individual OTUs across samples was also studied in which OTUs with number of sequences over 100,000 were considered highly abundant (*Fig. 6*). This criterion was met for a number of 18 out of the 146 OTUs (*Table A2; Fig. 6*). These selected OTUs are OTU1-OTU15, OTU86, OTU109 and OTU130. Description of these selected highly abundant OTUs in terms of taxonomy of their phyla, genera and/or species is shown in *Table A1*. After removal of 12 non-fungal OTUs (OTU2, OTU7, OTU9-OTU15, OTU86, OTU109 and OTU130), the results indicated that the highly abundant OTUs belong to *Pichia kudriavzevii* (OTU3), *Candida glabrata* (OTU4), *Saccharomyces cerevisiae* (OTU5) and OTU6 of phylum Ascomycota and *Rhodotorula mucilaginosa* of phylum Basidiomycota.
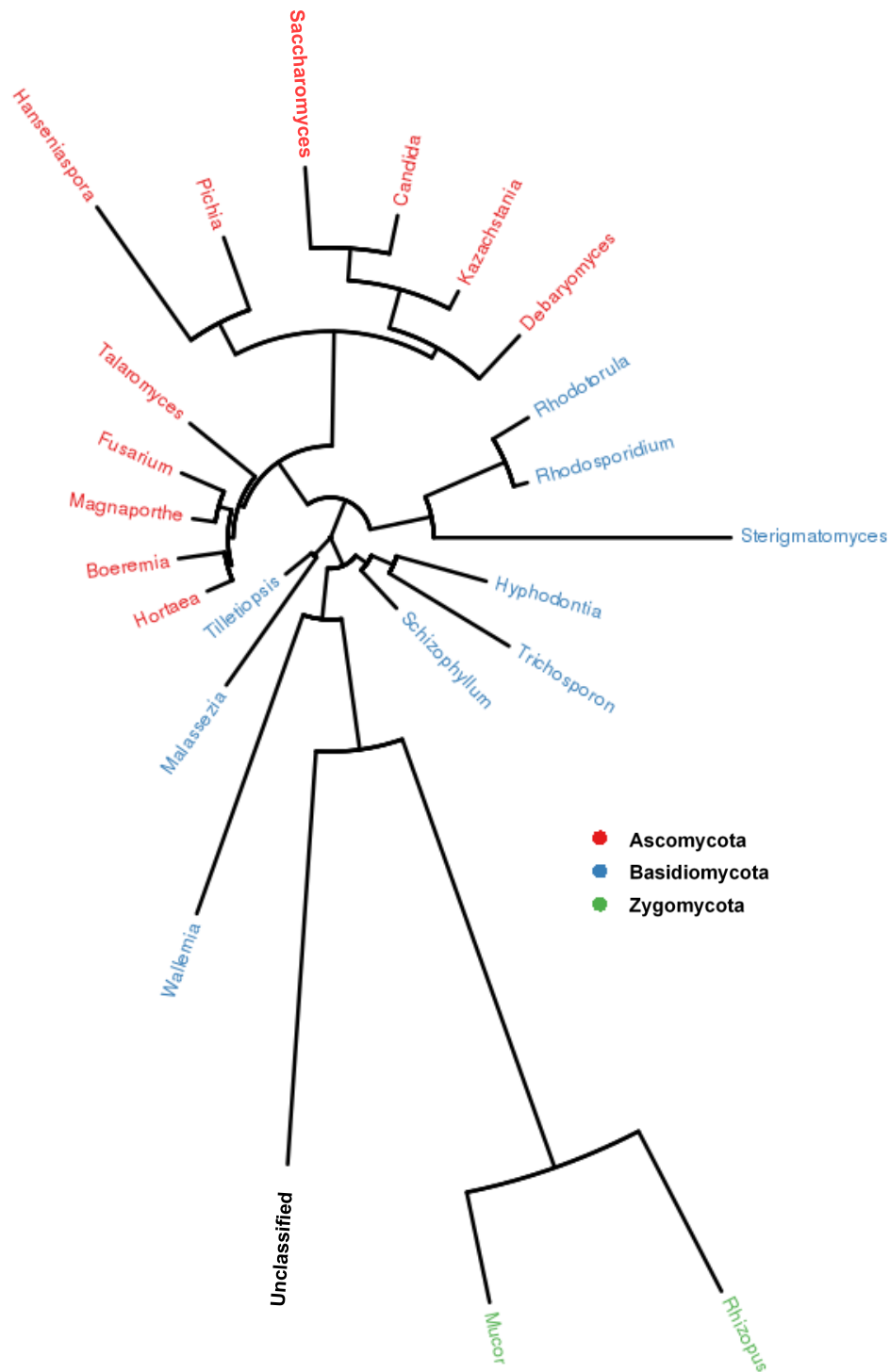
***Figure 5.** Genus level phylogenetic tree of gut mycobiome across healthy individuals and Entamoeba-infected patients. Genera with the same color belong to the same phylum*

Venn diagram indicated the existence of 71 common OTUs in both groups (*Fig. 7*; *Table A3*). The number of OTUs uniquely found in healthy subjects was 11, reduced to eight after removal of non-fungal taxa representing the taxonomic groups of the phyla

Ascomycota (5), Basidiomycota (2) and Zygomycota (1). The main taxa of Ascomycota are *Pichia kluyveri*, *Candida sake* and *Xylariales* spp., while those of Basidiomycota are *Malassezia obtusa* and unclassified genera of family Cystofilobasidiaceae and that of Zygomycota is *Rhizopus oryzae*. The number of OTUs uniquely found in healthy subjects was 64, reduced to 36 after removal of non-fungal taxa representing the taxonomic groups of the phyla Ascomycota (27), Basidiomycota (8) and Zygomycota (1).

The main taxa of Ascomycota are *Hortaea werneckii*, *Talaromyces purpurogenus*, *Debaryomyces hansenii*, *Kazachstania servazzii*, unclassified genera of the family Plectosphaerellaceae, *Hemibeltrania* sp. CL12WA, while those of Basidiomycota are *Schizophyllum radiatum*, *Hyphodontia rimosissima*, *Malassezia* spp., *Malassezia pachydermatis*, *Trichosporon coremiiforme*, *Rhodosporidium toruloides*, *Wallemia* spp. and that of Zygomycota is *Mucor racemosus*. None of these unique markers was consistent in either group in terms of existence as well as abundance. A larger number of samples might be required in order to reach a certain conclusion before we decide which of them can be considered as a biomarker for amebiasis.
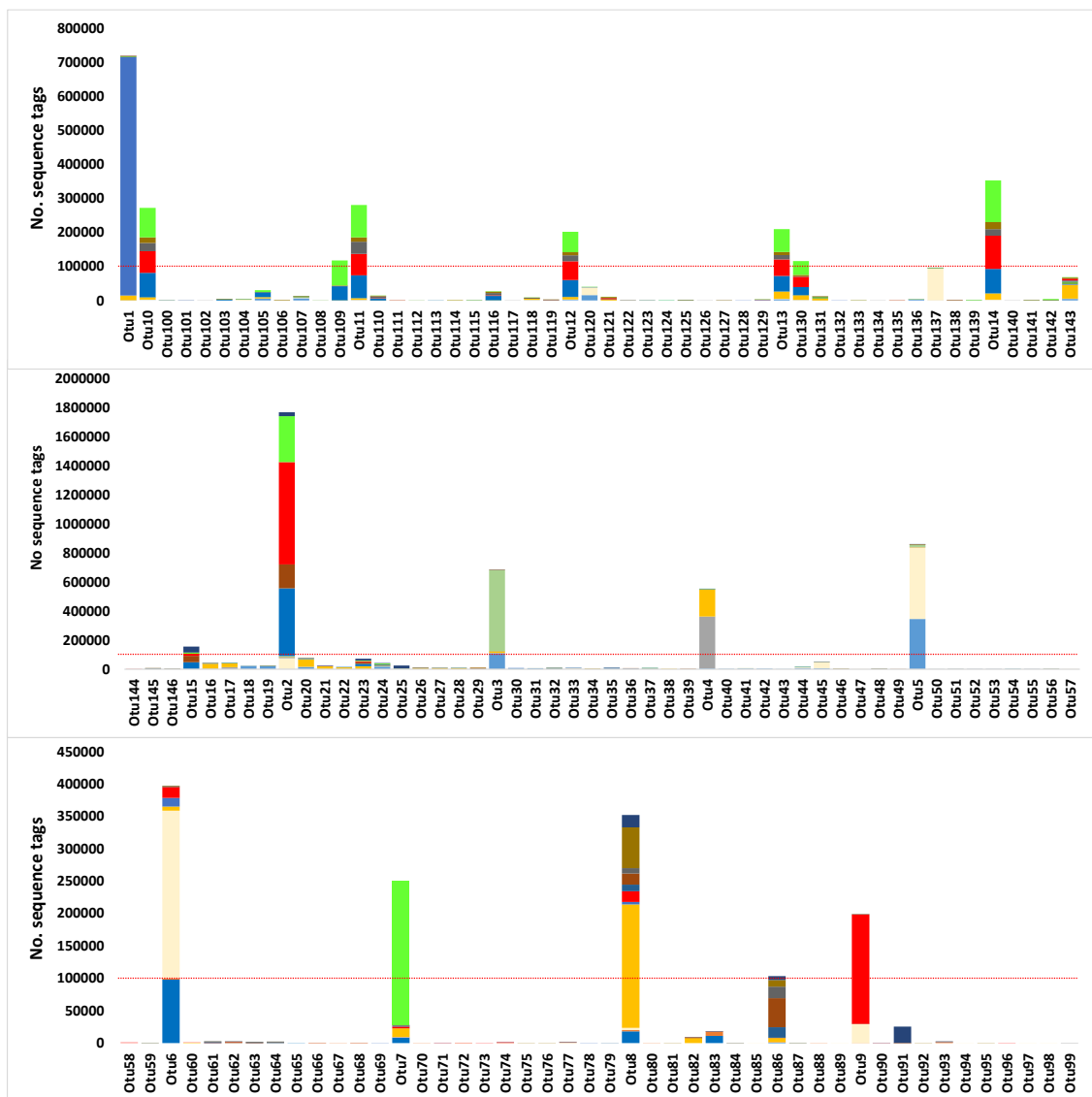


***Figure 6.*** *Stacked OTU abundance across healthy individuals and Entamoeba-infected patients*
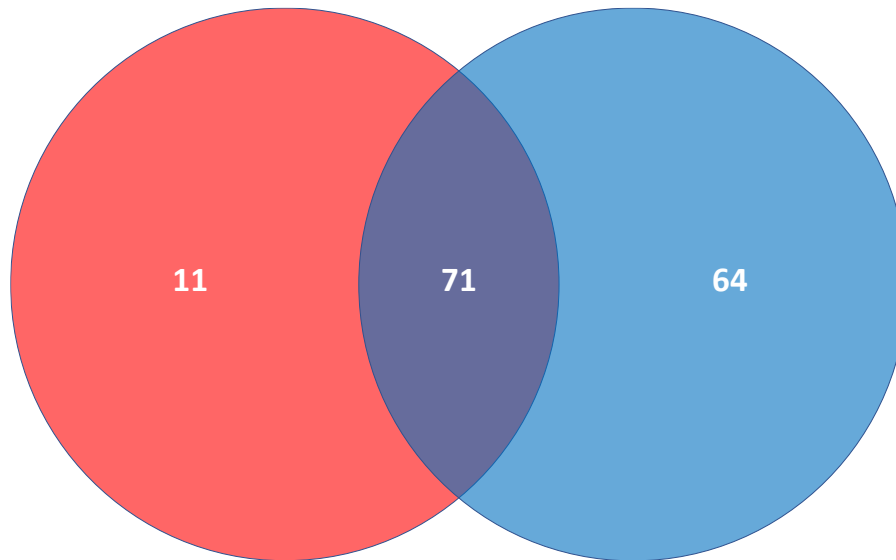
***Figure 7.*** *Venn diagram describing the unique (11 for healthy individuals and 64 Entamoeba-infected patients) and shared OTUs (71) between the two groups of samples. The OTUs of different groups are shown in Table A3 and description of these OTUs is shown on Tables A1 and A2. Healthy = red, Infected = blue*

***Table 2.*** *Taxonomy of gut mycobiome across healthy individuals and Entamoeba-infected patients*

| Phylum | Family | Genus |
|---|---|---|
| Ascomycota | Didymellaceae | *Boeremia* |
| | Saccharomycetaceae | *Kazachstania* |
| | | *Candida* |
| | | *Saccharomyces* |
| | | *Debaryomyces* |
| | | *Hanseniaspora* |
| | | *Pichia* |
| | Nectriaceae | *Fusarium* |
| | Teratosphaeriaceae | *Hortaea* |
| | Magnaporthaceae | *Magnaporthe* |
| | Trichocomaceae | *Talaromyces* |
| Basidiomycota | Hymenochaetaceae | *Hyphodontia* |
| | Sporidiobolaceae | *Rhodotorula* |
| | Ustilaginaceae | *Rhodosporidium* |
| | Schizophyllaceae | *Schizophyllum* |
| | Agaricostilbaceae | *Sterigmatomyces* |
| | N/A | *Tilletiopsis* |
| | Wallemiaceae | *Wallemia* |
| | Malasseziaceae | *Malassezia* |
| | Trichosporonaceae | *Trichosporon* |
| Zygomycota | Mucoraceae | *Mucor* |
| | | *Rhizopus* |

The analysis of weighted and unweighted unifrac diversity distances, as a beta diversity measures, describing the overall structures of gut mycobiomes of different healthy (H) and infected (P) subjects is shown in *Figure 8*. As indicated earlier, the weighted diversity considers the abundance of the different OTUs, while unweighted diversity considers presence or absence of a given OTU. *Figure 8* indicates that analysis based on weighted unifrac diversity distances successfully resulted in complete separation between the two groups, e.g., healthy subjects (H) and *Entamoeba*-infected patients (P). However, analysis based on unweighted unifrac diversity distances resulted in no separation between the two groups although a large number of group-specific OTUs was recovered. As indicated above, these unique or group-specific OTUs were not consistently shown in all subjects of the group. Thus, no particular gut mycobiome signature was detected for either group. As this inconsistency was shown for all unique OTUs, analysis was not able to make a complete discrimination between the two groups based on the unique OTUs.

The results shown in *Figure A2*, displaying the beta diversity heat maps of weighted and unweighted unifrac diversity distances between the two groups supports the latter conclusions. A heat map is a graphical representation of data where the subject distance values contained in a matrix are represented as colors. *Figure A2* also indicated that P4 followed by P1 are the most divergent within the infected subjects (P). Gut mycobiome signature of P1 was shown to be the closest to the signatures of the two healthy subjects. However, the most closely related infected subjects in terms of gut mycobiome signatures are P5, P6, P7 and P10. The differential signatures among infected subjects might reflect the differential severity of the disease.

## *Differential abundance of microbes due to amebiasis*

Differential abundance of microbes of different healthy (H) and infected (P) subjects was studied at the phylum, family, genus and species levels (*Fig. 9, 10, 11* and *12*, respectively). The results indicated that a number of four phyla, six families, nine genera and ten species showed considerable changes in gut mycobiome signatures within and among groups. The phyla Ascomycota and Basidiomycota were consistently shown in all subjects, albeit different in abundance, while Chytridiomycota and Cnidaria were only shown with relatively low abundances, in a single subject, e.g., P2 and H2, respectively. Therefore, no conclusions can be taken for the latter two phyla in terms of relative abundance (*Fig. 9*). The results at the family level indicated consistency in two out of the six families, e.g., Saccharomycetaceae and Trichocomaceae (*Fig. 10*). The two families belong to phylum Ascomycota. The results at the genus level indicated consistency in three out of the nine genera, e.g., *Candida* spp., *Malassezia* spp. and *Saccharomyces* spp. (*Fig. 11*). The first two genera were relatively highly abundant in the infected subjects as compared with the healthy subjects. Opposite conclusion was reached for the third genus, e.g., *Saccharomyces* spp., where it was relatively highly abundant in the healthy subjects as compared with the infected subjects (*Fig. 11*). The results at the species level indicated consistency in only one, e.g., *Saccharomyces cerevisiae*, out of the ten species (*Fig. 12*). Accordingly, it was decided to further analyze the results of the Ascomycota at the phylum level, Saccharomycetaceae and Trichocomaceae at the family level, *Malassezia* spp. and *Saccharomyces* spp. at the genus level and *Saccharomyces cerevisiae* at the species level (*Figs. 13* and *14*).
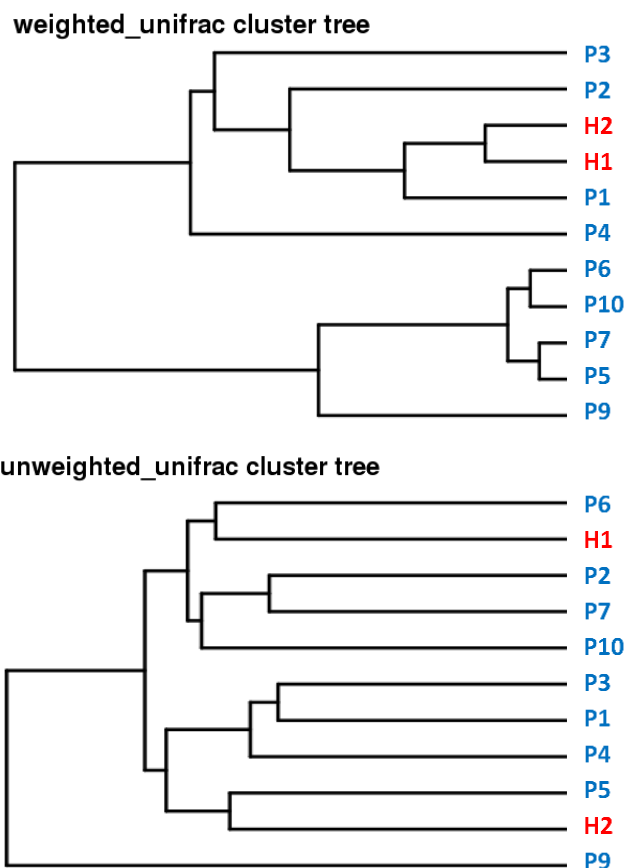
***Figure 8.*** *Dendrogram trees describing the weighted and unweighted unifrac diversity distances among healthy individuals (H) and Entamoeba-infected patients (P). Healthy = red, Infected = blue*

Among these taxa, unclassified species of the genus *Malassezia* belonging to phylum Basidiomycota showed high abundance in infected subjects as compared with healthy subjects (*Fig. 13*), while the rest of taxa belonging to phylum Ascomycota showed opposite results (*Fig. 14*).

Linear discriminant analysis (or LDA) effect size (LEFSE) is a software that determines the entire features (organisms, clades, operational taxonomic units, genes, or functions) likely explaining differences among groups. Such features were determined by coupling standard tests for statistical significance with additional tests encoding biological consistency and effect relevance. In the present study, LEFSE was based on determining the features in the different taxa in order to detect the ones that can conclusively be considered as biomarkers for amebiasis. The analysis shown in *Figure 15* indicated that the family Saccharomycetaceae, of which *Saccharomyces cerevisiae* is a downstream taxon, is the only taxon meeting the standard criteria set of the analysis. Interestingly, there is no particular OTU specific to this family, however, the most consistent OTU with this result, e.g., OTU107, belongs to the genus *Saccharomyces*, also a downstream taxon (*Table A4*). This taxon, referring to OTU107, along with its downstream species, referring to OTU5, share a negative biomarker of amebiasis (*Fig. 15*). This result supports that generated when studying the relative abundance at the phylum, family, genus and species levels. Considering the richness of

the two OTUs (*Table A2*), we claim that OTU5, referring to *Saccharomyces cerevisiae*, is more appropriate as the new negative biomarker for amebiasis.

## Discussion

Preliminary detection of the gut diseases usually involves microscopic observation with the fecal samples. Although this approach is not sensitive, it remains to be the gold standard utilized in almost all parasitological diagnostic laboratories (McHardy et al., 2014). Drawbacks of this approach arise when only few eggs exist in samples, thus, other methods such as serology-based, e.g., ELISA (Ungar, 1990), might be a better alternative. More recently, molecular approaches have been used as a promising tool for differentiating and detecting stool parasites (Tavares et al., 2011), with the first molecular study to detect mycobiome signatures published in 2008 (Scanlan and Marchesi, 2008).
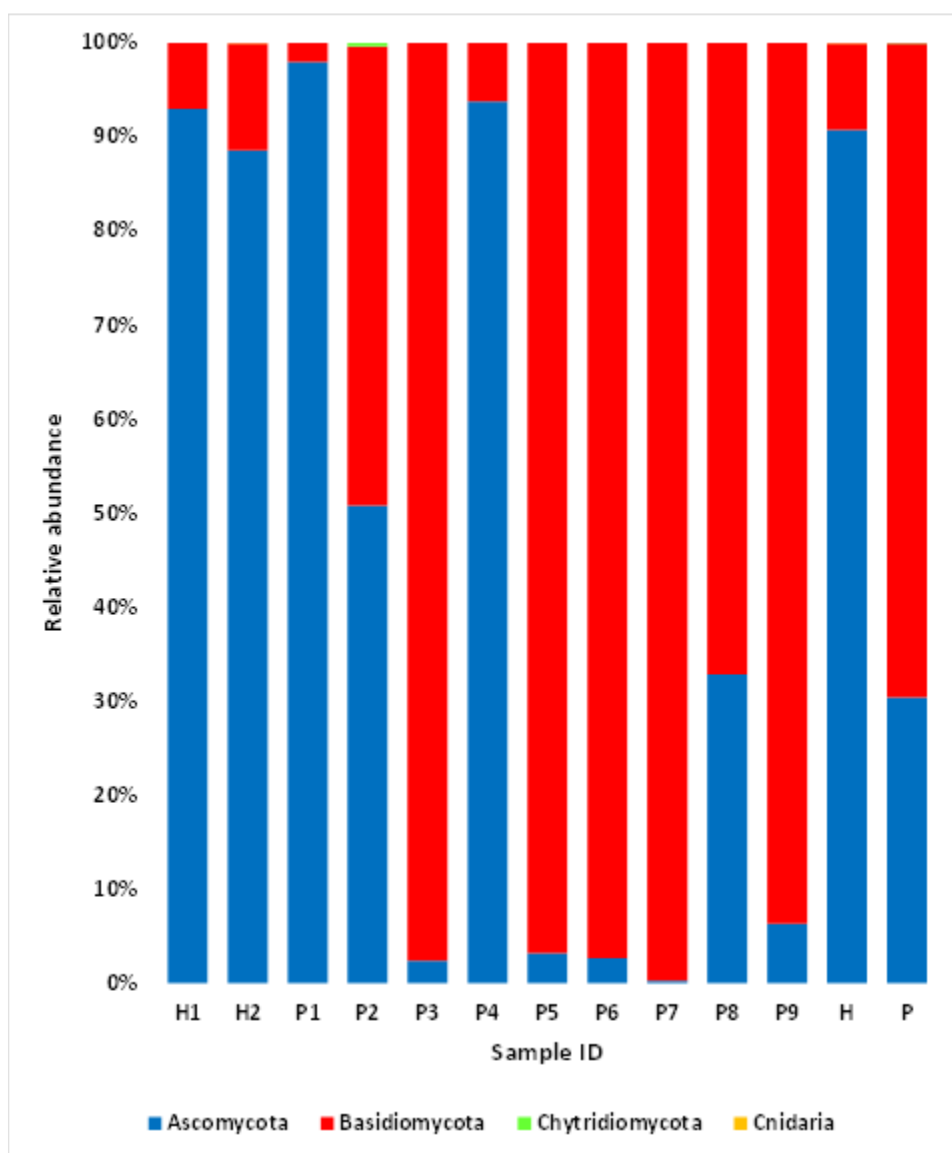


***Figure 9.*** *Relative abundance at the phylum level as measured by Metastats at the sample and group levels for healthy individuals (H) and Entamoeba-infected patients (P)*
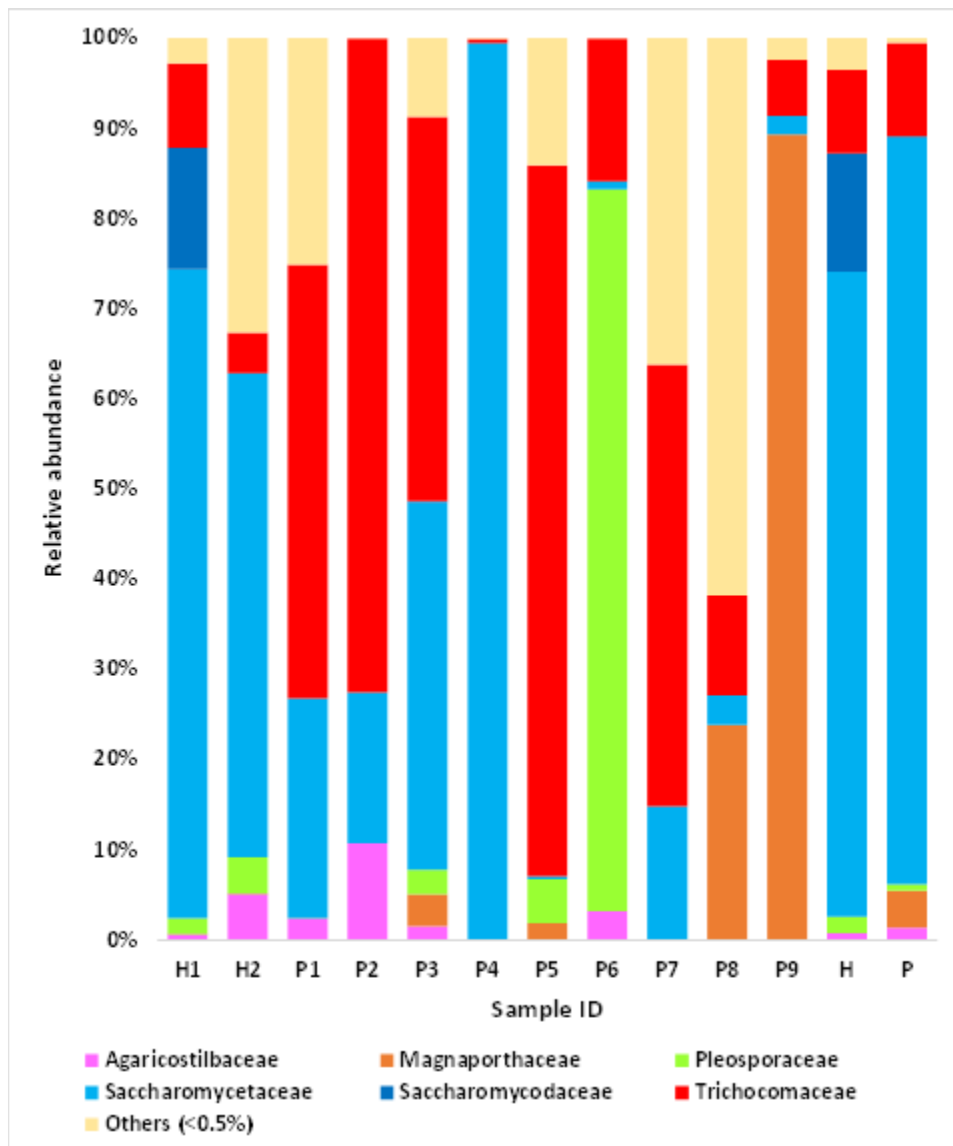
***Figure 10.*** *Relative abundance at the family level as measured by Metastats at the sample and group levels for healthy individuals (H) and Entamoeba-infected patients (P)*

Nowadays, next generation sequencing (NGS) has proven to be the ideal approach in detecting the microbiome signatures and providing an accurate figures of microbe intensity (Stensvold et al., 2010). Such high-throughput technologies are employed in detecting microbiome signatures in different human organs in relation to disease status, e.g., amebiasis (Gao et al., 2017; Jo et al., 2017). The approach is based on the use of 16S ribosomal RNA (rRNA) gene for bacteriome, while 18S or ITS (internal transcribed spacer) for fungal microbiome or mycobiome (Belda-Ferre et al., 2012; Ahn et al., 2015; Johansson et al., 2016; Zhou et al., 2016). Although, internal transcribed spacer (ITS) seems to be more appropriate than 18S rDNA for analyzing human gut mycobiome at the level of clades and species (Hamad et al., 2016), it still difficult to use it in making a proper fungal metagenomic analysis due to the absence of a well-established or rich fungal ITS database for OTU annotation (Tang et al., 2015).

There were two major issues to be considered before we approach analyzing mycobiome. The first is concerned with the influence of diet at the harvest time and the second is concerned with the lack of a database specific for mycobiome at the species level. The preliminary observation made during the analysis was the presence of a large number of non-fungal OTUs at the species level, while none at the levels of the phylum, family or genus (*Figs. 9–12*).
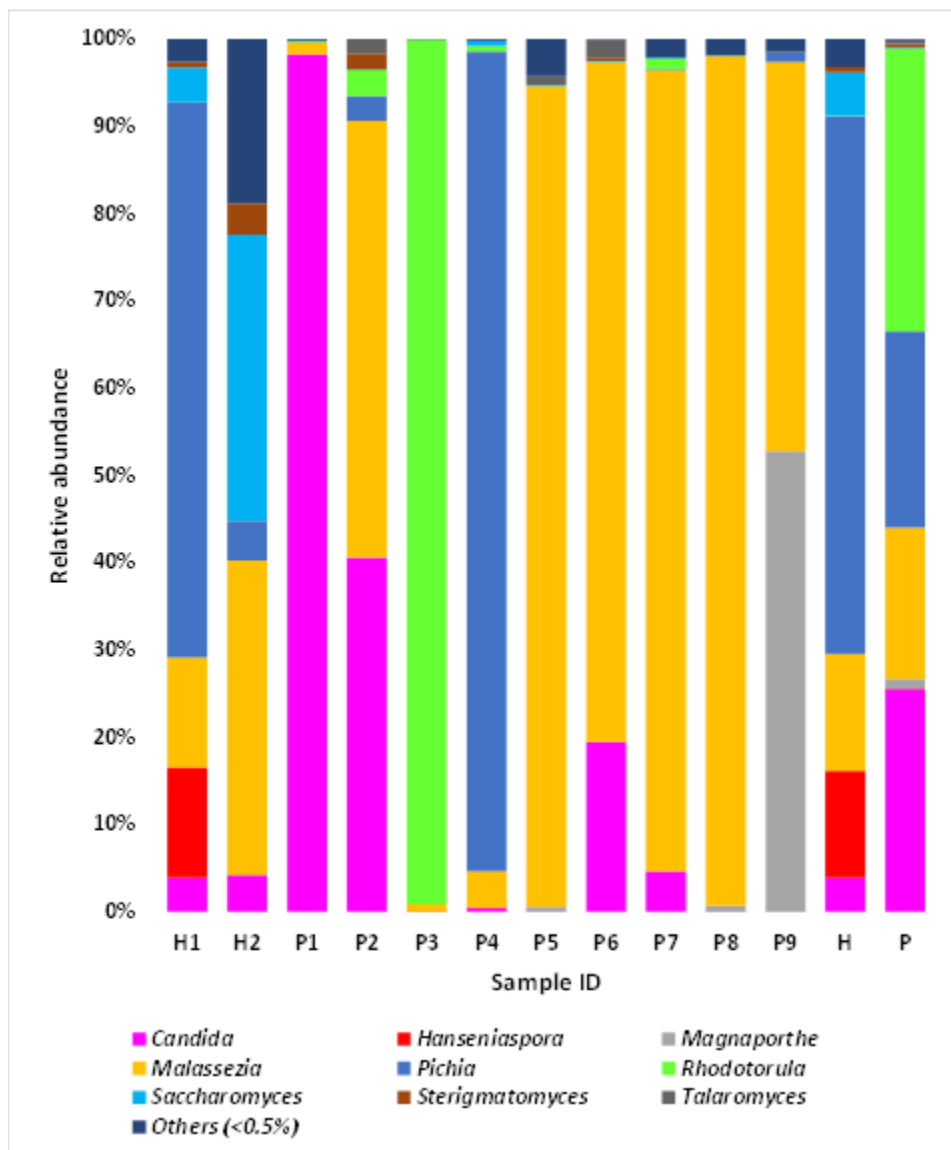


***Figure 11.*** *Relative abundance at the genus level as measured by Metastats at the sample and group levels for healthy individuals (H) and Entamoeba-infected patients (P)*

This indicates that the data recovered at the species level requires further removal of non-fungal OTUs in order to get accurate figure of the relative abundance at the species level. *Figure 16* indicates the comparison of relative abundance of mycobiomes of the two healthy subjects at the species level before and after removal of non-fungal species. The latter are four plant species namely *Oryza sativa* (rice), *Triticum aestivum* (bread wheat), *Sesamum indicum* (sesame) and *Prunus mume* (Japanese apricot). As shown in

the figure, fungal species showed a more accurate results for their relative abundance after correcting the data (*Fig. 16*). The datasets of OTU taxonomy also required further removal of uncharacterized fungi. In addition, OTUs for a single species, genus, family, order or phylum were merged to avoid redundancy in the taxon and getting a better estimate of its relative abundance, thus, can give a better estimate of the mycobiome signature of the *Entamoeba*-infected subject. The corrected number of OTUs became 46 out of the total 146 OTUs recovered after the initial annotation step. A recent study discussed the influence of diet and the environmental factors to gut mycobiome composition, and indicated that gut mycobiome is less stable than the bacteriome, and is subject to the two factors, thus, usually changes from day to day and over time (Hallen-Adams and Suhr, 2017). David et al. (2014) further indicated that dietary-induced changes in the gut microbiome rely on the type of food consumed (either plant- or animal-based). Hugon et al. (2017) indicated that bacterial composition responds to nutrient availability, while fungal composition appears to be driven by food colonization. The latter evidences were fully supported by gut mycobiome analysis of the present study.
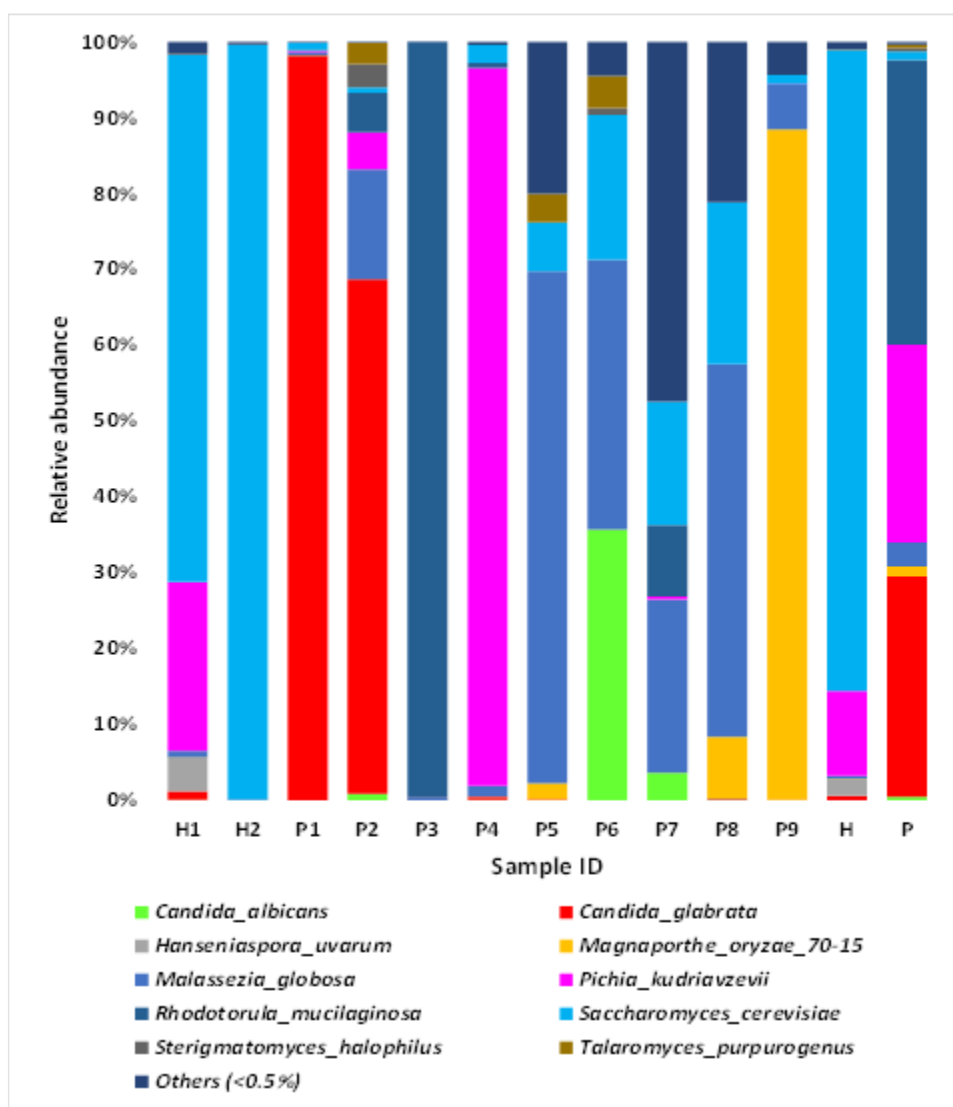


*Figure 12. Relative abundance at the species level as measured by Metastats at the sample and group levels for healthy individuals (H) and Entamoeba-infected patients (P)*
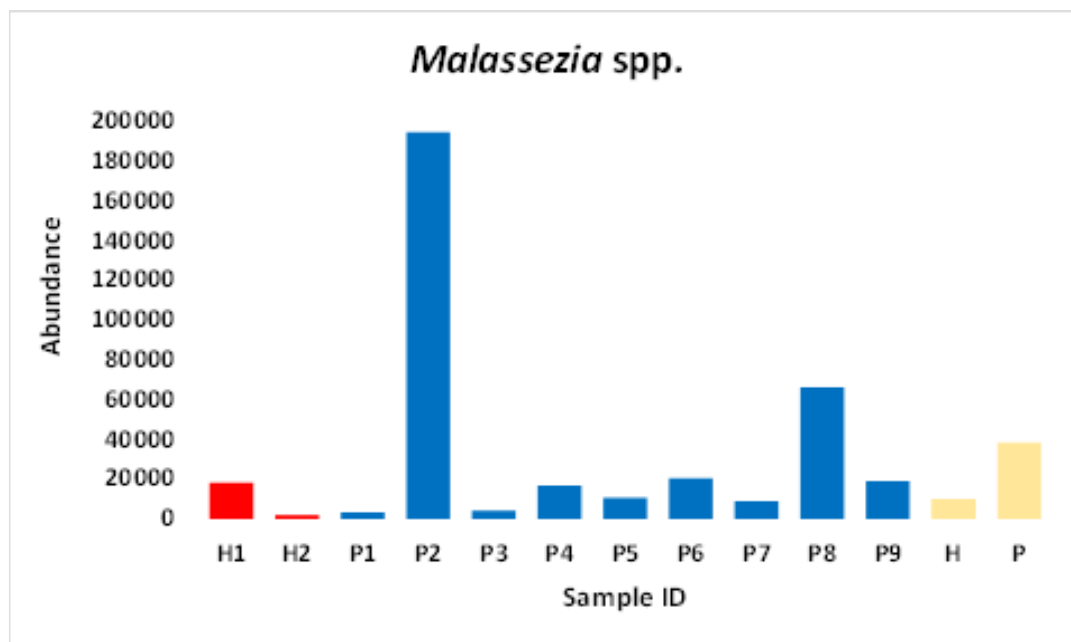
***Figure 13.*** *High abundance of the unassigned species of the genus Malassezia spp. within or across Entamoeba-infected patients (P) as compared with healthy individuals (H)*

Based on recent molecular phylogenetic classification, the taxonomy of human gut eukaryotes includes the five major groups of Amoebozoa, Opisthokonta (including fungi), Excavata, Sar and Archaeplastida (Adl et al., 2012; Hamad et al., 2016). Fungal species, including yeasts and filamentous fungi, are considered as the most abundant group of eukaryotes in the human gut. It is likely that most fungi species represent some sort of food contamination (Rajilic-Stojanovic and de Vos, 2014), thus, their role in the human gut microbiota remains unclear. While, eukaryotic parasites, e.g., protozoa and helminths, exist in the human gut in lower abundance (Rajilic-Stojanovic and de Vos, 2014). Gut mycobiome was proven to be comprised of the four major phyla of Ascomycota, Basidiomycota, Zygomycota and Microsporidia (Moran et al., 2012; Rajilic-Stojanovic and de Vos, 2014). In the present study, only the first three phyla were detected in the mycobiomes of healthy and infected subjects (*Table 2*). Other reports indicated that gut mycobiota in healthy subjects mainly comprises Ascomycota and Basidiomycota as the most abundant phyla, and the *Candida*, *Saccharomyces* and *Malassezia* and as the most abundant genera (Ott et al., 2008; Moran et al., 2012; Hoffmann et al., 2013; Rajilic-Stojanovic and de Vos, 2014; Nash et al., 2017). The results of the present study completely align with those of the latter reports. We also found that two species of *Candida* as well as *Saccharomyces cerevisiae*, followed by *Malassezia globosa* and *Pichia kudriavzevii* are the most dominating species across subjects (*Fig. 12*). The latter species are natural, asymptomatic fungi of the human microbiome, but can become opportunistic pathogens under specific circumstances and cause candidiasis or gut microbial dysbiosis (Vogel et al., 2007; Saleh et al., 2011; Rajilic-Stojanovic and de Vos, 2014). Several other species belonging to genera *Penicillium* and *Aspergillus* have also been detected in GI tract in many researches (Scanlan and Marchesi, 2008; Gouba et al., 2013, 2014). None of the two species was detected across subject in the present study.
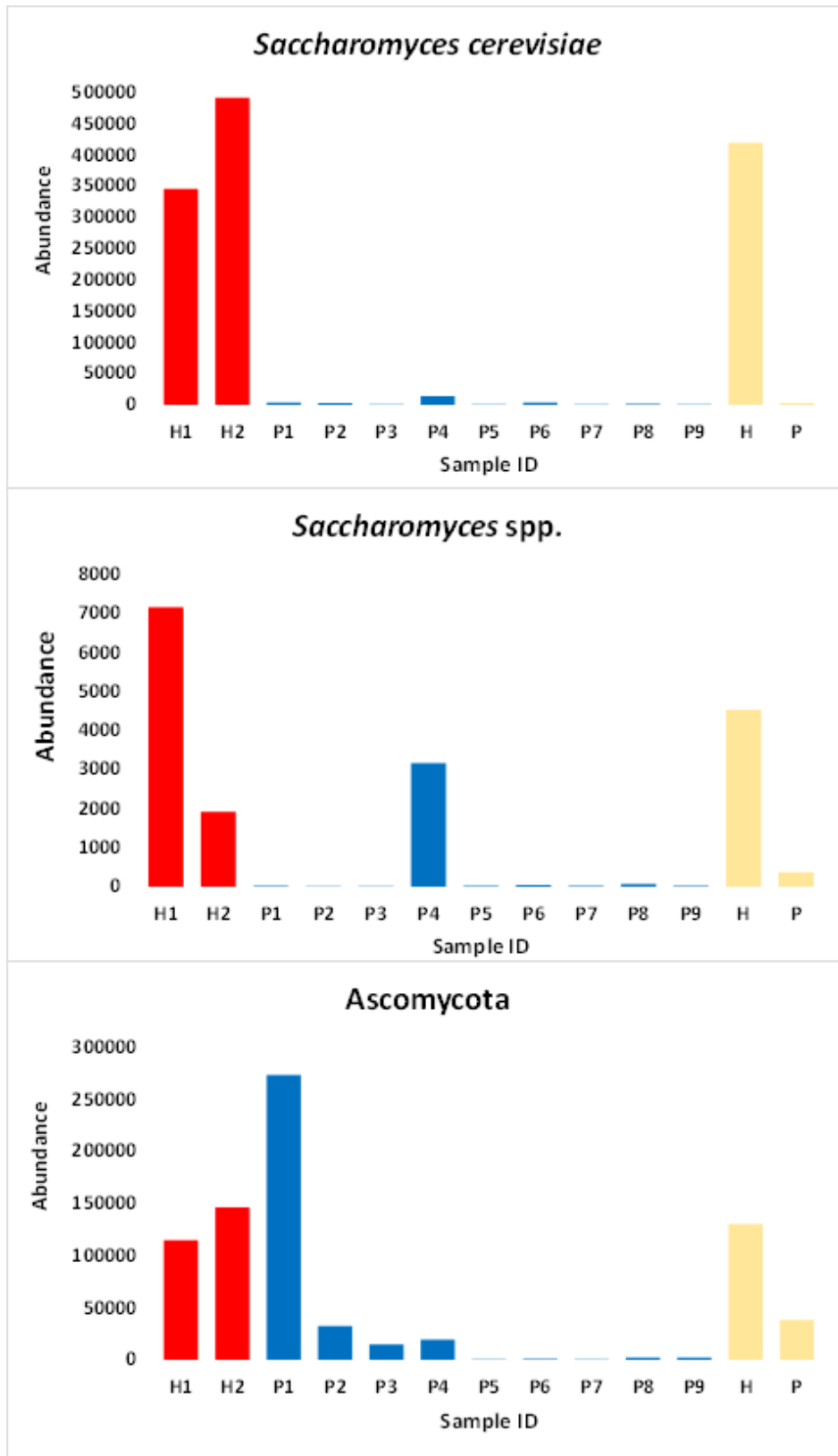
***Figure 14.*** *Low abundance of Saccharomyces cerevisiae, unassigned species of Sccharomyces spp. along with their phylum Ascomycota as well as member of the family Trycocomaceae within or across Entamoeba-infected patients (P) as compared with healthy individuals (H). Healthy = red, Infected = blue, Mean = beige*
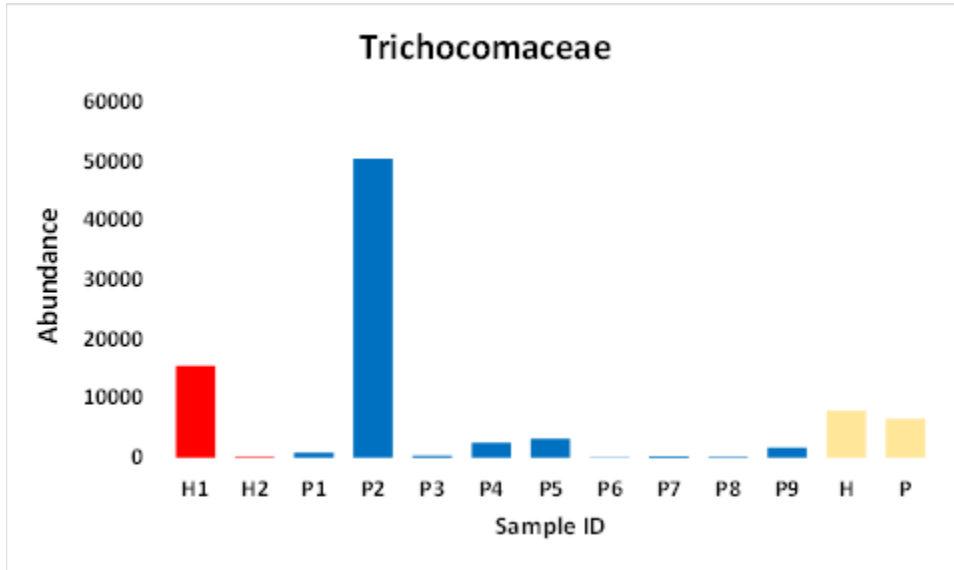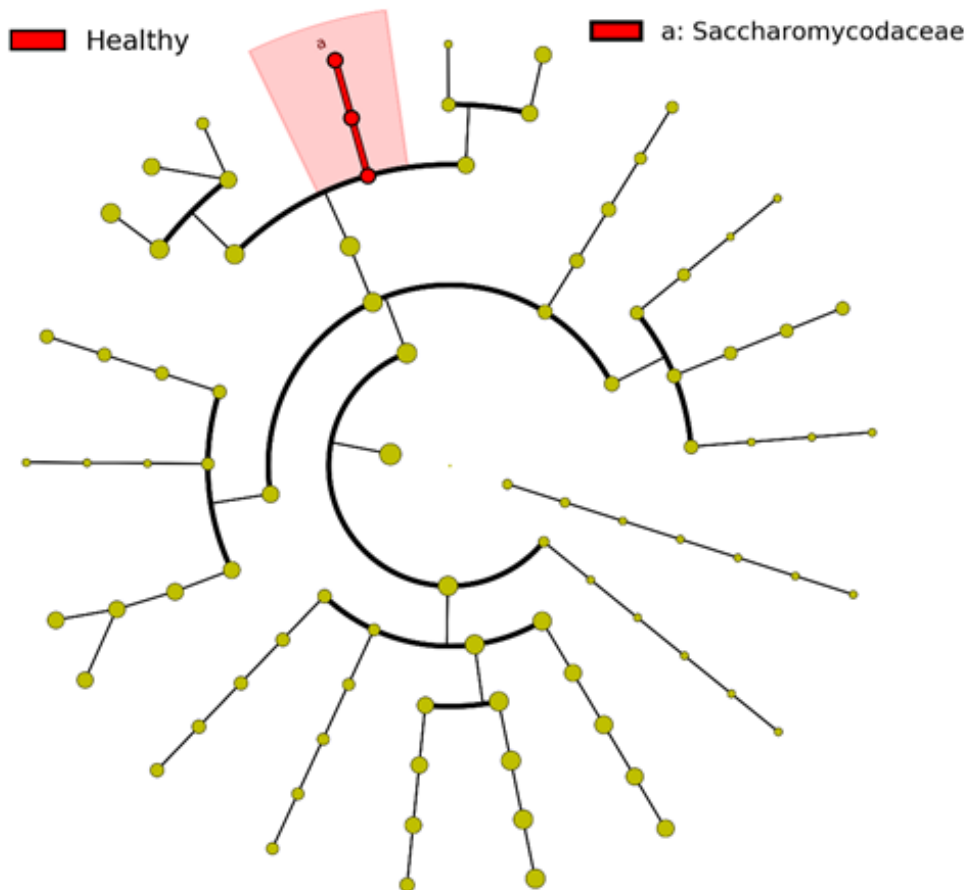
***Figure 14.*** *Continued*



***Figure 15.*** *LEFse cladogram describing a negative biomarkers for amebiasis. The new important microbe biomarker in the healthy group and the biomark name are listed in the upper right corner. The yellow notes represent the other biomarker which do not show any importance in either groups for healthy individuals (H) or Entamoeba-infected patients (P)*
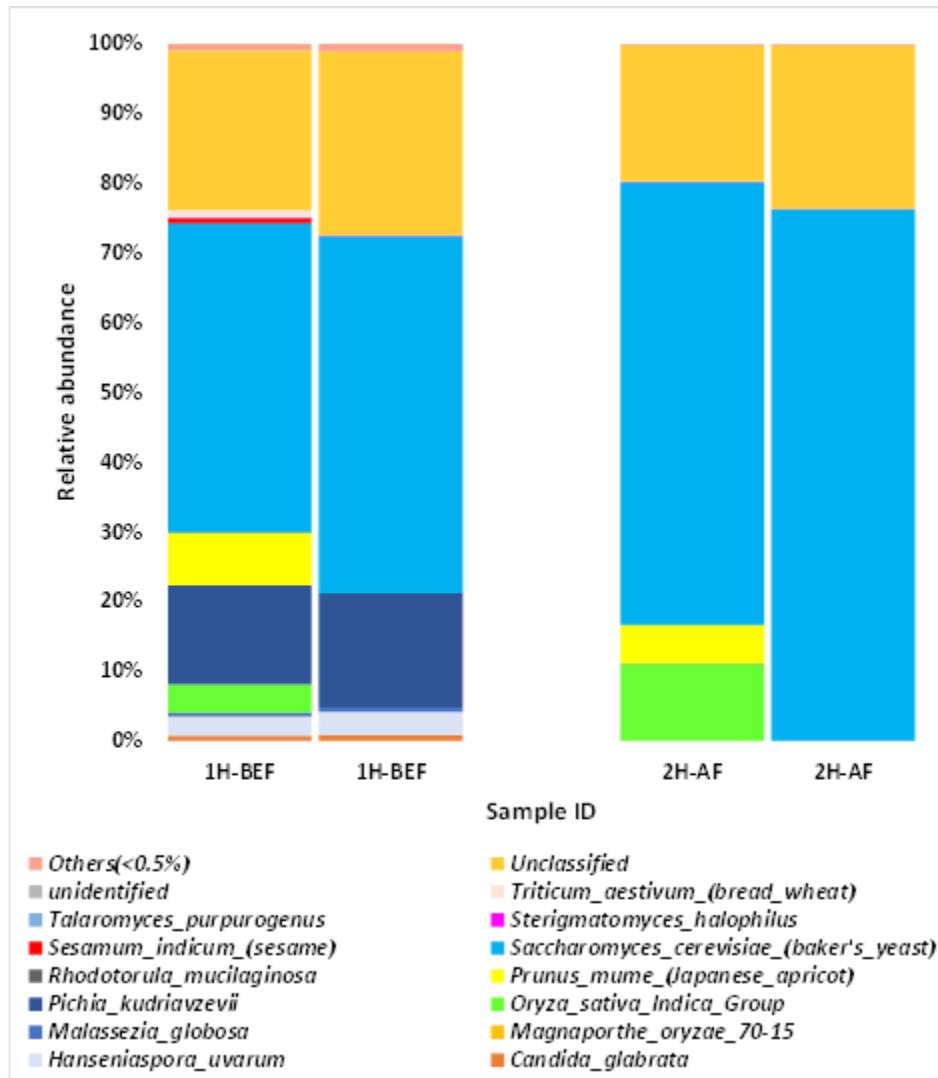
***Figure 16.** Differential relative abundance of the two healthy subjects (H) before (BEF) and after (AF) removal of non-fungal information. Description of OTUs for mycobiome species is shown in Table A4*

Species of the genus *Malassezia* are well-known for the inability to synthesize their own lipids, thus, they are completely dependent on their host for nutritional needs (Gupta et al., 2004). In the present study, abundance of this genus increased in mycobiomes of the infected subjects. In other words, there is a positive relationship in the present study in the abundances of *Malassezia* and *Entamoeba histolytica*. Upon infection, *E. histolytica* was reported to exploit lipids and carbohydrates present in the colonic environment (Thibeaux et al., 2013). This indicates that lipids are available for other microbes. Therefore, there is a possibility that *E. hystolytica* be indicative of lipids available for *Malassezia* (*Fig. 17*). The latter conclusion might require further experimentation and support as genes involved in lipids metabolism were recently reported to be upregulated under growth stress in the study of RNA-Seq data of *Entamoeba* (Naiyer et al., 2019b). Expectedly, the known niche for this genus is the skin that is rich of adipose cells, while, surprisingly, it was reported in significant abundance in the gut (Gouba et al., 2013; Cano et al., 2014; Suhr et al., 2016).

Therefore, there should be other resources of lipids for this genus in the gut. There are two other groups of fungi of the phylum Ascomycota that reported to colonize the healthy human GI tract although gut is not their primary colonizer (Hugon et al., 2017). They are *Cladosporium* and yeasts of the family Dipodascaceae (e.g., *Geotrichum* and *Galactomyces*). Several evidences suggest that the majority of the fungi in GI tract are indigenous (Iliev et al., 2012; Dollive et al., 2013). Family Trichocomaceae also belongs to phylum Ascomycota and includes species *Aspergillus nidulans*, an ascomycetous fungus that is used extensively for studying processes in fungi. This species was proven to be closely related to the *S. cerevisiae* (Class: Hemiascomycetes) in several biological processes, but little is known about the family it derived from (https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/trichocomaceae).

The dimorphic yeast *S. cerevisiae* is widely distributed in the environment and considered commensal upon being colonized in the gut, however, it can occasionally cause superficial and systemic infections (Byron et al., 1995; Munoz et al., 2005; De Llanos et al., 2006). This species is detected frequently in human fecal samples and abundance is directly associated with food ingestion prior to sampling (Scanlan and Marchesi, 2008). The ability of this species to grow at 37 °C makes it likely contributes to gut microbial ecology (Hugon et al., 2017). The potential for *S. cerevisiae* to persist and respond to the gut environment does not support their possible role in the gut, as its preferred niche is the open environmental (soil, air, plant matter). More recently, an article documented a protective role of *S. cerevisiae* inflammatory disorders (Jiang et al., 2017), while other article showed that *S. cerevisiae* reduces bacterial growth and colonization, and inhibits the adhesion of enterotoxigenic *E. coli* (ETEC) (Roussel et al., 2018). The gut mycobiome analysis in the present study indicated that *S. cerevisiae* proved to be a negative biomarker of amebiasis. At the relative abundance level, *S. cerevisiae* was shown to dominate in gut mycobiomes of the healthy subject. There are two possible speculations that can be reached based on the relative abundance of this species. The first is that the high abundance of *S. cerevisiae* might block the growth/spread of *E. histolytica* trophozoites by different possible ways, thus *S. cerevisiae* dominates in the healthy subjects. The second is that *E. histolytica* might propagate due to the low abundance of *S. cerevisiae*, thus, *E. histolytica* might dominate in the infected subjects. We speculate that the two species cross-talk via a number of processes basically relied on glucose level as well as growth condition (aerobic or anaerobic) in the gut, besides, the possible occurrence of oxidative stresses (*Fig. 17*).

Chitin formation is a process made by trophozoite during encystation. Chitin in the cyst wall of *E. histolytica* is synthesized by two chitin synthases, one of which is unique to it (EhCHS-1) and the other is analogue to those of insects, fungi and nematodes (EhCHS-2) (Tellam et al., 2000; Roncero, 2002; Ruiz-Herrera, 2002). In *S. cerevisiae*, chitin is produced by three specialized chitin synthases (Chs1p, Chs2p and Chs3p) (Bulawa, 1993; Choi and Cabib, 1994; Henar Valdivieso, 1999; Roncero, 2002). The first enzyme functions during cytokinesis, while the second makes the chitin in the primary septum and the third makes the bulk of cellular chitin. Yeasts transformed with the unique EhCHS-1 was proven to be functional and chitin was detected in their cells. Action of this enzyme seems to be extracellular towards the cell walls, thus, we speculate that yeast cells might compete with trophozoite for the use of the amoebic enzyme (EhCHS-1), which eventually retards the spread of the parasite by inhibiting the encystation process.
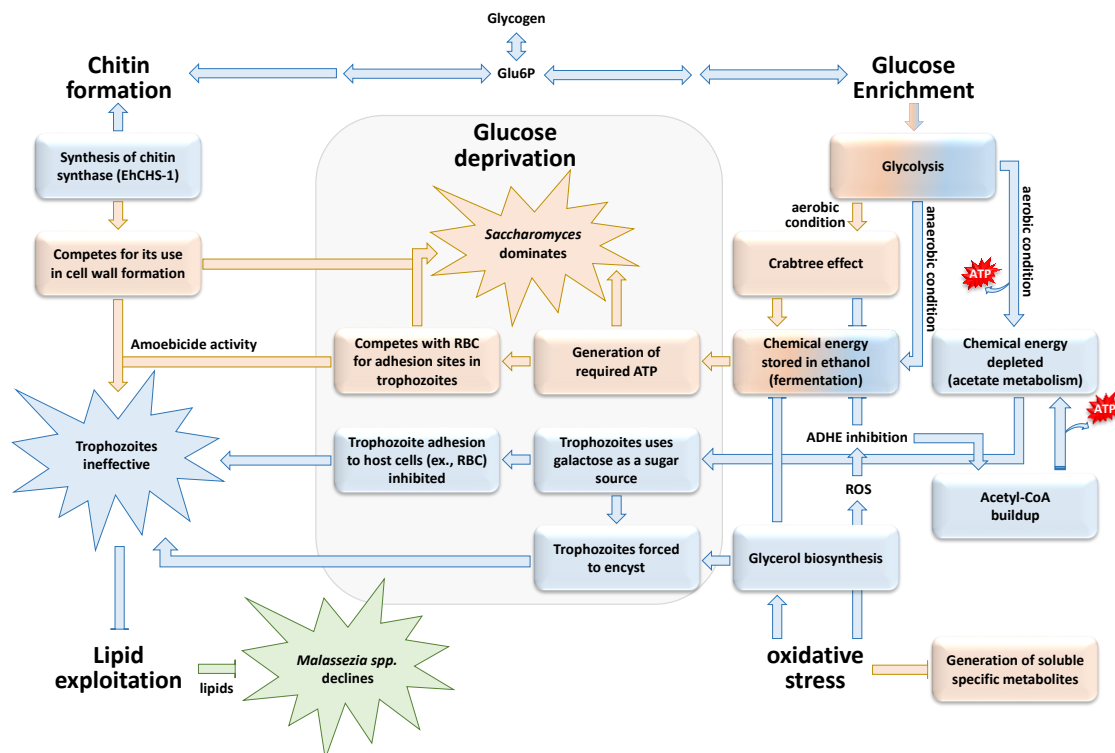
*Figure 17.* *Roles of the new negative biomarker for amebiasis, e.g., Saccharomyces cerevisiae. Main influence of the biomarker possibly occurs in relation to availability of glucose in the gut and/or during cell wall formation of Entamoeba histolytica. Abundance of Malassezia spp. is based on the availability of lipids proposed to be provided by E. histolytica. Cross-talking among the three taxa involves five major conditions/actions, e.g., glucose enrichment, glucose deprivation, oxidative stress, chitin formation and lipid exploitation. Processes or actions related to S. cerevisiae are shown in pink, while those of E. histolytica and Malassezia spp. are shown in light blue and light green, respectively*

It is well-established that glucose in the human gut perform four different destinations in amoebic cells (*Fig. 18*). They are glycolysis, chitin formation, fermentation and acetate biosynthesis. The first two destinations and the last two destination are very much connected by a number of redux reactions. The direction of the different reactions is dependent on the conditions at which parasite is facing, e.g., glucose level, oxygen presence, oxidative stress, etc. Glucose-6-phosphate (Glu6P) is considered as the main substrate lying upstream redux reactions of these four destinations (*Figs. 17* and *18*). Under the low level of glucose and anaerobic condition, it was reported earlier that adhesion of trophozoites to red blood cells (RBCs) were inhibited (López-Revilla and Cano-Mansera, 1982). The authors indicated that adhesion is dependent on amoebal cytoskeleton and affinity of both amoebal and RBC surface receptors. Another analysis on the effects of *Saccharomyces* on cecal amebiasis in rats indicated that the numbers of diseased rats and the severity of the infection were significantly reduced by the yeast treatment (Rigothier et al., 1990). A more recent report justified the inhibition of adhesion of trophozoite to RBCs as it indicated that pretreatment of amoebae with live yeast led to a significant reduction in the percentage of adhesion (Rigothier et al., 1994). The authors indicated that substances produced by *Saccharomyces* might compete with red blood cells, as an amoebicide, for adhesion

sites of the trophozoite (*Fig. 17*). Other reports indicated that the reptilian parasite *Entamoeba invadens* was induced to encyst under osmotic shock, glucose deprivation, or both (Vazquezdelara-Cisneros and Arroyo-Begovich, 1984; Avron et al., 1986; Sanchez et al., 1994; Van Dellen et al., 2006). Un-organized encystation threatens trophozoite's life or it can be a reason for the inefficiency of trophozoite to cause further action in the human body (*Fig. 17*). On the other hand, *Saccharomyces* will not suffer from glucose deprivation as it can consult its saved energy in ethanol that was synthesized by fermentation and use it for regular metabolic processes rather than being forced to make irregular metabolic processes.
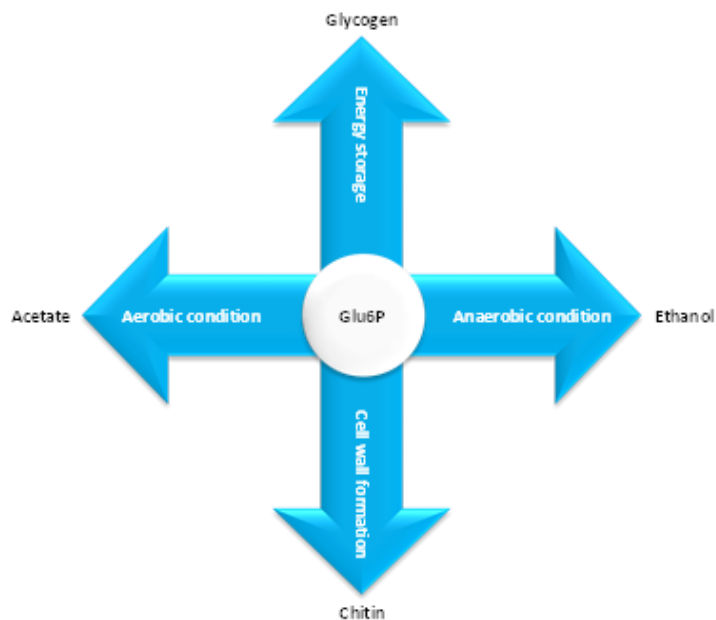


***Figure 18.*** *Four different destinations for glucose metabolism of which Glu6P is the core molecule upstream several redox reaction based on the condition at which parasite is facing*

As a response to glucose deprivation in the human gut, trophozoite has to utilize another source of carbon such as galactose. Frederick and Petri (2005) indicated that adhesion of the trophozoites to colonic mucins and host cells occurs mainly through the abundant galactose-binding lectin on the surface of the cells, however, this adhesion is almost completely inhibited by β-D-galactose (*Fig. 17*). We speculate that utilization of galactose as a carbon source due to glucose deprivation will negatively affect the trophozoite efficiency. On the other hand, *S. cerevisiae* simply transports galactose into the cell, when glucose level is low, by galactose permease enzyme to be converted to glucose-6-phosphate (Glu-6P) that drive the reaction normally towards glycolysis (Bro et al., 2005). More recently, Naiyer et al. (2019a) indicated that aspartate might be used as energy source by *Entamoeba* under glucose deprivation due to the regulation of the gene encoding aspartate aminotransferase. Galactose was also recently reported to act in inhibiting the formation of neutrophil extracellular traps (NET) that traps trophozoites, thus, makes trophozoites invisible by host neutrophil (Fonseca et al., 2019).

Husain et al. (2012) indicated that oxidative stress also resulted in reduced level of glycolysis and ethanol production and the redirection of metabolic flux towards glycerol production and chitin biosynthesis. Reduction of glycolysis and ethanol levels indicate

the reduced stored energy required for trophozoite actions and the promotion of chitin biosynthesis can stimulate the trophozoites to encyst, thus, further retard their actions. In addition, oxidative stress induces the production of reactive oxygen species (ROS). The latter inhibits the bifunctional aldehyde/alcohol dehydrogenase (ADHE) enzyme, important during ethanol production (Nozaki and Bhattacharya, 2015), thus, results in the buildup of Acetyl CoA that is the intermediate step towards acetate metabolism and depletion of energy. Chitin biosynthesis is a further response to oxidative stress. As indicated earlier, this process enforces trophozoites to encyst, thus, no further action is taken by the microbe in the gut. *Saccharomyces boulardii* has previously been used effectively to treat rotavirus (RV) diarrhea in human enterocytes by preventing RV-induced oxidative stress (Buccigrossi et al., 2014). The study was repeated in an organ culture model using human intestinal biopsies, and results indicated that *Saccharomyces* produces soluble metabolites that prevent oxidative stress in human intestine (*Fig. 17*). Interestingly, Naiyer et al. (2019a) indicated that genes upregulated during encystation include phospholipase D as well as genes involved in meiotic division such as *MND1*, *Hop2* and *Rad52*. Meiosis was not proven in *Entamoeba*, however, earlier reports indicated the occurrence of homologous recombination (Singh et al., 2013).

Glucose is known to be converted into pyruvate through glycolysis, a metabolic pathway employed in nearly all organisms (Berg et al., 2002). In the downstream reactions, pyruvate is further degraded either through the respiration or the fermentation pathway. The first is a pathway to release or exhaust ATP, while the second is a pathway to store ATP. Under conditions of glucose enrichment, trophozoites mainly perform fermentation under anaerobic condition, while acetate biosynthesis under aerobic condition (Nozaki and Bhattacharya, 2015). Aerobic condition is accompanied by the presence of oxygen. It was reported that trophozoites produce 75% ethanol and 25% acetate under anaerobic condition, while opposite ratios under aerobic condition. The latter ratios indicate that trophozoites cannot save appropriate amount of stored energy at the form of ATP via fermentation under aerobic or oxidative stress condition (*Fig. 17*). On the other hand, some yeast species such as *S. cerevisiae* use fermentation even in the presence of oxygen, albeit glucose concentration is sufficiently high. Such *Saccharomyces* species are able to choose the proper destination for glycolysis (respiration or fermentation) under high glucose levels and aerobic conditions, a property related to Crabtree effect (Crabtree, 1929; De Deken, 1966; Postma and Verduyn, 1989; Van Urk et al., 1990; Pfeiffer and Morley, 2014), of which the direction choices are based on cell's requirement for energy. This phenomenon indicates that excess sugar availability promotes cells to ferment it rather than going through respiration or oxidative phosphorylation. Reasons for this phenomenon includes the limited membrane space that harbors the enzymes involved in oxidative phosphorylation (Hagman et al., 2013).

## Conclusion

The main goal of the present study is the detection of gut mycobiome signature of *E. histolytica*-infected patients in a trial to improve our ability to diagnose and treat the disease. We suggest that the results on the discovery of new negative biomarker for amebiasis can provide new insights into the possible use of *Saccharomyces* in the future as a probiotic to treat diarrhea and downstream severity of the disease. The latter conclusion requires further experimentation way before we decide to use yeasts as a treatment in the market place.

# REFERENCES

[1] Abarenkov, K., Nilsson, R. H., Larsson, K.- H., et al. (2010): The UNITE database for molecular identification of fungi-recent updates and future perspectives. – New Phytol. 186: 281-285.

[2] Adl, S. M., Simpson, A. G., Lane, C. E., et al. (2012): The revised classification of eukaryotes. – J. Eukaryot. Microbiol. 59: 429-493.

[3] Ahn, J., Yang, L., Paster, B. J., et al. (2015): Oral microbiome profiles: 16S rRNA pyrosequencing and microarray assay comparison. – PLoS ONE 6: e22788.

[4] Ali, V., Nozaki, T. (2007): Current therapeutics, their problems, and sulfur-containing-amino-acid metabolism as a novel target against infections by 'amitochondriate' protozoan parasites. – Clin. Microbiol. Rev. 20: 164-187.

[5] Avron, B., Stolarsky, T., Chayen, A., Mirelman, D. (1986): Encystation of *Entamoeba invadens* IP-1 is induced by lowering the osmotic pressure and depletion of nutrients from the medium. – J. Protozool. 33: 522-525.

[6] Belda-Ferre, P., Alcaraz, L. D., Cabrera-Rubio, R., et al. (2012): The oral metagenome in health and disease. – ISME J. 6: 46-56.

[7] Benjamini, Y., Hochberg, Y. (1995): Controlling the false discovery rate: a practical and powerful approach to multiple testing. – J. Royal Stat. Soc. (Series B) 57: 289-300.

[8] Berg, A. H., Combs, T. P., Scherer, P. E. (2002): ACRP30/adiponectin: an adipokine regulating glucose and lipid metabolism. – Trends Endocrinol. Metab. 13: 84-89.

[9] Bliss, J. M., Basavegowda, K. P., Watson, W. J., et al. (2008): Vertical and horizontal transmission of *Candida albicans* in very low birth weight infants using DNA fingerprinting techniques. – Pediatr. Infect. Dis. J. 27: 231-235.

[10] Bro, C., Knudsen, S., Regenberg, B., et al. (2005): Improvement of galactose uptake in *Saccharomyces cerevisiae* through overexpression of phosphoglucomutase: example of transcript analysis as a tool in inverse metabolic engineering. – Appl. Environ. Microbiol. 71: 6465-6472.

[11] Buccigrossi, V., Laudiero, G., Russo, C., et al. (2014): Chloride secretion induced by rotavirus is oxidative stress-dependent and inhibited by *Saccharomyces boulardii* in human enterocytes. – PLoS ONE 9: e99830.

[12] Bulawa, C. E. (1993): Genetics and molecular biology of chitin synthesis in fungi. – Annu. Rev. Microbiol. 47: 505-534.

[13] Byron, J. K., Clemons, K. V., McCusker, J. H., et al. (1995): Pathogenicity of *Saccharomyces cerevisiae* in complement factor five-deficient mice. – Infect. Immun. 63: 478-485.

[14] Cano, R. J., Rivera-Perez, J., Toranzos, G. A., et al. (2014): Paleomicrobiology: revealing fecal microbiomes of ancient indigenous cultures. – PLoS ONE 9: e106833.

[15] Caporaso, J. G., Bittinger, K., Bushman, F. D., et al. (2010a): PyNAST: a flexible tool for aligning sequences to a template alignment. – Bioinformatics 26: 266-267.

[16] Caporaso, J. G., Justin, K., Jesse, S. (2010b): QIIME allows analysis of high-throughput community sequencing data. – Nature Methods 7: 335-336.

[17] Choi, W. J., Cabib, E. (1994): The use of divalent cations and pH for the determination of specific yeast chitin synthetases. – Anal. Biochem. 219: 368-372.

[18] Cole, J. R., Wang, Q., Fish, J. A., et al. (2014): Ribosomal Database Project: data and tools for high throughput rRNA analysis. – Nucleic Acids Res. 42: D633-D642.

[19] Crabtree, H. G. (1929): Observations on the carbohydrate metabolism of tumours. – Biochem. J. 23: 536-545.

[20] Cui, L., Morris, A., Ghedin, E. (2013): The human mycobiome in health and disease. – Genome Med. 5: 63.

[21] David, L. A., Maurice, C. F., Carmody, R. N., et al. (2014): Diet rapidly and reproducibly alters the human gut microbiome. – Nature 505: 559-563.

[22] De Deken, R. H. (1966): The Crabtree effect: a regulatory system in yeast. – J. Gen. Microbiol. 44: 149-156.

[23] De Llanos, R., Querol, A., Pemán, J., et al. (2006): Food and probiotic strains from the *Saccharomyces cerevisiae* species as a possible origin of human systemic infections. – Int. J. Food Microbiol. 110: 286-290.

[24] Dollive, S., Chen, Y.- Y., Grunberg, S., et al. (2013): Fungi of the murine gut: episodic variation and proliferation during antibiotic treatment. – PLoS ONE 8: e71806.

[25] Eaton, R. D. P., Meerovitch, E., Costerton, J. W. (1970): The functional morphology of pathogenicity in *Entamoeba histolytica*. – Ann. Trop. Med. Parasitol. 64: 299-304.

[26] Edgar, R. C. (2010): Search and clustering orders of magnitude faster than BLAST. – Bioinformatics 26: 2460-2461.

[27] Edgar, R. C. (2013): UPARSE: highly accurate OTU sequences from microbial amplicon reads. – Nature Methods 10: 996-998.

[28] Edgar, R. C., Haas, B. J., Clemente, J. C., et al. (2011): UCHIME improves sensitivity and speed of chimera detection. – Bioinformatics 27: 2194-2200.

[29] Ewing, B., Green, P. (1998): Base-calling of automated sequencer traces using Phred. II. Error probabilities. – Genome Res. 8: 186-194.

[30] Fonseca, Z., Uribe-Querol, E., Díaz-Godínez, C., et al. (2019): Pathogenic *Entamoeba histolytica*, but not *Entamoeba dispar*, induce neutrophil extracellular trap (NET) formation. – J. Leukoc. Biol. 105: 1167-1181.

[31] Frederick, J. R., Petri Jr., W. A. (2005): Roles for the galactose-/N-acetylgalactosamine-binding lectin of *Entamoeba* in parasite virulence and differentiation. – Glycobiol. 15: 53R-59R.

[32] Gao, R., Gao, Z., Huang, L., et al. (2017): Gut microbiota and colorectal cancer. – Eur. J. Clin. Microbiol. Infect. Dis. 36: 757-769.

[33] Gouba, N., Drancourt, M. (2015): Digestive tract mycobiota: a source of infection. – Med. Mal. Infect. 45: 9-16.

[34] Gouba, N., Raoult, D., Drancourt, M. (2013): Plant and fungal diversity in gut microbiota as revealed by molecular and culture investigations. – PLoS ONE 8: e59474.

[35] Gouba, N., Raoult, D., Drancourt, M. (2014): Eukaryote culturomics of the gut reveals new species. – PloS ONE 9: e106994.

[36] Gupta, A. K., Batra, R., Bluhm, R., et al. (2004): Skin diseases associated with *Malassezia* species. – J. Am. Acad. Dermatol. 51: 785-98.

[37] Gupta, S. S., Singh, O., Shukla, S., Raj, M. K. (2009): Acute fulminant necrotizing amoebic colitis: a rare and fatal complication of amoebiasis: a case report. – Cases J. 2: 6557.

[38] Hagman, A., Sall, T., Compagno, C., Piskur, J. (2013): Yeast "Make-accumulate-consume" life strategy evolved as a multi-step process that predates the whole genome duplication. – PloS ONE 8: e68734.

[39] Hallen-Adams, H. E., Suhr, M. J. (2017): Fungi in the healthy human gastrointestinal tract. – Virulence 8: 352-358.

[40] Hamad, I., Raoult, D., Bittar, F. (2016): Repertory of eukaryotes (eukaryome) in the human gastrointestinal tract: taxonomy and detection methods. – Parasite Immunol. 38: 12-36.

[41] Henar Valdivieso, M., Duran, A., Roncero, C. (1999): Chitin synthases in yeast and fungi. – EXS 87: 55-69.

[42] Hirata, K. K., Que, X., Melendez-Lopez, et al. (2007): A phagocytosis mutant of *Entamoeba histolytica* is less virulent due to deficient proteinase expression and release. – Exp. Parasitol. 115: 192-199.

[43] Hoffmann, C., Dollive, S., Grunberg, S., et al. (2013): Archaea and fungi of the human gut microbiome: correlations with diet and bacterial residents. – PLoS ONE 8: e66019.

[44] Huffnagle, G., Noverr, M. (2013): The emerging world of the fungal microbiome. – Trends Microbiol. 21: 334-341.

[45] Hugon, P., Lagier, J.- C., Colson, P., et al. (2017): Repertoire of human gut microbes. – Micro. Pathog. 106: 103-112.

[46] Husain, A., Sato, D., Jeelani, G., et al. (2012): Dramatic increase in glycerol biosynthesis upon oxidative stress in the anaerobic protozoan parasite *Entamoeba histolytica.* – PLoS Negl. Trop. Dis. 6: e1831.

[47] Ianiro, G., Bruno, G., Lopetuso, L., et al. (2014): Role of yeasts in healthy and impaired gut microbiota: the gut mycome. – Curr. Phar. Des. 20: 4565-4569.

[48] Iliev, I. D., Funari, V. A., Taylor, K. D., et al. (2012): Interactions between commensal fungi and the C-type lectin receptor Dectin-1 influence colitis. – Science 336: 1314-1317.

[49] Ishida, H. Inokuma, S., Murata, N., et al. (2003): Fulminant amoebic colitis with perforation successfully treated by staged surgery: a case report. – J. Gastroenterol. 38: 92-96.

[50] Iyer, L. R., Verma, A. K., Paul, J., et al. (2019): Phagocytosis of gut bacteria by *Entamoeba histolytica.* – Front. Cell. Infect. Microbiol. 9: Article 34.

[51] Jiang, T. T., Shao, T.- Y., Ang, W. G., et al. (2017): Commensal fungi recapitulate the protective benefits of intestinal bacteria. – Cell Host Microbe 22: 809-816.

[52] Jo, J. H., Kennedy, E. A., Kong, H. H. (2017): Topographical and physiological differences of the skin microbiome in health and diseases. – Virulence 8: 324-333.

[53] Johansson, I., Witkowska, E., Kaveh, B., et al. (2016): The microbiome in populations with a low and high prevalence of caries. – J. Dent. Res. 95: 80-86.

[54] Khomkhum, N., Leetachewa, S., Pawestri, A. R., et al. (2019): Host-antibody inductivity of virulent *Entamoeba histolytica* and non-virulent *Entamoeba moshkovskii* in a mouse model. – Parasites & Vectors 12: Article 101.

[55] Kirschner, R., Hsu, T., Tuan, N. N., et al. (2015): Characterization of fungal and bacterial components in gut/fecal microbiome. – Curr. Drug. Metab. 16: 272-283.

[56] López-Revilla, R., Cano-Mansera, R. (1982): Adhesion of *Entamoeba histolytica* trophozoites to human erythrocytes. – Infect. Immun. 37: 281-285.

[57] Lozano, R., Naghavi, M., Foreman, K., et al. (2012): Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. – Lancet 380: 2095-2128.

[58] Lozupone, C., Lladser, M. E., Knights, D., et al. (2011): UniFrac: an effective distance metric for microbial community comparison. – ISME J. 5: 169-172.

[59] Magoc, T., Salzberg, S. (2011): FLASH: fast length adjustment of short reads to improve genome assemblies. – Bioinformatics 27: 2957-63.

[60] McHardy, I. H., Wu, M. R., Shimizu-Cohen, R., et al. (2014): Detection of intestinal protozoa in the clinical laboratory. – J. Clin. Microbiol. 52: 712-720.

[61] Mi-Ichi, F., Yoshida, H., Hamano, S. (2016): *Entamoeba* Encystation: new targets to prevent the transmission of amebiasis. – PLOS Pathog. 12: e1005845.

[62] Moran, G., Coleman, D., Sullivan, D. (2012): An Introduction to the Medically Important Candida Species. – In: Calderone, R. A., Clancy, C. J. (eds.) Candida and Candidiasis. 2nd Ed. ASM Press, Washington, DC, pp. 11-25.

[63] Moyes, D. L., Naglik, J. R. (2012): The mycobiome: influencing IBD severity. – Cell Host Microbe 11: 551-552.

[64] Mukherjee, P. K., Sendid, B., Hoarau, G., et al. (2015): Mycobiota in gastrointestinal diseases. – Nat. Rev. Gastroenterol. Hepatol. 12: 77-87.

[65] Munoz, P., Bouza, E., Cuenca-Estrella, M., et al. (2005): *Saccharomyces cerevisiae* fungemia: an emerging infectious disease. – Clin. Infect. Dis. 40: 1625-1634.

[66] Naiyer, S., Bhattacharya, A., Bhattacharya, S. (2019a): Advances in *Entamoeba histolytica* biology through transcriptomic analysis. – Front. Microbiol. 10: Article 1921.

[67] Naiyer, S., Kaur, D., Ahamad, et al. (2019b): Transcriptomic analysis reveals novel downstream regulatory motifs and highly transcribed virulence factor genes of *Entamoeba histolytica*. – BMC Genomics 20: 206.

[68] Nash, A. K., Auchtung, T. A., Wong, M. C., et al. (2017): The gut mycobiome of the human microbiome project healthy cohort. – Microbiome 5: 153.

[69]    Nozaki, T., Bhattacharya, A. (2015): Amebiasis: Biology and Pathogenesis of *Entamoeba*. – Springer, Japan.

[70]    Ott, S. J., Kühbacher, T., Musfeldt, M., et al. (2008): Fungi and inflammatory bowel diseases: alterations of composition and diversity. – Scand. J. Gastrol. 43: 831-841.

[71]    Paulson, J. N., Pop, M., Bravo, H. C. (2011): Metastats: an improved statistical method for analysis of metagenomic data. – Genome Biol. 12(Suppl 1): P17.

[72]    Penders, J., Thijs, C., Vink, C., et al. (2006): Factors influencing the composition of the intestinal microbiota in early infancy. – Pediatrics 118: 511-521.

[73]    Pfeiffer, T., Morley, A. (2014): An evolutionary perspective on the Crabtree effect. – Front. Mol. Biosci. Article 17.

[74]    Postma, E., Verduyn, C. (1989): Enzymic analysis of the Crabtree effect in glucose-limited chemostat cultures of *Saccharomyces cerevisiae*. – Appl. Environ. Microbiol. 55: 468-477.

[75]    Quast, C., Pruesse, E., Yilmaz, P., et al. (2013): The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. – Nucl. Acids Res. 41: D590-D596.

[76]    Rajilic-Stojanovic, M., de Vos, W. M. (2014): The first 1000 cultured species of the human gastrointestinal microbiota. – FEMS Microbiol. Rev. 38: 996-1047.

[77]    Rajilic-Stojanovic, M., Smidt, H., de Vos, W. M. (2007): Diversity of the human gastrointestinal tract microbiota revisited. – Environ. Microbiol. 9: 2125-2136.

[78]    Ravdin, J. I., Croft, B. Y., Guerrant, R. L. (1980): Cytopathogenic mechanisms of *Entamoeba histolytica*. – J. Exp. Med. 152: 377-390.

[79]    Richard, M. L., Lamas, B., Liguori, G., et al. (2015): Gut fungal microbiota: the Yin and Yang of inflammatory bowel disease. – Inflamm. Bowel Dis. 21: 656-665.

[80]    Rigothier, M. C., Maccario, J., Vuong, P. N., Gayral, P. (1990): Effects of *Saccharomyces boulardii* yeast on trophozoites of *Entamoeba histolytica in vitro* and in cecal amebiasis in young rats. – Ann. Parasitol. Hum. Comp. 65: 51-60.

[81]    Rigothier, M. C., Maccario, J., Gayral, P. (1994): Inhibitory activity of *saccharomyces* yeasts on the adhesion of *Entamoeba histolytica* trophozoites to human erythrocytes in vitro. – Parasitol. Res. 80: 10-15.

[82]    Roncero, C. (2002): The genetic complexity of chitin synthesis in fungi. – Curr. Genet. 41: 367-378.

[83]    Roussel, C., Sivignon, A., de Vallée, A., et al. (2018): Anti-infectious properties of the probiotic *Saccharomyces cerevisiae* CNCM I-3856 on enterotoxigenic *E. coli* (ETEC) strain H10407. – Appl. Microbiol. Biotechnol. 102: 6175-6189.

[84]    Ruiz-Herrera, J., Gonzalez-Prieto, J. M., Ruiz-Medrano, R. (2002): Evolution and phylogenetic relationships of chitin synthases from yeasts and fungi. – FEMS Yeast Res. 1: 247-356.

[85]    Saleh, H. A., Moawad, A. A., El-Hariri, M., Refai, M. K. (2011): Prevalence of yeasts in human, animals and soil sample at El-Fayoum Governorate in Egypt. – Int. J. Microbiol. Res. 2: 233-239.

[86]    Sanchez, L., Enea, V., Eichinger, D. (1994): Identification of a developmentally regulated transcript expressed during encystation of *Entamoeba invadens*. – Mol. Biochem. Parasitol. 67: 125-135.

[87]    Scanlan, P. D., Marchesi, J. R. (2008): Micro-eukaryotic diversity of the human distal gut microbiota: qualitative assessment using culture-dependent and independent analysis of feces. – ISMA J. 2: 1183-1193.

[88]    Schloss, P. D., Westcott, S. L., Ryabin, T., et al. (2009): Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. – Appl. Environ. Microbiol. 75: 7537-7541.

[89]    Seed, P. C. (2014): The human mycobiome. – Cold Spring Harb. Perspect. Med. 5: a019810.

[90]    Singh, N., Bhattacharya, S., Paul, J. (2011): *Entamoeba invadens*: dynamics of DNA synthesis during differentiation from trophozoite to cyst. – Exp. Parasitol. 127: 329-333.

[91] Stensvold, C. R., Smith, H. V., Nagel, R., et al. (2010): Eradication of *Blastocystis* carriage with antimicrobials: reality or delusion? – J Clin Gastroenterol 44: 85-90.

[92] Soares, N. M., Azevedo, H. C., Pacheco, F. T. F., et al. (2019): A cross-sectional study of *Entamoeba histolytica*/*dispar*/*moshkovskii* complex in Salvador, Bahia, Brazil. – BioMed Res. Intl. Article ID 7523670.

[93] Stanley, S. L. (2003): Amebiasis. – Lancet 361: 1025-1034.

[94] Suhr, M. J., Hallen-Adams, H. E. (2015): The human gut mycobiome: pitfalls and potentials - a mycologist's perspective. – Mycologia 107: 1057-1073.

[95] Suhr, M. J., Banjara, N., Hallen-Adams, H. E. (2016): Sequence-based methods for detecting and evaluating the human gut mycobiome. – Lett. Appl. Microbiol. 62: 209-215.

[96] Tang, J., Iliev, I. D., Brown, J., et al. (2015): Mycobiome: approaches to analysis of intestinal fungi. – J. Immunol. Methods 421: 112-121.

[97] Tavares, R. G., Staggemeier, R., Borges, A. L. P. (2011): Molecular techniques for the study and diagnosis of parasite infection. – The Journal of Venomous Animals and Toxins including Tropical Diseases 17: 239-248.

[98] Tellam, R. L., Vuocolo, T., Johnson, S. E., et al. (2000): Insect chitin synthase cDNA sequence, gene organization and expression. – Eur. J. Biochem. RD 267: 6025-6043.

[99] Thibeaux, R., Weber, C., Hon, C.- C., et al. (2013): Identification of the virulence landscape essential for *Entamoeba histolytica* invasion of the human colon. – PLoS Pathog. 9: e1003824.

[100] Trissl, D., Martinez-Palomo, A., de la Torre, M., et al. (1978): Surface properties of *Entamoeba*: increased rates of human erythrocyte phagocytosis in pathogenic strains. – J. Exp. Med. 148: 1137-1145.

[101] Underhill, D., Iliev, I. (2014): The mycobiota: interactions between commensal fungi and the host immune system. – Nat. Rev. Immunol. 14: 405-416.

[102] Ungar, B. L. (1990): Enzyme-linked immunoassay for detection of *Cryptosporidium* antigens in fecal specimens. – J. Clin. Microbiol. 28: 2491-2495.

[103] Van Dellen, L., Bulik, D. A., Specht, C. A., et al. (2006): Heterologous expression of an *Entamoeba histolytica* chitin synthase in *Saccharomyces cerevisiae*. – Eukaryotic Cell 5: 203-206.

[104] Van Urk, H., Voll, W. S., Scheffers, W. A., Van Dijken, J. P. (1990): Transient-state analysis of metabolic fluxes in Crabtree-positive and Crabtree-negative yeasts. – Appl. Environ. Microbiol. 56: 281-287.

[105] Vazquezdelara-Cisneros, L. G., Arroyo-Begovich, A. (1984): Induction of encystation of *Entamoeba invadens* by removal of glucose from the culture medium. – J. Parasitol. 70: 629-633.

[106] Verkerke, H. P., Petri, W. A., Marie, C. S. (2012): The dynamic interdependence of amebiasis, innate immunity, and undernutrition. – Semin. Immunopathol. 34: 771-785.

[107] Vogel, C., Rogerson, A., Schatz, S., et al. (2007): Prevalence of yeasts in beach sand at three bathing beaches in South Florida. – Water Res. 41: 1915-1920.

[108] Wang, Z. K., Yang, Y. S., Stefka, A. T., et al. (2014): Review article: fungal microbiota and digestive diseases. – Aliment Pharmacol. Ther. 39: 751-766.

[109] WHO, PAHO, UNESCO (1997): Report: A Consultation with Experts on Amoebiasis. – WHO, Geneva; PAHO, Washington; UNESCO, Paris.

[110] Ximénez, C., Morán, P., Rojas, L., et al. (2009): Reassessment of the epidemiology of amebiasis: state of the art. – Infect. Genet. Evol. 9: 1023-1032.

[111] Zhou, J., Jiang, N., Wang, S., et al. (2016): Exploration of human salivary microbiomes-insights into the novel characteristics of microbial community structure in caries and caries-free subjects. – PLoS ONE 11: e0147039.
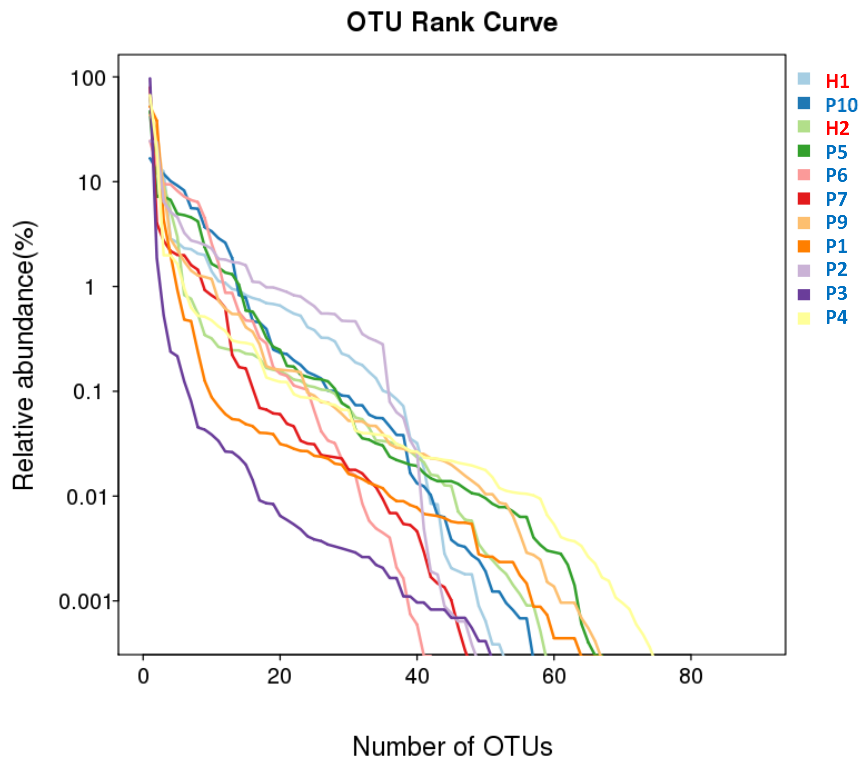
## APPENDIX



***Figure A1.*** *Number of OTUs in relation to relative abundance of OTUs for different healthy individuals (H) and Entamoeba-infected patients (P)*
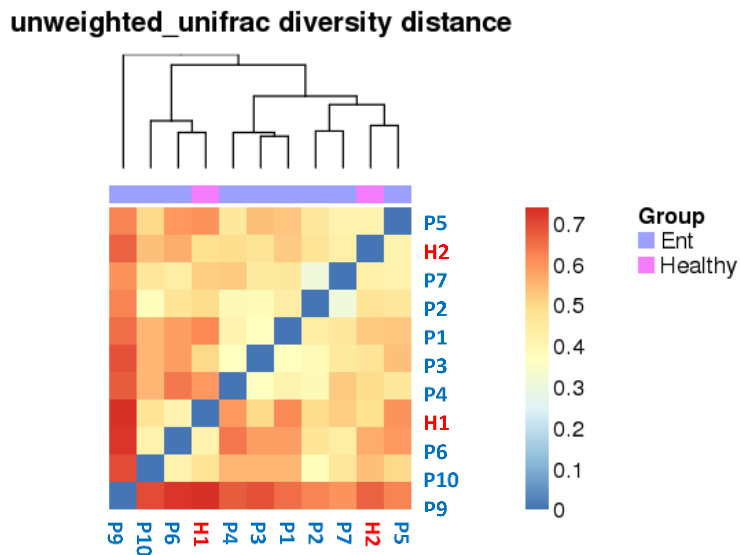


***Figure A2.*** *Beta diversity heat map describing weighted and unweighted unifrac diversity distances for different healthy individuals (H) and Entamoeba-infected patients (P)*

***Tables A1–A4:*** *See electronic appendix*