Mycobiome profiles in breast milk from healthy women depending on mode of delivery, geographic location and interactions with bacteria

Alba Boix-Amorós[1,2], Fernando Puente-Sánchez[3], Elloise du Toit[4], Kaisa M. Linderborg[5], Yumei Zhang[6], Baouru Yang[5], Seppo Salminen[7], Erika Isolauri[8], Javier Tamames[3], Alex Mira[2], Maria Carmen Collado[1].

1: Department of Biotechnology, Institute of Agrochemistry and Food Technology- National Research Council (IATA-CSIC). Valencia. Spain.

2: Department of Health and Genomics. Center for Advanced Research in Public Health, FISABIO Foundation. Valencia. Spain.

3: Systems Biology Program, Centro Nacional de Biotecnología, Consejo Superior de Investigaciones Científicas (CSIC), Madrid, Spain.

4: Division of Medical Microbiology, Department of Pathology, University of Cape Town. Cape Town, South Africa.

5: Food Chemistry and Food Development, Department of Biochemistry, University of Turku. Turku, Finland.

6: Department of Nutrition and Food Hygiene, School of Public Health, Peking University. Beijing, China.

7: Functional Foods Forum, Faculty of Medicine, University of Turku. Turku, Finland.

8: Department of Pediatrics, University of Turku. Turku, Finland.

Abstract

Recent studies reported the presence of fungal species in human breast milk from healthy mothers, suggesting a potential role on infant mycobiome development. In the present work, we aimed to characterize the influence of geographical location and mode of delivery on the healthy human breast milk mycobiota, as well as studying its interaction with bacterial profiles in the same samples. The mycobiome of 80 mature breast milk samples from 4 different locations were studied by using ITS Illumina sequencing. Basidiomycota and Ascomycota were found to be the dominant phyla, with *Malassezia* and *Davidiella* being the most prevalent genera across countries. A core formed by *Malassezia, Davidiella, Sistotrema* and *Penicillium* was shared in milk samples from all countries, although specific shifts on mycobiome composition and diversity were associated to geographic location and delivery mode. Network analysis of bacteria and fungi showed complex interactions that were influenced by geographical location, mode of delivery, maternal age and pre-gestational Body Mass Index. Those Mycobiome-bacteriome-host interactions in milk may have a significant impact on the colonization and development of the infant microbiota, as well as on the immunological and metabolic health programming, which should be explored.

Introduction

Early human microbial gut colonization is an essential step-wise process initiating the programming for later health by shaping both the microbiota development and immune system maturation[1,2,3]. Fungi residing in the human gut have been recognized to be an important part of the gut microbiota, and can have direct effects on human health status[1–7]. Although information about fungal communities in the infant is scarce, there is evidence that fungal species (mainly yeast-like) can be found in the gut since early in life[8,9,10] . Only a few reports have shown that fungal transfer from mothers to infants occur, although little is known on how the mycobiome is shaped during this period[11,12,13].

Breast milk is one of the most important sources of bacteria and oligosaccharides to the infant gut, contributing to the settlement of the gut microbiota and therefore also acquired immunity[14,15]. A recent study has suggested the presence of a wide diversity of fungal species in human breast milk from healthy mothers, including *Malassezia*, *Candida* and *Saccharomyces* as the most common genera detected by using multiple approaches that included high-throughput sequencing, microscopy and other culture-independent techniques[16]. Moreover, viable yeasts predominantly *Candida parapsilosis* and *Rhodotorula mucilaginosa* species, were isolated and characterized. This finding highlights the potential influence of breast milk on infant mycobiome development.

Complex interactions between bacteria and fungi have been reported in the human gut, oral cavity and skin[17–19] and may also occur in breast milk. Furthermore, accumulating evidence suggests that some environmental factors, such as geographic location or delivery mode, can influence breast milk bacterial composition[20–26], although little is known about their potential impact on the fungal fraction.

In the present study, we characterized the breast milk mycobiota in healthy beast-feeding mothers from four different countries (Spain, Finland, South Africa and China), in order

to investigate the potential influence of geographic location and the impact of delivery mode on its composition. In addition, co-occurrence networks between specific fungi and bacteria were studied to detect potential interactions and their variations depending on mode of delivery across the countries.

Material and Methods

Subjects and Sampling

Breast milk samples at 1-month post-partum were obtained from 80 healthy volunteers from 4 different geographical locations, including China (Beijing area), South Africa (Cape Town), Finland (Southwestern area), and Spain (Valencia, Mediterranean area).

All volunteers were practising exclusive breastfeeding. Subjects from each country (n= 20) were grouped according to mode of delivery: vaginal (n= 10 per country) and Caesarean-section (C-section) (n= 10 per country). Maternal characteristics such as age, weight and pre-gestational body mass index (BMI) were collected at the time of enrolment. All women who delivered via C-section received prophylactic antibiotics, except Finnish women where no prophylaxis is routinely used as per the hospital policy. All participants were given detailed oral and written information, and written informed consent was obtained. The study protocol was approved by the Ethics Committees of the respective participating institutions: Spain (Bioethics Committee of CSIC and the Regional Ethics Committee for Biomedical Research), Finland (Ethics Committee, Hospital District of Southwest Finland), China (Medical Research Board of Peking University) and South Africa (University of Cape Town, Human Research Ethics Committee).

All the samples were kept frozen at –20°C until they were delivered to the laboratory and then stored at –80 °C until further analysis. To avoid bias, all milk samples were collected

using the same standardised protocol, as previously described[25], and were processed and analysed in a single laboratory.

Microbial DNA Extraction and Sequencing

Breast milk samples (1.5 ml) were centrifuged at 14,000 rpm for 20 min at 4°C to remove fat, and pellets were used for total DNA extraction that involved mechanical and chemical cell lysis. Bead beating was carried out using FastPrep® (FP120-230, Bio 101 ThermoSavant, Holbrook, NY, USA), and the InviMag® Stool DNA kit (Stratec Molecular, Berlin, Germany) was used with the King Fisher magnetic particle processor (Thermo Fisher Scientific Oy, Vantaa, Finland). The DNA extraction protocol was also followed with water to use as negative controls. Isolated DNA concentrations were measured using a Qubit® 2.0 Fluorometer (Life Technology, Carlsbad, CA, USA).

Primers targeting the highly variable fungal internal transcriber spacer ITS1 of the fungal 18S ribosomal rRNA gene (forward: TAGAGGAAGTAAAAGTCGTAA, reverse: TTYRCTRCGTTCTTCATC)[27] with adaptors were used for sequencing on an Illumina Miseq platform. Sequencing was carried out at the Foundation for the Promotion of Health and Biomedical Research, FISABIO (Valencia, Spain). No-template controls (NTCs) and negative controls during DNA extraction were included to rule out potential contaminations at the time of DNA extraction or sequencing.

Data Analysis

ITS1 reads were pair-end joined using FLASH program[28] applying default parameters. Resulting sequences were end-trimmed in 20 bp sliding windows with average quality value >30, and length >50 bp, using the Prinseq-lite program[25]. Chimeric reads were eliminated using

UCHIME algorithm [29], resulting in a total of 9,797,578 reads. Taxonomy assignment of the remaining sequences was performed using Ribosomal Database Project classifier standalone tool[30] with the UNITE fungal ITS v 7.2 trainset[31], and an 80% confidence threshold. Sequences where clustered into operational taxonomical units (OTUs) based on 99% identity, and representative OTUs sequences were obtained using CD-hit software[32]. OTU tables were rarefied to 9200 sequences per sample to avoid variations in sequencing depth, and Shannon and Chao1 indexes were calculated using the "plyr" and "vegan" packages from R software (version 3.2.2)[33].

Statistical Analysis

SPSS 24.0 software (IBM Corp., Armonk, NY, USA) was implemented to perform two-way multivariate analysis of variance (two-way MANOVA) on normalised data (total sum normalization and square root transformation) for comparison of phylotypes at different levels. Wilks's lambda multivariate test was applied to study the statistical effect of country and delivery mode on the samples composition, and Bonferroni *post hoc* test and t-test analyses were applied to compare the main effects of the variables in the groups. Calypso software (version 8.2) was used to obtain Venn diagram for shared phylotypes; and Discriminant Analysis of Principal Components (DAPC) was performed at OTU level, using geographic location as factor[34]. Linear discriminant analysis effect size (LefSe)[35] algorithm was used to detect the most differentially abundant fungi between vaginal and C-section deliveries in each country. Other statistical analysis and graphs were performed using GraphPad PRISM^{R} 6 (GraphPad Software).

Analysis of interactions between bacteria and fungi

Sequences from the 16S rRNA gene of the same samples, from Kumar et al[25] were obtained from NCBI (SRA accession: SRP082263 and submission ID: SUB1772296). Quality filtering, chimera checking and OTU clustering were as followed for the ITS1 reads.

RDP classifier was used to taxonomically assign the bacterial (against RDP's 16S rRNA training set 16[36]) and fungal (against the UNITE v 07-04-2014 trainset[31]) representative OTU sequences. Samples with less than 1500 sequences were excluded from the analysis.

For the bacterial datasets, OTUs with a higher relative abundance in any of the two controls than in the breast milk samples were treated as putative contaminants and discarded. This procedure could not be performed on the fungal datasets, since the sequencing of the two controls yielded too few reads. Nevertheless, the low fraction of reads assigned to putative contaminants in the bacterial datasets (2% on average) leads us to believe that the samples were essentially contamination-free. Both the bacterial and fungal OTU tables were rarefied to 1500 sequences per sample. OTUs from both the bacterial and fungal datasets having an overall relative abundance higher than 1% of the total reads, or appearing in at least one sample with a relative abundance higher than 5%, were combined into a single table. Associations between pairs of bacterial and fungal OTUs were calculated using the Maximal Information Coefficient, as implemented in MICtools[37]. Pseudo p-values were obtained by generating 200,000 null matrices, and further transformed to Storey's Q-values to correct for multiple hypothesis testing with the Benjamini-Hochberg method. Correlations with a False Discovery Rate lower than 0.01 were deemed significant Further, we divided the samples into 8 groups according to the combination of the 4 countries and 2 delivery modes. We used linear regression to calculate correlations between pairs of OTUs and factors (Age, BMI) in a given group. For each group, only OTUs appearing in at least 4 samples and with a relative abundance higher than 2% in at least one sample were included. Correlations with a p-value lower than 0.05 were deemed significant. Network analysis was performed on Cytoscape [38].

Phylogenetic relationships between *Malassezia* reads

ITS sequences of the 20 most abundant OTUs assigned to the *Malassezia* genus by the RDP classifier were combined with those of known *Malassezia* representatives from the UNITE

v07-04-2014 database[31]. A multiple sequence alignment was constructed with MAFFT v7.313 [39]. *Cryptococcus neoformans* was selected as an outgroup, and its ITS sequence was added to the alignment using the *add* option from MAFFT. The resulting alignment was manually curated and further refined with MUSCLE v3.8.31 [40]. Phylogenetic trees were inferred with RaxML v8 [41] and MrBayes v3.2 [42], using 1000 replicates and 1,000,000 generations respectively. TreeGraph2 [43] was used to combine and visualize the maximum likelihood and bayesian inference trees.

Results

Subject Description

The characteristics of the subjects are listed in Table 1. No differences on maternal age nor BMI for all 80 participants were detected. Interestingly, women who delivered vaginally had lower mean BMI, 23.4 (SD ± 2.11); while women who delivered by C-section had mean BMI values of 24.7 (SD ± 2.8). This difference was only significant in South African C-section samples, that showed the highest BMI, 26.67 (SD ± 1.41) ($p<0.05$) compared to the other mothers.

Fungal composition of breast milk through geographic locations and impact of perinatal factors

A mean of 107,765 taxonomically assigned, clean and filtered sequences per sample (± 45,493 SD), with an average length of 301 bp were obtained.

All breast milk samples contained fungal DNA and they were dominated by two phyla: Basidiomycota (58.65%) and Ascomycota (41.03%). South African samples had significantly higher levels of Ascomycota and lower levels of Basidiomycota compared to the other countries

183 (*p*<0.05*)*; At genus level, breast milk samples were dominated by *Malassezia* (40.6% average
184 abundance), followed by *Davidiella* (9.0%), that were prevalent regardless of the location or the
185 donor's type of delivery **(Figure 1a)**.

186 A two-way MANOVA was conducted and reflected that milk mycobiota differed
187 significantly across geographic location (Wilks' lambda = 0.076, *p*=0.002), which affected
188 significantly the levels of *Malassezia* (F (3) = 3.65, *p*=0.016), and *Rhodotorula* (F (3) = 7.74,
189 *p*=0.000). Bonferroni *post hoc* tests showed that *Malassezia* abundances were statistically
190 higher in Finnish and Chinese samples (p<0.05); while *Rhodotorula* was higher in South African
191 and Spanish samples (p<0.05) compared to the rest of locations **(Figure 1b)**. Discriminant
192 Analysis of Principal Components (DAPC), which transforms data using a principal components
193 analysis (PCA) and subsequently identifies clusters using discriminant analysis (DA), showed that
194 South African samples clustered distanced from the other countries. (**Figure 1d**).

195 Despite the differences, a core of 4 genera shared across the four countries was
196 identified, including *Malassezia, Davidiella, Sistotrema* and *Penicillium*. *Wallemia* was only
197 found in samples from Finland, *Trichoderma* in breast milk from Chinese donors, and
198 *Debaromyces* and an unidentified Saccharomycetales in South African samples. *Rhodotorula*
199 was present in samples from other countries except China **(Figure 1c)**.

200 The impact of mode of delivery on mycobiota composition was not consistent across the
201 milk samples from different geographic origins **(Figure 2a)**. However, we found that *Candida*
202 was statistically higher in milk samples from vaginal deliveries (0.89 ± 1.42) compared to C-
203 section births (0.37 ± 0.60, *t* (78) = 2.13, *p* = 0.036) **(Figure 2b).** In addition, LefSe results showed
204 differentially abundant fungi between vaginal and C-section deliveries in each country, at OTU
205 level. In Chinese breast milk samples, *Candida smithsonii* was significantly more abundant in
206 vaginal deliveries; *Sistotrema sp.* in C-section Spanish samples; *Ascomycota sp* in Finnish vaginal

delivery samples: *Malasezzia restricta* in C-section samples; and *Malassezia restricta* and *Davidiella tassiana* in C-section South African samples (LefSe analysis, $p<0.05$) **(Figure 2c)**.

Indexes of alpha diversity and richness across the samples were similar and no statistical differences were observed between geographic locations. Taking into account the mode of delivery, Spanish mothers who delivered by C-section had decreased alpha-diversity (Shannon mean index=2.11, SD=1.0), although differences were not significant (**Figure 3**).

Fungal and bacterial interactions: a network analysis

Network analyses of the bacteria and fungi present in the breast milk samples showed complex interactions intra- and inter-domain, with different associations between organisms depending on the country of origin and delivery mode, some of which were also influenced by maternal features. For example, a *Malassezia* OTU (Fungi_1) correlated positively with a *Streptococcus* (Bact_6) from vaginal delivery samples from Finnish mothers, and with a *Streptococcus* (Bact_1) from C-section deliveries from Finnish samples, whose abundances were dependent of maternal age. The same *Malassezia* OTU correlated positively with several *Streptococcus* OTUs in samples from C-section deliveries from Chinese mothers, and also positively with an Unclassified Bacilli (Bact_2) from South African samples and vaginal deliveries. Significant influence of maternal age and BMI on specific bacterial and fungal organisms were also observed (**Figure 4**).

In order to study the diversity of the most common yeasts in our samples, a phylogenetic tree of the most prevalent *Malassezia* OTUs detected in this work across geographic locations was performed, including known members of the *Malassezia* genus as a reference (**Figure 5**). The tree shows a large diversity of *Malassezia* isolates with similarity to at least four known species, including OTUs which could potentially represent new species. With the exception of one OTU (Fungi 37, which was found to be uniquely present in China), all other sequences were

found in all countries and appear to be therefore ubiquitous. In relation to mode of delivery, all the OTUs were present in breasmilk from mothers with both delivery types **(Figure 5)**.

Discussion

The mycobiome, the fungal fraction of the human microbiome, is present in lower abundances and has been much less explored than the bacterial fraction. However, its importance for human health and disease has stimulated an increased interest on this field[5,6,7,10]. In the infant, fungal species can be detected since very early in life[9,11,44]. However, the infant mycobiome is almost unknown, and information about its development is scarce.

Breast milk is a continuous source of microbes that are transmitted, together with many nutrients and protective compounds. These are continuously delivered to the infant gut during breastfeeding where they play several physiological and immunomodulatory roles[14,15]. Although bacteria inhabiting human breast milk have been extensively studied, the presence of fungi in the fluid had not been assessed until recently, when a diversity of fungal phylotypes in breast milk from healthy Spanish mothers was reported [16]. In the present study, we have characterized breast milk mycobiome and confirmed the presence of diverse fungal communities in Spain, Finland, China, South Africa and Spain.

Fungi were detected in all breast milk samples through massive DNA sequencing, with the two phyla Ascomycota and Basidiomycota being the most prevalent and presenting reciprocal patterns of abundance in all countries except for South Africa, where Ascomycota levels were significantly higher and Basidiomycota lower compared to the other countries. At genus level, *Malassezia* dominated in all countries, followed by *Davidiella.* In our previous work reporting the presence of fungi in breast milk, *Malassezia* also represented the most abundant

254 genus[16]. Other genera found in the current manuscript such as *Alternaria*, *Rhodotorula,*
255 *Saccharomyces*, or *Candida* were also found in the mentioned study.

256 Our findings demonstrate that environmental factors such as geographic location and
257 delivery mode may affect breast milk fungal composition. Samples from South Africa clustered
258 distanced from the other countries in fungal composition. *Malassezia* and *Rhodotorula* genera
259 were significantly influenced by geographic location, with the first being more prevalent in China
260 and Finland, and the second being more prevalent in South Africa and Spain. Nevertheless, a
261 core constituted by four genera, *Malassezia*, *Davidiella*, *Sistotrema* and *Penicillium* was shared
262 among all countries.

263 Breast milk mycobiota profiles did not differ significantly across samples taking into
264 account mode of delivery, although specific fungi, like the genus *Candida* appeared to be more
265 prevalent among samples from vaginal deliveries. A decreased diversity (as measured by the
266 Shannon index) was observed in Spanish samples from C-section deliveries, which correlates
267 with the lower bacterial diversity found in the same samples as described by Kumar *et al*[25],
268 although this difference was not significant.

269 Although the origin of breast milk fungi is unknown, most of the organisms detected in
270 this study can be found in other human niches. *Malassezia* are yeasts whose primary niche is
271 the human body (can you give more specific info on normal presence – in the gut or skin or
272 elsewhere?) (and other animals). In healthy individuals they are part of the normal microbiota
273 where they predominantly colonize the seborrheic parts of the skin[45], and they are commonly
274 found in infants [9,46-48]. *Malassezia* has also been detected in significant abundance in adult[44,45]
275 and infant fecal samples[49], and therefore may play a role in the intestine, and has also been
276 described as an oral commensal[50]. Although *Malassezia* was detected in high proportions in
277 breast milk before, no viable cells could be recovered by classic culture methods[16] and further

efforts should be made to culture this fastidious organism, which has also been shown to be able to penetrate the cell and survive intracellularly.

Davidiella, the second most prevalent fungi found in the samples of this study, has been detected in the only published study about the characterization of vaginal microbiota and mycobiota of asymptomatic women[51]. In the same study, *Candida* was found to be the predominant genus. Therefore, they may play an important role in the early colonization of vaginally born infants. In our previous study on breast milk fungi, *Davidiella* could not be detected. However, *Davidiella* represents the sexual form of the *Cladosporium* genus[52]. Fungi can have an asexual (anamorph) and sexual (teleomorph) form that may be classified into different genera. This sexual dimorphism can be a significant problem when classifying fungal sequences and the use of different databases and/or sequencing of different genes can lead to conflicting classifications. In a study with paediatric Inflammatory bowel disease (IBD) patients, *Cladosporium cladosporiodes* abundance was decreased in IBD, while *Pichia jadinii* and *Candida parapsilosis* increased compared to controls[53].

Candida is probably the most ubiquitous genera of the human mycobiome. It is the major fungal genera detected in the adult oral cavity[54,55] , and has also been detected in the infant mouth, including several species as common inhabitants (*C. parapsilosis , C. tropicalis*, *C. orthopsilosis,* etc.*)* [9,56,57]. Several *Candida* species are also commonly present in the adult skin and fecal samples[5], and in the infant anus and fecal samples[9,58]. *Candida* is also the most prevalent fungi in the vaginal mycobiome of healthy women[51], although it can expand and cause vaginal infections [59]. Transmission of *Candida* from mother to infant likely occurs, as the same fingerprinting of the DNA has showed identity between maternal *Candida* from vagina, rectum, oral cavity and skin, and infant oral cavity and rectum[11].

Other prevalent fungi detected in our samples are also present in several body niches. *Saccharomyces* are among the most abundant fungi in the gut[5], and *Saccharomyces cerevisiae*

has been reported to be highly prevalent and abundant in the infant oral and anal mycobiome during the first month of life[9]. In a recent study, bacteria and fungi from fecal samples in children suffering atopic wheeze where analysed, and Saccharomycetales taxa appeared decreased in the atopic wheeze group, while the species *Pichia kudriavzevii* was increased compared to controls[10]. Others such as *Penicillium* or *Aspergillus* can also be detected in fecal samples, and *Debaromyces hansenii* represents one of the main species present in breastfed infants' gut[12]. In the present study, we have detected *Debaromyces* although none of the sequences have been classified as *D. hansenii*. However, DNA from this species was previously detected in breast milk[16].

The study of inter-species interactions within a population is necessary to better understand the microbiota's role. It is known that microorganisms can interact by competition and sometimes collaboration, thereby influencing microbiota composition and host's health. It has been demonstrated that cross-talk between bacteria and fungi can exist, modulating host defence mechanisms, protecting against infections or collaborating to cause them[60]. For example, synergies between oral *S. oralis* and *C. albicans* increased *C. albicans* invasion through activation of host enzymes that cleave epithelial junctions proteins[61]. On the contrary, *S. mutans* showed ability to modulate biofilm formation and to reduce *C. albicans* virulence in an animal model[61]. Some vaginal isolates of *Lactobacilli* strains have shown anti-fungal activity *in vitro* against *Candida spp*[10] and probiotic *L. rhamnosus* and *L. reuterii* strains showed *in vitro* efficacy against *C. albicans*[10], a pervasive causal agent of vaginal infections. To understand microbial relationships, microbial networks analyses are indispensable by identifying and representing the most influential members in a bacterial community and their interactions with other microorganisms[62]. In a recent work, bacterial interactions in colostrum and mature milk of Italian and Burundian mothers were analysed, and showed different bacterial networks among the two populations. The identified networks were complex and dynamic, changing from colostrum to mature milk[63]. In the present study, we have analysed interactions between bacteria and fungi

in breast milk, observing a complex network of interactions between fungi and bacteria, in addition to relationships within the same domain. Microbial interactions were dependent of delivery characteristics (mode of delivery and geographic location), and maternal features (BMI and age) influenced the prevalence of particular microorganisms. Interesting positive correlations were observed between several *Malassezia* the most prevalent fungi detected in breast milk by sequencing, and different streptococci, which reprensent one of the most common bacterial genera in breast milk[64]. Interestingly, in our previous study we observed a significant positive correlation between *Malassezia* and bacterial load[16], and further experimental research should analyse potential synergistic relationships between genera.

Our data support the potential role of breast milk on the initial seeding of fungal species to the infant gut mycobiome. A greater understanding of the influence of the environment on the bacterial and fungal communities and their metabolic potential is necessary, as this will allow us to design and customize new strategies to modulate and maintain homeostasis and direct an adequate intestinal colonization in children. This can be achieved through microbiome cohort-based studies of human populations across continents.

Data availability

All ITS1 sequences have been deposited in the European National Archive (ENA) server under the study ID PRJEB25581. Samples accession IDs: ERS2311788-2311867.

Acknowledgements

Competing interests

Conflict of interest. The authors declare that they have no conflict of interest.

References

1. Chervonsky, A. V. Influence of microbial environment on autoimmunity. *Nat. Immunol.* **11,** 28–35 (2010).
2. Tamburini, S., Shen, N., Wu, H. C. & Clemente, J. C. The microbiome in early life: implications for health outcomes. *Nat. Med.* **22,** 713–722 (2016).
3. Houghteling, P. D. & Walker, W. A. Why Is Initial Bacterial Colonization of the Intestine Important to Infants' and Children's Health? *J. Pediatr. Gastroenterol. Nutr.* **60,** 294–307 (2015).
4. Hatoum, R., Labrie, S. & Fliss, I. Antimicrobial and Probiotic Properties of Yeasts: From Fundamental to Novel Applications. *Front. Microbiol.* **3,** 421 (2012).
5. Underhill, D. M. & Iliev, I. D. The mycobiota: interactions between commensal fungi and the host immune system. *Nat. Rev. Immunol.* **14,** 405–16 (2014).
6. Cui, L., Morris, A. & Ghedin, E. The human mycobiome in health and disease. *Genome Med.* **5,** 63 (2013).
7. Sokol, H. *et al.* Fungal microbiota dysbiosis in IBD. *Gut* gutjnl-2015-310746 (2016). doi:10.1136/gutjnl-2015-310746
8. Ward, T. L., Knights, D. & Gale, C. A. Infant fungal communities: current knowledge and research opportunities. *BMC Med.* **15,** 30 (2017).
9. Ward, T. L. *et al.* Development of the Human Mycobiome over the First Month of Life and across Body Sites. *mSystems* **3,** e00140-17 (2018).
10. Parolin, C. *et al.* Isolation of Vaginal Lactobacilli and Characterization of Anti-Candida Activity. *PLoS One* **10,** e0131220 (2015).
11. Bliss, J. M., Basavegowda, K. P., Watson, W. J., Sheikh, A. U. & Ryan, R. M. Vertical and Horizontal Transmission of Candida albicans in Very Low Birth Weight Infants Using DNA Fingerprinting Techniques. *Pediatr. Infect. Dis. J.* **27,** 231–235 (2008).
12. Schei, K. *et al.* Early gut mycobiota and mother-offspring transfer. *Microbiome* **5,** 107 (2017).
13. Drell, T. *et al.* The Influence of Different Maternal Microbial Communities on the Development of Infant Gut and Oral Microbiota. *Sci. Rep.* **7,** 9940 (2017).

14. Jost, T., Lacroix, C., Braegger, C. P., Rochat, F. & Chassard, C. Vertical mother-neonate transfer of maternal gut bacteria via breastfeeding. *Environ. Microbiol.* **16,** 2891–2904 (2014).

15. Allan Walker, W. & Shuba Iyengar, R. Breast milk, microbiota, and intestinal immune homeostasis. (2014). doi:10.1038/pr.2014.160

16. Boix-Amorós, A., Martinez-Costa, C., Querol, A., Collado, M. C. & Mira, A. Multiple Approaches Detect the Presence of Fungi in Human Breastmilk Samples from Healthy Mothers. *Sci. Rep.* **7,** 13016 (2017).

17. Mason, K. L. *et al.* Candida albicans and Bacterial Microbiota Interactions in the Cecum during Recolonization following Broad-Spectrum Antibiotic Therapy. *Infect. Immun.* **80,** 3371–3380 (2012).

18. Peleg, A. Y., Hogan, D. A. & Mylonakis, E. Medically important bacterial–fungal interactions. *Nat. Rev. Microbiol.* **8,** 340–349 (2010).

19. Hoffmann, C. *et al.* Archaea and fungi of the human gut microbiome: correlations with diet and bacterial residents. *PLoS One* **8,** e66019 (2013).

20. Cabrera-Rubio, R. *et al.* The human milk microbiome changes over lactation and is shaped by maternal weight and mode of delivery. *Am J Clin Nutr* **96,** 544–51 (2012).

21. Cabrera-Rubio, R., Mira-Pascual, L., Mira, A. & Collado, M. C. Impact of mode of delivery on the milk microbiota composition of healthy women. *J. Dev. Orig. Health Dis.* **7,** 54–60 (2016).

22. Khodayar-Pardo, P., Mira-Pascual, L., Collado, M. C. & Martínez-Costa, C. Impact of lactation stage, gestational age and mode of delivery on breast milk microbiota. *J. Perinatol.* **34,** 599–605 (2014).

23. Hoashi, M. *et al.* Human Milk Bacterial and Glycosylation Patterns Differ by Delivery Mode. *Reprod. Sci.* **23,** 902–907 (2016).

24. Toscano, M. *et al.* Impact of delivery mode on the colostrum microbiota composition. *BMC Microbiol.* **17,** 205 (2017).

25. Kumar, H. *et al.* Distinct Patterns in Human Milk Microbiota and Fatty Acid Profiles Across Specific Geographic Locations. *Front. Microbiol.* **7,** 1619 (2016).

26. Li, S.-W. *et al.* Bacterial Composition and Diversity in Breast Milk Samples from Mothers Living in Taiwan and Mainland China. *Front. Microbiol.* **8,** 965 (2017).

27. Toju, H., Tanabe, A. S., Yamamoto, S. & Sato, H. High-Coverage ITS Primers for the DNA-Based Identification of Ascomycetes and Basidiomycetes in Environmental Samples. (2012). doi:10.1371/journal.pone.0040863

28. Magoc, T. & Salzberg, S. L. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27,** 2957–2963 (2011).

29. Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C. & Knight, R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27,** 2194–2200 (2011).

30. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naive Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Appl. Environ. Microbiol.* **73,** 5261–5267 (2007).

31. Kõljalg, U. *et al.* Towards a unified paradigm for sequence-based identification of fungi.

Mol. Ecol. **22,** 5271–5277 (2013).

32. Strati, F. *et al.* Age and Gender Affect the Composition of Fungal Population of the Human Gastrointestinal Tract. *Front. Microbiol.* **7,** 1227 (2016).

33. R Development Core Team. R: A Language and Environment for Statistical Computing.Vienna, Austria : the R Foundation for Statistical Computing. ISBN: 3-900051-07-0. Available online at http://www.R-project.org/. (2011).

34. Zakrzewski, M. *et al.* Calypso: a user-friendly web-server for mining and visualizing microbiome–environment interactions. *Bioinformatics* **33,** btw725 (2016).

35. Segata, N. *et al.* Metagenomic biomarker discovery and explanation. *Genome Biol.* **12,** R60 (2011).

36. Cole, J. R. *et al.* Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* **42,** D633–D642 (2014).

37. Albanese, D., Riccadonna, S., Donati, C. & Franceschi, P. A practical tool for Maximal Information Coefficient analysis. *bioRxiv* 215855 (2017). doi:10.1101/215855

38. Shannon, P. *et al.* Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **13,** 2498–2504 (2003).

39. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30,** 772–780 (2013).

40. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32,** 1792–1797 (2004).

41. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30,** 1312–1313 (2014).

42. Ronquist, F. *et al.* MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. *Syst. Biol.* **61,** 539–542 (2012).

43. Stöver, B. C. & Müller, K. F. TreeGraph 2: Combining and visualizing evidence from different phylogenetic analyses. *BMC Bioinforma. 2010 111* **11,** 7 (2010).

44. Seddik, H. A., Ceugniez, A., Bendali, F., Cudennec, B. & Drider, D. Yeasts isolated from Algerian infants's feces revealed a burden of Candida albicans species, non-albicans Candida species and Saccharomyces cerevisiae. *Arch. Microbiol.* **198,** 71–81 (2016).

45. Findley, K. *et al.* Topographic diversity of fungal and bacterial communities in human skin. *Nature* **498,** 367–370 (2013).

46. Jo, J.-H. *et al.* Diverse Human Skin Fungal Communities in Children Converge in Adulthood. *J. Invest. Dermatol.* **136,** 2356–2363 (2016).

47. Gouba, N., Raoult, D. & Drancourt, M. Plant and Fungal Diversity in Gut Microbiota as Revealed by Molecular and Culture Investigations. *PLoS One* **8,** e59474 (2013).

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

49. Strati, F. *et al.* Age and Gender Affect the Composition of Fungal Population of the Human Gastrointestinal Tract. *Front. Microbiol.* **7,** 1227 (2016).

50. Dupuy, A. K. *et al.* Redefining the human oral mycobiome with improved practices in

amplicon-based taxonomy: discovery of Malassezia as a prominent commensal. *PLoS One* **9,** e90899 (2014).

51. Drell, T. *et al.* Characterization of the Vaginal Micro- and Mycobiome in Asymptomatic Reproductive-Age Estonian Women. *PLoS One* **8,** e54379 (2013).

52. Schubert, K. *et al.* Biodiversity in the Cladosporium herbarum complex (Davidiellaceae, Capnodiales), with standardisation of methods for Cladosporium taxonomy and diagnostics. *Stud. Mycol.* **58,** 105–56 (2007).

53. Chehoud, C. *et al.* Fungal Signature in the Gut Microbiota of Pediatric Patients With Inflammatory Bowel Disease. *Inflamm. Bowel Dis.* **21,** 1948–56 (2015).

54. Ghannoum, M. A. & Mukherjee, P. K. The human mycobiome and its impact on health and disease. *Curr. Fungal Infect. Rep.* **7,** 345–350 (2013).

55. Kraneveld, E. A. *et al.* The Relation between Oral Candida Load and Bacterial Microbiome Profiles in Dutch Older Adults. *PLoS One* **7,** e42770 (2012).

56. Stecksén-Blicks, C., Granström, E., Silfverdal, S. A. & West, C. E. Prevalence of oral *Candida* in the first year of life. *Mycoses* **58,** 550–556 (2015).

57. Kleinegger, C. L., Lockhart, S. R., Vargas, K. & Soll, D. R. Frequency, intensity, species, and strains of oral Candida vary as a function of host age. *J. Clin. Microbiol.* **34,** 2246–54 (1996).

58. LaTuga, M. S. *et al.* Beyond Bacteria: A Study of the Enteric Microbial Consortium in Extremely Low Birth Weight Infants. *PLoS One* **6,** e27858 (2011).

59. Trama, J. P., Mordechai, E. & Adelson, M. E. Detection and identification of Candida species associated with Candida vaginitis by real-time PCR and pyrosequencing. *Mol. Cell. Probes* **19,** 145–152 (2005).

60. Ten Oever, J. & Netea, M. G. The bacteriome–mycobiome interaction and antifungal host defense. *Netea Eur. J. Immunol* **44,** 3182–3191 (2014).

61. Xu, H., Sobue, T., Bertolini, M., Thompson, A. & Dongari-Bagtzoglou, A. *Streptococcus oralis* and *Candida albicans* Synergistically Activate μ-Calpain to Degrade E-cadherin From Oral Epithelial Junctions. *J. Infect. Dis.* **214,** 925–934 (2016).

62. Layeghifard, M., Hwang, D. M. & Guttman, D. S. Disentangling Interactions in the Microbiome: A Network Perspective. *Trends Microbiol.* **25,** 217–228 (2017).

63. Drago, L. *et al.* Microbiota network and mathematic microbe mutualism in colostrum and mature milk collected in two different geographic areas: Italy versus Burundi. *ISME J.* **11,** 875–884 (2017).

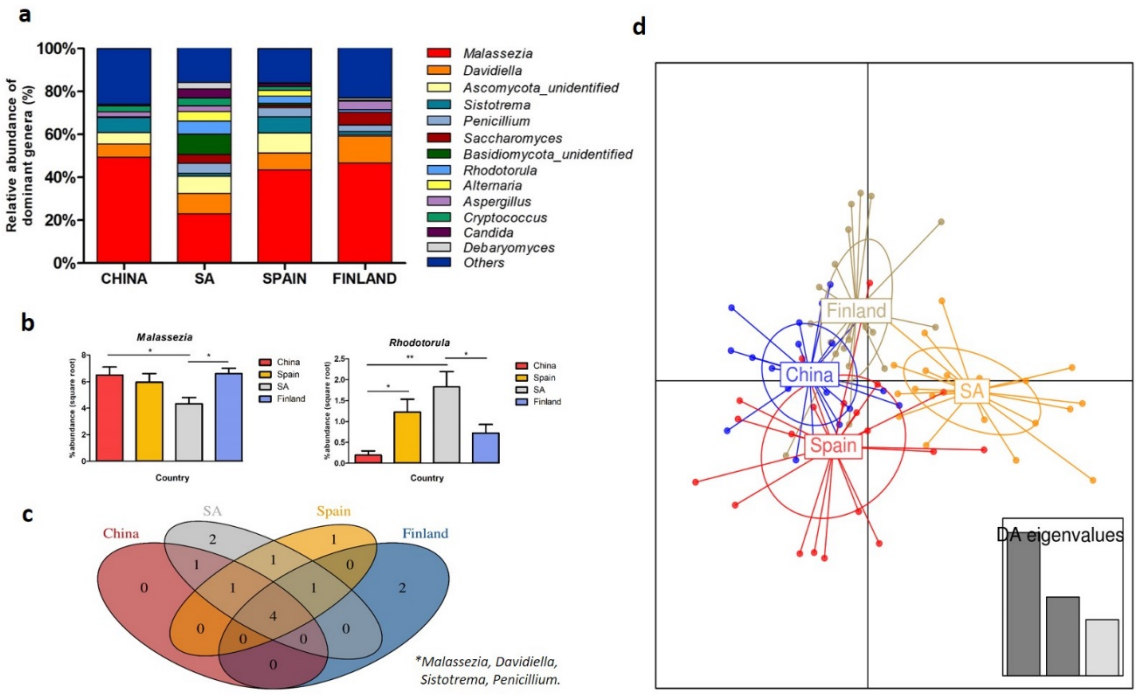
64. Jost, T., Lacroix, C., Braegger, C. & Chassard, C. Assessment of bacterial diversity in breast milk using culture-dependent and culture-independent approaches. *Br. J. Nutr.* **110,** 1253–1262 (2013).

Tables

Table 1. Clinical characteristics of donors providing samples for the study.

	Delivery mode	Age	P value	BMI ± SD	P value
Finland	C-section (10)	35.20 ± 4.07		24.70 ± 2.89	0.185
	Vaginal (10)	33.70 ± 6.02	0.820	23.41 ± 4.60	
	Total (20)	34.45 ± 5.06	ns	24.47 ± 6.46	ns
Spain	C-section (10)	34.50 ± 2.59		24.34 ± 1.47	
	Vaginal (10)	32.20 ± 5.16	0.288	24.25 ± 1.43	0.630
	Total (20)	33.35 ± 4.14	ns	24.30 ± 1.41	ns
South Africa	C-section (10)	36.60 ± 6.08		26.67 ± 1.41	
	Vaginal (10)	31.50 ± 5.76	0.944	24.81 ± 2.67	0.043
	Total (20)	34.05 ± 2.29	ns	25.75 ± 2.29	ns
China	C-section (10)	32.60 ± 2.95		21.49 ± 2.29	
	Vaginal (10)	31.90 ± 4.25	0.970	21.92 ± 1.54	0.449
	Total (20)	32.25 ± 3.58	ns	21.71 ± 1.97	0.004
All	C-section (10)	34.72 ± 4.25		24.70 ± 2.83	
	Vaginal (10)	32.32 ± 5.20	0.058	23.41 ± 2.11	0.072
	Total (20)	33.52 ± 4.87	ns	24.06 ± 3.85	ns

Figures



530

531

532 **Figure 1. Effect of geographical location on fungal composition in breast milk samples. (a)**
533 Fungal relative abundances at genus level across countries. Only genera present in more than
534 1% abundance in at least 20% of the samples are represented. **(b)** Genera significantly influenced
535 by geographic location. Corrected square root of genera abundances are represented in the y
536 axis. **(c)** Shared phylotypes across countries at genus level. *, core of four fungal genera shared
537 across geographic locations. Venn's diagram cut-off: 0.5 **(d)** DAPC analysis showing relationships
538 in fungal composition among samples from different locations. n=80 (n per country=20).

539

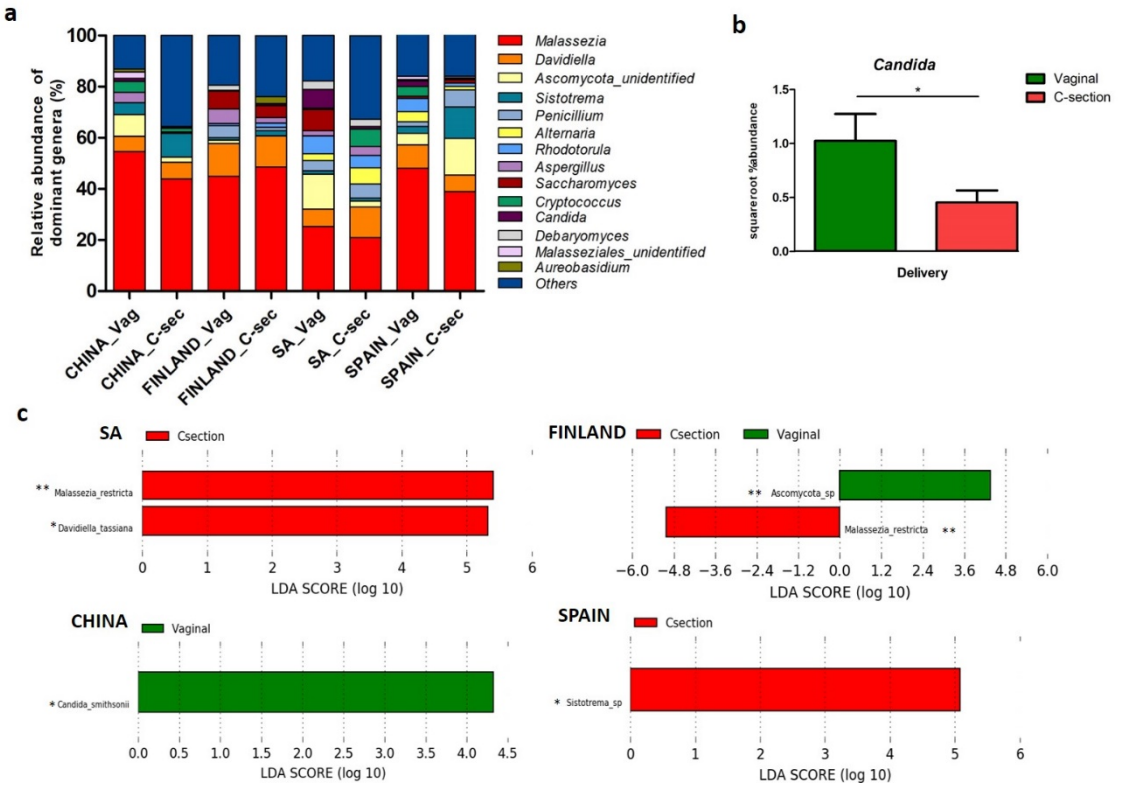
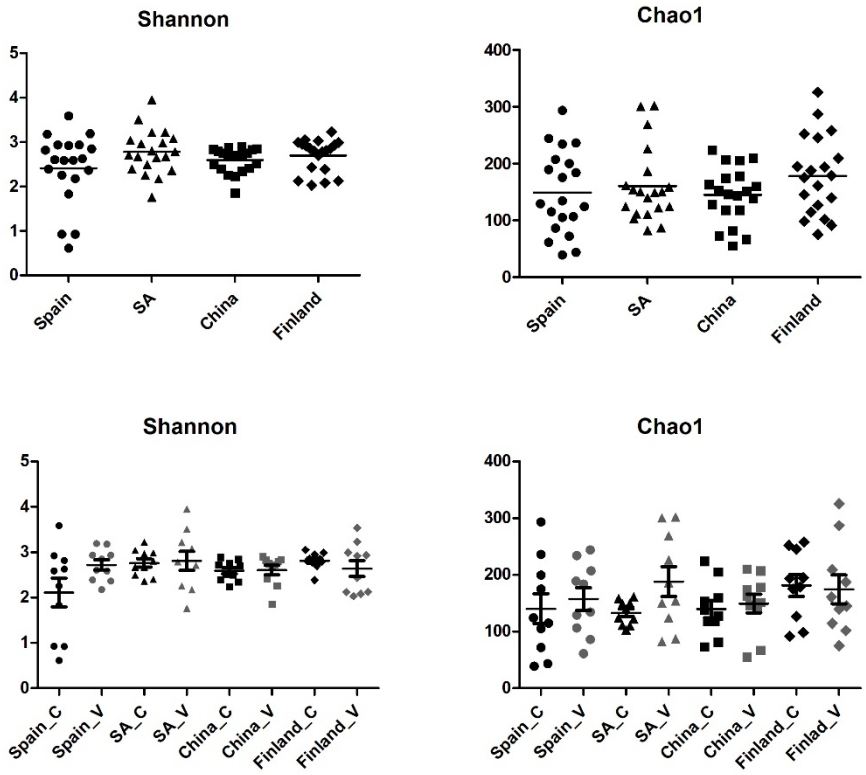


Figure 2. Differences in fungal composition in vaginal and C-section deliveries per country. (a) Comparison of fungal relative abundances per country and mode of delivery. Only genera present in more than 1% abundance in at least 20% of the samples are represented. **(b)** Corrected square root of genera abundances of the only genus found to be significantly influenced by mode of delivery. **(c)** Differentially abundant species in breast milk samples depending on delivery mode and geographic location, as detected by the LEfSe algorithm. The threshold for logarithmic discriminant analysis *(LDA)* score was 2. *, $P < 0.05$ and **, $P < 0.01$. n=80 (vaginal deliveries n=40, C-section deliveries n=40).



549

550 **Figure 3. Fungal diversity and richness in human breast milk samples.** Plots show Shannon
551 diversity and Chao1 richness indexes per country (upper panels), and by mode of delivery (lower
552 panels). Means and standard errors are included in the plots. SA= South Africa; V= vaginal
553 delivery; C= C-section delivery.

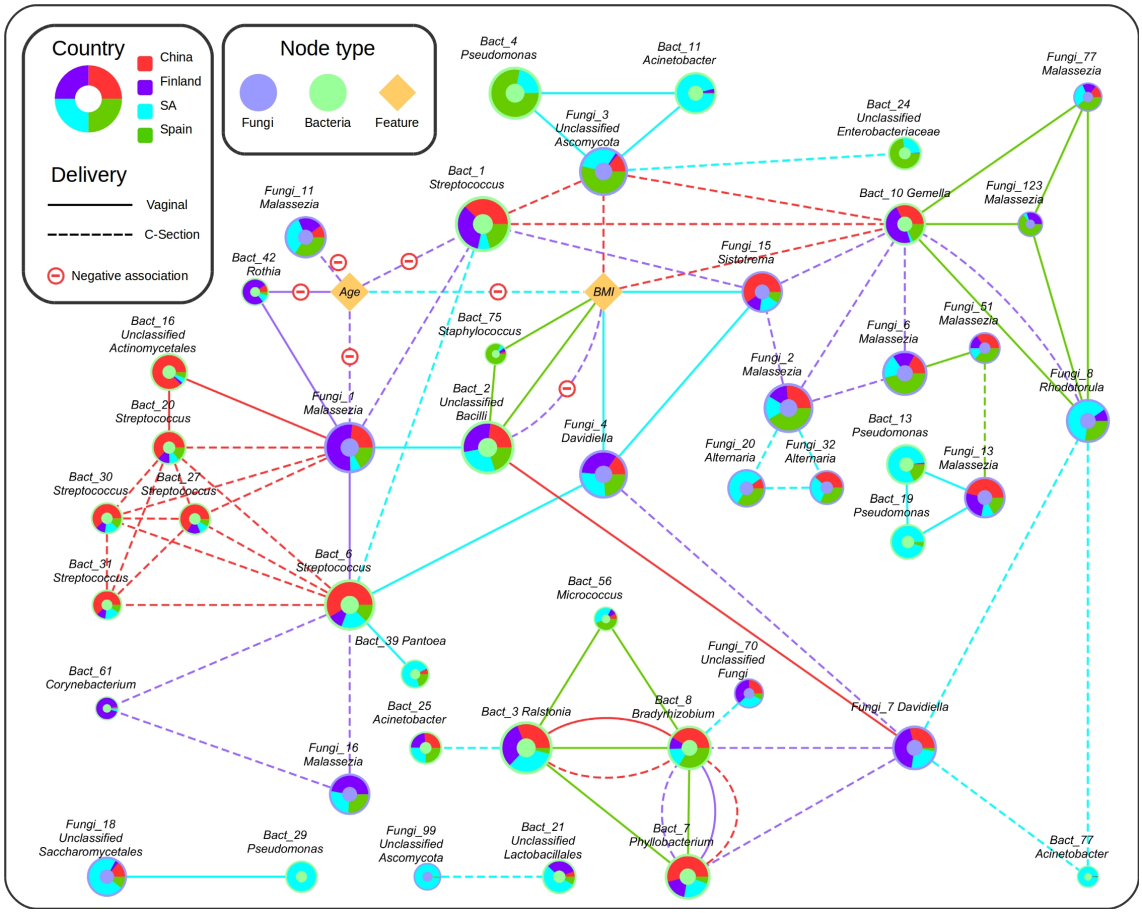
554

555

556

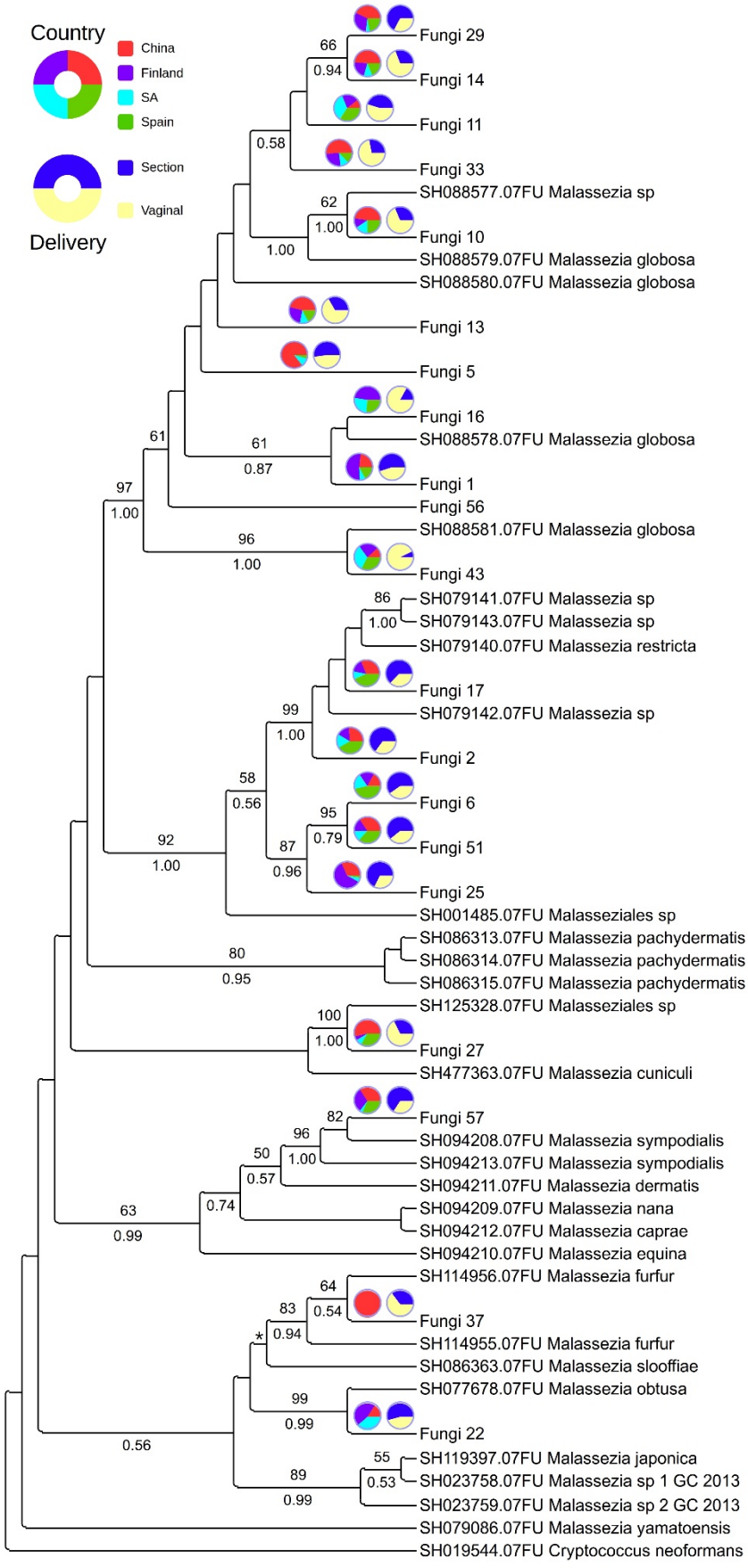
557

558



559

560 **Figure 4. Correlations between bacteria and fungi in breastmilk samples depending on**
561 **maternal features and delivery mode.** Green nodes represent bacterial OTUs, blue nodes
562 represent fungal OTUs, and yellow nodes represent features. Nodes size indicates OTU
563 abundance. Pie chart colours represent the overall distribution of each OTU across countries.
564 Each link indicates a significant ($p<0.05$) interaction between OTUs or features in samples from
565 a given combination of country and delivery mode (Vaginal, C-section). Link colour denotes the
566 country, and line type indicates delivery mode.


Country
China
Finland
SA
Spain
Section
Vaginal
Delivery
Fungi 29
Fungi 14
Fungi 11
Fungi 33
SH088577.07FU Malassezia sp
Fungi 10
SH088579.07FU Malassezia globosa
SH088580.07FU Malassezia globosa
Fungi 13
Fungi 5
Fungi 16
SH088578.07FU Malassezia globosa
Fungi 1
Fungi 56
SH088581.07FU Malassezia globosa
Fungi 43
SH079141.07FU Malassezia sp
SH079143.07FU Malassezia sp
SH079140.07FU Malassezia restricta
Fungi 17
SH079142.07FU Malassezia sp
Fungi 2
Fungi 6
Fungi 51
Fungi 25
SH001485.07FU Malasseziales sp
SH086313.07FU Malassezia pachydermatis
SH086314.07FU Malassezia pachydermatis
SH086315.07FU Malassezia pachydermatis
SH125328.07FU Malasseziales sp
Fungi 27
SH477363.07FU Malassezia cuniculi
Fungi 57
SH094208.07FU Malassezia sympodialis
SH094213.07FU Malassezia sympodialis
SH094211.07FU Malassezia dermatis
SH094209.07FU Malassezia nana
SH094212.07FU Malassezia caprae
SH094210.07FU Malassezia equina
SH114956.07FU Malassezia furfur
Fungi 37
SH114955.07FU Malassezia furfur
SH086363.07FU Malassezia slooffiae
SH077678.07FU Malassezia obtusa
Fungi 22
SH119397.07FU Malassezia japonica
SH023758.07FU Malassezia sp 1 GC 2013
SH023759.07FU Malassezia sp 2 GC 2013
SH079086.07FU Malassezia yamatoensis
SH019544.07FU Cryptococcus neoformans

567

Figure 5. Molecular phylogenetic tree inferred from a maximum likelihood analysis of ITS sequences of the *Malassezia* OTUs obtained in this work and known members of the genus *Malassezia*. ML support values > 50% over 10,000 replicates are shown above the branches. For branches that were also supported by Bayesian inference, the posterior probability is shown below the branches. Brackets surrounding posterior probability values show a conflict between the Bayesian inference and maximum likelihood analysis, in which *M. nana* clustered in the *M. restricta* branch in maximum likelihood analysis, but outside it in bayesian inference. The tree is rooted with *Cryptococcus neoformans.* Pie charts indicate prevalence of each OTU per country and mode of delivery. The 20 most prevalent *Malassezia* OTUs found in this work are included in the tree.