International PhD program in Biomolecular Sciences

Centre for Integrative Biology

XXIX Cycle

The microbiota-gut-brain axis: characterization of the gut microbiota in neurological disorders

Tutor:
Olivier Jousson
Centre for Integrative Biology
University of Trento

Supervisor:
Claudio Donati
Computational Biology Research Unit
Fondazione Edmund Mach

Ph.D. thesis of
Francesco Strati
University of Trento
Fondazione Edmund Mach

Academic Year 2015/2016

Declarations

I hereby confirm that this is my own work and the use of all material from external sources has been properly and fully acknowledged.

The research work for this Ph.D. thesis was carried out at and founded with a scholarship from Fondazione Edmund Mach and co-founded by Unifarm S.p.A.

Abstract

The human gut microbiota plays a crucial role in the functioning of the gastrointestinal tract and its alteration can lead to gastrointestinal abnormalities and inflammation. Additionally, the gut microbiota modulates central nervous system (CNS) activities affecting several aspect of host physiology. Motivated by the increasing evidences of the role of the gut microbiota in the complex set of interactions connecting the gut and the CNS, known as gut-brain axis, in this Ph.D. thesis we asked whether the gastrointestinal abnormalities and inflammation commonly associated with neurological disorders such as Rett syndrome (RTT) and Autism could be related to alterations of the bacterial and fungal intestinal microbiota.

First, since only few reports have explored the fungal component of the gut microbiota in health and disease, we characterized the gut mycobiota in a cohort of healthy individuals, in order to reduce the gap of knowledge concerning factors influencing the intestinal microbial communities. Next, we compared the gut microbiota of three cohorts of healthy, RTT and autistic subjects to investigate if these neurological disorders harbour alterations of the gut microbiota.

Culture-based and metataxonomics analysis of the faecal fungal populations of healthy volunteers revealed that the gut mycobiota differs in function of individuals' life stage in a gender-related fashion. Different fungal species were isolated showing phenotypic adaptation to the intestinal environment. High frequency of azoles resistance was also found, with potential clinical significance.

It was further observed that autistic subjects are characterized by a reduced incidence of *Bacteroidetes* and that *Collinsella*, *Corynebacterium*, *Dorea* and *Lactobacillus* were the taxa predominating in the gut microbiota of autistic subjects. Constipation has been associated with different bacterial patterns in autistic and neurotypical subjects, with constipated autistic individuals characterized by higher levels of *Escherichia/Shigella* and *Clostridium cluster XVIII* than constipated neurotypical subjects.

RTT is a neurological disorder caused by loss-of-function mutations of *MeCP2* and it is commonly associated with gastrointestinal dysfunctions and constipation. We showed that RTT subjects harbour bacterial and fungal microbiota altered from those of healthy controls, with a reduced microbial richness and dominated by *Bifidobacterium*, different Clostridia and *Candida*. The alterations of the gut microbiota observed did not depend on the constipation status of RTT subjects while this microbiota produced altered SCFAs profiles potentially contributing to the constipation itself.

Phenotypical and immunological characterizations of faecal fungal isolates from RTT subjects showed *Candida parapsilosis* as the most abundant species isolated in RTT, genetically unrelated to healthy controls' isolates and with elevated resistance to azoles. Furthermore these isolates induced high levels of IL-10 suggesting increased tolerance and persistence within the host.

Finally, the importance of multiple sequence alignment (MSA) accuracy in microbiome research was investigated comparing three implementations of the widely used NAST algorithm. By now, different implementations of NAST have been developed but no one tested the performances and the accuracy of the MSAs generated with these implementations. We showed that micca, a new bioinformatics pipeline for metataxonomics data improves the quality of NAST alignments by using a fast and memory efficient reimplementation of the NAST algorithm.

Contents

Overview of the Ph.D. thesis

The importance of the human gut microbiota in host physiology is now well known, ranging from the modulation of host metabolism to immune system homeostasis. Nevertheless the ability of the gut microbiota to affect complex physiological activities of the central nervous system still remains elusive. In particular its contribution to the pathophysiology of extra-intestinal, neurological disorders has been suggested but not completely unveiled. Several studies reported the importance that alterations of the fungal gut microbiota, together with its bacterial counterpart, can have in different pathologies. However, a comprehensive knowledge of what actually a "*healthy gut mycobiota*", both in terms of structure, diversity and functions, is still missing.

To lay the basis for a more complete understanding of the role of the fungal intestinal microbiota in health and disease, in the first chapter of this Ph.D. thesis we characterized the gut mycobiota in a cohort of healthy individuals by using a mixture of culture-dependent and culture-independent methods. Additionally, given the increasing appreciation of the role of the microbiome-gut-brain axis and its impact on neuronal and intestinal pathophysiology, in the second, third and fourth chapter of the Ph.D. thesis we studied the intestinal microbial communities in autism spectrum disorder and Rett syndrome, two neurological disorders with different aetiology but sharing some neurological and gastrointestinal symptoms (constipation, inflammation, restricted repetitive behaviour, *etc.*), in order to elucidate the role of the gut microbiota in the gastrointestinal abnormalities and inflammatory disorders commonly associated with these pathologies. To give a more complete picture of the role of the gut microbiota in these states, both fungal and bacterial components have been characterized through amplicon-based sequencing of the bacterial V3-V5 variable regions of the 16S rRNA gene and of the ITS1 region of the fungal rRNA genes. Finally, due to the significant impact on the calculation of phylogenetic and pairwise genetic distances for microbial ecology estimates, we investigated the importance of multiple sequence alignments (MSA) in microbiome research. MSA algorithms can rely on two approaches: *de novo* methods which are hardly applicable for metagenomics data analysis, both in terms of accuracy and computational performances (due to the type and size of datasets used), and reference-based methods which rely on the use of curated templates of 16S rRNA sequences. The reference-based methods are the preferred alignment methods for metagenomics studies since they allow the generation of accurate MSAs that incorporate the predicted secondary structure of the 16S rRNA gene. Therefore, in the fifth chapter of this thesis, we compared the performances and accuracy of different implementations of the widely used reference-based aligner NAST on synthetic and biological 16S rRNA sequence datasets. The thesis gathers together results recently published and manuscripts submitted or about to be published and it is

structured into five chapters: i) *Age and gender affect the composition of fungal population of the human gastrointestinal tract*; ii) *New evidences on the altered gut microbiota in autism spectrum disorders*; iii) *Altered gut microbiota in Rett syndrome;* iv) *Intestinal Candida isolates from Rett syndrome subjects bear potential virulent traits and capacity to persist within the host*; v) *The importance of multiple sequence alignments in microbiome research: comparison of NAST algorithm implementations.* Each chapter is preceded by a brief resume of the main findings and a personal statement regarding my involvement and contribution in each project.

Introduction

The human gut microbiota: ruler of host physiology

The human gastrointestinal (GI) tract is a complex ecological niche in which all the three domains of life (Archaea, Bacteria and Eukarya) and Virus co-exist in close association with the host [1-3]. This complex microbial community, known as the gut microbiota, has co-evolved with the host in a mutualistic relationship that influences many physiological functions such as host energy metabolism, development and function of the immune system, modulation of neurological and behavioural states of the host [4-6]. The development of the gut microbiota starts at birth with colonization by a low number of species from the vaginal and faecal microbiota of the mother and is characterized by many shifts in composition during infancy [7]. Furthermore the initial composition of the infant gut microbiota depends on the delivery mode with infants delivered by C-section harbouring a gut microbiota characteristics of the skin rather than of the vaginal tract [8]. Since the microbial community structure during the first year of life has a long-lasting influence on metabolic and immune homeostasis of the host it has been hypothesized that C-section delivery may increase the risk of certain diseases such as celiac disease, type 1 diabetes (T1D), and asthma due to aberrant T-helper responses [9]. Breast-feeding further influences the composition of the gut microbiota through transmission of milk-derived microbiota and selection of bacterial species adapted to the utilization of human milk oligosaccharides such as the bifidobacteria, prominent taxa of the infant gut microbiota [10, 11]. In this way, the infant gut microbiota is shaped from continuous exposures to the microbiota of parents, environmental factors and diet, becoming comparable to the adult microbiota by 2 years of age [12]. From an initial gut microbiota characterized by low ecological richness and diversity during infancy, the adult intestinal microbiota become a complex ecosystem composed of thousands species belonging principally to the bacterial phyla *Firmicutes* (principally *Clostridium*, *Ruminococcus*, *Faecalibacterium*, *Lactobacillus*) and *Bacteroidetes* (principally *Prevotella* and *Bacteroides*). *Actinobacteria*, *Fusobacteria*, *Proteobacteria* and *Verrucomicrobia* are also present but in small proportions [3], despite their importance in health and disease [10, 13]. Although the core gut microbiota in adulthood is quite stable (more in term of encoded functions rather that taxonomy [14], variables like age, sex, genetic and environmental factors, medical treatments and diet are accountable for inter-individual differences in the intestinal microbial structure. For example sex hormones determine some differences observed between genders [15] and diet can shift the microbiota composition in a sex-dependent manner [16]. Indeed it is exactly the diet that has the major impact on microbiota assembly. The switch from a high-fat/low-fibre diet to a low-fat/high-

fibre diet causes notable changes in the gut microbiota within 24 hr [17]. In a fine study on the effect of short-term consumption of an animal-based or plant-based diet it has been showed how the human gut microbiota changes drastically and rapidly in response to dietary lifestyles [18]. Such changes reflect trade-offs between carbohydrate and protein fermentation which favour (or not) the overgrowth of microbial taxa promoting or protecting against intestinal inflammation. Cross-sectional studies on different human populations have shown that there are evident differences between the gut microbiota composition of individuals from Western countries *e.g.* United States and Italy and from non-Westernised countries such as Burkina Faso, Malawi and the Amazonas states of Venezuela [19, 20]. Although these differences may also be ascribed to the distinct genetic pools of these populations, cultural factors related to diet, hygienic habits and antibiotic usage are critical in shaping the gut microbiota. Indeed the populations that have a diet rich in fibres tend to have a gut microbiota enriched in bacterial taxa known with enzymatic repertoires for the hydrolysis of complex plant polysaccharides (*e.g. Prevotella* or *Xylanibacter*). The fermentation of such fibres produces relevant amounts of Short Chain Fatty Acids (SCFAs) (especially acetate, propionate and butyrate) which are important for colon health because of their immune-regulatory activities [21] and are a primary energy source for colonic cells [22]. By contrast the gut microbiota of Western populations is deprived of these microorganism and tend to have less SCFAs [19]. Interestingly the incidence of Inflammatory Bowel Diseases (IBDs) and other immunological disorders is greater in industrialized populations than in communities in developing countries living as farmers or hunter gatherers. Intestinal cells interact with the gut microbiota secreting signalling molecules, such as gut hormones, chemokines and cytokines, in a fine interplay between the gut microbiota and the host that modulate the functionality of host tissues and organs maintaining the homeostatic control [23]. The subtle equilibrium between the gut microbiota and the host is a key element in human health. In fact, alterations in the composition of the microbial community structure, termed as "*dysbiosis*", have been associated to an increasing number of health conditions such as metabolic disorders (*e.g*. diabetes, obesity) [24-26], blood pressure and heart disease [27], autoimmune disorders [28] and Autism Spectrum Disorders (ASDs) [29].

Mechanisms of interaction between the gut microbiota and the immune system

The immune system and the gut microbiota start developing together since the beginning of birth and it has been hypothesized that their co-evolution maintains and selects mutualistic or symbiotic microorganisms within the GI niche [30]. This early co-existence is necessary to avoid phenomena of undesirable reactions against the *healthy* gut microbiota. Indeed the intestine and its associated immunological components have to deal with several, in some cases incompatible, tasks. Apart from

all the functions related to digestion and absorption of nutrients, the intestine has to be tolerant towards mutualistic/commensal microorganisms and to keep control over pathobionts (*i.e.* those resident microbes with pathogenic potential [31]), preventing microbial overgrowth and invasion of the epithelial intestinal barrier. In turn the gut microbiota has to modulate and regulate several aspects of host's immune system towards tolerance rather than responsiveness. The impact of the gut microbiota on intestinal and systemic immunity regulation has been demonstrated by studying germ-free (GF) mice models that develop an altered immune system and are more prone to microbial infections and colitis [32]. Before birth, when the gut of the foetus is not colonized by microbes, the lymphoid tissue inducer cells (LTi) stimulate the organization of the gut-associated lymphoid tissues (GALTs), including Payer's patches, crypt patches and isolated lymphoid follicles (ILFs) [33-35]. However, maturation and differentiation of these tissues depends on subsequent gut microbiota stimulations [36]. As a consequence GF mice have underdeveloped GALTs [37]. In addition to stimulating the development of lymphoid tissues the gut microbiota modulates the differentiation of several immune cell subsets. The immune and epithelial cells recognize the so-called pathogen-associated molecular patterns (PAMPs) (*i.e.* flagellin, chitin, LPS, lipoteichoic acid, mannans, microbial nucleic acids) of microorganisms through pattern recognition receptors (PRRs), among which the most relevant are the Toll-like receptors (TLRs), NOD-like receptors (NLRs) and RIG-I (retinoic acid-inducible gene I)-like receptors (RLRs) [5]. The stimulation of such receptors activates the innate immunity determining a cascade of events (expressions of chemokine and cytokines from neutrophils, macrophages, NK cells or dendritic cells) that leads to the differentiation of specific $CD4^+$ T-cells subpopulations, pro-inflammatory (Th1 and Th17) of regulatory (T-reg and T_R1), in function of the type of stimulus [5, 23]. Different members of the gut microbiota, both symbionts and pathobionts, as well as pathogens, drive specific T-cells responses. Segmented filamentous bacteria (SFB), Gram-positive Clostridia-related bacteria, specifically induce a Th17 responses [38-40]. Th17 cells, which preferentially accumulates in the intestine, are fundamental for host defence and their localization in the intestine is strongly affected by the microbiota, as observed in antibiotic-treated and GF mice models [38-40]. The adhesion of SFB to intestinal epithelial cells (IECs) of the host stimulates the production of IL-1β, IL-6 and IL-23 by lamina propria DCs and macrophages and subsequent expression of pro-inflammatory cytokines IL-17A and IL-22 by Th17 cells [40]. On the contrary the Gram-negative, polysaccharide A (PSA) producing *Bacteroides fragilis* as well as the Gram-positive *Clostridium* spp. belonging to the clusters IV and XIV elicit anti-inflammatory responses mediated by $FOXP3^+$ $CD4^+$ T-reg cells. This cell lineage produces principally IL-10, which promotes tolerance and maintains the immune homeostasis controlling/suppressing pro-inflammatory Th17 responses [41, 42]. Beside these $CD4^+$ Th cell lineages, innate lymphoid cells (ILCs) cover also

an important role in the homeostatic balance between immunity and microbiota in the gut sharing some functional characteristics with T-cells [43]. ILCs have been categorized in three groups: T-bet$^+$ ILCs (or ILC1s), GATA3$^+$ ILCs (or ILC2s) and RORγt$^+$ ILCs (or ILC3s) [44]. The latter in particular finely regulate microbiota composition producing IL-22. This cytokine promotes the production of antimicrobial peptides (AMPs) from IECs [45] and negatively regulates Th17 responses inhibiting, for example, the expansion of commensal SFB. IL-22 production by ILC3s is particularly dependent on aryl-hydrocarbon receptor (AhR) activation [46, 47]. In response to dietary tryptophan, commensal lactobacilli produce AhR ligands (in particular the indole-3-aldheide) inducing production of IL-22 by ILC3s cells [47] that, in turn, inhibits SFB intestinal expansion. In addition to regulation of T-cell responses, the intestinal microbiota modulates also gut-specific B-cell responses. IgAs are produced by plasma cells (differentiated B-cells) in the GALTs via TLR5 binding of bacterial flagellin by DCs and subsequent promotion of B-cell differentiation through secretion of B-cell activating factor (BAFF), inducible nitric oxygen synthase (iNOS) or Tumor Necrosis Factor (TNF) [48]. Intestinal barrier homeostasis is finely regulated by luminal secreted IgA as they coat commensals avoiding their binding to the epithelium and their translocation into the lamina propria [49]. All these mechanisms are summarized in **Figure 1**. Any disruption of the delicate balance between the gut microbiota and the immune system has potentially pathological consequences on the health status of the host. Dysbiosis of the gut microbiota can lead to chronic inflammation encompassing hyper activation of Th1 and Th17 immune responses including chronic inflammation as observed in Inflammatory Bowel Diseases (IBDs) [5]. IBDs appear to be caused by a dysregulated immune response to the commensal microorganisms harbouring virulence traits in their genome (the so-called pathobionts) in a genetically susceptible host [50]. Gut microbiota influences also various extra-intestinal (auto)immune disorders. Despite the role of gut microbes may vary from protection to stimulation of autoimmunity it has been shown that GF mice are unsusceptible to arthritis, T1D and even to experimental autoimmune encephalomyelitis (a study model for Multiple Sclerosis) [5].

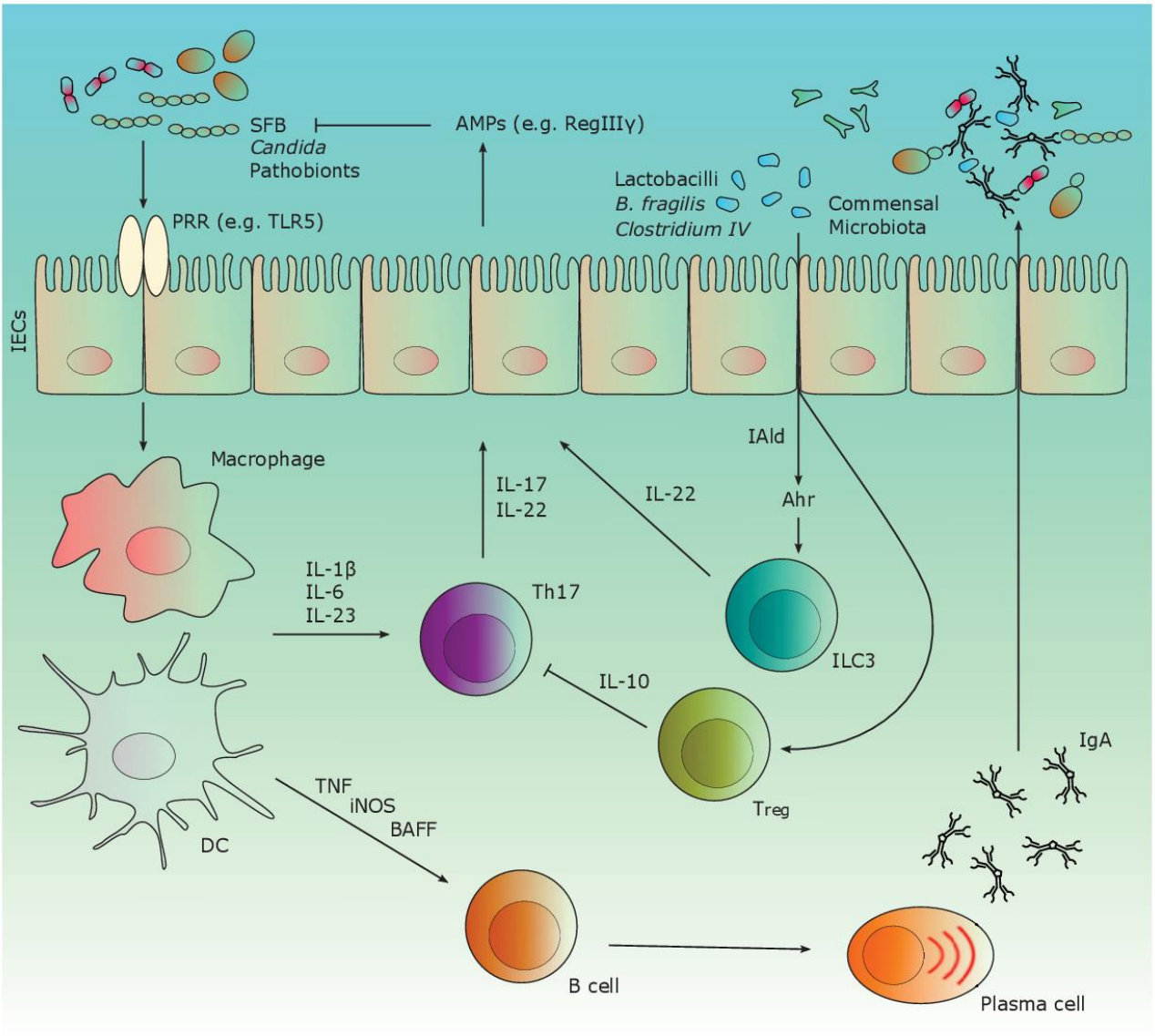


Figure 1: The recognition of specific PAMPs from pathobionts by PRRs induces antigen presenting cells (APCs) like macrophages and dendritic cells (DCs) to produce IL-1β, IL-6 and IL-23. Subsequently Th17 cells express the pro-inflammatory cytokines IL-17 and IL-22 that stimulates the production of AMPs by IECs, controlling the expansion of commensal microbes that could be harmful. Similarly, ILC3s contribute to control excessive expansion of microbiota by producing IL-22 in response to metabolites (*e.g.* indole-3-aldheyde) produced by commensal microorganisms (*i.e.* lactobacilli). On the other hand commensals such as Bifidobacteria, PSA^{+} *B. fragilis* and *Clostridium spp*. cluster XIV and IV stimulates APCs to promote anti-inflammatory IL-10 T-reg responses regulating Th17 responses. Furthermore lamina propria DCs promote B-cells differentiation in IgA-producing plasma cells secreting the B-cell activating factor (BAFF), inducible nitric oxygen synthase (iNOS) or Tumor Necrosis Factor (TNF). The gut microbiota composition is thus regulated by direct binding of the secreted IgAs on commensals and pathobionts, avoiding possible translocations of such microbes through IECs.

The microbiome-gut-brain axis: how the gut microbiota interacts with the nervous system

Several evidences demonstrated the bidirectional communication between the gut and the central nervous system (CNS) [6] and even more studies showed the impact of the gut microbiota on

neurological functions through varied and complex mechanisms. Well known pathways of interaction between the gut microbiota and the nervous system are i) the autonomous nervous system (ANS), in particular the enteric nervous system (ENS) and the vagal nerves [51], ii) the hypothalamic-pituitary-adrenal (HPA) axis [52], iii) tryptophan metabolism [53] and others bacterial metabolites [54], iv) the immune system [55].

Recent studies have shown that *Lactobacillus reuterii* and *Bacteroides fragilis* activate colonic and intestinal afferent neurons [56, 57] probably through the direct interaction of the gut microbiota with TLRs expressed in the ENS [58].

The HPA axis is one of the most relevant parts of the neuroendocrine system modulating stress responses and physiological processes, including digestion and immune regulation. GF mice exposed to stressful situations have been shown to present levels of adrenocorticotrophin and corticosterone higher than wild-type mice exposed to the same sources of stress [52], providing a first evidence of the direct interaction of the gut microbiota with the HPA. Furthermore cortisol release by the HPA axis regulates gut motility and integrity affecting the host-microbiota interaction at mucosal level [59].

The gut microbiota has an important effect on host metabolism affecting metabolic pathways directly related to the CNS signalling. Tryptophan metabolism in particular can affect CNS functions and impairments of serotonin and kynurenine pathways, two important branches of tryptophan metabolism, are implicated in different neurological disorders [60]. The serotonergic system is involved in mood regulation and it has an important role in GI secretion and signalling [61]. It has been observed that serotonin levels and related metabolites in GF mice are altered in the striatum and hippocampus [62, 63] while there is a 2.8 fold increase in plasma serotonin levels when GF mice are colonized by the gut microbiota [64]. Indeed the gut microbiota directly modulates serotonin biosynthesis [65, 66] and probiotic treatments have been shown to increase serotonin and kynurenine levels [67]. Moreover, dysfunctions of these pathways have been implicated in Autism Spectrum Disorders (ASDs) [60].

Some members of the gut microbiota, principally lactobacilli and bifidobacteria, can produce neuroactive molecules including neurotransmitters and neuromodulators such as γ-aminobutyric acid (GABA) [68, 69]. Interestingly, *in-vivo* administration of a probiotic strain of *Lactobacillus rhamnosus* increased CNS levels of GABA and diminished the anxiety-like behaviour of treated mice while vagotomised mice did not show behavioural changes [51]. Moreover the gut microbiota is the principal source of SCFAs that, even if they are not really neuroactive molecules, act also on neuronal physiology. Indeed SCFAs can pass through the blood-brain barrier (BBB) [70] and inhibit histone deacetylases (HDACs) in active neurons resulting in long-lasting effects on gene expression [71, 72].

Acetate may further modulate hippocampal activity altering the levels of GABA and anorectic neuropeptide suppressing the appetite stimulus [73]. Nevertheless cumulative, chronic delivery of non-physiological levels of propionate may have detrimental effects on CNS activities promoting ASD-like behaviours [74]. Finally it is worth noting that the gut microbiota impacts also on neuroplasticity as observed by the altered levels of hippocampal brain-derived neurotrophic factor (BDNF) in GF mice [75].

The impact of the gut microbiota on stress-related behaviours is one of the best studied effect on the microbiome:gut:brain axis. Alterations in the gut microbiota, both resulting from antibiotic treatments or infections, increase anxiety-like behaviour in wild-type mice whereas GF mice showed increased motor activity and reduced anxiety [62]. The gut microbiota affects the development of neural systems that govern the endocrine response to stress. Indeed GF mice showed exaggerate HPA responses and reduced BDNF levels compared to gnotobiotic mice while this phenotype was reversed by treatment with probiotics or by restoration of GF mice microbiota during early stage of development [52].

Alterations in the composition of the gut microbiota have been implicated in a wide variety of other neurological disorders, including ASDs [29]. ASDs are neurodevelopmental disorders characterized by alterations in social interactions and communication and by restricted and repetitive behaviour [76]. Further, ASD subjects commonly suffer of GI abnormalities [77-79]. When compared to healthy subjects, the faecal microbiota of ASD subjects is characterized by high abundance of Clostridia [80, 81], imbalances of the *Bacteroidetes* to *Firmicutes* ratio [82, 83] and reduced incidence of *Prevotella*, *Ruminococcus* and *Bifidobacterium* [82, 84]. The altered microbial community structure associated to ASDs and its metabolic impact on neurological and behavioural pathophysiology has been recently studied on the ASD-like maternal immune activation (MIA) mouse model [85]. Hsiao and colleagues observed that the MIA offspring harboured a dysbiotic gut microbiota, defects in intestinal permeability and increased levels of 4-ethylphenilsulphate (4EPS), a microbial metabolite related to *p*-cresol, a putative metabolic marker for autism [85]. Treatments with PSA^+ *Bacteroides fragilis* attenuated the stereotypic behaviours of the MIA offspring, restored intestinal barrier functions and reduced the levels of 4EPS [85] highlighting the impact that the gut microbiota may have on the onset of extra-intestinal, neurological disorders. Furthermore the gut microbiota influences the development of immune-mediated CNS disorders as Multiple Sclerosis. The gut microbiota may interact indirectly with the CNS through the peripheral immune system affecting neurological functions and behaviour [6]. The activation of the immune system mediated by PAMPs interaction with PRRs leads to the production of pro-inflammatory cytokines such as IL-1β, IL-6 and TNFα that can cross the BBB modulating neuronal physiology [86]. Indeed, the increase of peripheral

inflammation is a risk factor for depression due to diminished concentration of serotonin available to signal in the synapses [87] and a reduction of 5-HT receptors on neurons [88]. In the experimental autoimmune encephalomyelitis (EAE) model of multiple sclerosis the colonization of the gut microbiota is required for the induction of the EAE. GF mice showed less severe EAE whereas colonization with SFB resulted in dysregulation of Th17 responses which exacerbate the severity of the disease [89]. On the contrary, treatments with PSA^+ *B. fragilis* or *Bifidobacterium* alleviate EAE symptoms [90-92]. Additionally, immune dysregulation has been observed in several neurological disorder, including ASD, underlining the intricate crosstalk between immune system, CNS, metabolism and commensal microorganisms and the wide impact that the gut microbiota exerts at different levels of host physiology.

The human gut mycobiota: a "hidden" regulator of host physiology

The human gut microbiota has been extensively studied in several aspects of its interaction with the host in health and disease. Nevertheless the majority of the studies focused only on the bacterial component ignoring the mycobiota *i.e.* the fungal component of the gut microbiota. Despite fungi represent approximately 0.1% of total gut microorganisms [1, 93], in terms of biomass the gut mycobiota covers an important role in the GI niche since the average size of a fungal cell is 100 fold larger than the average bacterial cell [94]. Furthermore, the gut mycobiota, together with its bacterial counterpart, exerts key roles in maintaining microbial community structure, metabolic functions and it has strong immunomodulatory properties, being a main actor in host physiopathology [94]. Fungi interact with the immune system through mechanisms similar to those described in **Figure 1**. Fungal PAMPs are recognized by host cells' PRRs, among which C‑type lectin receptors (CLRs; *e.g.* dectin 1 and dectin 2, also known as CLEC7A and CLEC6A respectively) are central for fungal recognition and for the development of innate and adaptive immune responses [95]. Well balanced pro-inflammatory and tolerogenic responses are a prerequisite to avoid potential harmful inflammatory responses triggered by GI fungi. Different subsets of DCs equipped with different PRRs have the ability to initiate adaptive antifungal immune responses enabling fungi to switch from commensalism to infection [95]. The activation of signal transduction pathways affecting the balance of NF-kB, a pleiotropic regulator of cell responses [96], and ultimately the expression of IDO1, has a key role on plasticity of DCs activities in balancing between $CD4^+$ effector Th cells and T-reg cells [95]. Th1 and Th17 cell responses are crucial in antifungal immunity. Th1 cells, through the production of IFNγ and TNFα, are necessary for the activation and recruitment of phagocytes (macrophages, neutrophils) at potential sites of infection [95]. Although Th1 cell responses are central for host protection against fungi, patients with genetic defects in Th1 pathways do not have increased susceptibility to fungal

infections [97]. On the contrary, Th17 cells play a central role in protection against fungal infections [97, 98]. In fact the ability of IL-17 to mobilize neutrophils and induce AMPs contribute to an efficient control of the infection at different body sites [95]. However IL-10 producing T-reg cells are necessary to provide immune homeostasis and limiting immunopathology [95]. The shift between pro-inflammatory and tolerogenic DCs responses are mediated by the kynurenine pathway of tryptophan metabolism, in which IDO1 covers a central role [99]. IDO1 is widely recognized as a regulator of immune homeostasis and suppressor of inflammation by inducing IL-10 through the production of immune-active kynurenines that activate the AhR in lymphoid tissues [100] and promotes T-reg cells development [101]. The gut microbiota, in particular lactobacilli, modulates the immunological responses to *Candida* in the GI tract by providing tryptophan-derived AhR ligands (*i.e.* indole-3-aldheide) that stimulate ILC3s to produce IL-22 [47]. Together with IL-17, IL-22 avoids excessive proliferation of commensal fungi. GF mice are highly susceptible to *Candida* infections [102] and antibiotic-treated mice showed an overgrowth of the fungal population, with the genus *Candida* representing up to the 99% of total intestinal microbes [103-105]. Furthermore, prolonged antibiotic usage and subsequent colonization with *C. albicans* prevents from a rapid and full restoration of the bacterial community structure that results in altered abundances of *Bacteriodetes*, *Lactobacillaceae*, *Ruminococcaceae* and *Lachnospiraceae* [106]. At the mucosal site, fungal colonization induces the production of IL-17 and IFNγ which is a strong activator of IDO1 [107]. The activation of the IFNγ-IDO1 axis harmonizes the tolerogenic response against the fungal microbiota allowing commensalism [95]. Nevertheless commensal fungi, such as *C. albicans*, may shift IDO's activity from kynurenine towards 5-hydroxytriptophan, an inhibitor of Th17 host responses [108]; furthermore IDO's blockade induces the transition from yeast to hyphal morphology [107] resulting in a transition from commensalism to infection. The mechanisms through commensal *C. albicans* choses to shift its phenotype towards infection are not well understood but the disruption of the microbial community structure resulting in intestinal dysbiosis may be one of the reasons [109]. Although the GI tract is colonized by commensal fungi able to modulate immune responses, the composition and characteristics of the gut mycobiota in healthy hosts have been poorly explored. The bacterial gut microbiota evolves rapidly during the early stages of life but then remains relatively stable over time, more in term of functions than taxonomically [1]. On the contrary the gut mycobiota seems to be more variable with respect to its bacterial counterpart. It has been showed that the intestinal fungal population of a mouse model of study undergoes significant changes over 2.5 months while the bacterial population remained quite stable [105]. Despite some studies investigated the gut mycobiota using labour intensive experimental approaches [*i.e.* culture-based approaches, denaturing gradient gel electrophoresis (DGGE)], these are inadequate for the large scale characterization of

complex microbial communities. Only with the advent of high-throughput sequencing technologies it has become possible to characterize quantitatively the structure and composition of intestinal fungal communities. It has been found that the mammalian gut mycobiota is characterized by over 50-60 fungal genera (in particular *Candida*, *Saccharomyces*, *Cladosporium*, *Aspergillus* and *Penicillium)* [110-112]. Interestingly, the distribution of species belonging to the genus *Candida* changes in function of the host geographical origin. In fact, a remote population of Amerindian showed an elevated gut colonization of foodborne *Candida krusei* and *Saccharomyces cerevisiae* over *Candida albicans*, that is one of the most abundant GI commensal species in industrialized countries [113]. This study suggested that the exposition to different dietary habits and environmental factors has an important role in shaping the gut mycobiota. Indeed an animal-based dietary regimen has been associated with an expansion of *Penicillium* species [112] whereas carbohydrate-rich, plant-based diets correlated with high abundance of *Candida* [18, 112]. The complex relationship existing among commensal fungi, bacteria and the immune system results in a delicate equilibrium which disruption could contribute to pathologies ranging from metabolic disorders (obesity) to colorectal adenomas and IBDs. Since diet is a dynamic variable shaping the composition of the gut microbiota and consequently mucosal immunity it is not surprising that the gut mycobiota of obese individuals is different from non-obese subjects. In fact *Dipodascaceae*, *Saccharomycetaceae* and *Tremellomycetes* showed an increase of their relative abundances in obese individuals, further correlating with adiposity, inflammatory parameters and the occurrence of metabolic abnormalities [114]. Another recent report showed an altered fungal microbiota in biopsies from colorectal adenomas characterized by reduced fungal diversity and dominated by the opportunistic fungal pathogens *Phoma* and *Candida*, representing up to the 45% of the relative abundance. Furthermore different fungal patterns were found at different stages of adenomas, resulting in potential diagnostic biomarkers [115]. These studies well describe how pathophysiological changes of the host may affect also the composition of the gut mycobiota, but the best studied pathologies directly involving alterations of the intestinal fungal population are IBDs. Crohn's disease (CD) and ulcerative colitis (UC) are the two primary type of IBDs. Although the exact aetiology of IBDs remains elusive it probably involves altered immunological responses against the gut microbiota in genetically predisposed subjects, leading to extended intestinal inflammation [50]. Bacterial dysbiosis has been previously observed in IBDs [116], nevertheless the discovery of anti-*Saccharomyces cerevisiae* antibody (ASCA) in CD [117] highlighted the potential role of the gut mycobiota in IBDs pathophysiology. Preliminary studies based on DGGE analysis of fungal 18S rRNA revealed an increased abundance of fungal richness and diversity in IBD patients characterized by expansion of *Candida* spp., *Penicillium* spp., *Saccharomyces* spp., *Gibberella moniliformis*, *Alternaria brassicicola*, and *Cryptococcus*

neoformans [118, 119]. Moreover high-throughput sequencing analysis of the gut mycobiota in a DSS-induced colitis mouse model showed that dectin 1 deficiency (*Clec7a*$^{-/-}$) increases the proportion of opportunistic pathogenic fungi, including *Candida* and *Trichosporon*, in inflamed tissues, whereas non-pathogenic *Saccharomyces* decreases [111]. Similarly, a recent study on an IBD cohort showed an altered gut mycobiota with a decreased proportion of *S. cerevisiae* and an increased proportion of *C. albicans* compared to healthy subjects [120]. However the role of the gut mycobiota in the maintenance of health is still far from being well understood. Indeed fungi are highly underrepresented in sequence databases compared to bacteria making the correct understanding of fungal population and community structures more difficult due to under-detection of fungal sequences in high-throughput metagenomics analysis [94].

Investigating the gut microbiota: how culture-independent approaches revolutionised the study of complex microbial communities

To study the microbial intestinal communities, commensal microorganisms needs to be identified and quantified. Historically microbes were identified by microscopic observation and staining [121, 122]. The early observation that the number of organisms that can be observed under a microscope is much higher than the number of those that can be cultivated led to the deduction that microorganisms require particular conditions to grow, giving birth to the research field of microbial ecology [123]. Early studies on microbial communities relied on culturing microbes from various environmental sources, ignoring the full complexity of microbial communities. Molecular approaches such as restriction fragment length polymorphism (RFLP) analysis, amplified rDNA restriction analysis (ARDRA) and denaturing gradient gel electrophoresis (DGGE) were good at revealing more complexity than culturing methods but were unable to identify specific taxon. In the 70s and 80s the idea of using ribosomal RNA gene sequences as molecular markers for phylogenetic classification of (micro)organisms [124, 125], together with the development of pioneering methods in molecular biology such as PCR and Sanger sequencing [126], gave a great impulse to the application of culture-independent DNA-based methods for the study of microbial communities. The advent of new high-throughput sequencing technologies such as Roche 454 and Ilumina-Solexa then changed dramatically the way of how complex microbial communities were studied allowing to characterize microbial populations and their gene content, starting the so-called "metagenomics era" of microbiome research [127].

Analysis of microbial communities can rely on two metagenomics approaches, targeted or amplicon-based metagenomics (also known as metataxonomics [128]), and untargeted or shotgun metagenomics. By now, the most recent advancements in microbiome research are going to use whole

metagenomics shotgun because of their ability to infer genes, functional capabilities and metabolic potential of a microbial community [129]. Nevertheless shotgun sequencing is less sensitive with respect to metataxonomics which is still the preferred approach to investigate rare, less abundant components of a microbial community, as the fungal population of the gut microbiota [130]. Furthermore metataxonomics is still the most widely used approach in microbiome research due to the low sequencing costs and relying on the use of well-established bioinformatics pipelines.

The metataxonomics analysis of complex microbial communities is based on the amplification and analysis of taxonomically informative sequences from a microbial community, such as the 16S rRNA gene for prokaryotes or the intergenic transcribed spacers (ITS) for eukaryotes. Amplicon-based sequencing of targeted marker genes has been widely used thanks to its versatility, since it can be adapted to many different sequencing platforms. Furthermore one of the advantages of using ribosomal genes as phylogenetic markers for metataxonomics is the huge amounts of deposited annotated sequences (in particular for bacterial 16S rDNA) in many databases (including RDP [131], Greengenes [132] and SILVA [133]), allowing for easy taxonomy assignments even in complex ecological niches. Since bacteria often represent the majority part of a microbial community, most of the metataxonomics analytical pipelines have been developed for the analysis of the 16S rRNA gene. The 16S rRNA gene sequence is about 1.5 kb in length and contains highly conserved, variable and hypervariable regions that vary in function of evolutionary time and that can be used as molecular markers to infer the phylogenetic structure of a microbial community [125]. Sequence similarity provides the only unambiguous definition of bacterial taxa and 16S rRNA sequences having at least 97% identity are taken together as a proxy for species and defined as an operational taxonomic unit (OTU) [134, 135]. Horizontal gene transfer [136], multi-copy rDNA genes [137], sequencing errors [138] and other confounding factors represent the major pitfalls of this approach [139]. Indeed, such degree of sequence divergence (*i.e.* 97% or lower sequence similarity) is required for accurate taxonomic assignments although it can results in low resolution at species level [140]. Sequence assignments to OTUs, known as binning, is a crucial step in microbial ecology analyses and can use supervised and unsupervised methods [131, 141]. Bioinformatics pipelines for the analysis of metataxonomics data include the following steps: filtering of low quality sequences, OTUs clustering, multiple sequence alignment (MSA) and sample distance calculations for microbial ecology measures. Quality filtering, MSA and clustering rely on complex computational procedures that are critical in order to produce consistent data [142]. Filtering of low quality sequences [143] is necessary to avoid overestimation of samples complexity while the clustering algorithms have to be flexible and accurate enough in order to identify precisely the number of OTUs in a dataset without overestimating the number of different species present in the samples [144].

There are different algorithms to calculate clusters for binning sequences into OTUs. These can be categorized principally into hierarchical clustering methods and heuristic clustering methods. The hierarchical clustering methods calculate a distance matrix measuring the difference between each pair of sequences using then sequence dissimilarity to define OTUs within a pre-defined threshold (*e.g.* the 97%). Mothur [145] and ESPRIT [146] are two representative algorithms using hierarchical clustering methods. Nevertheless, aligning millions of reads and generating distance matrices are computationally expensive tasks. Therefore most of the clustering algorithms commonly used in microbiome research are based on heuristic approaches which are faster and require less memory then exhaustive searches. CD-HIT [147], UCLUST [148], and DNACLUST [149] are widely used heuristic clustering methods. These "greedy" clustering algorithms use the '*k*-mer' strategy to avoid the computationally intensive step of pairwise alignment selecting an input sequence as a seed for the initial cluster, and then examining each input sequence sequentially. If the distance between the query sequence and representative sequences of the existing clusters is within a pre-defined threshold, the input sequence is added to the corresponding cluster, otherwise, a new cluster is created and the query sequence is stored as a new seed [144]. Independently from the selected clustering method, it is important to consider that these may generate different estimates of OTUs, and this fact should be taken into account when comparing the ecological diversity.

Furthermore, the approach used to generate MSAs has a significant impact on downstream data analysis since it is a crucial step in order to calculate phylogenetic and pairwise genetic distances to be used for OTUs taxonomic assignments and phylogenetic-based measures for microbial ecology estimates [150] (*e.g.* UniFrac measures [151]). Popular aligner tools such as ClustalW [152], MUSCLE [153] and T-Coffee [154] are widely used to generate sequences alignments, nevertheless these *de novo* methods scale exponentially in space and time in function of the length and number of sequences and cannot provide the quality required for phylogenetic analysis, necessitating a manual curation step [155, 156]. On the other hand, large databases of curated alignments of 16S rRNA sequences exist, obtained taking secondary structure information into account, and reference-based methods have been proposed to use this pre-existing knowledge to guide the alignment step. With this approach the computation complexity is reduced, making the alignment process readily scalable by reducing also the excessive introduction of gaps within the informative, non-conserved positions of the alignment [155]. Currently RDP [131], SILVA [133] and Greengenes [132] are the most widely used 16S rRNA gene databases in microbiome research. All of them provides curated MSAs of 16S rRNA sequences and their own aligners which relies, for example, on the use of Hidden Markov Models such as in the case of Infernal, the RDP aligner [157], or on the use of *k*-mer approaches such as SINA [158] and NAST [159], the aligners provided within SILVA and Greengenes, respectively.

The NAST algorithm, used by Greengenes in its original implementation and in the more recent implementations provided by QIIME, mothur and micca [145, 160-162], is one of the most widely used aligner in microbiome research.

All these computational steps are fundamental in order to assess the microbial population diversity of an ecological niche. In addition, one important goal in microbial ecology is to compare different microbial populations estimating the between-sample taxonomic diversity, *i.e.* the *beta*-diversity. Different *beta*-diversity measures have been used to characterize communities, *e.g.* Bray-Curtis dissimilarity, unweighted UniFrac and weighted UniFrac, but these ecological estimators were designed with different purposes and have different pros and cons [163]. The Bray-Curtis dissimilarity quantifies the compositional dissimilarity between two different populations basing on the relative abundance of each binned OTU but is limited by the fact that the taxa are treated as equally related to one another. Alternatively, the *beta*-diversity can be analysed by using the UniFrac distance matrices which take into account the phylogenetic distributions of the samples by measuring the phylogenetic relatedness of the OTUs [151]. The UniFrac distance, in its first implementation *i.e.* the unweighted UniFrac [164], is a qualitative measure because duplicate sequences contribute no additional branch length to the phylogenetic tree and therefore the unweighted UniFrac does not take into account for changes in the relative abundance of different taxa between ecological communities [163]. On the contrary the more recent weighted UniFrac [165] is a quantitative *beta*-diversity measure because it detects changes in how many sequences from each lineage are present, as well as changes in which taxa are present [163]. UniFrac distances depend on the quality of the input phylogenetic tree and this, in turn, relies on the quality of the MSA. Therefore it is clear that the MSA step is crucial for the correct interpretation of the microbial community structures.

Although the sequencing methods and bioinformatics pipelines for the study of bacteria in complex ecological communities are well-established, the analysis of fungi in these communities presents some difficulties. The metataxonomics analysis of one of the two internal transcribed spacer (ITS) regions of the fungal rDNA [130] is best way to study complex fungal populations. The ITS1 and the ITS2 are located respectively between the 18S and 5.8S genes and between the 5.8S and the 28S genes. Both regions are variable enough to classify fungi down to genus or species level. Nevertheless they have some intrinsic limitations with respect to the bacterial 16S rDNA. First of all, fungal ITS have variable sequence length in function of the fungal species of origin [166, 167] while the bacterial 16S rRNA gene has always the same length. The variable length often creates difficulties in preparing amplicon libraries necessary for the different steps of high-throughput sequencing. Furthermore, unlike bacterial 16S data, ITS sequences cannot rely on well-annotated databases of fungal sequences with which to compare the sequencing data, requiring the use of *de novo* methods for the MSA step

and therefore leading to the introduction of further uncertainty in the metataxonomics analysis. It has been estimated that the deposited fungal sequences represent only a small fraction of the currently fungal species discovered and often are misclassified or unclassified [167, 168]. Indeed fungal taxonomy is quite complicated since fungi may be classified in two different taxa based on their sexual dimorphism even if belonging to the same fungal species, making difficult OTUs taxonomy assignments [94]. However the development of new pipelines and curated DBs for fungal metagenomics is further improving our comprehensive understanding of microbiome interactions in their ecological niches, from environmental samples to the human gut.

References

1. Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, Fernandes GR, Tap J, Bruls T, Batto JM *et al*: **Enterotypes of the human gut microbiome**. *Nature* 2011, **473**(7346):174-180.
2. Reyes A, Haynes M, Hanson N, Angly FE, Heath AC, Rohwer F, Gordon JI: **Viruses in the faecal microbiota of monozygotic twins and their mothers**. *Nature* 2010, **466**(7304):334-338.
3. Human Microbiome Project C: **Structure, function and diversity of the healthy human microbiome**. *Nature* 2012, **486**(7402):207-214.
4. Maslowski KM, Mackay CR: **Diet, gut microbiota and immune responses**. *Nature immunology* 2011, **12**(1):5-9.
5. Kamada N, Seo SU, Chen GY, Nunez G: **Role of the gut microbiota in immunity and inflammatory disease**. *Nature reviews Immunology* 2013, **13**(5):321-335.
6. Sampson TR, Mazmanian SK: **Control of brain development, function, and behavior by the microbiome**. *Cell host & microbe* 2015, **17**(5):565-576.
7. Palmer C, Bik EM, DiGiulio DB, Relman DA, Brown PO: **Development of the human infant intestinal microbiota**. *PLoS biology* 2007, **5**(7):e177.
8. Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N, Knight R: **Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns**. *Proceedings of the National Academy of Sciences of the United States of America* 2010, **107**(26):11971-11975.
9. Rautava S, Luoto R, Salminen S, Isolauri E: **Microbial contact during pregnancy, intestinal colonization and human disease**. *Nature reviews Gastroenterology & hepatology* 2012, **9**(10):565-576.
10. Ventura M, Turroni F, Motherway MO, MacSharry J, van Sinderen D: **Host-microbe interactions that facilitate gut colonization by commensal bifidobacteria**. *Trends in microbiology* 2012, **20**(10):467-476.
11. Rogier EW, Frantz AL, Bruno ME, Wedlund L, Cohen DA, Stromberg AJ, Kaetzel CS: **Secretory antibodies in breast milk promote long-term intestinal homeostasis by regulating the gut microbiota and host gene expression**. *Proceedings of the National Academy of Sciences of the United States of America* 2014, **111**(8):3074-3079.
12. Koenig JE, Spor A, Scalfone N, Fricker AD, Stombaugh J, Knight R, Angenent LT, Ley RE: **Succession of microbial consortia in the developing infant gut microbiome**. *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108 Suppl 1**:4578-4585.
13. Derrien M, Belzer C, de Vos WM: **Akkermansia muciniphila and its role in regulating host functions**. *Microbial pathogenesis* 2016.

14. Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R: **Diversity, stability and resilience of the human gut microbiota**. *Nature* 2012, **489**(7415):220-230.
15. Markle JG, Frank DN, Mortin-Toth S, Robertson CE, Feazel LM, Rolle-Kampczyk U, von Bergen M, McCoy KD, Macpherson AJ, Danska JS: **Sex differences in the gut microbiome drive hormone-dependent regulation of autoimmunity**. *Science* 2013, **339**(6123):1084-1088.
16. Bolnick DI, Snowberg LK, Hirsch PE, Lauber CL, Org E, Parks B, Lusis AJ, Knight R, Caporaso JG, Svanback R: **Individual diet has sex-dependent effects on vertebrate gut microbiota**. *Nature communications* 2014, **5**:4500.
17. Wu GD, Chen J, Hoffmann C, Bittinger K, Chen YY, Keilbaugh SA, Bewtra M, Knights D, Walters WA, Knight R *et al*: **Linking long-term dietary patterns with gut microbial enterotypes**. *Science* 2011, **334**(6052):105-108.
18. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, Ling AV, Devlin AS, Varma Y, Fischbach MA *et al*: **Diet rapidly and reproducibly alters the human gut microbiome**. *Nature* 2014, **505**(7484):559-563.
19. De Filippo C, Cavalieri D, Di Paola M, Ramazzotti M, Poullet JB, Massart S, Collini S, Pieraccini G, Lionetti P: **Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa**. *Proceedings of the National Academy of Sciences of the United States of America* 2010, **107**(33):14691-14696.
20. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, Magris M, Hidalgo G, Baldassano RN, Anokhin AP *et al*: **Human gut microbiome viewed across age and geography**. *Nature* 2012, **486**(7402):222-227.
21. Le Poul E, Loison C, Struyf S, Springael JY, Lannoy V, Decobecq ME, Brezillon S, Dupriez V, Vassart G, Van Damme J *et al*: **Functional characterization of human receptors for short chain fatty acids and their role in polymorphonuclear cell activation**. *The Journal of biological chemistry* 2003, **278**(28):25481-25489.
22. Scheppach W: **Effects of short chain fatty acids on gut morphology and function**. *Gut* 1994, **35**(1 Suppl):S35-38.
23. Sommer F, Backhed F: **The gut microbiota--masters of host development and physiology**. *Nature reviews Microbiology* 2013, **11**(4):227-238.
24. Holmes E, Loo RL, Stamler J, Bictash M, Yap IK, Chan Q, Ebbels T, De Iorio M, Brown IJ, Veselkov KA *et al*: **Human metabolic phenotype diversity and its association with diet and blood pressure**. *Nature* 2008, **453**(7193):396-400.
25. Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE, Sogin ML, Jones WJ, Roe BA, Affourtit JP *et al*: **A core gut microbiome in obese and lean twins**. *Nature* 2009, **457**(7228):480-484.
26. Larsen N, Vogensen FK, van den Berg FW, Nielsen DS, Andreasen AS, Pedersen BK, Al-Soud WA, Sorensen SJ, Hansen LH, Jakobsen M: **Gut microbiota in human adults with type 2 diabetes differs from non-diabetic adults**. *PloS one* 2010, **5**(2):e9085.
27. Blaser MJ: **Who are we? Indigenous microbes and the ecology of human diseases**. *EMBO reports* 2006, **7**(10):956-960.
28. Dicksved J, Halfvarson J, Rosenquist M, Jarnerot G, Tysk C, Apajalahti J, Engstrand L, Jansson JK: **Molecular analysis of the gut microbiota of identical twins with Crohn's disease**. *The ISME journal* 2008, **2**(7):716-727.
29. Mayer EA, Padua D, Tillisch K: **Altered brain-gut axis in autism: comorbidity or causative mechanisms?** *BioEssays : news and reviews in molecular, cellular and developmental biology* 2014, **36**(10):933-939.
30. McFall-Ngai M, Hadfield MG, Bosch TC, Carey HV, Domazet-Loso T, Douglas AE, Dubilier N, Eberl G, Fukami T, Gilbert SF *et al*: **Animals in a bacterial world, a new imperative for the life sciences**. *Proceedings of the National Academy of Sciences of the United States of America* 2013, **110**(9):3229-3236.

31. Chow J, Tang H, Mazmanian SK: **Pathobionts of the gastrointestinal microbiota and inflammatory disease**. *Current opinion in immunology* 2011, **23**(4):473-480.
32. Jain N, Walker WA: **Diet and host-microbial crosstalk in postnatal intestinal immune homeostasis**. *Nature reviews Gastroenterology & hepatology* 2015, **12**(1):14-25.
33. Gordon HA, Bruckner-Kardoss E, Wostmann BS: **Aging in germ-free mice: life tables and lesions observed at natural death**. *Journal of gerontology* 1966, **21**(3):380-387.
34. Hamada H, Hiroi T, Nishiyama Y, Takahashi H, Masunaga Y, Hachimura S, Kaminogawa S, Takahashi-Iwanaga H, Iwanaga T, Kiyono H *et al*: **Identification of multiple isolated lymphoid follicles on the antimesenteric wall of the mouse small intestine**. *Journal of immunology* 2002, **168**(1):57-64.
35. Bouskra D, Brezillon C, Berard M, Werts C, Varona R, Boneca IG, Eberl G: **Lymphoid tissue genesis induced by commensals through NOD1 regulates intestinal homeostasis**. *Nature* 2008, **456**(7221):507-510.
36. Renz H, Brandtzaeg P, Hornef M: **The impact of perinatal immune development on mucosal homeostasis and chronic inflammation**. *Nature reviews Immunology* 2012, **12**(1):9-23.
37. Round JL, Mazmanian SK: **The gut microbiota shapes intestinal immune responses during health and disease**. *Nature reviews Immunology* 2009, **9**(5):313-323.
38. Ivanov, II, Frutos Rde L, Manel N, Yoshinaga K, Rifkin DB, Sartor RB, Finlay BB, Littman DR: **Specific microbiota direct the differentiation of IL-17-producing T-helper cells in the mucosa of the small intestine**. *Cell host & microbe* 2008, **4**(4):337-349.
39. Gaboriau-Routhiau V, Rakotobe S, Lecuyer E, Mulder I, Lan A, Bridonneau C, Rochet V, Pisi A, De Paepe M, Brandi G *et al*: **The key role of segmented filamentous bacteria in the coordinated maturation of gut helper T cell responses**. *Immunity* 2009, **31**(4):677-689.
40. Ivanov, II, Atarashi K, Manel N, Brodie EL, Shima T, Karaoz U, Wei D, Goldfarb KC, Santee CA, Lynch SV *et al*: **Induction of intestinal Th17 cells by segmented filamentous bacteria**. *Cell* 2009, **139**(3):485-498.
41. Maynard CL, Harrington LE, Janowski KM, Oliver JR, Zindl CL, Rudensky AY, Weaver CT: **Regulatory T cells expressing interleukin 10 develop from Foxp3+ and Foxp3- precursor cells in the absence of interleukin 10**. *Nature immunology* 2007, **8**(9):931-941.
42. Atarashi K, Tanoue T, Shima T, Imaoka A, Kuwahara T, Momose Y, Cheng G, Yamasaki S, Saito T, Ohba Y *et al*: **Induction of colonic regulatory T cells by indigenous Clostridium species**. *Science* 2011, **331**(6015):337-341.
43. Sonnenberg GF, Artis D: **Innate lymphoid cell interactions with microbiota: implications for intestinal health and disease**. *Immunity* 2012, **37**(4):601-610.
44. Walker JA, Barlow JL, McKenzie AN: **Innate lymphoid cells--how did we miss them?** *Nature reviews Immunology* 2013, **13**(2):75-87.
45. Vaishnava S, Yamamoto M, Severson KM, Ruhn KA, Yu X, Koren O, Ley R, Wakeland EK, Hooper LV: **The antibacterial lectin RegIIIgamma promotes the spatial segregation of microbiota and host in the intestine**. *Science* 2011, **334**(6053):255-258.
46. Qiu J, Guo X, Chen ZM, He L, Sonnenberg GF, Artis D, Fu YX, Zhou L: **Group 3 innate lymphoid cells inhibit T-cell-mediated intestinal inflammation through aryl hydrocarbon receptor signaling and regulation of microflora**. *Immunity* 2013, **39**(2):386-399.
47. Zelante T, Iannitti RG, Cunha C, De Luca A, Giovannini G, Pieraccini G, Zecchi R, D'Angelo C, Massi-Benedetti C, Fallarino F *et al*: **Tryptophan catabolites from microbiota engage aryl hydrocarbon receptor and balance mucosal reactivity via interleukin-22**. *Immunity* 2013, **39**(2):372-385.
48. Uematsu S, Fujimoto K, Jang MH, Yang BG, Jung YJ, Nishiyama M, Sato S, Tsujimura T, Yamamoto M, Yokota Y *et al*: **Regulation of humoral and cellular gut immunity by**

lamina propria dendritic cells expressing Toll-like receptor 5. *Nature immunology* 2008, **9**(7):769-776.

49. Fagarasan S, Kawamoto S, Kanagawa O, Suzuki K: **Adaptive immune regulation in the gut: T cell-dependent and T cell-independent IgA synthesis**. *Annual review of immunology* 2010, **28**:243-273.
50. Baumgart DC, Carding SR: **Inflammatory bowel disease: cause and immunobiology**. *Lancet* 2007, **369**(9573):1627-1640.
51. Bravo JA, Forsythe P, Chew MV, Escaravage E, Savignac HM, Dinan TG, Bienenstock J, Cryan JF: **Ingestion of Lactobacillus strain regulates emotional behavior and central GABA receptor expression in a mouse via the vagus nerve**. *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108**(38):16050-16055.
52. Sudo N, Chida Y, Aiba Y, Sonoda J, Oyama N, Yu XN, Kubo C, Koga Y: **Postnatal microbial colonization programs the hypothalamic-pituitary-adrenal system for stress response in mice**. *The Journal of physiology* 2004, **558**(Pt 1):263-275.
53. O'Mahony SM, Clarke G, Borre YE, Dinan TG, Cryan JF: **Serotonin, tryptophan metabolism and the brain-gut-microbiome axis**. *Behavioural brain research* 2015, **277**:32-48.
54. Tan J, McKenzie C, Potamitis M, Thorburn AN, Mackay CR, Macia L: **The role of short-chain fatty acids in health and disease**. *Advances in immunology* 2014, **121**:91-119.
55. Rook GA, Lowry CA, Raison CL: **Lymphocytes in neuroprotection, cognition and emotion: is intolerance really the answer?** *Brain, behavior, and immunity* 2011, **25**(4):591-601.
56. Kunze WA, Mao YK, Wang B, Huizinga JD, Ma X, Forsythe P, Bienenstock J: **Lactobacillus reuteri enhances excitability of colonic AH neurons by inhibiting calcium-dependent potassium channel opening**. *Journal of cellular and molecular medicine* 2009, **13**(8B):2261-2270.
57. Mao YK, Kasper DL, Wang B, Forsythe P, Bienenstock J, Kunze WA: **Bacteroides fragilis polysaccharide A is necessary and sufficient for acute activation of intestinal sensory neurons**. *Nature communications* 2013, **4**:1465.
58. Brun P, Giron MC, Qesari M, Porzionato A, Caputi V, Zoppellaro C, Banzato S, Grillo AR, Spagnol L, De Caro R *et al*: **Toll-like receptor 2 regulates intestinal inflammation by controlling integrity of the enteric nervous system**. *Gastroenterology* 2013, **145**(6):1323-1333.
59. Cryan JF, Dinan TG: **Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour**. *Nature reviews Neuroscience* 2012, **13**(10):701-712.
60. Schwartz CE: **Aberrant tryptophan metabolism: the unifying biochemical basis for autism spectrum disorders?** *Biomarkers in medicine* 2014, **8**(3):313-315.
61. Mawe GM, Hoffman JM: **Serotonin signalling in the gut--functions, dysfunctions and therapeutic targets**. *Nature reviews Gastroenterology & hepatology* 2013, **10**(8):473-486.
62. Diaz Heijtz R, Wang S, Anuar F, Qian Y, Bjorkholm B, Samuelsson A, Hibberd ML, Forssberg H, Pettersson S: **Normal gut microbiota modulates brain development and behavior**. *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108**(7):3047-3052.
63. Clarke G, Grenham S, Scully P, Fitzgerald P, Moloney RD, Shanahan F, Dinan TG, Cryan JF: **The microbiome-gut-brain axis during early life regulates the hippocampal serotonergic system in a sex-dependent manner**. *Molecular psychiatry* 2013, **18**(6):666-673.
64. Wikoff WR, Anfora AT, Liu J, Schultz PG, Lesley SA, Peters EC, Siuzdak G: **Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites**. *Proceedings of the National Academy of Sciences of the United States of America* 2009, **106**(10):3698-3703.

65. Yano JM, Yu K, Donaldson GP, Shastri GG, Ann P, Ma L, Nagler CR, Ismagilov RF, Mazmanian SK, Hsiao EY: **Indigenous bacteria from the gut microbiota regulate host serotonin biosynthesis**. *Cell* 2015, **161**(2):264-276.
66. Reigstad CS, Salmonson CE, Rainey JF, 3rd, Szurszewski JH, Linden DR, Sonnenburg JL, Farrugia G, Kashyap PC: **Gut microbes promote colonic serotonin production through an effect of short-chain fatty acids on enterochromaffin cells**. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 2015, **29**(4):1395-1403.
67. Desbonnet L, Garrett L, Clarke G, Bienenstock J, Dinan TG: **The probiotic Bifidobacteria infantis: An assessment of potential antidepressant properties in the rat**. *Journal of psychiatric research* 2008, **43**(2):164-174.
68. Lyte M: **Microbial endocrinology in the microbiome-gut-brain axis: how bacterial production and utilization of neurochemicals influence behavior**. *PLoS pathogens* 2013, **9**(11):e1003726.
69. Barrett E, Ross RP, O'Toole PW, Fitzgerald GF, Stanton C: **gamma-Aminobutyric acid production by culturable bacteria from the human intestine**. *Journal of applied microbiology* 2012, **113**(2):411-417.
70. Mitchell RW, On NH, Del Bigio MR, Miller DW, Hatch GM: **Fatty acid transport protein expression in human brain and potential role in fatty acid transport across human brain microvessel endothelial cells**. *Journal of neurochemistry* 2011, **117**(4):735-746.
71. Davie JR: **Inhibition of histone deacetylase activity by butyrate**. *The Journal of nutrition* 2003, **133**(7 Suppl):2485S-2493S.
72. Peleg S, Sananbenesi F, Zovoilis A, Burkhardt S, Bahari-Javan S, Agis-Balboa RC, Cota P, Wittnam JL, Gogol-Doering A, Opitz L *et al*: **Altered histone acetylation is associated with age-dependent memory impairment in mice**. *Science* 2010, **328**(5979):753-756.
73. Frost G, Sleeth ML, Sahuri-Arisoylu M, Lizarbe B, Cerdan S, Brody L, Anastasovska J, Ghourab S, Hankir M, Zhang S *et al*: **The short-chain fatty acid acetate reduces appetite via a central homeostatic mechanism**. *Nature communications* 2014, **5**:3611.
74. Frye RE, Melnyk S, Macfabe DF: **Unique acyl-carnitine profiles are potential biomarkers for acquired mitochondrial disease in autism spectrum disorder**. *Translational psychiatry* 2013, **3**:e220.
75. Bercik P, Denou E, Collins J, Jackson W, Lu J, Jury J, Deng Y, Blennerhassett P, Macri J, McCoy KD *et al*: **The intestinal microbiota affect central levels of brain-derived neurotropic factor and behavior in mice**. *Gastroenterology* 2011, **141**(2):599-609, 609 e591-593.
76. Association AP: **Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)**: American Psychiatric Pub; 2013.
77. Furuta GT, Williams K, Kooros K, Kaul A, Panzer R, Coury DL, Fuchs G: **Management of constipation in children and adolescents with autism spectrum disorders**. *Pediatrics* 2012, **130 Suppl 2**:S98-105.
78. McElhanon BO, McCracken C, Karpen S, Sharp WG: **Gastrointestinal symptoms in autism spectrum disorder: a meta-analysis**. *Pediatrics* 2014, **133**(5):872-883.
79. Hsiao EY: **Gastrointestinal issues in autism spectrum disorder**. *Harvard review of psychiatry* 2014, **22**(2):104-111.
80. Finegold SM, Dowd SE, Gontcharova V, Liu C, Henley KE, Wolcott RD, Youn E, Summanen PH, Granpeesheh D, Dixon D *et al*: **Pyrosequencing study of fecal microflora of autistic and control children**. *Anaerobe* 2010, **16**(4):444-453.
81. Song Y, Liu C, Finegold SM: **Real-time PCR quantitation of clostridia in feces of autistic children**. *Applied and environmental microbiology* 2004, **70**(11):6459-6465.
82. Williams BL, Hornig M, Buie T, Bauman ML, Cho Paik M, Wick I, Bennett A, Jabado O, Hirschberg DL, Lipkin WI: **Impaired carbohydrate digestion and transport and mucosal**

dysbiosis in the intestines of children with autism and gastrointestinal disturbances. *PloS one* 2011, **6**(9):e24585.

83. Tomova A, Husarova V, Lakatosova S, Bakos J, Vlkova B, Babinska K, Ostatnikova D: **Gastrointestinal microbiota in children with autism in Slovakia**. *Physiology & behavior* 2015, **138**:179-187.
84. Kang DW, Park JG, Ilhan ZE, Wallstrom G, Labaer J, Adams JB, Krajmalnik-Brown R: **Reduced incidence of Prevotella and other fermenters in intestinal microflora of autistic children**. *PloS one* 2013, **8**(7):e68322.
85. Hsiao EY, McBride SW, Hsien S, Sharon G, Hyde ER, McCue T, Codelli JA, Chow J, Reisman SE, Petrosino JF *et al*: **Microbiota modulate behavioral and physiological abnormalities associated with neurodevelopmental disorders**. *Cell* 2013, **155**(7):1451-1463.
86. Dantzer R, Konsman JP, Bluthe RM, Kelley KW: **Neural and humoral pathways of communication from the immune system to the brain: parallel or convergent?** *Autonomic neuroscience : basic & clinical* 2000, **85**(1-3):60-65.
87. Zhu CB, Blakely RD, Hewlett WA: **The proinflammatory cytokines interleukin-1beta and tumor necrosis factor-alpha activate serotonin transporters**. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology* 2006, **31**(10):2121-2131.
88. Cai W, Khaoustov VI, Xie Q, Pan T, Le W, Yoffe B: **Interferon-alpha-induced modulation of glucocorticoid and serotonin receptors as a mechanism of depression**. *Journal of hepatology* 2005, **42**(6):880-887.
89. Lee YK, Menezes JS, Umesaki Y, Mazmanian SK: **Proinflammatory T-cell responses to gut microbiota promote experimental autoimmune encephalomyelitis**. *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108 Suppl 1**:4615-4622.
90. Ochoa-Reparaz J, Mielcarz DW, Ditrio LE, Burroughs AR, Begum-Haque S, Dasgupta S, Kasper DL, Kasper LH: **Central nervous system demyelinating disease protection by the human commensal Bacteroides fragilis depends on polysaccharide A expression**. *Journal of immunology* 2010, **185**(7):4101-4108.
91. Kobayashi T, Kato I, Nanno M, Shida K, Shibuya K, Matsuoka Y, Onoue M: **Oral administration of probiotic bacteria, Lactobacillus casei and Bifidobacterium breve, does not exacerbate neurological symptoms in experimental autoimmune encephalomyelitis**. *Immunopharmacology and immunotoxicology* 2010, **32**(1):116-124.
92. Kwon HK, Kim GC, Kim Y, Hwang W, Jash A, Sahoo A, Kim JE, Nam JH, Im SH: **Amelioration of experimental autoimmune encephalomyelitis by probiotic mixture is mediated by a shift in T helper cell immune response**. *Clinical immunology* 2013, **146**(3):217-227.
93. Huffnagle GB, Noverr MC: **The emerging world of the fungal microbiome**. *Trends in microbiology* 2013, **21**(7):334-341.
94. Underhill DM, Iliev ID: **The mycobiota: interactions between commensal fungi and the host immune system**. *Nature reviews Immunology* 2014, **14**(6):405-416.
95. Romani L: **Immunity to fungal infections**. *Nature reviews Immunology* 2011, **11**(4):275-288.
96. Shih VF, Tsui R, Caldwell A, Hoffmann A: **A single NFkappaB system for both canonical and non-canonical signaling**. *Cell research* 2011, **21**(1):86-102.
97. Puel A, Picard C, Cypowyj S, Lilic D, Abel L, Casanova JL: **Inborn errors of mucocutaneous immunity to Candida albicans in humans: a role for IL-17 cytokines?** *Current opinion in immunology* 2010, **22**(4):467-474.
98. Vinh DC, Sugui JA, Hsu AP, Freeman AF, Holland SM: **Invasive fungal disease in autosomal-dominant hyper-IgE syndrome**. *The Journal of allergy and clinical immunology* 2010, **125**(6):1389-1390.

99. Mellor AL, Munn DH: **IDO expression by dendritic cells: tolerance and tryptophan catabolism**. *Nature reviews Immunology* 2004, **4**(10):762-774.
100. Opitz CA, Litzenburger UM, Sahm F, Ott M, Tritschler I, Trump S, Schumacher T, Jestaedt L, Schrenk D, Weller M *et al*: **An endogenous tumour-promoting ligand of the human aryl hydrocarbon receptor**. *Nature* 2011, **478**(7368):197-203.
101. Fallarino F, Grohmann U, You S, McGrath BC, Cavener DR, Vacca C, Orabona C, Bianchi R, Belladonna ML, Volpi C *et al*: **The combined effects of tryptophan starvation and tryptophan catabolites down-regulate T cell receptor zeta-chain and induce a regulatory phenotype in naive T cells**. *Journal of immunology* 2006, **176**(11):6752-6761.
102. Naglik JR, Fidel PL, Jr., Odds FC: **Animal models of mucosal Candida infection**. *FEMS microbiology letters* 2008, **283**(2):129-139.
103. Mason KL, Erb Downward JR, Mason KD, Falkowski NR, Eaton KA, Kao JY, Young VB, Huffnagle GB: **Candida albicans and bacterial microbiota interactions in the cecum during recolonization following broad-spectrum antibiotic therapy**. *Infection and immunity* 2012, **80**(10):3371-3380.
104. Noverr MC, Falkowski NR, McDonald RA, McKenzie AN, Huffnagle GB: **Development of allergic airway disease in mice following antibiotic therapy and fungal microbiota increase: role of host genetics, antigen, and interleukin-13**. *Infection and immunity* 2005, **73**(1):30-38.
105. Dollive S, Chen YY, Grunberg S, Bittinger K, Hoffmann C, Vandivier L, Cuff C, Lewis JD, Wu GD, Bushman FD: **Fungi of the murine gut: episodic variation and proliferation during antibiotic treatment**. *PloS one* 2013, **8**(8):e71806.
106. Erb Downward JR, Falkowski NR, Mason KL, Muraglia R, Huffnagle GB: **Modulation of post-antibiotic bacterial community reassembly and host response by Candida albicans**. *Scientific reports* 2013, **3**:2191.
107. Bozza S, Fallarino F, Pitzurra L, Zelante T, Montagnoli C, Bellocchio S, Mosci P, Vacca C, Puccetti P, Romani L: **A crucial role for tryptophan catabolism at the host/Candida albicans interface**. *Journal of immunology* 2005, **174**(5):2910-2918.
108. Cheng SC, van de Veerdonk F, Smeekens S, Joosten LA, van der Meer JW, Kullberg BJ, Netea MG: **Candida albicans dampens host defense by downregulating IL-17 production**. *Journal of immunology* 2010, **185**(4):2450-2457.
109. Neville BA, d'Enfert C, Bougnoux ME: **Candida albicans commensalism in the gastrointestinal tract**. *FEMS yeast research* 2015, **15**(7).
110. Dollive S, Peterfreund GL, Sherrill-Mix S, Bittinger K, Sinha R, Hoffmann C, Nabel CS, Hill DA, Artis D, Bachman MA *et al*: **A tool kit for quantifying eukaryotic rRNA gene sequences from human microbiome samples**. *Genome biology* 2012, **13**(7):R60.
111. Iliev ID, Funari VA, Taylor KD, Nguyen Q, Reyes CN, Strom SP, Brown J, Becker CA, Fleshner PR, Dubinsky M *et al*: **Interactions between commensal fungi and the C-type lectin receptor Dectin-1 influence colitis**. *Science* 2012, **336**(6086):1314-1317.
112. Hoffmann C, Dollive S, Grunberg S, Chen J, Li H, Wu GD, Lewis JD, Bushman FD: **Archaea and fungi of the human gut microbiome: correlations with diet and bacterial residents**. *PloS one* 2013, **8**(6):e66019.
113. Angebault C, Djossou F, Abelanet S, Permal E, Ben Soltana M, Diancourt L, Bouchier C, Woerther PL, Catzeflis F, Andremont A *et al*: **Candida albicans is not always the preferential yeast colonizing humans: a study in Wayampi Amerindians**. *The Journal of infectious diseases* 2013, **208**(10):1705-1716.
114. Mar Rodriguez M, Perez D, Javier Chaves F, Esteve E, Marin-Garcia P, Xifra G, Vendrell J, Jove M, Pamplona R, Ricart W *et al*: **Obesity changes the human gut mycobiome**. *Scientific reports* 2015, **5**:14600.

115. Luan C, Xie L, Yang X, Miao H, Lv N, Zhang R, Xiao X, Hu Y, Liu Y, Wu N *et al*: **Dysbiosis of fungal microbiota in the intestinal mucosa of patients with colorectal adenomas**. *Scientific reports* 2015, **5**:7980.
116. Morgan XC, Tickle TL, Sokol H, Gevers D, Devaney KL, Ward DV, Reyes JA, Shah SA, LeLeiko N, Snapper SB *et al*: **Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment**. *Genome biology* 2012, **13**(9):R79.
117. Standaert-Vitse A, Jouault T, Vandewalle P, Mille C, Seddik M, Sendid B, Mallet JM, Colombel JF, Poulain D: **Candida albicans is an immunogen for anti-Saccharomyces cerevisiae antibody markers of Crohn's disease**. *Gastroenterology* 2006, **130**(6):1764-1775.
118. Ott SJ, Kuhbacher T, Musfeldt M, Rosenstiel P, Hellmig S, Rehman A, Drews O, Weichert W, Timmis KN, Schreiber S: **Fungi and inflammatory bowel diseases: Alterations of composition and diversity**. *Scandinavian journal of gastroenterology* 2008, **43**(7):831-841.
119. Li Q, Wang C, Tang C, He Q, Li N, Li J: **Dysbiosis of gut fungal microbiota is associated with mucosal inflammation in Crohn's disease**. *Journal of clinical gastroenterology* 2014, **48**(6):513.
120. Sokol H, Leducq V, Aschard H, Pham HP, Jegou S, Landman C, Cohen D, Liguori G, Bourrier A, Nion-Larmurier I *et al*: **Fungal microbiota dysbiosis in IBD**. *Gut* 2016.
121. Beveridge TJ: **Use of the gram stain in microbiology**. *Biotechnic & histochemistry : official publication of the Biological Stain Commission* 2001, **76**(3):111-118.
122. Blevins SM, Bronze MS: **Robert Koch and the 'golden age' of bacteriology**. *International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases* 2010, **14**(9):e744-751.
123. McFall-Ngai M: **Are biologists in 'future shock'? Symbiosis integrates biology across domains**. *Nature reviews Microbiology* 2008, **6**(10):789-792.
124. Woese CR, Fox GE: **Phylogenetic structure of the prokaryotic domain: the primary kingdoms**. *Proceedings of the National Academy of Sciences of the United States of America* 1977, **74**(11):5088-5090.
125. Olsen GJ, Lane DJ, Giovannoni SJ, Pace NR, Stahl DA: **Microbial ecology and evolution: a ribosomal RNA approach**. *Annual review of microbiology* 1986, **40**:337-365.
126. Sanger F, Nicklen S, Coulson AR: **DNA sequencing with chain-terminating inhibitors**. *Proceedings of the National Academy of Sciences of the United States of America* 1977, **74**(12):5463-5467.
127. Escobar-Zepeda A, Vera-Ponce de Leon A, Sanchez-Flores A: **The Road to Metagenomics: From Microbiology to DNA Sequencing Technologies and Bioinformatics**. *Frontiers in genetics* 2015, **6**:348.
128. Marchesi JR, Ravel J: **The vocabulary of microbiome research: a proposal**. *Microbiome* 2015, **3**:31.
129. Segata N, Boernigen D, Tickle TL, Morgan XC, Garrett WS, Huttenhower C: **Computational meta'omics for microbial community studies**. *Molecular systems biology* 2013, **9**:666.
130. Tang J, Iliev ID, Brown J, Underhill DM, Funari VA: **Mycobiome: Approaches to analysis of intestinal fungi**. *Journal of immunological methods* 2015, **421**:112-121.
131. Wang Q, Garrity GM, Tiedje JM, Cole JR: **Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy**. *Applied and environmental microbiology* 2007, **73**(16):5261-5267.
132. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL: **Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB**. *Applied and environmental microbiology* 2006, **72**(7):5069-5072.

133. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glockner FO: **The SILVA ribosomal RNA gene database project: improved data processing and web-based tools**. *Nucleic acids research* 2013, **41**(Database issue):D590-596.
134. Wayne L, Brenner D, Colwell R, Grimont P, Kandler O, Krichevsky M, Moore L, Moore W, Murray R, Stackebrandt E: **Report of the ad hoc committee on reconciliation of approaches to bacterial systematics**. *International Journal of Systematic and Evolutionary Microbiology* 1987, **37**(4):463-464.
135. Stackebrandt E, Goebel B: **Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology**. *International Journal of Systematic and Evolutionary Microbiology* 1994, **44**(4):846-849.
136. Bodilis J, Nsigue-Meilo S, Besaury L, Quillet L: **Variable copy number, intra-genomic heterogeneities and lateral transfers of the 16S rRNA gene in Pseudomonas**. *PloS one* 2012, **7**(4):e35647.
137. Acinas SG, Marcelino LA, Klepac-Ceraj V, Polz MF: **Divergence and redundancy of 16S rRNA sequences in genomes with multiple rrn operons**. *Journal of bacteriology* 2004, **186**(9):2629-2635.
138. Logares R, Haverkamp TH, Kumar S, Lanzen A, Nederbragt AJ, Quince C, Kauserud H: **Environmental microbiology through the lens of high-throughput DNA sequencing: synopsis of current platforms and bioinformatics approaches**. *Journal of microbiological methods* 2012, **91**(1):106-113.
139. Achtman M, Wagner M: **Microbial diversity and the genetic nature of microbial species**. *Nature reviews Microbiology* 2008, **6**(6):431-440.
140. Petrosino JF, Highlander S, Luna RA, Gibbs RA, Versalovic J: **Metagenomic pyrosequencing and microbial identification**. *Clinical chemistry* 2009, **55**(5):856-866.
141. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ *et al*: **Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities**. *Applied and environmental microbiology* 2009, **75**(23):7537-7541.
142. Weinstock GM: **Genomic approaches to studying the human microbiota**. *Nature* 2012, **489**(7415):250-256.
143. Haas BJ, Gevers D, Earl AM, Feldgarden M, Ward DV, Giannoukos G, Ciulla D, Tabbaa D, Highlander SK, Sodergren E *et al*: **Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons**. *Genome research* 2011, **21**(3):494-504.
144. Chen W, Zhang CK, Cheng Y, Zhang S, Zhao H: **A comparison of methods for clustering 16S rRNA sequences into OTUs**. *PloS one* 2013, **8**(8):e70837.
145. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ: **Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities**. *Applied and environmental microbiology* 2009, **75**(23):7537-7541.
146. Sun Y, Cai Y, Liu L, Yu F, Farrell ML, McKendree W, Farmerie W: **ESPRIT: estimating species richness using large collections of 16S rRNA pyrosequences**. *Nucleic acids research* 2009, **37**(10):e76.
147. Li W, Godzik A: **Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences**. *Bioinformatics* 2006, **22**(13):1658-1659.
148. Edgar RC: **Search and clustering orders of magnitude faster than BLAST**. *Bioinformatics* 2010, **26**(19):2460-2461.
149. White JR, Navlakha S, Nagarajan N, Ghodsi MR, Kingsford C, Pop M: **Alignment and clustering of phylogenetic markers--implications for microbial diversity studies**. *Bmc Bioinformatics* 2010, **11**:152.

150. Schloss PD: **The effects of alignment quality, distance calculation method, sequence filtering, and region on the analysis of 16S rRNA gene-based studies**. *PLoS computational biology* 2010, **6**(7):e1000844.
151. Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R: **UniFrac: an effective distance metric for microbial community comparison**. *The ISME journal* 2011, **5**(2):169-172.
152. Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD: **Multiple sequence alignment with the Clustal series of programs**. *Nucleic acids research* 2003, **31**(13):3497-3500.
153. Edgar RC: **MUSCLE: multiple sequence alignment with high accuracy and high throughput**. *Nucleic acids research* 2004, **32**(5):1792-1797.
154. Notredame C, Higgins DG, Heringa J: **T-Coffee: A novel method for fast and accurate multiple sequence alignment**. *Journal of molecular biology* 2000, **302**(1):205-217.
155. Schloss PD: **A high-throughput DNA sequence aligner for microbial ecology studies**. *PloS one* 2009, **4**(12):e8230.
156. Gardner PP, Wilm A, Washietl S: **A benchmark of multiple sequence alignment programs upon structural RNAs**. *Nucleic acids research* 2005, **33**(8):2433-2439.
157. Nawrocki EP, Kolbe DL, Eddy SR: **Infernal 1.0: inference of RNA alignments**. *Bioinformatics* 2009, **25**(10):1335-1337.
158. Pruesse E, Peplies J, Glöckner FO: **SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes**. *Bioinformatics* 2012, **28**(14):1823-1829.
159. DeSantis T, Hugenholtz P, Keller K, Brodie E, Larsen N, Piceno Y, Phan R, Andersen GL: **NAST: a multiple sequence alignment server for comparative analysis of 16S rRNA genes**. *Nucleic acids research* 2006, **34**(suppl 2):W394-W399.
160. Caporaso JG, Bittinger K, Bushman FD, DeSantis TZ, Andersen GL, Knight R: **PyNAST: a flexible tool for aligning sequences to a template alignment**. *Bioinformatics* 2010, **26**(2):266-267.
161. Albanese D, Fontana P, De Filippo C, Cavalieri D, Donati C: **MICCA: a complete and accurate software for taxonomic profiling of metagenomic data**. *Scientific reports* 2015, **5**:9743.
162. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI: **QIIME allows analysis of high-throughput community sequencing data**. *Nature methods* 2010, **7**(5):335-336.
163. Lozupone CA, Knight R: **Species divergence and the measurement of microbial diversity**. *FEMS microbiology reviews* 2008, **32**(4):557-578.
164. Lozupone C, Knight R: **UniFrac: a new phylogenetic method for comparing microbial communities**. *Applied and environmental microbiology* 2005, **71**(12):8228-8235.
165. Lozupone CA, Hamady M, Kelley ST, Knight R: **Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities**. *Applied and environmental microbiology* 2007, **73**(5):1576-1585.
166. Santamaria M, Fosso B, Consiglio A, De Caro G, Grillo G, Licciulli F, Liuni S, Marzano M, Alonso-Alemany D, Valiente G *et al*: **Reference databases for taxonomic assignment in metagenomics**. *Briefings in bioinformatics* 2012, **13**(6):682-695.
167. Abarenkov K, Henrik Nilsson R, Larsson KH, Alexander IJ, Eberhardt U, Erland S, Hoiland K, Kjoller R, Larsson E, Pennanen T *et al*: **The UNITE database for molecular identification of fungi--recent updates and future perspectives**. *The New phytologist* 2010, **186**(2):281-285.
168. Nilsson RH, Ryberg M, Kristiansson E, Abarenkov K, Larsson KH, Koljalg U: **Taxonomic reliability of DNA sequences in public sequence databases: a fungal perspective**. *PloS one* 2006, **1**:e59.

Chapter 1

Age and gender affect the composition of fungal population of the human gastrointestinal tract

This chapter has been reprinted* from:

Strati F, Di Paola M, Stefanini I, Albanese D, Rizzetto L, Lionetti P, Calabrò A, Jousson O, Donati C, Cavalieri D and De Filippo C (2016). *Age and gender affect the composition of fungal population of the human gastrointestinal tract*. Frontiers in Microbiology 7:1227, doi: 10.3389/fmicb.2016.01227.

The human gut microbiota exerts key roles in host energy metabolism and the maintaining of homeostatic controls over the immune system [1]. The term microbiota usually refers to the bacterial component of a microbial community although virus, unicellular and multicellular eukaryotes co-exists and co-evolved within and with the host. The fungal part of the microbiota is usually referred as the "mycobiota" and, although fungi represent only a small part of the microbiota compared to bacteria, its role in regulating host physiology is gaining great attention [2]. Despite the importance of the human mycobiota on host physiology, the prevalent interest in describing fungi as aetiological factors in infectious diseases lead to the underestimation of its role in the health status [2]. The aim of the study presented in this chapter was to characterize the microbial community structure of the human gut mycobiota by mean of culture-dependent and culture-independent methods in order to describe how the intestinal fungal microbiota varies in function of the life stage and gender of the host. In addition, we evaluated the ability of the isolated intestinal fungi to survive to gastrointestinal tract-like challenges and their susceptibility to antifungals in order to evaluate fungal adaptation to the gastrointestinal niche and virulence potentiality. The combination of metataxonomics and fungal cultivation allowed an in-depth understanding of the fungal intestinal community structure associated with the healthy status and the commensalism-related traits of the isolated fungi revealing that the human gut mycobiota changes during the lifetime of individuals and differs in a gender-related fashion. Interestingly, we found a high fraction of azole-resistant isolates from healthy subjects that had not been treated by antifungals, suggesting that the spread of antifungal resistance could represent a not yet recognized epidemiological problem. My personal contribution to this work concerned every aspect of the research project. After the conception of the study, I designed and performed all the experiments under the supervision and thanks to the expertise of all the co-authors of this paper. In particular I increased my past expertise in microbiology applying methods of classical microbiology to yeasts and moulds, deepening my knowledge in fungal and yeast biology. In addition I learned how to manage and analyse high-throughput sequencing data by applying different bioinformatics pipelines and their application in microbiome research by the integration of different type of data (*e.g.* ITS1 sequences, demographic or clinical data). The use of the R statistical programming language and the correct understanding of the different statistical tests used to fully support my discoveries gave a great impulse to this research project and to all the research work presented in this Ph.D. thesis. Finally, I wrote the paper and generated all the figures and tables.

References

1. Kamada N, Seo SU, Chen GY, Nunez G: **Role of the gut microbiota in immunity and inflammatory disease**. *Nature reviews Immunology* 2013, **13**(5):321-335.

2. Underhill DM, Iliev ID: **The mycobiota: interactions between commensal fungi and the host immune system**. *Nature reviews Immunology* 2014, **14**(6):405-416.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 03 August 2016
doi: 10.3389/fmicb.2016.01227

Age and Gender Affect the Composition of Fungal Population of the Human Gastrointestinal Tract

Francesco Strati[1,2], Monica Di Paola[3], Irene Stefanini[1], Davide Albanese[1], Lisa Rizzetto[1], Paolo Lionetti[3], Antonio Calabrò[4], Olivier Jousson[2], Claudio Donati[1], Duccio Cavalieri[1,5] and Carlotta De Filippo[6]**

[1] Department of Computational Biology, Research and Innovation Centre, Fondazione Edmund Mach, San Michele all' Adige, Italy, [2] Centre for Integrative Biology, University of Trento, Trento, Italy, [3] Department of Neuroscience, Psychology, Drug Research and Child Health, Meyer Children's Hospital, University of Florence, Florence, Italy, [4] Gastroenterology Unit, Department of Experimental and Clinical Biomedical Sciences, University of Florence, Florence, Italy, [5] Department of Biology, University of Florence, Sesto Fiorentino, Florence, Italy, [6] Institute of Biometeorology, National Research Council, Florence, Italy

OPEN ACCESS

Edited by:
Agostinho Carvalho, University of Minho, Portugal

Reviewed by:
*Norman Pavelka, Singapore Immunology Network (A*STAR), Singapore*
Teresa Rinaldi, Sapienza University of Rome, Italy

***Correspondence:**
Duccio Cavalieri
duccio.cavalieri@unifi.it
Carlotta De Filippo
c.de.filippo@ibimet.cnr.it

Specialty section:
This article was submitted to Fungi and Their Interactions, a section of the journal Frontiers in Microbiology

Received: *25 May 2016*
Accepted: *22 July 2016*
Published: *03 August 2016*

Citation:
Strati F, Di Paola M, Stefanini I, Albanese D, Rizzetto L, Lionetti P, Calabrò A, Jousson O, Donati C, Cavalieri D and De Filippo C (2016) Age and Gender Affect the Composition of Fungal Population of the Human Gastrointestinal Tract. Front. Microbiol. 7:1227. doi: 10.3389/fmicb.2016.01227

The fungal component of the human gut microbiota has been neglected for long time due to the low relative abundance of fungi with respect to bacteria, and only recently few reports have explored its composition and dynamics in health or disease. The application of metagenomics methods to the full understanding of fungal communities is currently limited by the under representation of fungal DNA with respect to the bacterial one, as well as by the limited ability to discriminate passengers from colonizers. Here, we investigated the gut mycobiota of a cohort of healthy subjects in order to reduce the gap of knowledge concerning fungal intestinal communities in the healthy status further screening for phenotypical traits that could reflect fungi adaptation to the host. We studied the fecal fungal populations of 111 healthy subjects by means of cultivation on fungal selective media and by amplicon-based ITS1 metagenomics analysis on a subset of 57 individuals. We then characterized the isolated fungi for their tolerance to gastrointestinal (GI) tract-like challenges and their susceptibility to antifungals. A total of 34 different fungal species were isolated showing several phenotypic characteristics associated with intestinal environment such as tolerance to body temperature (37°C), to acidic and oxidative stress, and to bile salts exposure. We found a high frequency of azoles resistance in fungal isolates, with potential and significant clinical impact. Analyses of fungal communities revealed that the human gut mycobiota differs in function of individuals' life stage in a gender-related fashion. The combination of metagenomics and fungal cultivation allowed an in-depth understanding of the fungal intestinal community structure associated to the healthy status and the commensalism-related traits of isolated fungi. We further discussed comparatively the results of sequencing and cultivation to critically evaluate the application of metagenomics-based approaches to fungal gut populations.

Keywords: commensal fungi, human gut mycobiota, antifungal resistance, fungal metagenomics, fungi-host interactions

INTRODUCTION

The human gut is a complex ecological niche in which archaea, bacteria, protozoa, fungi, and viruses co-exist in close association with the host (Reyes et al., 2010; Arumugam et al., 2011; Human Microbiome Project Consortium, 2012). Even if it has been estimated that the number of bacteria hugely outreaches the number of fungi in the gastrointestinal (GI) tract (Huffnagle and Noverr, 2013), fungi play a relevant role in the physiology of the human host (Oever and Netea, 2014; Underhill and Iliev, 2014). Recent studies showed that, while the composition of the bacterial community is relatively stable over time, the fungal population inhabiting the murine gut undergoes significant changes during the animal's lifetime (Dollive et al., 2013). This brought to the conclusion that gut fungal populations are more variable than bacterial ones and that their composition may be influenced by environmental fungi (Underhill and Iliev, 2014). Despite evidence that fungi inhabit the mammalian GI tract and interact with the host immune system (Romani, 2011; Rizzetto et al., 2014; Underhill and Iliev, 2014), the composition and characteristics of the mycobiota in healthy hosts have been poorly explored. The prevalent interest in describing pathogenic fungi, their phenotypes and the process by which they establish the infection is one of the major cause that brought to neglect the harmless part of the commensal fungal population. Despite this topic has been only marginally explored to date, it has been shown that mucosal fungi are able to modulate both the innate and adaptive immune responses (Romani, 2011; Rizzetto et al., 2014; Underhill and Iliev, 2014) thus supporting the need to further study the whole gut mycobiota. Furthermore, alterations of the gut mycobiota have been associated to different pathologies ranging from metabolic disorders (obesity) to colorectal adenomas and Inflammatory Bowel Diseases (IBDs) (Luan et al., 2015; Mar Rodriguez et al., 2015; Sokol et al., 2016). A recent study showed the association of IBDs to alteration of the gut mycobiota. In particular Sokol and colleagues showed that IBD patients bear a smaller proportion of *Saccharomyces cerevisiae* and higher of *Candida albicans* compared to healthy subjects. In addition, they highlighted the existence in Crohn's disease of interconnected alterations between bacterial and fungal communities (Sokol et al., 2016). However, the role of the gut mycobiota in the maintenance of health it is still far from being well-understood because the studies carried out so far focused on disease-causing taxa. Nevertheless, some yeasts have been clinically prescribed for a long time because of their potential probiotic properties, suggesting a beneficial role of some fungi for host health. A great example of "beneficial" fungus is represented by *S. cerevisiae* var. *boulardii,* used for the relief of gastroenteritis (Hatoum et al., 2012). In order to reduce the gap of knowledge concerning the gut mycobiota and its interplay with the host, we characterized the gut mycobiota composition of a cohort of healthy subjects by means of metagenomics, fungal cultivation, and phenotypic assays.

MATERIALS AND METHODS

Study Participants

Fecal samples were collected from 111 Italian healthy volunteers (49 male and 62 female, average age, 10 ± 8.2; **Table 1**) and analyzed within 24 h. Written informed consent has been obtained from all the enrolled subjects or tutors in accordance with the guidelines and regulations approved by the Research Ethical Committees of the Meyer Children's Hospital and the Azienda Ospedaliera Careggi, Florence. All the subjects enrolled were non-smokers, followed a Mediterranean-based diet and they did not take antibiotics, antifungals or probiotics in the 6 months prior to sample collection. None of the participants had any history of GI abnormalities.

Isolation and Identification of Cultivable Fungal Species from Feces

Stool samples were diluted in sterile Ringer's solution and plated on solid YPD medium (1% Yeast extract, 2% Bacto-peptone, 2% D-glucose, 2% agar) supplemented with 25 U/ml of penicillin, 25 μg/ml of streptomycin (Sigma-Aldrich) and incubated aerobically at 27°C for 3–5 days. All fungal isolates grown on the selective medium were further isolated to obtain single-cell pure colonies. Genomic DNA was extracted from pure cultures of isolated colonies as previously described (Hoffman and Winston, 1987). Strains were identified by amplification and sequencing of the ribosomal Internal Transcribed Spacer (ITS) region, using ITS1 (5′-GTTTCCGTAGGTGAACTTGC-3′) and ITS4 (5′-TCCTCCGCTTATTGATATGC-3′) primers, as previously described (Sebastiani et al., 2002). Fungal isolates were identified by using the BLAST algorithm in the NCBI database (minimum 97% sequence similarity and 95% coverage with a described species).

Phenotypical Characterization of Fungal Isolates

Fungal isolates were tested for phenotypical features that could be related to the ability of colonization and persistence in the human gut. Cell growth in liquid media was monitored by optical density measurement at 630 nm with a microplate reader (Synergy2, BioTek, USA) after 48 h of incubation under tested conditions. Three independent replicates were performed for each test.

Growth at Supra Optimal Temperatures

Fungal isolates ($\sim 10^5$ cells/ml) were grown at supra optimal temperatures in liquid YPD medium (40, 42, 44, and 46°C).

pH Impact on Growth

Fungal isolates ($\sim 10^5$ cells/ml) were grown at 37°C in liquid YPD medium at pH 2.0 and pH 3.0 adding hydrochloric acid/potassium chloride and citrate buffers, respectively, to test their ability to resist to the acidic environments encountered during GI tract passage.

TABLE 1 | Characteristics of the study participants.

Age group (year)		Infants (0–2)	Children (3–10)		Adolescents (11–17)	Adults (≥18)		All subjects
Number of subjects		18	48		24	21		111
% with fungi		88.9	83.3		70.8	76.2		80.2
Subject ID	**Gender**	**Age (year)**	**Subject ID**	**Gender**	**Age (year)**	**Subject ID**	**Gender**	**Age (year)**
HS1	M	5	HS38*	F	25	HS75	F	1
HS2	M	5	HS39*	F	27	HS76*	M	1
HS3	M	14	HS40*	M	27	HS77	F	4
HS4	M	1	HS41*	F	24	HS78*	M	12
HS5	F	20	HS42*	F	24	HS79*	M	0.1
HS6	F	20	HS43*	M	26	HS80	F	0.1
HS7	F	20	HS44*	F	24	HS81	F	7
HS8	M	5	HS45*	F	6	HS82*	M	10
HS9	M	14	HS46*	F	6	HS83*	M	12
HS10*	F	2	HS47*	F	10	HS84	F	6
HS11	M	16	HS48*	F	2.5	HS85	F	10
HS12	M	15	HS49*	M	2.5	HS86*	M	7
HS13*	F	18	HS50*	F	1.5	HS87*	M	9
HS14	F	0.3	HS51*	F	8	HS88*	M	7
HS15*	F	11	HS52*	F	23	HS89*	M	12
HS16	M	14	HS53*	F	23	HS90	F	8
HS17	M	15	HS54	M	2	HS91	F	2
HS18	M	11	HS55*	M	2	HS92	F	12
HS19	F	3	HS56*	M	2	HS93	F	4
HS20*	F	4	HS57	F	12	HS94	F	4
HS21*	F	5	HS58	F	3	HS95	F	10
HS22*	F	15	HS59*	M	5	HS96	F	12
HS23*	F	11	HS60	F	3	HS97*	M	6
HS24	M	15	HS61*	M	2	HS98	F	16
HS25	M	7	HS62	F	4	HS99	F	3
HS26	M	3	HS63*	M	5	HS100*	M	0.1
HS27*	F	9	HS64	F	3	HS101*	M	4
HS28	M	5	HS65*	M	5	HS102	F	13
HS29*	F	16	HS66*	M	0.1	HS103*	M	7
HS30*	F	12	HS67	F	1	HS104*	M	4
HS31*	F	24	HS68	F	4	HS105	F	8
HS32*	F	32	HS69*	M	6	HS106	F	5
HS33*	F	32	HS70	F	11	HS107	M	13
HS34*	F	25	HS71*	M	1	HS108	M	4.5
HS35*	F	26	HS72	F	10	HS109	M	1
HS36*	M	20	HS73	F	4	HS110	M	12
HS37*	F	28	HS74*	M	6	HS111	M	18

**Samples analyzed also by mean of amplicon-based ITS1metagenomics.*

Tolerance to Bile Acids

Fungal isolates ($\sim 10^5$ cells/ml) were grown in liquid YPD medium at 37°C in the presence of three different concentrations of bile [Ox-bile, Sigma-Aldrich; 0.5, 1, and 2% (w/v)] mimicking the physiological intestinal settings (Noriega et al., 2004).

Resistance to Oxidative Stress

Fungal resistance to oxidative stress was evaluated by measuring the inhibition halo induced by the treatment of fungal strains ($\sim 10^7$ cells/ml) grown on YPD solid medium with 0.5 mM hydrogen peroxide (H_2O_2). The percentage of sensitivity to oxidative stress was calculated as the deviation of the inhibition

halo diameter (Ø) from that of the environmental, oxidative stress sensitive M28-4D *S. cerevisiae* strain (Cavalieri et al., 2000) according to the following formula: [(Ø sample—Ø M28-4D)/Ø M284D]*100.

Invasive Growth

The ability of fungal strains to penetrate the YPD solid medium was tested as previously described (Vopalenska et al., 2005). M28-4D and BY4742 *S. cerevisiae* strains, known to be invasive and non-invasive, respectively, have been used as controls. The strain invasiveness was assigned with scores from 3 (highly invasive) to 0 (non-invasive).

Hyphal Formation

Fungal cells ($\sim 10^5$ cells/ml) were grown for 7 days in liquid YPD and YNB media [0.67% Yeast Nitrogen Base w/o aminoacids and $(NH_4)_2SO_4$ (Sigma-Aldrich), 2% glucose], both at 27 and 37°C in order to evaluate hyphae or pseudohyphae formation. Formation of hyphae was inspected by optical microscope observation with a Leica DM1000 led instrument (magnification 40x and 100x).

Antifungal Susceptibility Testing

All fungal isolates were tested for susceptibility to fluconazole, itraconazole, and 5-flucytosine (Sigma-Aldrich) by Minimum Inibitory Concentration (MIC) assays according to the European Committee on Antimicrobial Susceptibility Testing (EUCAST) recommendations (Rodriguez-Tudela et al., 2008a,b). Clinical and Laboratory Standards Institute (CLSI) clinical breakpoints (CBPs) were used to evaluate the antifungal resistance (Pfaller and Diekema, 2012; Castanheira et al., 2014). CBPs have not been established for non-*Candida* yeasts and non-*Aspergillus* molds, however have been used as a proxy for the evaluation of antifungals susceptibility in such isolates.

DNA Extraction and PCR Amplification of Fungal ITS1 rDNA Region

DNA extraction from fecal samples (250 mg) was performed using the FastDNA™ SPIN Kit for Feces (MP-Biomedicals, USA) following manufacturer's instructions. DNA quality was checked on 1% agarose gel TAE 1X and quantified with a NanoDrop® spectrophotometer. For each sample, fungal ITS1 rDNA region was amplified using a specific fusion primer set coupled with forward primer 18SF (5′-GTAAAAGTCGTAACAAGGTTTC-3′) and reverse primer 5.8S1R (5′-GTTCAAAGAYTCGATGATTCAC-3′; Findley et al., 2013) containing adaptors, key sequence and barcode (Multiple IDentifier) sequences as described by the 454 Sequencing System Guidelines for Amplicon Experimental Design (Roche, Switzerland). The PCR reaction mix contained 1X FastStart High Fidelity PCR buffer, 2 mM $MgCl_2$, 200 μM of dNTPs, 0.4 μM of each primer (PRIMM, Italy), 2.5 U of FastStart High Fidelity Polymerase Blend, and 100 ng of gDNA as template. Thermal cycling conditions used were 5 min at 95°C, 35 cycles of 45 s at 95°C, 45 s at 56°C, and 1.30 min at 72°C followed by a final extension of 10 min at 72°C. All PCR experiments were carried out in triplicates using a Veriti® Thermal Cycler (Applied Biosystems, USA).

Library Construction and Pyrosequencing

The PCR products obtained were analyzed by gel electrophoresis and cleaned using the AMPure XP beads kit (Beckman Coulter, USA) following the manufacturer's instructions, quantified via quantitative PCR using the Library quantification kit—Roche 454 titanium (KAPA Biosystems, USA) and pooled in equimolar way in a final amplicon library. The 454 pyrosequencing was carried out on the GS FLX+ system using the XL+ chemistry following the manufacturer's recommendations (Roche, Switzerland).

Data Analysis

Pyrosequencing resulted in a total of 1.337.184 reads with a mean of 19.379 ± 13.334 sequences *per* sample. Raw 454 files were demultiplexed using the Roche's sff file software and submitted to the European Nucleotide Archive with accession number PRJEB11827 (http://www.ebi.ac.uk/ena/data/view/PRJEB11827). Sample accessions and metadata are available in **Supplementary Table S1**. Reads were pre-processed using the MICCA pipeline (Albanese et al., 2015) (http://www.micca.org). Forward and reverse primers trimming and quality filtering were performed using micca-preproc. *De-novo* sequence clustering, chimera filtering, and taxonomy assignment were performed by micca-otu-*denovo*: Operational Taxonomic Units (OTUs) were assigned by clustering the sequences with a threshold of 97% pairwise identity and their representative sequences were classified using the RDP classifier version 2.8 (Wang et al., 2007) against the UNITE fungal ITS database (Koljalg et al., 2013). *De novo* multiple sequence alignment was performed using T-Coffee (Notredame et al., 2000). Fungal taxonomy assignments were then manually curated using BLASTn against the GenBank's database for accuracy. High quality fungal sequences were detected in all samples. Furthermore, the sequences belonging to Agaricomycetes [unlikely to be residents of the human gut due to their ecology Hibbett, 2006] were manually filtered out.

The phylogenetic tree was inferred by using micca-phylogeny (Price et al., 2010). Rarefaction analysis resulted in a sequencing depth adequate to capture the ecological diversity of the samples up to saturation. Sampling heterogeneity was reduced by rarefaction. *Alpha* and *beta*-diversity estimates were computed using the phyloseq R package (McMurdie and Holmes, 2013). PERMANOVA (Permutational multivariate analysis of variance) was performed using the adonis() function of the *vegan* R package with 999 permutations. Permutations have been constrained within age groups (corresponding to 0–2, 3–10, 11–17, and >18 y/o) or gender to reduce possible biases related to the unequal age and gender distributions among subjects using the "strata" argument within the adonis() function. Two-sided, unpaired Welch t-statistics were computed using the function mt() in the phyloseq library and the *p*-values were adjusted for multiple comparison controlling the family-wise Type I error rate (minP procedure; Westfall and Young, 1993). Wilcoxon rank-sum tests and Spearman's correlations were performed using the R software (Team, 2014) through the *stats* R package (version 3.1.2) and the *psych* R package (Revelle, 2013), respectively. *p*-values have been corrected for multiple comparison by using the false discovery rate correction (Benjamini and Hochberg, 1995).

RESULTS

Cultivable Gut Mycobiota

The cultivable gut mycobiota of 111 healthy volunteers was investigated through isolation in selective media. Fungi were detected in more than 80% of subjects leading to the identification of 349 different isolates (**Supplementary Table S2**). Thirty-four different fungal species were detected at different frequencies of isolation (**Table 2**) among which *Aspergillus glaucus, Candida albicans, Candida deformans, Candida fermentati, Candida glabrata, Candida intermedia, Candida lusitaniae, Candida metapsilosis, Candida parapsilosis, Candida pararugosa, Candida tropicalis, Candida zelanoydes, Cryptococcus saitoi, Lichtheimia ramosa, Mucor circinelloides, Pleurostomophora richardsiae, Rhodotorula mucilaginosa, Trichosporon asahii, Yarrowia lipolytica*. These species were previously found in different human body sites, including the GI tract as commensal or opportunistic pathogens (Araujo et al., 2007; Johnson, 2009; Alastruey-Izquierdo et al., 2010; Kurtzman et al., 2011; Levenstadt et al., 2012; Gouba et al., 2014; Lee et al., 2014; Rizzetto et al., 2014). We also isolated the environmental fungi *Aspergillus pseudoglaucus, Eurotium amstelodami, Eurotium rubrum, Penicillium brevicompactum, Penicillium paneum, Penicillium crustosum, Pichia caribbica, Pichia fermentans, Pichia kluyveri, Pichia manshurica, Rhodosporidium kratochvilovae, Saccharomyces cerevisiae, Starmerella bacillaris*, and *Torulaspora delbrueckii*. Such species were previously found in fermentations, oenological samples (Chitarra et al., 2004; Butinar et al., 2005; Kurtzman et al., 2011; Barata et al., 2012; Bezerra-Bussoli et al., 2013; Tristezza et al., 2013; Vardjan et al., 2013; Belda et al., 2015; de Melo Pereira et al., 2014; Santini et al., 2014; Wang et al., 2014) and rarely found in clinical samples (de la Camara et al., 1996; Kaygusuz et al., 2003; Butinar et al., 2005; Rizzetto et al., 2014). The 39.8% of subjects showed at least one *C. albicans* isolate, which resulted in the most common yeast species found in our samples, in line with previous reports on the gut mycobiota of healthy subjects (Khatib et al., 2001; Bougnoux et al., 2006).

Population level analysis of the cultivable gut mycobiota revealed significant gender-related differences, with female subjects showing a higher number of fungal isolates ($p < 0.005$, Wilcoxon rank-sum test; **Figure 1A**) and fungal species ($p < 0.05$, Wilcoxon rank-sum test; **Figure 1B**) compared to male subjects (not related to individual's age) while we did not observed significant differences in the fungal population among the investigated age groups (**Figures 1C,D**). Finally, no species *per se* was responsible for these differences, as indicated by the fact that we did not find significant differences between individual species abundances in male and female subjects for any investigated age group.

TABLE 2 | Fungal isolates and frequencies of isolation.

Species	%	Species	%
Candida albicans	39.8	*Rhodosporidium kratochvilovae*	0.57
Rhodotorula mucilaginosa	12.6	*Trichosporon asahii*	0.57
Candida parapsilosis	12.3	*Yarrowia lipolytica*	0.57
Torulaspora delbrueckii	6.59	*Aspergillus cristatus*	0.28
Pichia fermentans	4.29	*Candida deformans*	0.28
Penicillium brevicompactum	3.72	*Candida fermentati*	0.28
Pichia manshurica	3.43	*Candida glabrata*	0.28
Pichia kluyveri	2.86	*Candida intermedia*	0.28
Candida lusitaniae	2.58	*Candida metapsilosis*	0.28
Pennicillium crustosum	1.43	*Candida tropicalis*	0.28
Saccharomyces cerevisiae	1.14	*Candida zelanoydes*	0.28
Penicillium paneum	0.58	*Eurotium amstelodami*	0.28
Aspergillus glaucus	0.57	*Eurotium rubrum*	0.28
Aspergillus pseudoglaucus	0.57	*Lichtheimia ramose*	0.28
Candida pararugosa	0.57	*Pichia carribica*	0.28
Cryptococcus saitoi	0.57	*Pleurostomophora richardsiae*	0.28
Mucor circinelloides	0.57	*Starmerella bacillaris*	0.28

Fungal Gut Metagenomics

To better characterize the intestinal fungal community structure associated to our cohort of healthy subjects we further analyzed a subset of these subjects (57 subjects, 29 females, and 28 males, average age 12 ± 9.5) by means of amplicon-based ITS1 targeted metagenomics, looking at gender and age groups differences. The analysis led to the identification of 68 fully classified (to the genus level) fungal taxa and 26 taxa only partially classified (of which 2 classified to the phylum level, 5 classified to the order level, 9 classified to the class level, and 9 classified to the family level). Measurements of the fungal richness within each sample i.e., the *alpha*-diversity (see Materials and Methods), revealed no significant differences among male and female subjects (**Figure 2A**), differently from the above finding based on the culture-based analysis in which we observed an increased number of intestinal fungal species in females compared to males (**Figure 1B**). Furthermore, we observed that infants and children harbor a higher fungal richness compared to adults as indicated by the number of the observed OTUs ($p < 0.05$, Wilcoxon rank-sum test, **Figure 2B**). The analysis of *beta*-diversity identified significant differences in the composition of the gut mycobiota among gender and age groups. PCoA (Principal Coordinates Analysis) revealed that samples cluster by gender, based on the unweighted UniFrac distance and the Bray-Curtis dissimilarity ($p < 0.05$, PERMANOVA; **Figures 2C,D**, **Supplementary Table S3**) and by age groups, based on the unweighted UniFrac distance ($p < 0.05$, PERMANOVA; **Figure 2C**, **Supplementary Table S3**). We calculated PERMANOVAs constraining permutations within levels (gender or age groups) to avoid biases related to the unequal distribution of genders among age groups and *vice-versa*. Genus level analysis showed *Penicillium*, *Aspergillus*, and *Candida* as the most abundant genera in this subset of subjects (22.3, 22.2, and 16.9%, respectively; **Figure 3**, **Supplementary Table S4**). We further observed that *Aspergillus* and Tremellomycetes_unidentified_1 were significantly more abundant in male than female subjects ($p < 0.05$, Welch *t*-test) and in children than adults ($p < 0.05$, Welch *t*-test). To note, the latter result could be biased by the unbalanced distribution of male and female subjects in children and adults

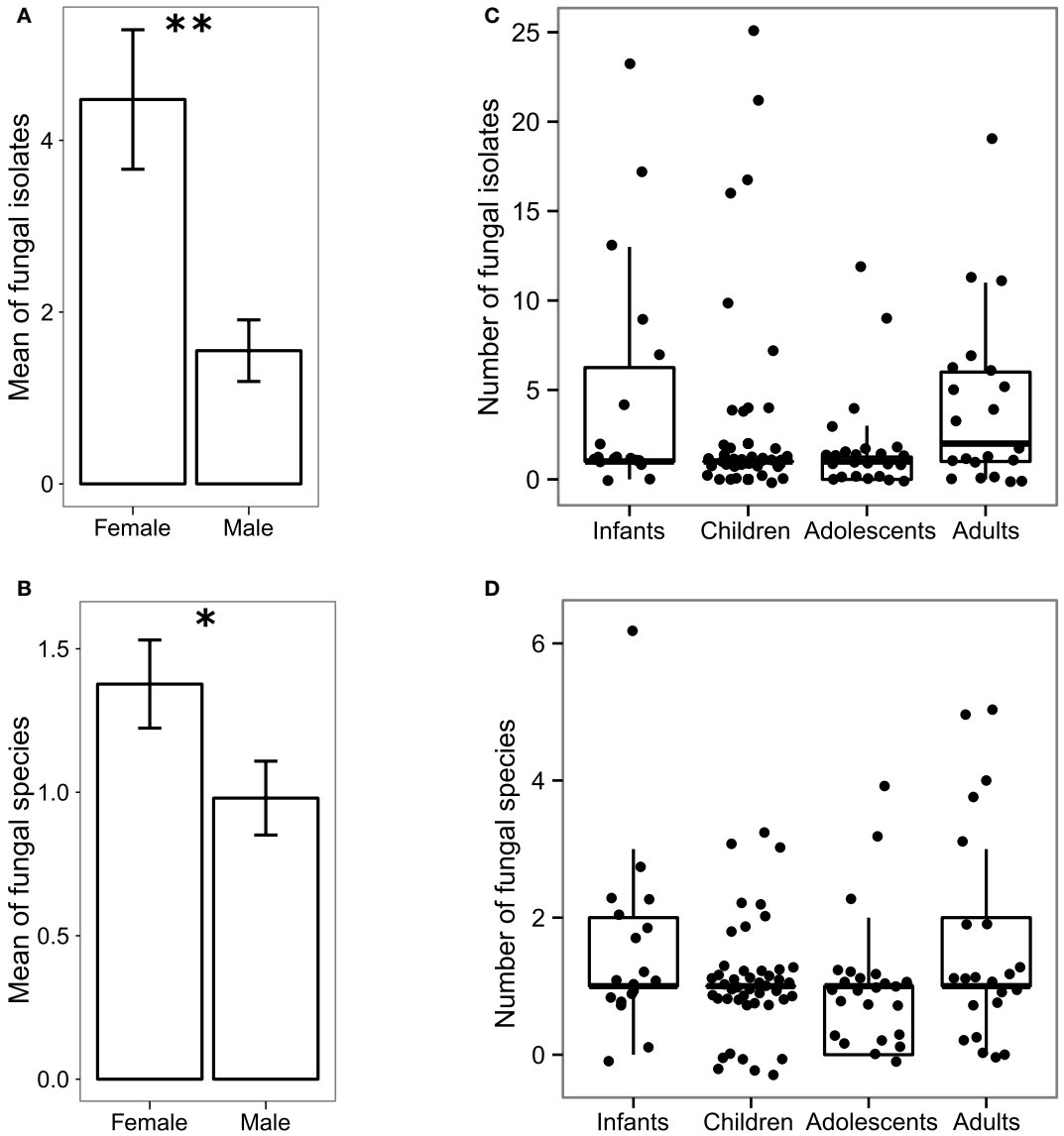


FIGURE 1 | Gender-related and age-related differences in the gut mycobiota of 111 healthy volunteers. Histogram of the mean of **(A)** abundances and **(B)** the richness ± standard error of fungal isolates in female and male subjects; box-plot representation of the **(C)** abundance and **(D)** richness of fungal isolates in different age groups i.e., infants (0–2 years old), children (3–10 years old), adolescents (11–17 years old) and adults (≥18 years old). $^{**}p < 0.005$, $^{*}p < 0.05$, Wilcoxon rank-sum test.

groups (14/22 male children and 3/17 male adults). Furthermore, the genus *Penicillium* was significantly more abundant in infants than adults ($p < 0.05$, Welch *t*-test). Interestingly, we identified sequences belonging to the single-cell protozoa *Blastocystis*, eukaryotes abundant in the human gut microbiota (Scanlan and Marchesi, 2008), only in adolescent and adult females (**Figure 3**, **Supplementary Table S4**) that could potentially be due to exposure to animals (Scanlan et al., 2014).

Phenotyping the Gut Mycobiota

The characterization of phenotypic features of the isolates related to the ability to survive and colonize the human gut was performed to estimate if such isolates were commensals adapted to this ecological niche or passengers introduced through the diet and delivered with the feces.

We therefore investigated the isolates' resistance by a series of assays mimicking the conditions that fungal isolates face during passage through the human GI tract. In addition to the fact that the human body temperature (37°C) is higher than the optimum for most fungal species, in the GI tract fungi are also exposed to acidic and oxidative environments and to bile salts, produced by the liver and secreted into the duodenum, exposing the microorganisms to oxidative stress and DNA damage (Kandell and Bernstein, 1991).

The majority of the isolates were found to tolerate acidic conditions (58.9 and 94.8% of isolates were able to grow at pH

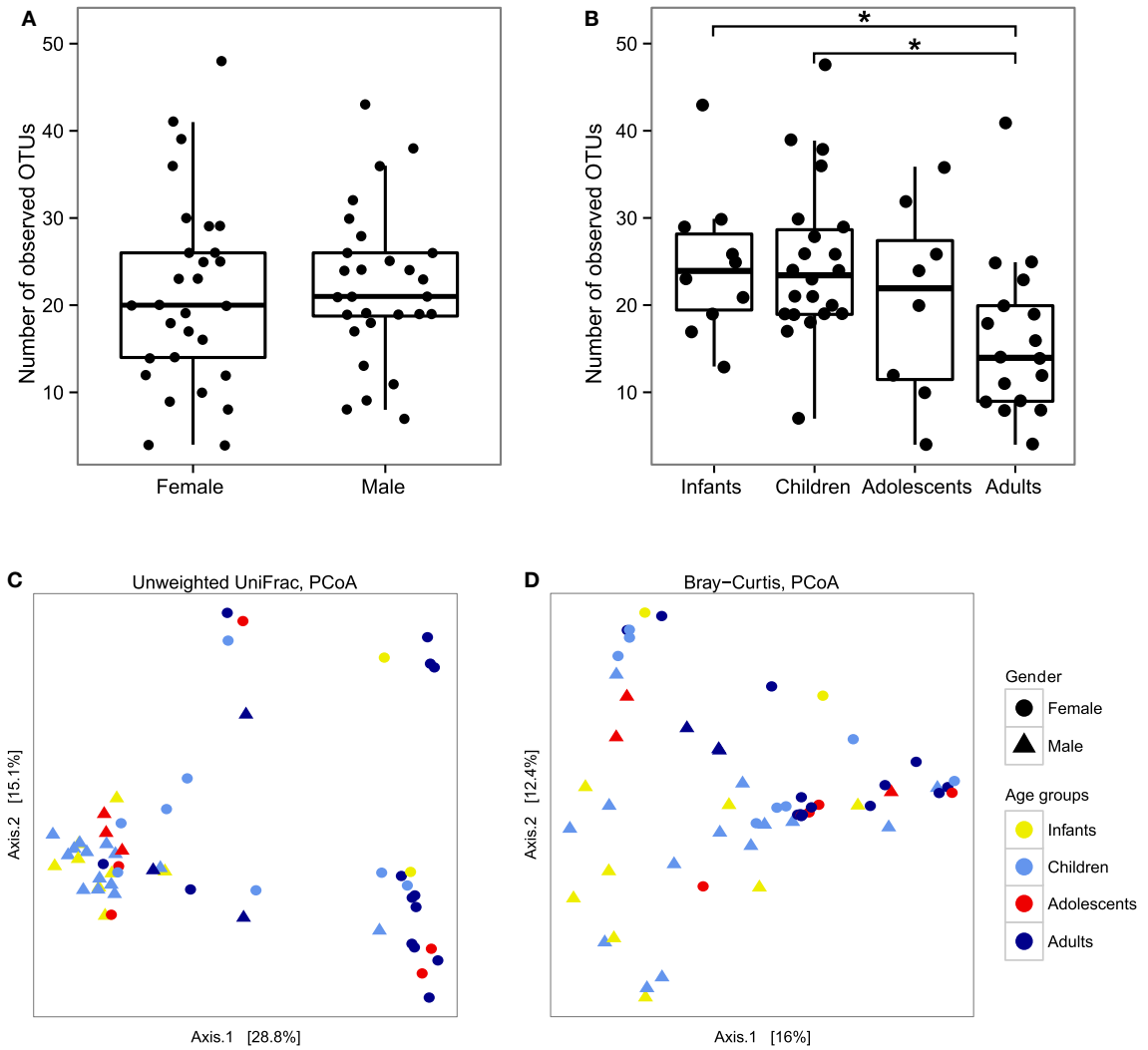


FIGURE 2 | Box-plot representation of fungal *alpha*-diversity measures using the number of observed OTUs between (A) genders or (B) age groups and measures of fungal *beta*-diversity by PCoA of the between samples distances measured using (C) the unweighted UniFrac distance and (D) the Bray-Curtis dissimilarity. *$p < 0.05$, Wilcoxon rank-sum test.

2 and pH 3, respectively) and oxidative stress (85.7% of the isolates showed higher tolerance compared to environmental M28 *S. cerevisiae* strain), both conditions are characteristic of the gut environment. Tolerance to physiological concentrations of bile acids was also observed (89.8, 87.5, and 85.7% of fungal isolates were able to grow in presence of ox-bile 0.5, 1, and 2%, respectively) as well as the ability to grow at *supra* optimal temperatures with almost all the isolates (99.4%) being able to grow at 37°C (**Supplementary Table S2**). The comparison of the growth ability of such isolates at pH 3 and at growing concentrations of ox-bile (i.e., 0.5, 1.0, and 2.0% ox-bile) with respect to the control growth condition (37°C, no bile, pH 6.5) revealed that these stressful conditions do not significantly affect the growth ability of the fungal isolates (**Figure 4**). By contrast, a significant growth reduction was observed when comparing the isolated grown at pH 2 with respect to the control growth condition ($p < 0.0001$, Wilcoxon rank-sum test; **Figure 4**). As expected, a progressive reduction of growth ability was observed in correspondence of incubation temperature increase (i.e., from 40 to 46°C) for all the tested isolates ($p < 0.0005$, Wilcoxon rank-sum test; **Figure 4**).

In addition to the ability of fungal isolates to tolerate the intestinal environmental stresses, we also explored their ability to undergo phenotypic changes favoring their persistence within the human gut. Among these, we assessed the formation of hyphae and the ability to penetrate the solid growth medium, thus to adhere to host tissues. The 56.9% of fungal isolates was able to form hyphae or pseudohyphae (**Supplementary Table S2**). In addition, the morphotype switch to hyphae and pseudohyphae was related to the isolates' invasiveness, with hyphae and pseudohyphae-forming isolates being the most invasive (**Figure 5A**), suggesting that such isolates may be able to

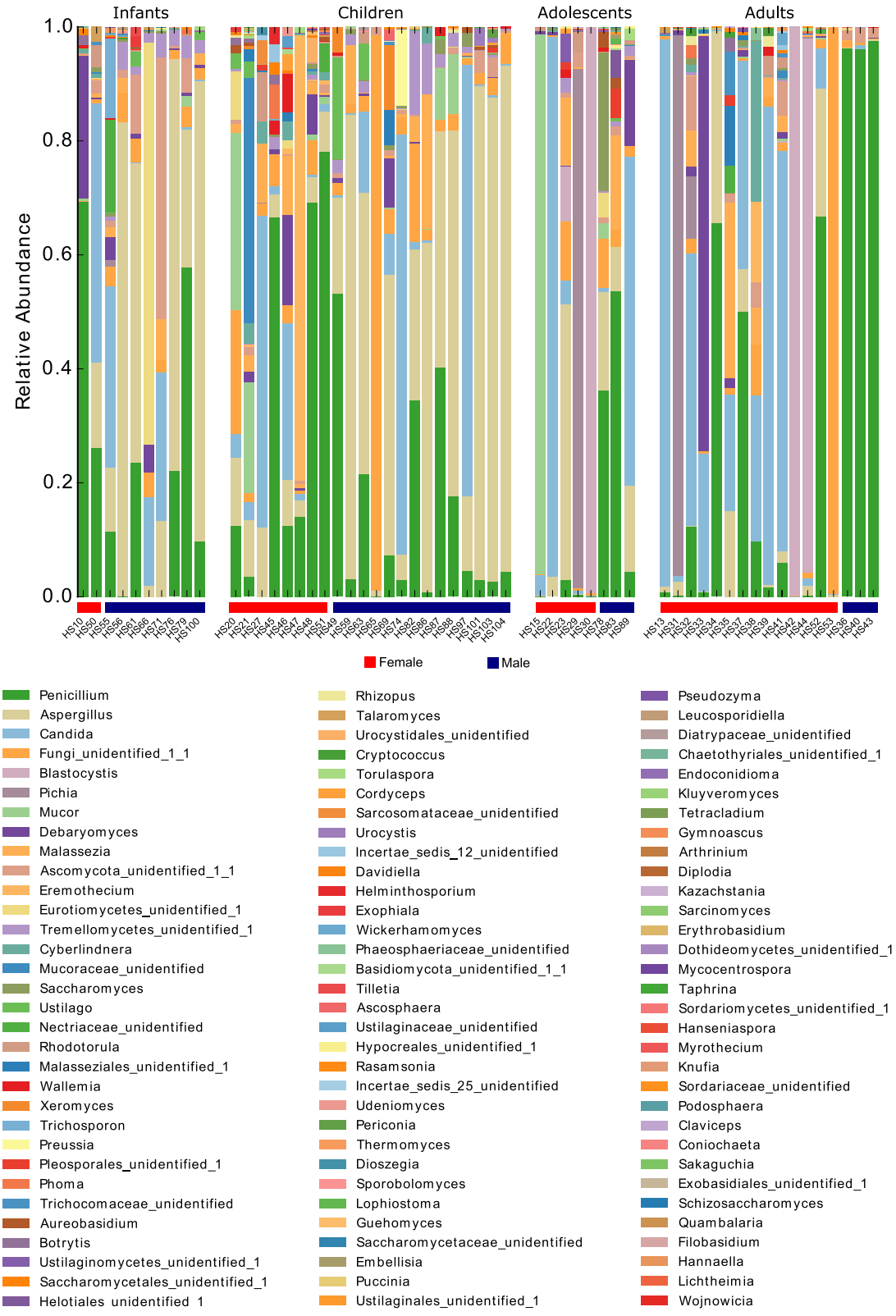
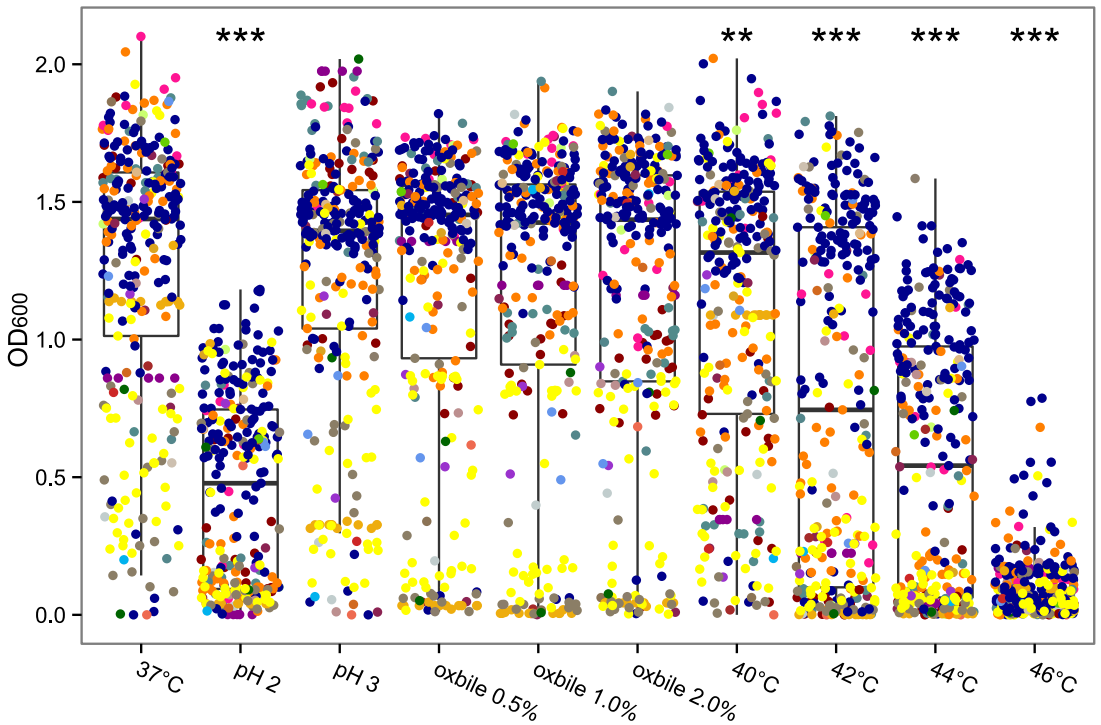


FIGURE 3 | Stacked bar-plot representation of the relative abundances at the genus level of the fecal mycobiota of healthy subjects from metagenomics analysis distributed according to individuals' life stage and gender.



Species

Aspergillus cristatus	*Lichtheimia ramosa*
Aspergillus glaucus	*Mucor circinelloides*
Aspergillus pseudoglaucus	*Penicillium brevicompactum*
Candida albicans	*Penicillium paneum*
Candida deformans	*Pennicillium crustosum*
Candida fermentati	*Pichia fermentans*
Candida intermedia	*Pichia kluyveri*
Candida lusitaniae	*Pichia manshurica*
Candida metapsilosis	*Pleurostomophora richardsiae*
Candida parapsilosis	*Rhodosporidium kratochvilovae*
Candida pararugosa	*Rhodotorula mucilaginosa*
Cryptococcus saitoi	*Saccharomyces cerevisiae*
Eurotium amstelodami	*Torulaspora delbrueckii*
Eurotium rubrum	*Trichosporon asahii*

FIGURE 4 | Box-plot representation of the comparison of fungal isolates growth ability at 37°C (control condition) vs. different stressful conditions mimicking the gastrointestinal tract challenges. $^{**}p < 0.0005$, $^{***}p < 0.0001$, Wilcoxon rank-sum test.

adhere to or invade the host tissues. Furthermore, we observed that hyphae-forming isolates are significantly more resistant to itraconazole than pseudohyphae-forming isolates and isolates unable to form hyphae ($p < 0.05$, Wilcoxon rank-sum test; **Figure 5B**). These phenotypic traits in conditions of altered immune system or in association with intestinal dysbiosis, could represent a pathogenic potential for the host.

It is now recognized that inappropriate antifungal use contributes to the increase in microbial antifungal resistance, complicating therapeutic intervention, and the eventual

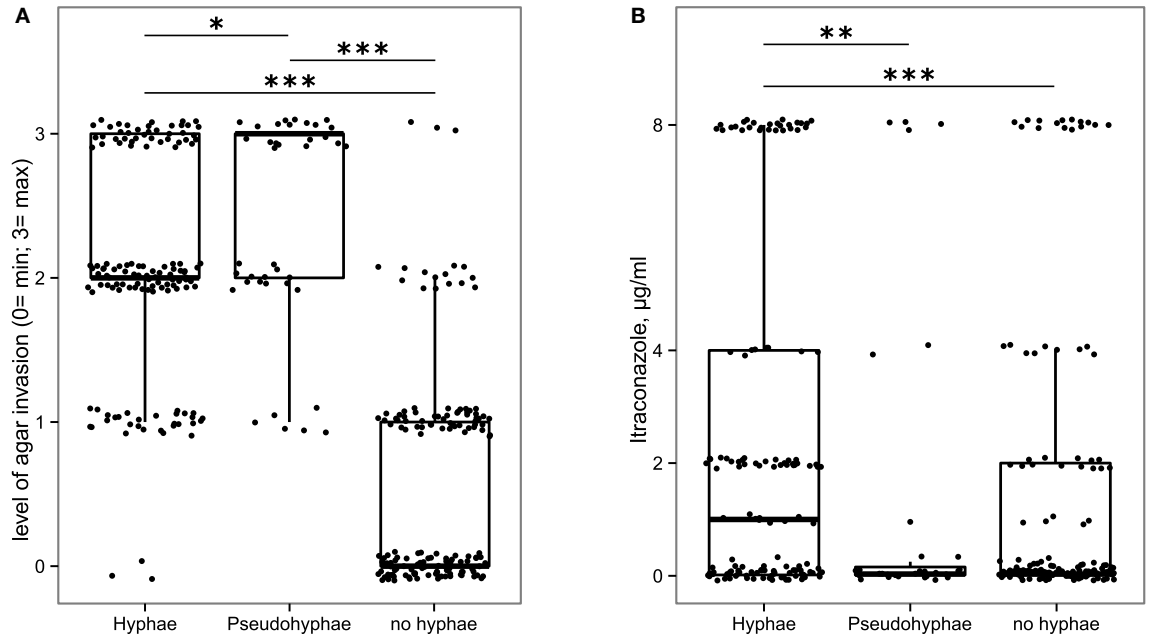


FIGURE 5 | Box-plot representation of fungal isolates able (or not) to produce hyphae or pseudohyphae in relationship with (A) their ability to be invasive on YPD solid medium, (B) their resistance to itraconazole. *$p < 0.05$, **$p < 0.005$, ***$p < 0.001$, Wilcoxon rank-sum test.

eradication of pathogens (Chen et al., 2010; Arendrup et al., 2011). Due to the relevance of such aspect and its impact on clinical studies, we tested all fungal isolates for their susceptibility to the widely therapeutically used azoles, fluconazole, and itraconazole (Martin, 2000) as well as the non-azole antifungal 5-flucytosine (Vermes et al., 2000). A total of 31.5% of the isolates were resistant to fluconazole and, as expected, similar levels of itraconazole resistance were found (for 39.2% of the isolates the MIC was $\geq$1 µg/ml; **Supplementary Table S2**). Previous studies have indeed suggested that cross-resistance may occur between fluconazole and other azole compounds (i.e., itraconazole) (Pfaller et al., 2006) and we further confirmed such observations with the finding of a significant positive correlation between the isolates resistance to these two antifungals (Spearman's $r = 0.43$, $p < 0.05$; **Figure 6**). Most of the isolates (99.34%) showed high susceptibility to 5-flucytosine with most MIC values $\leq$0.125 µg/ml (**Supplementary Table S2**). Among the 9 most abundant species (at least 6 isolates per species), *C. albicans*, *Pichia* spp. and *Rhodotorula mucillaginosa* showed the highest resistance to fluconazole, with $MIC_{90} > 64$ µg/ml (**Table 3**). Furthermore, it is worth to note that resistance to tested antifungals is positively correlated with the ability of strains to grow under stressful conditions, such as *supra* optimal temperature, acidic conditions, and bile salts exposure ($p < 0.05$, Spearman's *r* correlation; **Figure 6**).

DISCUSSION

The vast majority of fungal species inhabiting our body are commensals and opportunistic pathogens that could turn into potential threats depending on strain virulence traits and on the status of the host's immune system. In this perspective to discover a pathogenic infection it seems crucial to define exactly which species are normally present in a given body district.

The human GI tract is known to contain variable communities of bacteria but also fungi have an important role in this ecological niche (Underhill and Iliev, 2014). Nevertheless, the phylogenetic characterization of fungal microorganisms and their specific role as part of the GI niche have not yet been studied extensively.

The advent of sanitation and food globalization has reduced the possibility for humans to come across with the richness of fungal species present in traditional fermented foods. Fungal infections are an ever increasing problem either as side effects of antibiotics use, high dose chemotherapy, and of the spread of immunosuppressive diseases. Estimates of global mortality rates suggest that fungi are responsible for more deaths than either tuberculosis or malaria (Brown et al., 2012). Most of this mortality is caused by species belonging to four fungal genera: *Aspergillus*, *Candida*, *Cryptococcus*, and *Pneumocystis* that are rapidly becoming resistant to most antifungal drugs (Brown et al., 2012; Denning and Bromley, 2015). The information on these fungi so far derives from the study of lung infections, while little is known on the gut mycobiota composition and its role in health and disease. The knowledge on the gut mycobiota is currently limited to few studies making it difficult to assess the significance of differences found in the intestinal fungal populations of diseases such as IBDs due to the lack of information on what the healthy mycobiota is. Here we aimed at defining the "healthy" gut mycobiota, showing that the intestinal fungal community of a cohort of Italian healthy volunteers is a variegate ecosystem

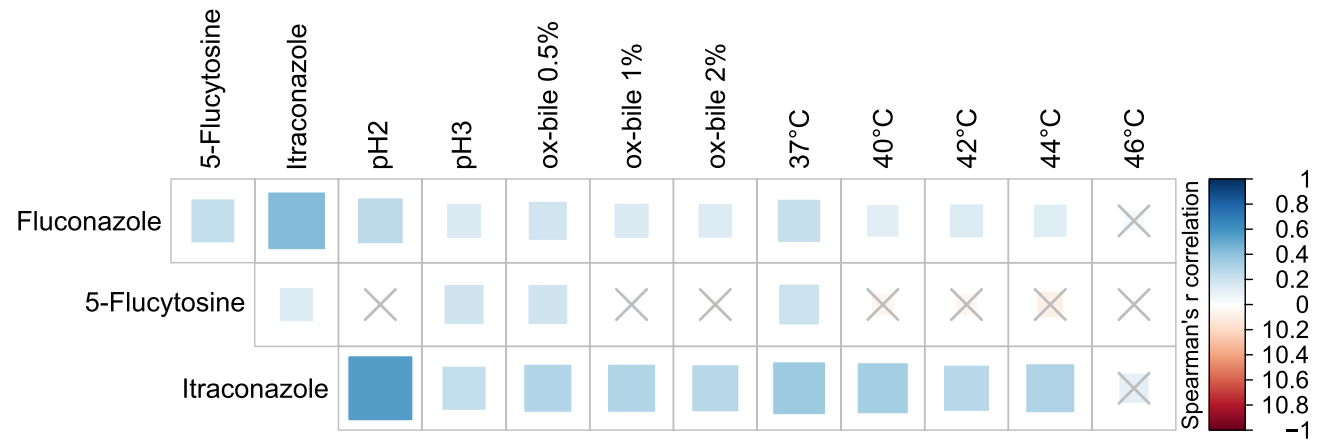


FIGURE 6 | Spearman's *r* correlation analysis between antifungals susceptibility and growth ability of the tested fungal isolates under different stress conditions. Solid squares represent the degree of correlation among the variables taken into account. Crossed squares indicate non-significant correlations; significant results with $p < 0.05$.

that differs in function of individuals' life stage in a gender-dependent manner. We identified 34 fungal species of different ecological origins. While the majority of our fungal isolates has been previously described as inhabitants of the mammalian GI tract (Kurtzman et al., 2011; Rizzetto et al., 2014), some of the isolates belong to species so far identified in environmental samples only. Environmental fungi, in particular putative food-borne fungi, have been previously observed to be able to survive the transition through the GI tract possibly being metabolically active in the gut (David et al., 2014). The phenotypic properties of fungi isolated in this study suggested that these isolates are able to survive in the human GI tract, prompting the hypothesis of an ecological selection and potential ability to colonize this niche (David et al., 2014). Indeed the phenotypic features of the fungal isolates identified endow such isolates with an excellent ecological fitness in the human GI tract. We observed that approximately half of the isolates form hyphae or pseudohyphae, which are known to be involved in the adhesion to or penetration within the GI mucosa (Staab et al., 2013), consolidation of the colony, nutrient intake and formation of 3-dimensional matrices (Brand, 2011). A key factor of *C. albicans* commensalism/pathogenicity is its ability to switch between different morphologies, comprising cellular, pseudohyphae, and hyphae forms. As reported for *C. albicans,* the reversible transition to filamentous growth as a response to environmental cues (Sudbery, 2011) and phenotypic switching is essential for mucosal fungal colonization (Vautier et al., 2015).

Previous studies have also shown that *C. albicans* over-expresses a wide range of genes involved in resistance to high temperature and pH, oxidative stress, and hyphae formation during ileum and colon commensal colonization of BALB/c mice (Pierce et al., 2013). Similarly, the fungal isolates of this study, showing resistance to oxidative, high temperature, bile acids, and pH stresses may hold the potential to colonize the human gut. It is plausible that fecal fungal isolates with specific characteristics (such as high resistance to acidic pH and bile salts) survived to the gut environment, and that these traits make them able to colonize the gut. Thus, we can hypothesize a long process of evolution, selection or adaptation of environmental and food-borne strains to the human host, suggesting that pathogenic strains of commensal species can evolve through a repeated process of evolution and selection, depending on the immune status of the host (De Filippo et al., 2014). These findings encourage for in-depth, strain-level extensive studies on human gut mycobiota and the integration of such data with immunology to further establish the relevance of fungi in host physiology and host-microbe interaction. Furthermore, fungi may train host's immune system simply when passengers, rather than necessarily persisting only as continuous colonizers (Rizzetto et al., 2016).

We discovered that several fungal isolates displayed different levels of antifungal resistance. About 20 years ago, azole-sensitive *C. albicans* dominated infections, with other *Candida* species rarely observed. Actually *C. glabrata* is the second most-commonly isolated *Candida* species in the European Union and United States and has high rates of antifungal resistance (Slavin et al., 2015). Inappropriate antifungal use has contributed to the increase in antifungal resistance, causing objective complications for the treatment of invasive fungal infections that nowadays represent a severe cause of morbidity and mortality among immunocompromised individuals, neonates and elderly (Brown et al., 2012). Recent studies indicated that fungal infections may originate from individual's own commensal strains suggesting that the ability of a commensal microorganism to promote disease is not merely a consequence of impaired host immunity (Odds et al., 2006), suggesting that rural and other commercial uses of azole could be the culprit for the emergence of these resistant strains (Snelders et al., 2012). This underlines the risk that the increase of antifungal usage outside of the clinic could also lead to increased resistance to antifungals of individual's own commensal strains representing an important epidemiological problem in the future and remarking the importance to increase the investment in antifungal research.

TABLE 3 | Antifungal activity against the most abundant fungal species.

#Species (Number of tested)	Antifungal	MIC (μg/ml)		*CBPs		
		MIC_{50}	MIC_{90}	%S	%SDD	%R
Candida albicans (123)	Fluconazole	0.5	>64	65.6	0.8	33.4
	Itraconazole	2	>8	29.3	5.7	65
	5-Flucytosine	0.125	0.5	98.4	0.8	0.8
Candida lusitaniae (6)	Fluconazole	0.125	0.5	100	0	0
	Itraconazole	0.0156	0.125	100	0	0
	5-Flucytosine	0.125	0.125	100	0	0
Candida parapsilosis (40)	Fluconazole	0.5	2	92.5	0	7.5
	Itraconazole	0.031	>8	75	5	20
	5-Flucytosine	0.125	0.125	100	0	0
Penicillium brevicompactum (13)*	Fluconazole	0.125	0.125	100	0	0
	Itraconazole	0.0156	0.0156	92.5	0	7.5
	5-Flucytosine	0.125	0.125	100	0	0
Pichia fermentans (15)*	Fluconazole	32	>64	15.4	0	84.6
	Itraconazole	0.25	4	44.7	20	33.3
	5-Flucytosine	0.5	2	92.3	7.7	0
Pichia kluyveri (9)*	Fluconazole	32	32	11.1	0	88.9
	Itraconazole	0.125	0.125	88.9	11.1	0
	5-Flucytosine	0.5	0.5	100	0	0
Pichia manshurica (9)*	Fluconazole	0.25	>64	77.8	0	22.2
	Itraconazole	0.0156	>8	77.8	0	22.2
	5-Flucytosine	0.125	8	77.8	11.1	11.1
Rhodotorula mucilaginosa (41)*	Fluconazole	0.5	>64	63.4	0	36.6
	Itraconazole	0.0156	2	75.6	2.4	22
	5-Flucytosine	0.125	0.125	100	0	0
Torulaspora delbrueckii (23)*	Fluconazole	0.125	8	87	0	13
	Itraconazole	0.0156	2	69.6	4.3	26.1
	5-Flucytosine	0.125	0.125	100	0	0

**species-specific CBPs are available only for Candida and Aspergillus spp.; for those non-Candida and non-Asperigillus isolates Candida and Asperigillus' CBPs have been used as a proxy; #MIC_{50}, MIC_{90}, and CBPs have been calculated only for those species with number of isolates >5; S, sensible; SDD, Sensibility Dose-Dependent or Intermediate; R, resistant. MIC ranges: Fluconazole 0.125–64 μg/ml; Itraconazole 0.0156–8 μg/ml; 5-Flucytosine 0.125–64 μg/ml.*

It should be noted that all the samples analyzed by metagenomics resulted in high quality fungal sequences, indicating that all the fecal samples studied had fungal DNA. So far, the estimated ratio fungi/bacteria of 1:10000 (Huffnagle and Noverr, 2013), discourages an approach based on whole metagenome shotgun sequencing (Underhill and Iliev, 2014). We thus performed amplicon-based ITS1 metagenomics on a subset of healthy donors identifying more than 90 different fungal taxa. The first striking evidence was that metagenomics detected also sequences belonging to Agaricomycetes, among which several edible fungi, thus suggesting that dietary fungal intake is a potential confounding effect when studying the gut mycobiota. On the contrary 34 different fungal species were isolated using the culture-based approach. Both methods detected in any case differences in the diverse groups of study (**Supplementary Figure S1**). The discrepancies observed between culture-dependent and culture-independent approaches on the description of fungal populations could be attributed to the methodological differences of the two procedures applied suggesting that several of the fungal taxa identified by the metagenomics approach are not cultivable, either because we lack the proper culture conditions or because these belong to DNA from dead cells, environmental or food-borne fungi that cannot survive the passage through the GI tract, but whose DNA is still detectable. Furthermore, the DNA extraction method used in this study could not be suited to extract all the fungal DNA from the stool samples since the rare taxa *Yarrowia*, *Starmerella*, *Rhodosporidium*, and *Pleurostomophora* have been found only by the culture-based approach. On the other hand the culture condition that we used might be responsible for some of the discrepancies observed between the two methods. In our experience most of the commensal fungi commonly found in the human gut can be cultivated in YPD, yet other fungi that we were not able to cultivate might need different culture conditions from those we used in this work.

Although, for example, *S. cerevisiae* is often found in fermented food, it has been shown that it can survive GI tract challenges being a commensal of the human GI tract (Rizzetto

et al., 2014) educating also adaptive immunity (Rizzetto et al., 2016). *S. cerevisiae* has been introduced in the human intestine through diet and fermented beverages and it has accompanied human evolution for at least the past 5150 years (Cavalieri et al., 2003). Our evidence, together with previous results, including a recent description of *S. cerevisiae* in IBDs (Sokol et al., 2016) showed that this microorganism is a potential commensal of the human intestine. The overall reduction of the amount and diversity of fungi introduced through consumption of fermented beverages suggests that the human gut mycobiota could be in dynamic change and certain potentially beneficial species could be lost as a result of modern food processing procedures, cultural changes, and food globalization. Ongoing studies on microbial anthropology in human populations consuming traditional fermented foods, hold the promise to shed light on the evolution of the fungal microbiota as associated to the evolution of diet. On the contrary the edible fungi belonging to Agaricomycetes cannot settle in the human gut due to their ecology (Hibbett, 2006) so we filtered-out these sequences for downstream analyses to reduce statistical noises on ecological measures, improving our results on the characterization of intestinal fungal communities. We are aware that other taxa identified by our analyses having environmental and food-borne origin may not be able to settle in the human gut, however little is known about these taxa while the Agaricomycetes sequences that we retrieved had a very low prevalence in the dataset and mostly belonged to edible fungi such as *Boletus, Suillus or Agrocybe*.

We further observed that amplicon-based ITS1 metagenomics cannot confidently describe fungal populations at a deeper level than genus overlooking species level information provided by the fungal cultivation approach (see **Figures 1B**, **2A**). On the other hand, metagenomics analysis detected community structure differences that fungal cultivation did not identified (see **Figures 2B–D**). Nevertheless, the analysis of *alpha*-diversity from cultivation data on the subset of subjects used for the metagenomics analysis revealed no significant differences among genders remarking that the different sample sizes used in this work are an additional factor in the discrepancies observed between the two methods. Although the major limitation of culture-based methods for the study of microbial communities is the loss of ecological information due to the inability to cultivate most microorganisms by standard culturing techniques, fungi included, culture-based analysis of the human gut mycobiota is fundamental to discern fungal phenotypes that would be otherwise lost by metagenomics.

However, population level analyses with both approaches revealed interesting cues. As occurs for the bacterial microbiota, the intestinal mycobiota is shaped by host's age, gender, diet, and geographical environment (Yatsunenko et al., 2012; Hoffmann et al., 2013; David et al., 2014). Previous studies have shown that the development of the gut bacterial microbiota starts at birth with colonization by a low number of species from the vaginal and fecal microbiota of the mother and is characterized by many shifts in composition during infancy (Yatsunenko et al., 2012). Similarly, the mycobiota may show the same fate, but we observed an inverted trend in which the richness of the gut mycobiota of infants (0–2 years old) and children (3–10 years old) was higher than adults (≥18 years old). It has been shown that suppression of the bacterial microbiota upon treatment with antibiotics results in the outgrowth of the gut mycobiota (Dollive et al., 2013) probably as a consequence of reduced ecological competition. Similarly, a weak bacterial competition, in particular during infancy when the bacterial microbiota is less stable (Koenig et al., 2011; Lozupone et al., 2012), could be the reason why we observed an increased fungal *alpha*-diversity during the early stages of life or this could be due to the different interactions between intestinal fungi and diet (Hoffmann et al., 2013; David et al., 2014) which is peculiar during infancy. We also found that female subjects had a higher number of fungal isolates and different fungal species compared to male subjects and that female mycobiota cluster apart from male mycobiota. This may be ascribed to the role of sex hormones in modulating microbiota composition (Markle et al., 2013) and of diet in shifting the microbiota composition in a gender-dependent manner (Bolnick et al., 2014). Furthermore, the higher relative abundance of *Candida* in the fecal samples from female than male subjects could be also attributed to the prevalence of *Candida* species in the vaginal mycobiota (Drell et al., 2013) due to the anatomical proximity of the two districts. To the best of our knowledge, this is the first time that gender-related differences are described in the human gut mycobiota.

In conclusion we can state that culture-independent approaches are very promising for future investigation of the mycobiota, but yet require significant improvements in the selection of markers for amplicon-based metagenomics and the reference databases. Additionally development of markers targeting pathogenicity traits, including the genes involved in host invasion or evasion of immune defenses, or markers detecting resistance to azoles or other antifungals, is required to thoughtfully apply metagenomics to fungal infections, discriminating the healthy mycobiota from an altered one. Such improvement can be achieved only through systematic sequencing efforts of the cultivable mycobiota, paralleling what happened for the prokaryotic microbiota. In our experience, currently, the combination of the two methods compensated the methodological limits intrinsic in both approaches avoiding to overlook significant differences present in the gut mycobiota of healthy subjects.

AVAILABILITY OF SUPPORTING DATA

Raw sequences are available in the European Nucleotide Archive (ENA) with accession number PRJEB11827 (http://www.ebi.ac.uk/ena/data/view/PRJEB11827).

AUTHOR CONTRIBUTIONS

FS designed and performed the experiments, analyzed the data, and wrote the manuscript. IS and MD performed the experiments. IS, DA, and CD supervised and contributed to data analysis. PL and AC recruited subjects and collected specimens.

IS, MD, LR, OJ, and CD critically reviewed the manuscript. DC and CDF conceived the study and approved the manuscript.

FUNDING

This work was supported by the Accordo di Programma Integrato "MetaFoodLabs" funded by the research office of the Provincia Autonoma di Trento (Italy; PAT Prot. S216/2012/537723).

ACKNOWLEDGMENTS

The authors would like to thank Daniela Renzi, Maddalena Sordo, Massimo Pindo, and the lab staff of the Sequencing Platform Daniela Nicolini, Erika Stefani, and Simone Larger for technical support. Furthermore, a special thanks to Unifarm S.p.A. for supporting FS.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2016.01227

Supplementary Table S1 | Correspondences between deposited metagenomics data and samples.

Supplementary Table S2 | Phenotypic characteristics and antifungals susceptibility of fungal isolates. *calculated as the deviation of the inhibition halo diameter (Ø) from that of the M28-4D *S. cerevisiae* strain, according to the following formula: (Ø sample−Ø M284D strain)/Ø M284D strain *100. #0, non-invasive; 1, poor invasive; 2, invasive; 3, very invasive. −, no growth as measured by $OD_{630} \leq 0.2$ or cfu/ml $\leq 10^5$; +, poor growth as measured by $0.2 < OD_{630} \leq 0.7$ or $10^5 <$ cfu/ml $\leq 10^6$; ++ good growth as measured by $0.7 < OD_{630} \leq 1.2$ or $10^6 <$ cfu/ml $\leq 10^7$; +++, very good growth as measured by $OD_{630} > 1.2$ or cfu/ml $> 10^7$. na, not applicable; nd, not detected.

Supplementary Table S3 | Permutational multivariate analysis of variance (PERMANOVA) tests on unweighted and weighted UniFrac distances and Bray-Curtis dissimilarity.

Supplementary Table S4 | Mean relative abundance (%) of OTUs at the genus level of fungal gut microbiota of healthy subjects from metagenomics analysis.

Supplementary Figure S1 | Mean relative abundances of the gut mycobiota in the different groups of study measured according to (A) the culture-based approach (at species level) and (B) the amplicon-based ITS1 metagenomics approach (at genus level). In panel **(B)** are shown the most abundant genera (with relative abundances >0.1%) while all the other less abundant genera were grouped together and labeled as "*others*".

REFERENCES

Alastruey-Izquierdo, A., Hoffmann, K., De Hoog, G. S., Rodriguez-Tudela, J. L., Voigt, K., Bibashi, E., et al. (2010). Species recognition and clinical relevance of the zygomycetous genus Lichtheimia (syn. Absidia pro parte, Mycocladus). *J. Clin. Microbiol.* 48, 2154–2170. doi: 10.1128/JCM.01744-09

Albanese, D., Fontana, P., De Filippo, C., Cavalieri, D., and Donati, C. (2015). MICCA: a complete and accurate software for taxonomic profiling of metagenomic data. *Sci. Rep.* 5:9743. doi: 10.1038/srep09743

Araujo, R., Pina-Vaz, C., and Rodrigues, A. G. (2007). Susceptibility of environmental versus clinical strains of pathogenic Aspergillus. *Int. J. Antimicrob. Agents* 29, 108–111. doi: 10.1016/j.ijantimicag.2006.09.019

Arendrup, M. C., Sulim, S., Holm, A., Nielsen, L., Nielsen, S. D., Knudsen, J. D., et al. (2011). Diagnostic issues, clinical characteristics, and outcomes for patients with fungemia. *J. Clin. Microbiol.* 49, 3300–3308. doi: 10.1128/JCM.00179-11

Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D. R., et al. (2011). Enterotypes of the human gut microbiome. *Nature* 473, 174–180. doi: 10.1038/nature09944

Barata, A., Malfeito-Ferreira, M., and Loureiro, V. (2012). The microbial ecology of wine grape berries. *Int. J. Food Microbiol.* 153, 243–259. doi: 10.1016/j.ijfoodmicro.2011.11.025

Belda, I., Navascues, E., Marquina, D., Santos, A., Calderon, F., and Benito, S. (2015). Dynamic analysis of physiological properties of Torulaspora delbrueckii in wine fermentations and its incidence on wine quality. *Appl. Microbiol. Biotechnol.* 99, 1911–1922. doi: 10.1007/s00253-014-6197-2

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Stat. Soc. Ser. B* 57, 289–300.

Bezerra-Bussoli, C., Baffi, M. A., Gomes, E., and Da-Silva, R. (2013). Yeast diversity isolated from grape musts during spontaneous fermentation from a Brazilian winery. *Curr. Microbiol.* 67, 356–361. doi: 10.1007/s00284-013-0375-9

Bolnick, D. I., Snowberg, L. K., Hirsch, P. E., Lauber, C. L., Org, E., Parks, B., et al. (2014). Individual diet has sex-dependent effects on vertebrate gut microbiota. *Nat. Commun.* 5, 4500. doi: 10.1038/ncomms5500

Bougnoux, M. E., Diogo, D., Francois, N., Sendid, B., Veirmeire, S., Colombel, J. F., et al. (2006). Multilocus sequence typing reveals intrafamilial transmission and microevolutions of *Candida albicans* isolates from the human digestive tract. *J. Clin. Microbiol.* 44, 1810–1820. doi: 10.1128/JCM.44.5.1810-1820.2006

Brand, A. (2011). Hyphal growth in human fungal pathogens and its role in virulence. *Int. J. Microbiol.* 2012:517529. doi: 10.1155/2012/517529

Brown, G. D., Denning, D. W., Gow, N. A., Levitz, S. M., Netea, M. G., and White, T. C. (2012). Hidden killers: human fungal infections. *Sci. Transl. Med.* 4, 165rv113. doi: 10.1126/scitranslmed.3004404

Butinar, L., Zalar, P., Frisvad, J. C., and Gunde-Cimerman, N. (2005). The genus Eurotium - members of indigenous fungal community in hypersaline waters of salterns. *FEMS Microbiol. Ecol.* 51, 155–166. doi: 10.1016/j.femsec.2004.08.002

Castanheira, M., Messer, S. A., Jones, R. N., Farrell, D. J., and Pfaller, M. A. (2014). Activity of echinocandins and triazoles against a contemporary (2012) worldwide collection of yeast and moulds collected from invasive infections. *Int. J. Antimicrob. Agents* 44, 320–326. doi: 10.1016/j.ijantimicag.2014.06.007

Cavalieri, D., McGovern, P. E., Hartl, D. L., Mortimer, R., and Polsinelli, M. (2003). Evidence for *S.* cerevisiae fermentation in ancient wine. *J. Mol. Evol.* 57(Suppl. 1), S226–S232. doi: 10.1007/s00239-003-0031-2

Cavalieri, D., Townsend, J. P., and Hartl, D. L. (2000). Manifold anomalies in gene expression in a vineyard isolate of *Saccharomyces cerevisiae* revealed by DNA microarray analysis. *Proc. Natl. Acad. Sci. U.S.A.* 97, 12369–12374. doi: 10.1073/pnas.210395297

Chen, S. C., Playford, E. G., and Sorrell, T. C. (2010). Antifungal therapy in invasive fungal infections. *Curr. Opin. Pharmacol.* 10, 522–530. doi: 10.1016/j.coph.2010.06.002

Chitarra, G. S., Abee, T., Rombouts, F. M., Posthumus, M. A., and Dijksterhuis, J. (2004). Germination of *Penicillium paneum* Conidia is regulated by 1-octen-3-ol, a volatile self-inhibitor. *Appl. Environ. Microbiol.* 70, 2823–2829. doi: 10.1128/AEM.70.5.2823-2829.2004

David, L. A., Maurice, C. F., Carmody, R. N., Gootenberg, D. B., Button, J. E., Wolfe, B. E., et al. (2014). Diet rapidly and reproducibly alters the human gut microbiome. *Nature* 505, 559–563. doi: 10.1038/nature12820

De Filippo, C., Di Paola, M., Stefanini, I., Rizzetto, L., Berná, L., Ramazzotti, M., et al. (2014). Population genomics of *Saccharomyces cerevisiae* human isolates: passengers, colonizers, invaders. *bioRxiv*. doi: 10.1101/001891

de la Camara, R., Pinilla, I., Munoz, E., Buendia, B., Steegmann, J. L., and Fernandez-Ranada, J. M. (1996). *Penicillium brevicompactum* as the cause of

a necrotic lung ball in an allogeneic bone marrow transplant recipient. *Bone Marrow Transplant.* 18, 1189–1193.

de Melo Pereira, G. V., Soccol, V. T., Pandey, A., Medeiros, A. B., Andrade Lara, J. M., Gollo, A. L., et al. (2014). Isolation, selection and evaluation of yeasts for use in fermentation of coffee beans by the wet process. *Int. J. Food Microbiol.* 188, 60–66. doi: 10.1016/j.ijfoodmicro.2014.07.008

Denning, D. W., and Bromley, M. J. (2015). Infectious Disease. How to bolster the antifungal pipeline. *Science* 347, 1414–1416. doi: 10.1126/science.aaa6097

Dollive, S., Chen, Y. Y., Grunberg, S., Bittinger, K., Hoffmann, C., Vandivier, L., et al. (2013). Fungi of the murine gut: episodic variation and proliferation during antibiotic treatment. *PLoS ONE* 8:e71806. doi: 10.1371/journal.pone.0071806

Drell, T., Lillsaar, T., Tummeleht, L., Simm, J., Aaspollu, A., Vain, E., et al. (2013). Characterization of the vaginal micro- and mycobiome in asymptomatic reproductive-age Estonian women. *PLoS ONE* 8:e54379. doi: 10.1371/journal.pone.0054379

Findley, K., Oh, J., Yang, J., Conlan, S., Deming, C., Meyer, J. A., et al. (2013). Topographic diversity of fungal and bacterial communities in human skin. *Nature* 498, 367–370. doi: 10.1038/nature12171

Gouba, N., Raoult, D., and Drancourt, M. (2014). Eukaryote culturomics of the gut reveals new species. *PLoS ONE* 9:e106994. doi: 10.1371/journal.pone.0106994

Hatoum, R., Labrie, S., and Fliss, I. (2012). Antimicrobial and probiotic properties of yeasts: from fundamental to novel applications. *Front. Microbiol.* 3:421. doi: 10.3389/fmicb.2012.00421

Hibbett, D. S. (2006). A phylogenetic overview of the Agaricomycotina. *Mycologia* 98, 917–925. doi: 10.3852/mycologia.98.6.917

Hoffman, C. S., and Winston, F. (1987). A 10-Minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia Coli*. *Gene* 57, 267–272. doi: 10.1016/0378-1119(87)90131-4

Hoffmann, C., Dollive, S., Grunberg, S., Chen, J., Li, H., Wu, G. D., et al. (2013). Archaea and fungi of the human gut microbiome: correlations with diet and bacterial residents. *PLoS ONE* 8:e66019. doi: 10.1371/journal.pone.0066019

Huffnagle, G. B., and Noverr, M. C. (2013). The emerging world of the fungal microbiome. *Trends Microbiol.* 21, 334–341. doi: 10.1016/j.tim.2013.04.002

Human Microbiome Project Consortium (2012). Structure, function and diversity of the healthy human microbiome. *Nature* 486, 207–214. doi: 10.1038/nature11234

Johnson, E. (2009). Rare and emerging Candida species. *Curr. Fungal Infect. Rep.* 3, 152–159. doi: 10.1007/s12281-009-0020-z

Kandell, R. L., and Bernstein, C. (1991). Bile salt/acid induction of DNA damage in bacterial and mammalian cells: implications for colon cancer. *Nutr. Cancer* 16, 227–238. doi: 10.1080/01635589109514161

Kaygusuz, I., Mulazimoglu, L., Cerikcioglu, N., Toprak, A., Oktay, A., and Korten, V. (2003). An unusual native tricuspid valve endocarditis caused by *Candida colliculosa*. *Clin. Microbiol. Infect.* 9, 319–322. doi: 10.1046/j.1469-0691.2003.00511.x

Khatib, R., Riederer, K. M., Ramanathan, J., and Baran, J. Jr. (2001). Faecal fungal flora in healthy volunteers and inpatients. *Mycoses* 44, 151–156. doi: 10.1046/j.1439-0507.2001.00639.x

Koenig, J. E., Spor, A., Scalfone, N., Fricker, A. D., Stombaugh, J., Knight, R., et al. (2011). Succession of microbial consortia in the developing infant gut microbiome. *Proc. Natl. Acad. Sci. U.S.A.* 108(Suppl. 1), 4578–4585. doi: 10.1073/pnas.1000081107

Koljalg, U., Nilsson, R. H., Abarenkov, K., Tedersoo, L., Taylor, A. F., Bahram, M., et al. (2013). Towards a unified paradigm for sequence-based identification of fungi. *Mol. Ecol.* 22, 5271–5277. doi: 10.1111/mec.12481

Kurtzman, C. P., Fell, J. W., and Boekhout, T., (eds.). (2011). "The yeasts, a taxonomic study volume 1 fifth edition preface," in *Yeasts: A Taxonomic Study,* Vol. 1–3, 5th Edn (Elsevier Science), 13.

Lee, S. C., Billmyre, R. B., Li, A., Carson, S., Sykes, S. M., Huh, E. Y., et al. (2014). Analysis of a food-borne fungal pathogen outbreak: virulence and genome of a *Mucor circinelloide*s isolate from yogurt. *Mbio* 5, e01390–e01314. doi: 10.1128/mBio.01390-14

Levenstadt, J. S., Poutanen, S. M., Mohan, S., Zhang, S., and Silverman, M. (2012). *Pleurostomophora richardsiae* - an insidious fungus presenting in a man 44 years after initial inoculation: a case report and review of the literature. *Can. J. Infect. Dis. Med. Microbiol.* 23, 110–113. doi: 10.1155/2012/406982

Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K., and Knight, R. (2012). Diversity, stability and resilience of the human gut microbiota. *Nature* 489, 220–230. doi: 10.1038/nature11550

Luan, C., Xie, L., Yang, X., Miao, H., Lv, N., Zhang, R., et al. (2015). Dysbiosis of fungal microbiota in the intestinal mucosa of patients with colorectal adenomas. *Sci. Rep.* 5:7980. doi: 10.1038/srep07980

Markle, J. G., Frank, D. N., Mortin-Toth, S., Robertson, C. E., Feazel, L. M., Rolle-Kampczyk, U., et al. (2013). Sex differences in the gut microbiome drive hormone-dependent regulation of autoimmunity. *Science* 339, 1084–1088. doi: 10.1126/science.1233521

Mar Rodriguez, M., Perez, D., Javier Chaves, F., Esteve, E., Marin-Garcia, P., Xifra, G., et al. (2015). Obesity changes the human gut mycobiome. *Sci. Rep.* 5:14600. doi: 10.1038/srep14600

Martin, M. V. (2000). The use of fluconazole and itraconazole in the treatment of *Candida albicans* infections: a review. *J. Antimicrob. Chemother.* 45, 555. doi: 10.1093/oxfordjournals.jac.a020880

McMurdie, P. J., and Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* 8:e61217. doi: 10.1371/journal.pone.0061217

Noriega, L., Gueimonde, M., Sanchez, B., Margolles, A., and De Los Reyes-Gavilan, C. G. (2004). Effect of the adaptation to high bile salts concentrations on glycosidic activity, survival at low PH and cross-resistance to bile salts in Bifidobacterium. *Int. J. Food Microbiol.* 94, 79–86. doi: 10.1016/j.ijfoodmicro.2004.01.003

Notredame, C., Higgins, D. G., and Heringa, J. (2000). T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* 302, 205–217. doi: 10.1006/jmbi.2000.4042

Odds, F. C., Davidson, A. D., Jacobsen, M. D., Tavanti, A., Whyte, J. A., Kibbler, C. C., et al. (2006). *Candida albicans* strain maintenance, replacement, and microvariation demonstrated by multilocus sequence typing. *J. Clin. Microbiol.* 44, 3647–3658. doi: 10.1128/JCM.00934-06

Oever, J. T., and Netea, M. G. (2014). The bacteriome-mycobiome interaction and antifungal host defense. *Eur. J. Immunol.* 44, 3182–3191. doi: 10.1002/eji.201344405

Pfaller, M. A., and Diekema, D. J. (2012). Progress in antifungal susceptibility testing of Candida spp. by use of Clinical and Laboratory Standards Institute broth microdilution methods, 2010 to 2012. *J. Clin. Microbiol.* 50, 2846–2856. doi: 10.1128/JCM.00937-12

Pfaller, M. A., Diekema, D. J., and Sheehan, D. J. (2006). Interpretive breakpoints for fluconazole and Candida revisited: a blueprint for the future of antifungal susceptibility testing. *Clin. Microbiol. Rev.* 19, 435–447. doi: 10.1128/CMR.19.2.435-447.2006

Pierce, J. V., Dignard, D., Whiteway, M., and Kumamoto, C. A. (2013). Normal adaptation of *Candida albicans* to the murine gastrointestinal tract requires Efg1p-dependent regulation of metabolic and host defense genes. *Eukaryot. Cell* 12, 37–49. doi: 10.1128/EC.00236-12

Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2–approximately maximum-likelihood trees for large alignments. *PLoS ONE* 5:e9490. doi: 10.1371/journal.pone.0009490

Revelle, W. (2013). *Psych: Procedures for Psychological, Psychometric, and Personality Research. R Package Version 1.3.10.* Evanston, IL: Northwestern University.

Reyes, A., Haynes, M., Hanson, N., Angly, F. E., Heath, A. C., Rohwer, F., et al. (2010). Viruses in the faecal microbiota of monozygotic twins and their mothers. *Nature* 466, 334–338. doi: 10.1038/nature09199

Rizzetto, L., De Filippo, C., and Cavalieri, D. (2014). Richness and diversity of mammalian fungal communities shape innate and adaptive immunity in health and disease. *Eur. J. Immunol.* 44, 3166–3181. doi: 10.1002/eji.201344403

Rizzetto, L., Ifrim, D. C., Moretti, S., Tocci, N., Cheng, S. C., Quintin, J., et al. (2016). Fungal chitin induces trained immunity in human monocytes during cross-talk of the host with *Saccharomyces cerevisiae*. *J. Biol. Chem.* 291, 7961–7972. doi: 10.1074/jbc.M115.699645

Rodriguez-Tudela, J., Arendrup, M., Arikan, S., Barchiesi, F., Bille, J., Chryssanthou, E., et al. (2008a). EUCAST DEFINITIVE DOCUMENT E. DEF 9.1: Method for the determination of broth dilution minimum inhibitory concentrations of antifungal agents for conidia forming moulds. *Clin. Microbiol. Infect.* 14, 982–984.

Rodriguez-Tudela, J., Arendrup, M., Barchiesi, F., Bille, J., Chryssanthou, E., Cuenca-Estrella, M., et al. (2008b). EUCAST Definitive Document EDef 7.1: method for the determination of broth dilution MICs of antifungal agents for fermentative yeasts: subcommittee on Antifungal Susceptibility Testing (AFST) of the ESCMID European Committee for Antimicrobial Susceptibility Testing (EUCAST)∗. *Clin. Microbiol. Infect.* 14, 398–405. doi: 10.1111/j.1469-0691.2007.01935.x

Romani, L. (2011). Immunity to fungal infections. *Nat. Rev. Immunol.* 11, 275–288. doi: 10.1038/nri2939

Santini, A., Mikušová, P., Sulyok, M., Krska, R., Labuda, R., and Šrobárová, A. (2014). Penicillium strains isolated from Slovak grape berries taxonomy assessment by secondary metabolite profile. *Mycotoxin Res.* 30, 213–220. doi: 10.1007/s12550-014-0205-3

Scanlan, P. D., and Marchesi, J. R. (2008). Micro-eukaryotic diversity of the human distal gut microbiota: qualitative assessment using culture-dependent and -independent analysis of faeces. *ISME J.* 2, 1183–1193. doi: 10.1038/ismej.2008.76

Scanlan, P. D., Stensvold, C. R., Rajilic-Stojanovic, M., Heilig, H. G., De Vos, W. M., O'toole, P. W., et al. (2014). The microbial eukaryote Blastocystis is a prevalent and diverse member of the healthy human gut microbiota. *FEMS Microbiol. Ecol.* 90, 326–330. doi: 10.1111/1574-6941.12396

Sebastiani, F., Barberio, C., Casalone, E., Cavalieri, D., and Polsinelli, M. (2002). Crosses between *Saccharomyces cerevisiae* and Saccharomyces bayanus generate fertile hybrids. *Res. Microbiol.* 153, 53–58. doi: 10.1016/S0923-2508(01)01286-4

Slavin, M., Van Hal, S., Sorrell, T. C., Lee, A., Marriott, D. J., Daveson, K., et al. (2015). Invasive infections due to filamentous fungi other than Aspergillus: epidemiology and determinants of mortality. *Clin. Microbiol. Infect.* 21, e491–490. doi: 10.1016/j.cmi.2014.12.021

Snelders, E., Camps, S. M., Karawajczyk, A., Schaftenaar, G., Kema, G. H., Van Der Lee, H. A., et al. (2012). Triazole fungicides can induce cross-resistance to medical triazoles in Aspergillus fumigatus. *PLoS ONE* 7:e31801. doi: 10.1371/journal.pone.0031801

Sokol, H., Leducq, V., Aschard, H., Pham, H. P., Jegou, S., Landman, C., et al. (2016). Fungal microbiota dysbiosis in IBD. *Gut.* doi: 10.1136/gutjnl-2015-310746. [Epub ahead of print].

Staab, J. F., Datta, K., and Rhee, P. (2013). Niche-specific requirement for hyphal wall protein 1 in virulence of *Candida albicans*. *PLoS ONE* 8:e80842. doi: 10.1371/journal.pone.0080842

Sudbery, P. E. (2011). Growth of *Candida albicans* hyphae. *Nat. Rev. Microbiol.* 9, 737–748. doi: 10.1038/nrmicro2636

Team, R. C. (2014). *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing*. Vienna: Austria. 2012. ISBN 3-900 051-07-0.

Tristezza, M., Vetrano, C., Bleve, G., Spano, G., Capozzi, V., Logrieco, A., et al. (2013). Biodiversity and safety aspects of yeast strains characterized from vineyards and spontaneous fermentations in the Apulia Region, Italy. *Food Microbiol.* 36, 335–342. doi: 10.1016/j.fm.2013.07.001

Underhill, D. M., and Iliev, I. D. (2014). The mycobiota: interactions between commensal fungi and the host immune system. *Nat. Rev. Immunol.* 14, 405–416. doi: 10.1038/nri3684

Vardjan, T., Mohar Lorbeg, P., Rogelj, I., and Canzek Majhenic, A. (2013). Characterization and stability of lactobacilli and yeast microbiota in kefir grains. *J. Dairy Sci.* 96, 2729–2736. doi: 10.3168/jds.2012-5829

Vautier, S., Drummond, R. A., Chen, K., Murray, G. I., Kadosh, D., Brown, A. J., et al. (2015). *Candida albicans* colonization and dissemination from the murine gastrointestinal tract: the influence of morphology and Th17 immunity. *Cell. Microbiol.* 17, 445–450. doi: 10.1111/cmi.12388

Vermes, A., Guchelaar, H. J., and Dankert, J. (2000). Flucytosine: a review of its pharmacology, clinical indications, pharmacokinetics, toxicity and drug interactions. *J. Antimicrob. Chemother.* 46, 171–179. doi: 10.1093/jac/46.2.171

Vopalenska, I., Hulkova, M., Janderova, B., and Palkova, Z. (2005). The morphology of *Saccharomyces cerevisiae* colonies is affected by cell adhesion and the budding pattern. *Res. Microbiol.* 156, 921–931. doi: 10.1016/j.resmic.2005.05.012

Wang, C., Esteve-Zarzoso, B., and Mas, A. (2014). Monitoring of *Saccharomyces cerevisiae*, Hanseniaspora uvarum, and *Starmerella bacillaris* (synonym Candida zemplinina) populations during alcoholic fermentation by fluorescence *in situ* hybridization. *Int. J. Food Microbiol.* 191, 1–9. doi: 10.1016/j.ijfoodmicro.2014.08.014

Wang, Q., Garrity, G. M., Tiedje, J. M., and Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* 73, 5261–5267. doi: 10.1128/AEM.00062-07

Westfall, P. H., and Young, S. S. (1993). *Resampling-Based Multiple Testing: Examples and Methods for p-value Adjustment*. Vol. 279. New York, NY: John Wiley & Sons.

Yatsunenko, T., Rey, F. E., Manary, M. J., Trehan, I., Dominguez-Bello, M. G., Contreras, M., et al. (2012). Human gut microbiome viewed across age and geography. *Nature* 486, 222–227. doi: 10.1038/nature11053

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Supplementary material chapter 1

Supplementary Table 1: Correspondences between deposited metagenomics data and samples.

Supplementary Table 2: Phenotypic characteristics and antifungals susceptibility of fungal isolates. *calculated as the deviation of the inhibition halo diameter (Ø) from that of the M28-4D *S. cerevisiae* strain, according to the following formula: (Ø sample – Ø M284D strain) / Ø M284D strain *100. #, 0= non-invasive; 1= poor invasive; 2= invasive; 3= very invasive. -, no growth as measured by $OD_{630} \leq 0.2$ or cfu/ml $\leq 10^5$; +, poor growth as measured by $0.2 < OD_{630} \leq 0.7$ or $10^5 <$ cfu/ml$\leq 10^6$; ++ good growth as measured by $0.7 < OD_{630} \leq 1.2$ or $10^6 <$ cfu/ml$\leq 10^7$; +++, very good growth as measured by $OD_{630} > 1.2$ or cfu/ml$> 10^7$. na, not applicable; nd, not detected.

Supplementary Table 3: Permutational multivariate analysis of variance (PERMANOVA) tests on unweighted and weighted UniFrac distances and Bray-Curtis dissimilarity.

Supplementary Table 4: Mean relative abundance (%) of OTUs at the genus level of fungal gut microbiota of healthy subjects from metagenomics analysis.

Supplementary Figure 1: Mean relative abundances of the gut mycobiota in the different groups of study measured according to **A)** the culture-based approach (at species level) and **B)** the amplicon-based ITS1 metagenomics approach (at genus level). In panel **B** are shown the most abundant genera (with relative abundances > 0.1%) while all the other less abundant genera were grouped together and labeled as "*others*".

Table S1: Correspondences between deposited metagenomics data and samples

Study_accession	Secondary_accession	Sample_accession	Experiment_accession	Run_accession	Sample_unique_name	File_name	Age	Gender	Age_group
PRJEB11827	SAMEA3670898	ERS978047	ERX1221136	ERR1142295	HS10	healthy_subject_10.fastq	2	Female	Infants
PRJEB11827	SAMEA3670899	ERS978048	ERX1221137	ERR1142296	HS13	healthy_subject_13.fastq	18	Female	Adults
PRJEB11827	SAMEA3670900	ERS978049	ERX1221138	ERR1142297	HS15	healthy_subject_15.fastq	11	Female	Adolescents
PRJEB11827	SAMEA3670901	ERS978050	ERX1221139	ERR1142298	HS20	healthy_subject_20.fastq	4	Female	Children
PRJEB11827	SAMEA3670902	ERS978051	ERX1221140	ERR1142299	HS21	healthy_subject_21.fastq	5	Female	Children
PRJEB11827	SAMEA3670903	ERS978052	ERX1221141	ERR1142300	HS22	healthy_subject_22.fastq	15	Female	Adolescents
PRJEB11827	SAMEA3670904	ERS978053	ERX1221142	ERR1142301	HS23	healthy_subject_23.fastq	11	Female	Adolescents
PRJEB11827	SAMEA3670905	ERS978054	ERX1221143	ERR1142302	HS27	healthy_subject_27.fastq	9	Female	Children
PRJEB11827	SAMEA3670906	ERS978055	ERX1221144	ERR1142303	HS29	healthy_subject_29.fastq	16	Female	Adolescents
PRJEB11827	SAMEA3670907	ERS978056	ERX1221145	ERR1142304	HS30	healthy_subject_30.fastq	12	Female	Adolescents
PRJEB11827	SAMEA3670908	ERS978057	ERX1221146	ERR1142305	HS31	healthy_subject_31.fastq	24	Female	Adults
PRJEB11827	SAMEA3670909	ERS978058	ERX1221147	ERR1142306	HS32	healthy_subject_32.fastq	32	Female	Adults
PRJEB11827	SAMEA3670910	ERS978059	ERX1221148	ERR1142307	HS33	healthy_subject_33.fastq	32	Female	Adults
PRJEB11827	SAMEA3670911	ERS978060	ERX1221149	ERR1142308	HS34	healthy_subject_34.fastq	25	Female	Adults
PRJEB11827	SAMEA3670912	ERS978061	ERX1221150	ERR1142309	HS35	healthy_subject_35.fastq	26	Female	Adults
PRJEB11827	SAMEA3670913	ERS978062	ERX1221151	ERR1142310	HS36	healthy_subject_36.fastq	20	Male	Adults
PRJEB11827	SAMEA3670914	ERS978063	ERX1221152	ERR1142311	HS37	healthy_subject_37.fastq	28	Female	Adults
PRJEB11827	SAMEA3670915	ERS978064	ERX1221153	ERR1142312	HS38	healthy_subject_38.fastq	25	Female	Adults
PRJEB11827	SAMEA3670916	ERS978065	ERX1221154	ERR1142313	HS39	healthy_subject_39.fastq	27	Female	Adults
PRJEB11827	SAMEA3670917	ERS978066	ERX1221155	ERR1142314	HS40	healthy_subject_40.fastq	27	Male	Adults
PRJEB11827	SAMEA3670918	ERS978067	ERX1221156	ERR1142315	HS41	healthy_subject_41.fastq	24	Female	Adults
PRJEB11827	SAMEA3670919	ERS978068	ERX1221157	ERR1142316	HS42	healthy_subject_42.fastq	24	Female	Adults
PRJEB11827	SAMEA3670920	ERS978069	ERX1221158	ERR1142317	HS43	healthy_subject_43.fastq	26	Male	Adults
PRJEB11827	SAMEA3670921	ERS978070	ERX1221159	ERR1142318	HS44	healthy_subject_44.fastq	24	Female	Adults
PRJEB11827	SAMEA3670922	ERS978071	ERX1221160	ERR1142319	HS45	healthy_subject_45.fastq	6	Female	Children
PRJEB11827	SAMEA3670923	ERS978072	ERX1221161	ERR1142320	HS46	healthy_subject_46.fastq	6	Female	Children
PRJEB11827	SAMEA3670924	ERS978073	ERX1221162	ERR1142321	HS47	healthy_subject_47.fastq	10	Female	Children

PRJEB11827	SAMEA3670925	ERS978074	ERX1221163	ERR1142322	HS48	healthy_subject_48.fastq	2.5	Female	Children
PRJEB11827	SAMEA3670926	ERS978075	ERX1221164	ERR1142323	HS49	healthy_subject_49.fastq	2.5	Male	Children
PRJEB11827	SAMEA3670927	ERS978076	ERX1221165	ERR1142324	HS50	healthy_subject_50.fastq	1.5	Female	Infants
PRJEB11827	SAMEA3670928	ERS978077	ERX1221166	ERR1142325	HS51	healthy_subject_51.fastq	8	Female	Children
PRJEB11827	SAMEA3670929	ERS978078	ERX1221167	ERR1142326	HS52	healthy_subject_52.fastq	23	Female	Adults
PRJEB11827	SAMEA3670930	ERS978079	ERX1221168	ERR1142327	HS53	healthy_subject_53.fastq	23	Female	Adults
PRJEB11827	SAMEA3670931	ERS978080	ERX1221169	ERR1142328	HS55	healthy_subject_55.fastq	2	Male	Infants
PRJEB11827	SAMEA3670932	ERS978081	ERX1221170	ERR1142329	HS56	healthy_subject_56.fastq	2	Male	Infants
PRJEB11827	SAMEA3670933	ERS978082	ERX1221171	ERR1142330	HS59	healthy_subject_59.fastq	5	Male	Children
PRJEB11827	SAMEA3670934	ERS978083	ERX1221172	ERR1142331	HS61	healthy_subject_61.fastq	2	Male	Infants
PRJEB11827	SAMEA3670935	ERS978084	ERX1221173	ERR1142332	HS63	healthy_subject_63.fastq	5	Male	Children
PRJEB11827	SAMEA3670936	ERS978085	ERX1221174	ERR1142333	HS65	healthy_subject_65.fastq	6	Male	Children
PRJEB11827	SAMEA3670937	ERS978086	ERX1221175	ERR1142334	HS66	healthy_subject_66.fastq	0.1	Male	Infants
PRJEB11827	SAMEA3670938	ERS978087	ERX1221176	ERR1142335	HS69	healthy_subject_69.fastq	6	Male	Children
PRJEB11827	SAMEA3670939	ERS978088	ERX1221177	ERR1142336	HS71	healthy_subject_71.fastq	1	Male	Infants
PRJEB11827	SAMEA3670940	ERS978089	ERX1221178	ERR1142337	HS74	healthy_subject_74.fastq	6	Male	Children
PRJEB11827	SAMEA3670941	ERS978090	ERX1221179	ERR1142338	HS76	healthy_subject_76.fastq	1	Male	Infants
PRJEB11827	SAMEA3670942	ERS978091	ERX1221180	ERR1142339	HS78	healthy_subject_78.fastq	12	Male	Adolescents
PRJEB11827	SAMEA3670943	ERS978092	ERX1221181	ERR1142340	HS79	healthy_subject_79.fastq	0.1	Male	Infants
PRJEB11827	SAMEA3670944	ERS978093	ERX1221182	ERR1142341	HS82	healthy_subject_82.fastq	10	Male	Children
PRJEB11827	SAMEA3670945	ERS978094	ERX1221183	ERR1142342	HS83	healthy_subject_83.fastq	12	Male	Adolescents
PRJEB11827	SAMEA3670946	ERS978095	ERX1221184	ERR1142343	HS86	healthy_subject_86.fastq	7	Male	Children
PRJEB11827	SAMEA3670947	ERS978096	ERX1221185	ERR1142344	HS87	healthy_subject_87.fastq	9	Male	Children
PRJEB11827	SAMEA3670948	ERS978097	ERX1221186	ERR1142345	HS88	healthy_subject_88.fastq	7	Male	Children
PRJEB11827	SAMEA3670949	ERS978098	ERX1221187	ERR1142346	HS89	healthy_subject_89.fastq	12	Male	Adolescents
PRJEB11827	SAMEA3670950	ERS978099	ERX1221188	ERR1142347	HS97	healthy_subject_97.fastq	6	Male	Children
PRJEB11827	SAMEA3670951	ERS978100	ERX1221189	ERR1142348	HS100	healthy_subject_100.fastq	0.1	Male	Infants
PRJEB11827	SAMEA3670952	ERS978101	ERX1221190	ERR1142349	HS101	healthy_subject_101.fastq	4	Male	Children
PRJEB11827	SAMEA3670953	ERS978102	ERX1221191	ERR1142350	HS103	healthy_subject_103.fastq	7	Male	Children
PRJEB11827	SAMEA3670954	ERS978103	ERX1221192	ERR1142351	HS104	healthy_subject_104.fastq	4	Male	Children

Table S2: Phenotypic characteristics and antifungals susceptibility of fungal isolates.

Isolate ID	Subject	Sex	Age	Fluconazole (MIC µg/ml)	5-Flucytosine (MIC µg/ml)	Itraconazole (MIC µg/ml)	pH2	pH3	oxbile 0.5%	oxbile 1%	oxbile 2%	*Ox stress resistance	# Agar invasivity	Hyphae and pseudohyphae	37°C	40°C	42°C	44°C	46°C	*Species*
YHS1	HS1	M	5	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	*Candida albicans*
YHS2	HS2	M	5	0. 25	0. 125	0. 0156	-	-	-	-	-	-33.3	0	nd	++	+	-	+	-	*Saccharomyces cerevisiae*
YHS3	HS3	M	14	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	*Candida albicans*
YHS4	HS4	M	1	64	0. 125	0. 5	-	++	+++	+++	+++	-26.7	na	na	++	++	-	-	-	*Candida albicans*
YHS5	HS4	M	1	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	*Yarrowia lipolytica*
YHS6	HS7	F	20	> 64	0. 125	> 8	++	+++	+++	+++	+++	-66.7	1	nd	+++	+++	+++	+++	+	*Saccharomyces cerevisiae*
YHS7	HS8	M	5	64	0. 125	1	-	++	+++	+++	+++	-20	0	nd	++	++	+	-	-	*Saccharomyces cerevisiae*
YHS8	HS9	M	14	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	*Candida parapsilosis*
YHS9	HS10	F	2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	*Candida albicans*
YHS10	HS11	M	16	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	*Candida albicans*
YHS11	HS12	M	15	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	*Candida albicans*
YHS12	HS13	F	18	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	*Pichia carribica*
YHS13	HS14	F	0.3	0. 125	0. 25	0. 25	+	+++	++	++	++	-40	0	nd	+++	++	++	-	-	*Pichia fermentans*
YHS14	HS14	F	0.3	0. 125	0. 5	0. 125	+	+++	+++	+++	++	-26.67	0	nd	+++	++	+	-	-	*Pichia kluyveri*
YHS15	HS14	F	0.3	32	0. 5	0. 031	-	+++	+++	+++	++	-26.67	1	nd	+++	+	-	-	-	*Pichia fermentans*
YHS16	HS14	F	0.3	32	1	1	+	++	+++	+++	+++	-33.33	1	nd	+++	+	-	-	-	*Pichia fermentans*
YHS17	HS14	F	0.3	> 64	4	4	+	++	++	++	++	-26.67	0	nd	+++	+	-	-	-	*Pichia fermentans*
YHS18	HS14	F	0.3	> 64	2	4	+	+++	++	++	++	-26.67	1	nd	+++	++	-	-	-	*Pichia fermentans*
YHS19	HS14	F	0.3	32	0. 5	0. 125	+	+++	+++	++	++	-26.67	1	nd	+++	++	-	-	-	*Pichia fermentans*
YHS20	HS14	F	0.3	32	0. 5	0. 25	-	+++	+++	++	++	-26.67	1	nd	+++	++	-	-	-	*Pichia fermentans*
YHS21	HS14	F	0.3	> 64	1	4	+	+++	+++	++	++	-33.33	1	nd	+++	+	-	-	-	*Pichia fermentans*
YHS22	HS14	F	0.3	32	0. 5	0. 25	+	+++	+++	++	++	-33.33	1	nd	+++	+	-	-	-	*Pichia fermentans*
YHS23	HS14	F	0.3	32	0. 5	0. 125	++	+++	+++	++	++	-26.67	1	nd	+++	+	-	-	-	*Pichia fermentans*
YHS24	HS14	F	0.3	> 64	2	> 8	-	++	+++	++	++	-26.67	1	nd	+++	+++	-	-	-	*Pichia fermentans*
YHS25	HS14	F	0.3	32	0. 5	0. 125	-	+++	+++	++	++	-33.33	0	nd	+++	+	+	-	-	*Pichia fermentans*
YHS26	HS14	F	0.3	32	0. 5	0. 125	-	+++	+++	++	++	-33.33	0	nd	+++	+	-	-	-	*Pichia fermentans*
YHS27	HS14	F	0.3	32	0. 5	0. 125	-	+++	+++	++	++	-26.67	0	nd	+++	+	-	-	-	*Pichia fermentans*
YHS28	HS14	F	0.3	32	0. 5	0. 125	-	+++	+++	++	++	-26.67	0	nd	+++	+	-	-	-	*Pichia kluyveri*
YHS29	HS14	F	0.3	32	0. 5	0. 125	-	+++	+++	++	++	-26.67	0	nd	+++	+	-	-	-	*Pichia kluyveri*
YHS30	HS14	F	0.3	32	0. 5	0. 125	-	+++	+++	+++	++	-26.67	0	nd	+++	+	-	-	-	*Pichia kluyveri*
YHS31	HS14	F	0.3	32	0. 5	0. 125	-	+++	+++	++	++	-26.67	0	nd	+++	+	-	-	-	*Pichia kluyveri*
YHS32	HS14	F	0.3	32	0. 5	0. 25	-	+++	+++	++	++	-26.67	0	nd	+	+	-	-	-	*Pichia kluyveri*

YHS33	HS14	F	0.3	32	0. 5	0. 125	-	+++	+++	++	++	-33.33	0	nd	+++	-	-	-	-	*Pichia kluyveri*
YHS34	HS14	F	0.3	32	0. 5	0. 125	-	+++	+++	++	++	-26.67	0	nd	+++	+	-	-	-	*Pichia kluyveri*
YHS35	HS14	F	0.3	32	0. 5	0. 125	-	+++	+++	++	++	-26.67	0	nd	+++	+	-	-	-	*Pichia kluyveri*
YHS36	HS16	M	14	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	*Candida zelanoydes*
YHS37	HS17	M	15	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	*Candida tropicalis*
YHS38	HS18	M	11	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	*Candida albicans*
YHS39	HS19	F	3	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	*Candida lusitaniae*
YHS40	HS21	F	5	0. 5	0. 125	1	++	+++	+++	+++	+++	-9.09	2	nd	+++	+++	+++	++	-	*Torulaspora delbrueckii*
YHS41	HS22	F	15	1	0. 125	2	++	+++	+++	+++	+++	-18.18	2	nd	+++	+++	+++	++	-	*Candida albicans*
YHS42	HS22	F	15	0. 5	0. 125	0. 125	++	+++	+++	+++	+++	-18.18	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS43	HS22	F	15	na	na	na	-	na	na	na	na	na	na	na	na	na	na	na	na	*Candida albicans*
YHS44	HS24	M	15	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	*Cryptococcus saitoi*
YHS45	HS25	M	7	na	na	na	-	na	na	na	na	na	na	na	na	na	na	na	na	*Candida albicans*
YHS46	HS26	M	3	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	*Candida lusitaniae*
YHS47	HS27	F	9	na	na	na	na	+++	na	na	na	-50	2	Pseudohyphae	+++	+++	+++	+++	-	*Candida albicans*
YHS48	HS28	M	5	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	*Rhodotorula mucilaginosa*
YHS49	HS29	F	16	0. 5	0. 125	0. 0156	-	+++	+++	+++	+++	-45.45	3	Pseudohyphae	+++	+++	+	+	-	*Candida parapsilosis*
YHS50	HS29	F	16	0. 125	0. 125	0. 0156	+	+++	+++	+++	+++	-36.36	3	Pseudohyphae	+++	+++	++	+	-	*Candida parapsilosis*
YHS51	HS29	F	16	0. 5	0. 125	0. 0156	+	+++	+++	+++	+++	-36.36	3	Pseudohyphae	+++	+++	+++	++	-	*Candida parapsilosis*
YHS52	HS29	F	16	0. 5	0. 125	0. 0156	++	+++	+++	+++	+++	-36.36	3	Pseudohyphae	+++	+++	+++	++	-	*Candida parapsilosis*
YHS53	HS29	F	16	0. 5	0. 125	0. 0156	-	+++	+++	+++	+++	-63.64	3	Pseudohyphae	+++	+++	++	+	-	*Candida parapsilosis*
YHS54	HS29	F	16	0. 5	0. 25	0. 0156	-	+++	+++	+++	+++	-63.64	3	Pseudohyphae	+++	+++	+	+	-	*Candida parapsilosis*
YHS55	HS29	F	16	0. 5	0. 125	1	++	+++	+++	+++	+++	-63.64	3	Pseudohyphae	+++	+++	+++	++	-	*Rhodotorula mucilaginosa*
YHS56	HS29	F	16	0. 25	0. 125	0. 0156	-	+++	+++	+++	+++	-45.45	3	Pseudohyphae	+++	+++	+++	++	-	*Rhodotorula mucilaginosa*
YHS57	HS29	F	16	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	*Pichia kluyveri*
YHS58	HS31	F	24	2	1	0. 0156	-	+++	+++	++	++	-46.7	3	Pseudohyphae	+++	+++	+++	+	-	*Candida parapsilosis*
YHS59	HS31	F	24	0. 5	4	0. 0156	-	+++	+++	+++	+++	-56.79	3	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS60	HS31	F	24	na	na	na	na	+++	na	na	na	-100	3	Pseudohyphae	+++	-	-	-	-	*Candida parapsilosis*
YHS61	HS31	F	24	0. 25	64	0. 0156	+	+++	+++	+++	+++	-53.3	0	nd	+++	+++	+++	++	-	*Candida albicans*
YHS62	HS31	F	24	na	na	na	na	-	na	na	na	-100	0	nd	+++	+++	+++	-	-	*Candida albicans*
YHS63	HS31	F	24	na	na	na	na	-	na	na	na	-100	0	nd	+++	+++	+++	-	-	*Candida albicans*
YHS64	HS31	F	24	1	0. 125	0. 031	-	++	+	+	+	-62.96	0	nd	+++	+++	+++	-	-	*Rhodotorula mucilaginosa*
YHS65	HS31	F	24	na	na	na	na	+++	na	na	na	-100	0	nd	+++	+++	+++	-	-	*Candida lusitaniae*
YHS66	HS31	F	24	0. 5	0. 5	0. 0156	++	+++	+++	+++	+++	-40	0	nd	+++	+++	+++	+	+	*Rhodotorula mucilaginosa*
YHS67	HS31	F	24	0. 5	0. 125	0. 0156	-	+++	+++	+++	+++	-66.7	0	nd	+++	+	-	-	-	*Rhodosporidium kratochvilovae*
YHS68	HS31	F	24	na	na	na	na	-	na	na	na	-75.31	1	Hyphae	+++	-	-	-	-	*Candida albicans*

YHS69	HS31	F	24	na	na	na	na	-	na	na	na	-56.79	0	nd	+++	-	-	-	-	*Candida albicans*
YHS70	HS31	F	24	na	na	na	na	+++	na	na	na	-100	1	nd	+++	+++	+++	-	-	*Candida albicans*
YHS71	HS31	F	24	0. 125	0. 125	0. 0156	-	-	-	-	-	16.7	0	nd	+++	+++	+++	-	-	*Candida albicans*
YHS72	HS31	F	24	0. 25	0. 5	0. 0156	-	+++	+++	+++	++	-73.3	0	nd	+++	+++	+++	-	-	*Candida lusitaniae*
YHS73	HS31	F	24	0. 125	0. 125	0. 0156	-	+	-	-	-	6.7	1	nd	+++	+++	+++	-	-	*Candida albicans*
YHS74	HS31	F	24	na	na	na	na	-	na	na	na	-100	1	nd	+++	+++	+++	-	-	*Candida albicans*
YHS75	HS31	F	24	na	na	na	na	+++	na	na	na	-100	3	Pseudohyphae	+++	+++	+++	-	-	*Candida parapsilosis*
YHS76	HS31	F	24	na	na	na	na	-	na	na	na	-100	1	nd	+++	+++	+++	-	-	*Candida albicans*
YHS77	HS32	F	32	0. 125	0. 5	0. 0156	-	+++	+++	+++	+++	13.3	0	nd	+++	+++	+++	-	-	*Pichia manshurica*
YHS78	HS32	F	32	0. 25	> 64	0. 0156	-	+++	+++	+++	+++	16.7	2	nd	+++	+++	+++	++	-	*Pichia manshurica*
YHS79	HS32	F	32	na	na	na	na	+++	na	na	na	-75.31	2	Hyphae	+++	+++	+++	-	-	*Pichia manshurica*
YHS80	HS32	F	32	0. 25	8	0. 0156	+	+++	+++	+++	+++	-66.7	0	nd	+++	+++	+++	+	-	*Pichia manshurica*
YHS81	HS32	F	32	> 64	0. 125	> 8	++	+++	+++	+++	++	-33.3	0	nd	+++	+++	+++	++	-	*Pichia manshurica*
YHS82	HS32	F	32	na	na	na	na	+++	na	na	na	-50.62	0	nd	+++	+++	+++	-	-	*Pichia manshurica*
YHS83	HS32	F	32	0. 5	0. 5	0. 0156	-	+++	+++	+++	++	16.7	2	nd	+++	+++	+++	-	-	*Pichia manshurica*
YHS84	HS32	F	32	0. 25	0. 125	0. 0156	++	+++	+++	+++	+++	0	1	nd	+++	+++	+++	++	-	*Pichia manshurica*
YHS85	HS32	F	32	0. 25	0. 125	0. 0156	++	+++	+++	+++	+++	0	2	Pseudohyphae	+++	+++	+++	++	-	*Pichia manshurica*
YHS86	HS32	F	32	> 64	0. 125	> 8	+	+++	+++	+++	+++	-66.7	1	nd	+++	+++	+++	+++	+	*Pichia manshurica*
YHS87	HS32	F	32	na	na	na	na	+++	na	na	na	-81.48	2	nd	+++	+++	+++	-	-	*Pichia manshurica*
YHS88	HS33	F	32	na	na	na	na	+++	na	na	na	-44.44	0	nd	+++	+++	+++	-	-	*Rhodotorula mucilaginosa*
YHS89	HS33	F	32	> 64	0. 125	> 8	+	+++	+++	+++	+++	-26.7	1	nd	+++	+++	+++	++	-	*Candida albicans*
YHS90	HS33	F	32	> 64	0. 125	> 8	+	+++	+++	+++	+++	-66.7	3	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS91	HS33	F	32	0. 25	0. 125	0. 0156	+	+++	+++	+++	+++	-20	3	Hyphae	+++	+++	+++	+	-	*Candida albicans*
YHS92	HS33	F	32	0. 5	0. 125	0. 0156	++	+++	+++	+++	+++	-20	3	Hyphae	+++	+++	+	+	-	*Candida albicans*
YHS93	HS33	F	32	0. 25	0. 125	0. 0156	++	+++	+++	+++	+++	-26.7	2	nd	+++	+++	+++	+	-	*Pichia manshurica*
YHS94	HS33	F	32	> 64	0. 125	> 8	+	+++	+++	+++	+++	-26.7	1	nd	+++	+++	+++	++	-	*Candida albicans*
YHS95	HS33	F	32	0. 25	0. 125	0. 0156	++	+++	+++	+++	+++	0	3	nd	+++	+++	+++	++	-	*Candida albicans*
YHS96	HS33	F	32	0. 25	0. 125	0. 0156	++	+++	+++	+++	+++	-13.3	3	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS97	HS33	F	32	0. 25	0. 125	0. 0156	++	+++	+++	+++	+++	-50.62	3	Hyphae	+++	+++	+++	++	+	*Candida albicans*
YHS98	HS33	F	32	> 64	0. 125	> 8	++	+++	+++	+++	+++	-26.7	0	Hyphae	+++	+++	+	+	-	*Rhodosporidium kratochvilovae*
YHS99	HS34	F	25	> 64	0. 125	> 8	++	+++	+++	+++	+++	-66.7	3	nd	+++	+++	+++	+++	+	*Candida albicans*
YHS100	HS36	M	20	0. 125	0. 125	0. 0156	-	+++	-	-	-	-100	0	Hyphae	+++	+++	+++	+++	-	*Candida albicans*
YHS101	HS36	M	20	0. 125	0. 125	0. 0156	-	+++	+	-	-	-100	1	Hyphae	+++	+++	+++	+++	-	*Torulaspora delbrueckii*
YHS102	HS36	M	20	0. 125	0. 125	0. 0156	-	+++	+	-	-	-100	0	nd	+++	+++	+++	+++	-	*Candida albicans*
YHS103	HS36	M	20	0. 125	0. 125	0. 0156	-	+++	-	-	-	-100	0	nd	+++	+++	+++	+++	-	*Candida glabrata*
YHS104	HS36	M	20	0. 125	0. 125	2	+	+++	+++	+++	+++	-7.41	3	Hyphae	+++	+++	+++	+++	-	*Candida parapsilosis*

YHS105	HS36	M	20	> 64	0. 5	2	++	+++	+++	+++	+++	-100	3	Hyphae	+++	+++	+++	+++	-	*Starmerella bacillaris*
YHS106	HS37	F	28	0. 25	0. 5	1	+	+++	+++	+++	+++	4.94	3	Hyphae	+++	+++	+++	+++	-	*Candida albicans*
YHS107	HS37	F	28	0. 25	0. 5	2	++	+++	+++	+++	+++	4.94	1	Hyphae	+++	+++	+++	+++	-	*Candida albicans*
YHS108	HS37	F	28	0. 25	0. 5	2	++	+++	+++	+++	+++	23.46	1	Hyphae	+++	+++	+++	+++	-	*Candida albicans*
YHS109	HS37	F	28	0. 5	0. 5	2	++	+++	+++	+++	+++	17.28	1	Hyphae	+++	+++	+++	+++	-	*Torulaspora delbrueckii*
YHS110	HS37	F	28	0. 5	0. 5	2	++	+++	+++	+++	+++	17.28	2	Hyphae	+++	+++	+++	+++	-	*Candida albicans*
YHS111	HS38	F	25	0. 125	0. 125	0. 0156	-	+++	-	-	-	-33.33	1	nd	++	+	+	+	-	*Candida deformans*
YHS112	HS39	F	27	> 64	0. 125	> 8	+	+++	+++	+++	+++	-100	0	nd	+++	+++	+++	+++	-	*Candida albicans*
YHS113	HS39	F	27	> 64	0. 125	0. 125	+	+++	+++	+++	+++	-13.58	3	Hyphae	+++	+++	+++	+++	-	*Candida albicans*
YHS114	HS39	F	27	> 64	0. 125	> 8	+	+++	+++	+++	+++	-100	0	nd	+++	+++	+++	+++	-	*Candida albicans*
YHS115	HS39	F	27	> 64	0. 125	0. 125	+	+++	+++	+++	+++	-7.41	3	Hyphae	+++	+++	+++	+++	-	*Candida albicans*
YHS116	HS39	F	27	> 64	0. 125	> 8	+	+++	+++	+++	+++	-100	0	nd	+++	+++	+++	+++	-	*Candida albicans*
YHS117	HS39	F	27	32	0. 125	1	+	+++	+++	+++	+++	-13.58	1	nd	+++	+++	+++	+++	-	*Candida albicans*
YHS118	HS40	M	27	0. 125	0. 125	0. 0156	-	+++	+++	+++	+++	20	3	Hyphae	++	+	+	-	-	*Pennicillium crustosum*
YHS119	HS40	M	27	0. 125	0. 125	0. 0156	-	+++	+++	+++	+++	20	3	Hyphae	++	+	+	-	-	*Pennicillium crustosum*
YHS120	HS40	M	27	0. 125	0. 125	0. 0156	-	+++	+++	+++	+++	20	3	Hyphae	++	+	+	-	-	*Pennicillium crustosum*
YHS121	HS40	M	27	0. 125	0. 125	0. 0156	-	+++	+++	+++	+++	20	3	Hyphae	++	+	+	-	-	*Pennicillium crustosum*
YHS122	HS40	M	27	0. 125	0. 125	0. 0156	-	+++	+++	+++	+++	20	3	Hyphae	++	+	+	-	-	*Pennicillium crustosum*
YHS123	HS41	F	24	0. 25	0. 125	0. 0156	-	++	+++	+++	+++	-33.33	2	Hyphae	+++	+++	+	-	-	*Candida parapsilosis*
YHS124	HS41	F	24	1	0. 125	0. 0156	-	+++	+++	+++	+++	-53.33	2	Hyphae	+++	++	-	-	-	*Candida parapsilosis*
YHS125	HS41	F	24	> 64	0. 125	0. 0156	-	-	+	+	+	20	3	Hyphae	+	+	+	+	-	*Aspergillus glaucus*
YHS126	HS43	M	26	0. 5	0. 125	0. 0156	-	+++	+++	+++	+++	-100	2	nd	+++	-	-	-	-	*Candida parapsilosis*
YHS127	HS44	F	24	0. 125	0. 125	0. 125	-	+++	+++	+++	++	-7.41	1	Hyphae	+++	+++	+++	+++	-	*Candida albicans*
YHS128	HS44	F	24	0. 125	0. 125	0. 0156	-	+++	+++	+++	+++	-7.41	2	nd	+++	+++	+++	+++	-	*Candida albicans*
YHS129	HS44	F	24	0. 125	0. 125	0. 25	-	+++	+++	+++	+++	4.94	0	Hyphae	+++	+++	+++	+++	-	*Torulaspora delbrueckii*
YHS130	HS44	F	24	0. 5	0. 125	0. 25	+	+++	+++	+++	+++	-7.41	0	nd	+++	+++	+++	+++	-	*Candida albicans*
YHS131	HS44	F	24	0. 5	0. 125	0. 25	+	+++	-	-	-	-100	2	Hyphae	+++	+++	+++	+++	-	*Candida albicans*
YHS132	HS44	F	24	0. 125	0. 125	0. 0156	+	+++	+++	+++	+++	-7.41	3	Pseudohyphae	+++	+++	+++	+++	-	*Candida parapsilosis*
YHS133	HS44	F	24	> 64	0. 5	0. 125	+	+++	+++	+++	+++	-7.41	2	nd	+++	+++	+++	+++	-	*Candida parapsilosis*
YHS134	HS46	F	6	0. 5	0. 125	2	++	-	+++	+++	+++	-18.18	3	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS135	HS46	F	6	0. 5	0. 125	2	++	+++	+++	+++	+++	-27.27	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS136	HS46	F	6	0. 5	0. 125	2	++	+++	+++	+++	+++	-27.27	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS137	HS46	F	6	0. 5	0. 125	2	+	+++	+++	+++	+++	-27.27	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS138	HS47	F	10	0. 125	0. 125	0. 0156	-	+++	-	-	-	0	2	Hyphae	+++	+++	+++	++	-	*Torulaspora delbrueckii*
YHS139	HS47	F	10	0. 25	0. 125	1	++	+++	+++	+++	+++	-18.18	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS140	HS47	F	10	0. 25	0. 125	0. 125	++	+++	+++	+++	+++	-54.55	2	Hyphae	+++	+++	+++	++	-	*Torulaspora delbrueckii*

YHS141	HS47	F	10	0. 25	0. 125	1	++	+++	+++	+++	+++	-9.09	3	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS142	HS50	F	1.5	0. 125	0. 125	0. 0156	-	+++	+	-	-	27.27	2	nd	+++	+++	+++	++	-	*Candida intermedia*
YHS143	HS50	F	1.5	0. 125	0. 125	0. 0156	-	+++	+++	+++	+++	-18.18	1	nd	+++	+++	+++	++	-	*Candida lusitaniae*
YHS144	HS50	F	1.5	0. 5	0. 125	0. 0156	++	+++	+++	+++	+++	-18.18	1	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS145	HS50	F	1.5	0. 25	0. 125	0. 0156	++	++	+++	+++	+++	63.64	1	Pseudohyphae	+++	++	+++	++	-	*Rhodotorula mucilaginosa*
YHS146	HS50	F	1.5	0. 25	0. 125	0. 0156	++	-	++	++	+	-18.18	1	nd	+++	+++	+++	+	-	*Candida lusitaniae*
YHS147	HS50	F	1.5	0. 5	0. 125	0. 125	-	+++	+++	+++	+++	-54.55	1	nd	+++	+++	+++	++	-	*Candida parapsilosis*
YHS148	HS50	F	1.5	0. 5	0. 125	0. 125	++	+++	+++	+++	+++	36.36	1	nd	+++	+++	+++	++	-	*Candida lusitaniae*
YHS149	HS50	F	1.5	0. 5	0. 125	0. 0156	-	+++	+++	+++	+++	-27.27	3	Pseudohyphae	+++	+++	+++	+	-	*Candida parapsilosis*
YHS150	HS50	F	1.5	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	*Yarrowia lipolytica*
YHS151	HS51	F	8	na	na	na	na	+++	na	na	na	36.36	1	nd	+++	+++	+++	++	-	*Rhodotorula mucilaginosa*
YHS152	HS52	F	23	0. 25	0. 125	0. 0156	-	+++	+++	+++	+++	-7.41	1	Hyphae	+++	+++	+++	+++	-	*Candida albicans*
YHS153	HS52	F	23	0. 25	0. 125	0. 0156	+	+++	+++	+++	+++	-7.41	1	Hyphae	+++	+++	+++	+++	-	*Candida albicans*
YHS154	HS52	F	23	0. 125	0. 125	0. 0156	++	+++	+++	+++	+++	4.94	1	Hyphae	+++	+++	+++	+++	-	*Candida albicans*
YHS155	HS52	F	23	0. 5	0. 125	0. 0156	-	+++	+++	+++	+++	-7.41	3	Hyphae	+++	+++	+++	+++	-	*Candida albicans*
YHS156	HS53	F	23	0. 5	0. 125	0. 0156	++	+++	+++	+++	+++	-27.27	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS157	HS53	F	23	0. 5	0. 125	2	++	+++	+++	+++	+++	-18.18	1	nd	+++	+++	+++	++	-	*Candida albicans*
YHS158	HS54	M	2	0. 125	0. 125	0. 0156	-	+++	+++	+++	+++	-40	1	Pseudohyphae	+++	+++	+++	+++	+	*Candida albicans*
YHS159	HS54	M	2	0. 5	0. 125	> 8	+	+++	+++	+++	+++	-53.33	2	Hyphae	+++	+++	+	-	-	*Candida parapsilosis*
YHS160	HS54	M	2	0. 25	0. 125	1	+	+++	na	na	na	-43.33	2	Hyphae	++	+++	+	+	+	*Candida albicans*
YHS161	HS54	M	2	0. 25	0. 125	2	+	+++	+++	+++	+++	-46.67	1	nd	+++	+++	+	+	+	*Candida albicans*
YHS162	HS54	M	2	0. 5	0. 125	4	+	+++	+++	+++	+++	-33.33	1	nd	+++	+++	+++	+++	+	*Candida albicans*
YHS163	HS54	M	2	0. 5	0. 125	2	+	++	+++	+++	+++	-46.67	2	Hyphae	+++	+++	+++	++	+	*Candida albicans*
YHS164	HS54	M	2	0. 25	0. 125	2	+	na	++	++	++	-33.33	0	nd	na	na	na	na	na	*Lichtheimia ramosa*
YHS165	HS55	M	2	> 64	0. 125	0. 0156	-	+	+++	+++	+++	na	1	Hyphae	+++	++	-	-	-	*Aspergillus glaucus*
YHS166	HS56	M	2	0. 25	0. 5	2	+	+++	+++	+++	+++	-20	0	nd	+++	++	++	++	+	*Candida albicans*
YHS167	HS57	F	12	0. 125	0. 125	0. 0156	-	+	-	-	-	-26.67	3	Hyphae	++	++	-	-	-	*Penicillium brevicompactum*
YHS168	HS57	F	12	0. 125	0. 125	0. 0156	-	+	-	-	-	-26.67	3	Hyphae	++	++	-	-	-	*Penicillium brevicompactum*
YHS169	HS57	F	12	0. 125	0. 125	0. 0156	-	+	-	-	-	-26.67	3	Hyphae	++	++	-	-	-	*Penicillium brevicompactum*
YHS170	HS57	F	12	0. 125	0. 125	0. 0156	-	+	-	-	-	-26.67	3	Hyphae	++	++	-	-	-	*Penicillium brevicompactum*
YHS171	HS57	F	12	0. 125	0. 125	0. 0156	-	+	-	-	-	-26.67	3	Hyphae	++	++	-	-	-	*Penicillium brevicompactum*
YHS172	HS57	F	12	0. 125	0. 125	0. 0156	-	+	-	-	-	-26.67	3	Hyphae	++	++	-	-	-	*Penicillium brevicompactum*
YHS173	HS57	F	12	0. 125	0. 125	0. 0156	-	+	-	-	-	-26.67	3	Hyphae	++	++	-	-	-	*Penicillium brevicompactum*
YHS174	HS57	F	12	0. 125	0. 125	0. 0156	-	+	-	-	-	-26.67	3	Hyphae	++	++	-	-	-	*Penicillium brevicompactum*
YHS175	HS57	F	12	0. 125	0. 125	0. 0156	-	+	-	-	-	-26.67	3	Hyphae	++	++	-	-	-	*Penicillium brevicompactum*
YHS176	HS57	F	12	0. 125	0. 125	0. 0156	-	+	-	-	-	-26.67	3	Hyphae	++	++	-	-	-	*Penicillium brevicompactum*

YHS177	HS57	F	12	0. 125	0. 125	0. 0156	-	+	-	-	-	-26.67	3	Hyphae	++	++	-	-	-	*Penicillium brevicompactum*
YHS178	HS57	F	12	0. 125	0. 125	0. 0156	-	+	-	-	-	-26.67	3	Hyphae	++	++	-	-	-	*Penicillium brevicompactum*
YHS179	HS58	F	3	0. 125	0. 125	0. 0156	-	+++	+++	+++	+++	-33.33	0	nd	+++	+++	++	++	+	*Candida albicans*
YHS180	HS60	F	3	0. 5	0. 125	0. 0625	+	+++	+++	+++	+++	-23.33	1	nd	++	++	++	++	+	*Candida albicans*
YHS181	HS61	M	2	0. 5	0. 125	4	+	++	+++	++	++	-26.67	1	Pseudohyphae	+++	++	++	++	+	*Candida parapsilosis*
YHS182	HS61	M	2	0. 5	0. 125	4	+	+++	+++	+++	+++	-33.33	2	Pseudohyphae	+++	++	++	+	-	*Candida parapsilosis*
YHS183	HS61	M	2	0. 25	0. 125	0. 0625	-	+++	+++	++	++	-20	1	nd	++	++	++	++	+	*Candida parapsilosis*
YHS184	HS61	M	2	0. 25	0. 125	0. 031	-	+++	+++	+++	+++	-33.33	2	nd	+++	++	+	-	-	*Candida parapsilosis*
YHS185	HS61	M	2	0. 25	0. 125	0. 031	-	++	+++	+++	+++	-40	2	Pseudohyphae	+++	++	+	-	-	*Candida parapsilosis*
YHS186	HS61	M	2	0. 25	0. 125	0. 031	-	+++	++	++	++	-53.33	2	Pseudohyphae	+++	+++	+	+	-	*Candida parapsilosis*
YHS187	HS61	M	2	0. 125	0. 125	0. 0156	-	+	++	++	++	-20	1	nd	++	++	+	++	-	*Torulaspora delbrueckii*
YHS188	HS61	M	2	0. 125	0. 125	0. 0156	-	+++	+++	+++	+++	-40	1	nd	+++	++	-	-	-	*Candida parapsilosis*
YHS189	HS61	M	2	0. 5	0. 125	4	+	++	++	++	++	-40	0	nd	++	++	++	++	-	*Torulaspora delbrueckii*
YHS190	HS61	M	2	0. 25	0. 125	0. 031	-	+++	+++	+++	+++	-26.67	1	nd	+++	++	-	-	-	*Candida parapsilosis*
YHS191	HS61	M	2	0. 25	0. 125	0. 031	-	+++	+++	+++	+++	-33.33	2	Pseudohyphae	+++	++	++	++	+	*Candida parapsilosis*
YHS192	HS61	M	2	0. 25	0. 125	0. 031	-	+++	+++	+++	+++	-53.33	2	Pseudohyphae	+++	++	-	-	-	*Candida parapsilosis*
YHS193	HS61	M	2	0. 25	0. 125	0. 031	-	++	+++	+++	+++	-100	2	Pseudohyphae	+++	+	-	-	-	*Candida parapsilosis*
YHS194	HS62	F	4	0. 5	0. 125	0. 0625	-	-	++	++	++	-33.33	3	Hyphae	+++	+	+	+	-	*Trichosporon asahii*
YHS195	HS62	F	4	0. 5	0. 125	0. 0625	-	+++	++	++	++	-20	3	Hyphae	++	+	+	+	-	*Trichosporon asahii*
YHS196	HS62	F	4	2	0. 125	0. 031	-	+++	+++	+++	+++	-60	1	Pseudohyphae	+++	++	++	++	+	*Candida parapsilosis*
YHS197	HS62	F	4	0. 125	0. 125	0. 0156	-	+++	+++	+++	+++	-60	3	Pseudohyphae	+++	+	+	-	-	*Candida parapsilosis*
YHS198	HS62	F	4	0. 125	0. 125	0. 0156	-	+	+	+	-	-53.33	1	nd	+++	++	-	-	-	*Rhodotorula mucilaginosa*
YHS199	HS62	F	4	0. 5	0. 125	2	+	+++	+++	+++	+++	-36.67	0	nd	++	++	-	-	-	*Rhodotorula mucilaginosa*
YHS200	HS62	F	4	0. 125	0. 125	0. 0156	-	+++	+++	+++	+++	-60	3	Pseudohyphae	+++	++	+	+	+	*Candida parapsilosis*
YHS201	HS62	F	4	0. 5	0. 125	0. 0625	+	+++	+++	+++	+++	-60	3	Pseudohyphae	+++	+	+	+	-	*Candida parapsilosis*
YHS202	HS62	F	4	0. 125	0. 125	0. 0156	-	+	+++	++	++	-33.33	0	nd	++	-	-	-	-	*Rhodotorula mucilaginosa*
YHS203	HS62	F	4	0. 125	0. 125	0. 0156	-	+++	++	++	++	-40	3	Pseudohyphae	++	+	+	+	-	*Rhodotorula mucilaginosa*
YHS204	HS62	F	4	0. 5	0. 125	0. 0625	-	+++	++	++	++	-100	3	Pseudohyphae	+++	++	+	-	-	*Candida parapsilosis*
YHS205	HS62	F	4	0. 125	0. 125	0. 0156	-	+	-	-	-	20	0	nd	+	++	-	-	-	*Rhodotorula mucilaginosa*
YHS206	HS62	F	4	0. 5	0. 125	0. 031	+	+++	+	+	-	-26.67	0	nd	+	+++	-	-	-	*Rhodotorula mucilaginosa*
YHS207	HS62	F	4	1	0. 125	0. 0156	-	+++	+++	+++	+++	-60	3	Pseudohyphae	+++	+	+	+	-	*Candida parapsilosis*
YHS208	HS62	F	4	0. 125	0. 125	0. 0156	-	+	+	+	-	20	0	nd	+	+++	-	-	-	*Rhodotorula mucilaginosa*
YHS209	HS62	F	4	0. 125	0. 125	0. 0156	-	-	++	++	++	-6.67	1	nd	++	++	++	++	-	*Rhodotorula mucilaginosa*
YHS210	HS64	F	3	0. 125	0. 125	0. 0156	-	+++	++	++	++	-33.33	1	nd	+++	+++	-	-	-	*Pichia fermentans*
YHS211	HS64	F	3	0. 5	0. 125	0. 0156	+	+++	+++	+++	+++	-26.67	1	Hyphae	+++	+++	+++	+++	-	*Candida albicans*
YHS212	HS64	F	3	0. 5	0. 125	0. 0156	+	+++	+++	+++	+++	-26.67	1	nd	+++	+++	++	++	-	*Candida albicans*

YHS213	HS64	F	3	0. 5	0. 125	0. 25	+	+++	+++	+++	+++	-33.33	1	nd	+++	+++	+++	++	-	*Candida albicans*
YHS214	HS64	F	3	0. 5	0. 125	0. 031	+	+++	+++	+++	+++	-13.33	1	nd	+++	+++	+++	++	++	*Candida albicans*
YHS215	HS64	F	3	0. 5	0. 125	0. 25	+	+++	+++	+++	+++	-26.67	1	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS216	HS64	F	3	0. 5	0. 125	0. 031	+	+++	+++	+++	+++	-13.33	1	nd	+++	+++	+++	++	-	*Candida albicans*
YHS217	HS65	M	5	0. 5	0. 125	0. 125	-	+++	+++	+++	+++	-46.67	1	nd	+++	+++	+++	++	-	*Candida lusitaniae*
YHS218	HS65	M	5	0. 5	0. 125	0. 125	+	+++	+++	+++	+++	-50	1	nd	+++	+++	+++	++	-	*Candida lusitaniae*
YHS219	HS65	M	5	0. 5	0. 125	0. 125	-	+	++	++	++	20	3	nd	++	-	-	-	-	*Candida pararugosa*
YHS220	HS65	M	5	0. 125	0. 125	0. 0156	-	++	+	+	+	20	0	nd	++	-	-	-	-	*Rhodotorula mucilaginosa*
YHS221	HS65	M	5	0. 25	0. 125	2	+	+	++	++	-	-13.33	0	nd	+++	+++	-	-	-	*Rhodotorula mucilaginosa*
YHS222	HS65	M	5	0. 125	0. 125	0. 0156	-	+	+++	+++	+++	-3.33	0	nd	+	-	-	-	-	*Rhodotorula mucilaginosa*
YHS223	HS65	M	5	0. 5	0. 125	0. 031	+	+++	+++	+++	+++	-13.33	0	nd	+++	+++	++	++	-	*Rhodotorula mucilaginosa*
YHS224	HS65	M	5	0. 125	0. 125	0. 0156	-	+++	+++	+++	+++	20	0	nd	+	+	-	-	-	*Rhodotorula mucilaginosa*
YHS225	HS65	M	5	0. 125	0. 125	0. 0156	-	+	++	++	++	-53.33	0	nd	++	++	++	++	-	*Rhodotorula mucilaginosa*
YHS226	HS65	M	5	0. 125	0. 125	0. 0156	-	+	+	+	-	20	0	nd	+	-	-	-	-	*Rhodotorula mucilaginosa*
YHS227	HS67	F	1	0. 125	0. 125	> 8	++	+++	-	-	-	-33.33	0	nd	+++	+++	-	-	-	*Torulaspora delbrueckii*
YHS228	HS67	F	1	0. 125	0. 125	0. 0156	-	+++	+++	-	-	6.67	0	nd	+	-	-	-	-	*Torulaspora delbrueckii*
YHS229	HS67	F	1	0. 125	0. 125	0. 0156	-	+	+	-	-	20	1	nd	+	-	-	-	-	*Rhodotorula mucilaginosa*
YHS230	HS67	F	1	> 64	0. 5	2	+	+++	-	-	-	-33.33	1	nd	+	++	-	-	-	*Rhodotorula mucilaginosa*
YHS231	HS67	F	1	0. 125	0. 125	0. 0156	-	+++	+++	+++	+++	-20	2	nd	++	++	++	++	-	*Torulaspora delbrueckii*
YHS232	HS67	F	1	0. 125	0. 125	0. 0156	-	+++	+++	+++	+++	-3.33	2	nd	++	++	++	++	-	*Torulaspora delbrueckii*
YHS233	HS67	F	1	0. 125	0. 125	0. 0156	+	+++	+++	+++	+++	-13.33	1	nd	+	+	-	-	-	*Torulaspora delbrueckii*
YHS234	HS67	F	1	0. 125	0. 125	0. 0156	-	-	-	-	-	-10	0	nd	-	-	-	-	-	*Torulaspora delbrueckii*
YHS235	HS67	F	1	0. 125	0. 125	0. 0156	-	-	-	-	-	0	0	nd	-	-	-	-	-	*Torulaspora delbrueckii*
YHS236	HS67	F	1	0. 125	0. 125	0. 0156	-	+++	+++	-	-	-13.33	0	nd	+++	+++	-	-	-	*Torulaspora delbrueckii*
YHS237	HS67	F	1	> 64	0. 125	2	+	+++	-	-	-	13.33	0	nd	+	+++	-	-	-	*Torulaspora delbrueckii*
YHS238	HS67	F	1	> 64	0. 125	0. 0156	-	+++	-	-	-	6.67	1	nd	+	+++	-	-	-	*Torulaspora delbrueckii*
YHS239	HS67	F	1	0. 125	0. 125	0. 0156	-	+++	+++	+++	+++	0	1	nd	+	++	-	-	-	*Torulaspora delbrueckii*
YHS240	HS67	F	1	0. 5	0. 125	4	+	+++	+++	+++	+++	-33.33	1	nd	+	++	-	-	-	*Rhodotorula mucilaginosa*
YHS241	HS67	F	1	0. 25	0. 125	0. 0156	-	+	+++	+++	+++	-20	1	nd	+	-	-	-	-	*Torulaspora delbrueckii*
YHS242	HS67	F	1	8	0. 125	2	+	+	-	-	-	0	0	nd	+++	+++	-	-	-	*Torulaspora delbrueckii*
YHS243	HS67	F	1	0. 125	0. 125	0. 0156	-	++	+	+	+	13.33	0	nd	+	+	-	-	-	*Torulaspora delbrueckii*
YHS244	HS68	F	4	0. 125	0. 125	0. 0156	-	++	+++	-	-	0	0	nd	+	-	-	-	-	*Torulaspora delbrueckii*
YHS245	HS68	F	4	0. 125	0. 125	0. 0156	-	++	++	++	++	-10	1	Hyphae	++	++	++	++	-	*Penicillium paneum*
YHS246	HS69	M	6	> 64	0. 125	2	+	+++	-	-	-	-26.67	1	Hyphae	+++	+++	+++	+++	-	*Aspergillus pseudoglaucus*
YHS247	HS70	F	11	> 64	0. 125	2	+	+++	-	-	-	-20	3	Hyphae	+++	+++	-	-	-	*Penicillium paneum*
YHS248	HS71	M	1	> 64	0. 125	2	+	+++	+++	+++	+++	-23.33	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*

YHS249	HS73	F	4	> 64	0. 25	1	+	++	+++	+++	+++	-33.33	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS250	HS74	M	6	> 64	0. 125	2	+	++	+++	+++	+++	-33.33	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS251	HS75	F	1	> 64	0. 125	2	+	++	+++	+++	+++	-26.67	2	Hyphae	+++	+++	+++	++	+	*Candida albicans*
YHS252	HS76	M	1	> 64	0. 125	8	++	++	+++	+++	+++	-6.67	1	Hyphae	+++	+++	+++	++	-	*Aspergillus pseudoglaucus*
YHS253	HS77	F	4	0. 5	0. 25	1	+	+++	+++	+++	+++	-20	2	nd	+++	+++	++	++	-	*Candida albicans*
YHS254	HS79	M	0	64	0. 5	8	+	+++	+++	+++	+++	-33.33	2	Hyphae	+++	+++	+	+	+	*Candida albicans*
YHS255	HS80	F	0	> 64	0. 5	2	+	+++	+++	+++	+++	-20	2	Hyphae	+++	+++	+	++	++	*Candida albicans*
YHS256	HS81	F	7	0. 5	0. 5	1	+	+++	+++	+++	+++	-26.67	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS257	HS84	F	6	> 64	0. 5	4	+	+++	+++	+++	+++	-20	2	Hyphae	+++	+++	++	++	-	*Candida albicans*
YHS258	HS85	F	10	0. 25	0. 125	> 8	++	+++	+++	+++	+++	-33.33	1	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS259	HS85	F	10	0. 125	0. 125	8	++	+++	+++	+++	+++	-20	1	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS260	HS85	F	10	0. 25	0. 125	> 8	++	++	+++	+++	+++	-20	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS261	HS85	F	10	0. 25	0. 125	> 8	++	++	+++	+++	+++	-23.33	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS262	HS85	F	10	0. 25	0. 125	> 8	+	++	+++	+++	+++	-23.33	1	Hyphae	+++	+++	+	+	+	*Candida albicans*
YHS263	HS85	F	10	4	0. 125	0. 25	+	+++	+++	+++	+++	-20	2	Hyphae	+++	+++	+	+	+	*Candida albicans*
YHS264	HS85	F	10	1	0. 125	1	++	+++	+++	+++	+++	-23.33	1	Hyphae	++	+++	+++	+++	-	*Candida albicans*
YHS265	HS85	F	10	0. 125	0. 125	2	++	+++	+++	+++	+++	-26.67	1	Hyphae	+++	+++	++	++	-	*Candida albicans*
YHS266	HS85	F	10	0. 125	0. 125	> 8	++	+++	+++	+++	+++	-6.67	2	Hyphae	+++	+++	++	++	-	*Candida albicans*
YHS267	HS85	F	10	0. 125	0. 125	8	++	+++	+++	+++	+++	-13.33	2	Hyphae	+++	+++	++	++	-	*Candida albicans*
YHS268	HS85	F	10	0. 125	0. 125	> 8	++	+++	+++	+++	+++	-13.33	2	Hyphae	+++	+++	++	++	-	*Candida albicans*
YHS269	HS85	F	10	> 64	0. 125	> 8	++	+++	+++	+++	+++	-20	2	Hyphae	+++	+++	+++	+++	-	*Candida albicans*
YHS270	HS85	F	10	0. 25	0. 125	> 8	+	+++	+++	+++	+++	-26.67	1	nd	+++	+++	++	+	-	*Candida albicans*
YHS271	HS85	F	10	0. 125	0. 125	> 8	++	+++	+++	+++	+++	-26.67	1	nd	+++	+++	++	+++	-	*Candida albicans*
YHS272	HS85	F	10	4	0. 125	0. 0625	++	+++	+++	+++	+++	-26.67	2	Hyphae	+++	+++	+++	++	-	*Candida fermentati*
YHS273	HS85	F	10	0. 125	0. 125	> 8	++	+++	+++	+++	+++	-26.67	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS274	HS85	F	10	0. 125	0. 125	8	++	+++	+++	+++	+++	-26.67	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS275	HS85	F	10	0. 125	0. 125	> 8	++	+++	+++	+++	+++	-20	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS276	HS85	F	10	0. 25	0. 125	2	++	+++	+++	+++	+++	-46.67	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS277	HS85	F	10	0. 5	0. 125	0. 0625	+	+++	+++	+++	+++	-33.33	2	Hyphae	+++	+++	+++	+	-	*Candida albicans*
YHS278	HS85	F	10	0. 25	0. 125	> 8	++	+++	+++	+++	+++	-20	2	Hyphae	+++	+++	+++	+++	-	*Candida albicans*
YHS279	HS85	F	10	0. 125	0. 125	0. 25	+	+++	+++	+++	+++	-20	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS280	HS85	F	10	0. 125	0. 125	0. 125	++	+++	+++	+++	+++	-23.33	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS281	HS85	F	10	0. 25	0. 125	2	++	+++	+++	+++	+++	-26.67	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS282	HS85	F	10	0. 5	0. 125	2	++	+++	+++	+++	+++	-6.67	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS283	HS86	M	7	0. 5	0. 5	0. 125	++	+++	+++	+++	+++	-33.33	2	Hyphae	+++	+++	+++	+++	-	*Candida albicans*
YHS284	HS87	M	9	0. 125	0. 125	0. 031	-	-	+++	+++	+++	20	1	nd	+	+	+	-	-	*Cryptococcus saitoi*

YHS285	HS87	M	9	> 64	0. 125	0. 0156	-	-	+++	+++	+++	-13.33	2	Hyphae	++	++	++	+	-	*Eurotium rubrum*
YHS286	HS88	M	7	> 64	0. 125	2	++	+++	+++	+++	+++	-26.67	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS287	HS89	M	12	0. 5	0. 125	0. 125	++	+++	+++	+++	+++	-46.67	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS288	HS89	M	12	0. 5	0. 125	2	+	+++	+++	+++	+++	-33.33	1	nd	+++	+++	+++	++	-	*Candida albicans*
YHS289	HS90	F	8	0. 25	0. 5	2	++	+++	+++	+++	+++	-26.67	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS290	HS91	F	2	0. 25	0. 5	0. 25	++	+++	+++	+++	+++	-40	2	Pseudohyphae	+++	+++	+++	+	-	*Candida parapsilosis*
YHS291	HS91	F	2	0. 5	0. 125	0. 25	-	+++	+++	+++	+++	-46.67	2	Pseudohyphae	+++	+++	+++	++	-	*Candida parapsilosis*
YHS292	HS91	F	2	> 64	0. 125	> 8	++	++	+++	+++	+++	13.33	2	Hyphae	+++	++	+	-	-	*Eurotium amstelodami*
YHS293	HS91	F	2	0. 25	0. 125	> 8	-	+++	++	++	++	-46.67	2	Pseudohyphae	+++	++	+	-	-	*Candida parapsilosis*
YHS294	HS92	F	12	0. 5	0. 125	> 8	+	++	+++	+++	+++	-6.67	3	Pseudohyphae	+++	+++	+++	+	-	*Saccharomyces cerevisiae*
YHS295	HS92	F	12	0. 25	1	4	++	+++	+++	+++	+++	-20	2	Hyphae	+++	+++	++	++	-	*Penicillium brevicompactum*
YHS296	HS93	F	4	> 64	0. 5	2	+	+++	+++	+++	+++	-33.33	3	Hyphae	+++	+++	+++	++	-	*Aspergillus cristatus*
YHS297	HS94	F	4	0. 5	0. 125	1	+	+++	+++	+++	+++	-33.33	1	nd	+++	++	+++	++	-	*Candida albicans*
YHS298	HS95	F	10	0. 125	0. 125	4	++	+++	+++	+++	+++	-26.67	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS299	HS96	F	12	0. 5	0. 125	1	+	+++	+++	+++	+++	-26.67	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS300	HS97	M	6	> 64	0. 5	2	++	+++	+++	+++	+++	-26.67	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS301	HS98	F	16	> 64	1	> 8	+	+++	+++	+++	+++	-20	0	nd	+++	++	+++	++	-	*Candida parapsilosis*
YHS302	HS99	F	3	0. 5	0. 125	> 8	++	++	+++	+++	+++	-53.33	2	Pseudohyphae	+++	++	+	-	-	*Candida parapsilosis*
YHS303	HS99	F	3	0. 5	0. 125	0. 031	++	+++	+++	+++	+++	-53.33	2	Pseudohyphae	+++	+++	+	+	-	*Candida parapsilosis*
YHS304	HS100	M	0	0. 5	0. 5	2	++	+++	+++	+++	+++	4.94	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS305	HS101	M	4	0. 125	0. 125	0. 0156	-	+++	+++	+++	+++	23.46	3	Hyphae	+++	+++	-	-	-	*Candida albicans*
YHS306	HS102	F	13	0. 25	0. 125	> 8	++	++	-	-	-	-60	0	nd	++	+	-	-	-	*Rhodotorula mucilaginosa*
YHS307	HS102	F	13	0. 125	0. 125	4	++	+++	+	+	+	-56.67	3	Hyphae	+++	+++	+	-	-	*Candida pararugosa*
YHS308	HS102	F	13	0. 5	0. 125	0. 0625	+	++	+	+	+	-100	3	Hyphae	+++	++	+	-	-	*Mucor circinelloides*
YHS309	HS102	F	13	4	0. 125	0. 0156	+	+	+++	+++	+++	20	3	Hyphae	++	+	+	-	-	*Pleurostomophora richardsiae*
YHS310	HS103	M	7	> 64	0. 125	0. 031	+	++	++	++	++	-100	3	Hyphae	+++	++	+++	++	-	*Mucor circinelloides*
YHS311	HS104	M	4	> 64	1	4	++	+++	+++	+++	+++	17.28	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS312	HS105	F	8	> 64	0. 125	> 8	+	+++	+++	+++	+++	-33.33	1	Pseudohyphae	+++	+++	+++	++	-	*Candida parapsilosis*
YHS313	HS105	F	8	> 64	0. 125	4	++	+++	+++	+++	+++	-33.33	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS314	HS105	F	8	> 64	0. 125	2	+	+++	+++	+++	+++	-33.33	1	Hyphae	+++	+++	+++	+++	-	*Candida albicans*
YHS315	HS105	F	8	> 64	0. 125	4	+	+++	+++	+++	+++	-33.33	0	nd	+++	+++	+++	++	-	*Candida albicans*
YHS316	HS105	F	8	> 64	0. 125	4	+	+++	+++	+++	+++	-26.67	0	nd	+++	+++	+++	++	-	*Candida albicans*
YHS317	HS105	F	8	> 64	0. 125	0. 25	+	+++	+++	+++	+++	-100	0	nd	+++	+++	+++	+	-	*Candida metapsilosis*
YHS318	HS105	F	8	> 64	0. 125	2	++	+++	+++	+++	+++	-20	0	nd	+++	+++	+++	++	-	*Candida albicans*
YHS319	HS105	F	8	> 64	0. 125	> 8	+	+++	+++	+++	++	-20	2	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS320	HS105	F	8	> 64	0. 125	8	++	+++	+++	+++	+++	-33.33	1	Hyphae	+++	+++	+++	++	-	*Candida albicans*

YHS321	HS105	F	8	32	0. 25	> 8	++	+++	+++	+++	++	-26.67	2	Hyphae	+++	+++	+++	++	+	*Candida albicans*
YHS322	HS105	F	8	> 64	0. 125	2	+	+++	+++	+++	+++	-20	2	Hyphae	+++	+++	+++	+++	-	*Candida albicans*
YHS323	HS105	F	8	> 64	0. 125	2	+	+++	+++	+++	+++	-33.33	1	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS324	HS105	F	8	> 64	0. 125	2	+	+++	+++	+++	+++	-26.67	1	Hyphae	+++	+++	+++	+++	-	*Candida albicans*
YHS325	HS105	F	8	> 64	0. 125	4	+	+++	+++	+++	+++	-33.33	1	Hyphae	+++	+++	+++	+++	-	*Candida albicans*
YHS326	HS105	F	8	> 64	0. 125	8	+	+++	+++	+++	+++	-33.33	1	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS327	HS105	F	8	> 64	1	8	+	+++	+++	+++	+++	-33.33	1	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS328	HS105	F	8	> 64	0. 125	8	++	+++	+++	+++	+++	-26.67	1	Hyphae	+++	+++	+++	++	-	*Candida albicans*
YHS329	HS105	F	8	64	0. 125	8	++	+++	+++	+++	+++	-26.67	2	Hyphae	+++	+++	++	++	-	*Candida albicans*
YHS330	HS105	F	8	32	0. 125	4	+	+++	+++	+++	+++	-26.67	2	Hyphae	+++	+++	+++	+++	-	*Candida albicans*
YHS331	HS105	F	8	> 64	0. 125	0. 0156	++	+++	+++	+++	+++	-33.33	2	Hyphae	+++	+++	+++	+++	-	*Candida albicans*
YHS332	HS105	F	8	> 64	0. 125	0. 0156	++	+++	+++	+++	+++	-33.33	2	Hyphae	+++	+++	+++	+++	-	*Candida albicans*
YHS333	HS106	F	5	> 64	0. 125	8	++	+++	++	++	++	-33.33	0	nd	+	+++	+	-	-	*Rhodotorula mucilaginosa*
YHS334	HS106	F	5	0. 125	0. 125	0. 031	-	+	+	+	+	-100	0	nd	+	++	+	-	-	*Rhodotorula mucilaginosa*
YHS335	HS106	F	5	0. 25	0. 125	0. 031	++	+	+	+	+	-100	0	nd	+	+	+	-	-	*Rhodotorula mucilaginosa*
YHS336	HS106	F	5	0. 125	0. 125	0. 0156	-	+	+	+	+	-100	0	nd	+	+	+	-	-	*Rhodotorula mucilaginosa*
YHS337	HS106	F	5	> 64	0. 125	4	+	++	-	-	-	-100	0	nd	+	+	-	-	-	*Rhodotorula mucilaginosa*
YHS338	HS106	F	5	> 64	0. 125	0. 0156	-	++	+	+	+	-100	0	nd	+++	+++	+	+	-	*Rhodotorula mucilaginosa*
YHS339	HS106	F	5	> 64	0. 125	0. 0156	++	+++	++	++	++	-100	0	nd	+	+	+	-	-	*Rhodotorula mucilaginosa*
YHS340	HS106	F	5	> 64	0. 125	0. 0156	+	+	++	++	++	-100	0	nd	+	+	+	-	-	*Rhodotorula mucilaginosa*
YHS341	HS106	F	5	> 64	0. 125	0. 0156	++	+	++	++	++	-100	0	nd	++	+	+	+	-	*Rhodotorula mucilaginosa*
YHS342	HS106	F	5	> 64	0. 5	0. 25	-	+	++	++	++	-100	0	nd	+	+	+	-	-	*Rhodotorula mucilaginosa*
YHS343	HS106	F	5	> 64	0. 125	0. 0156	++	+	+	+	+	-100	0	nd	++	+	+	-	-	*Rhodotorula mucilaginosa*
YHS344	HS106	F	5	> 64	0. 125	0. 0156	-	+	+	+	+	-100	0	nd	+++	+++	-	-	-	*Rhodotorula mucilaginosa*
YHS345	HS106	F	5	> 64	0. 125	0. 0156	-	++	++	++	++	-100	0	nd	+	-	-	-	-	*Rhodotorula mucilaginosa*
YHS346	HS106	F	5	> 64	0. 125	0. 0156	-	+++	++	++	++	-100	0	nd	+++	+	+	-	-	*Rhodotorula mucilaginosa*
YHS347	HS106	F	5	> 64	0. 125	2	++	++	++	++	++	-100	0	nd	++	+	-	-	-	*Rhodotorula mucilaginosa*
YHS348	HS106	F	5	> 64	0. 125	0. 0156	-	+	++	++	++	-100	0	nd	++	+	+	-	-	*Rhodotorula mucilaginosa*
YHS349	HS106	F	5	> 64	0. 125	0. 0156	-	++	++	++	++	-100	0	nd	++	++	+	-	-	*Rhodotorula mucilaginosa*

*calculated as the deviation of the inhibition halo diameter (Ø) from that of the M28-4D *S. cerevisiae* strain, according to the following formula: (Ø sample – Ø M284D strain) / Ø M284D strain *100. [#], 0= non-invasive; 1= poor invasive; 2= invasive; 3= very invasive. -, no growth as measured by $OD_{630} \leq 0.2$ or cfu/ml $\leq 10^5$; +, poor growth as measured by $0.2 < OD_{630} \leq 0.7$ or $10^5 <$ cfu/ml$\leq 10^6$; ++ good growth as measured by $0.7 < OD_{630} \leq 1.2$ or $10^6 <$ cfu/ml$\leq 10^7$; +++, very good growth as measured by $OD_{630} > 1.2$ or cfu/ml$> 10^7$.
na, not applicable; nd, not detected

Table S3: Permutational multivariate analysis of variance (PERMANOVA) tests on unweighted and weighted UniFrac distances and Bray-Curtis dissimilarity.

	Metric	**F**	**R^2**	***p*-value**
between genders	unweighted Unifrac	3.15	0.054	0.033
	weighted Unifrac	2.48	0.043	0.07
	Bray-Curtis	4.80	0.08	0.001
among age groups	unweighted Unifrac	1.79	0.092	0.031
	weighted Unifrac	1.27	0.067	0.268
	Bray-Curtis	1.69	0.087	0.11

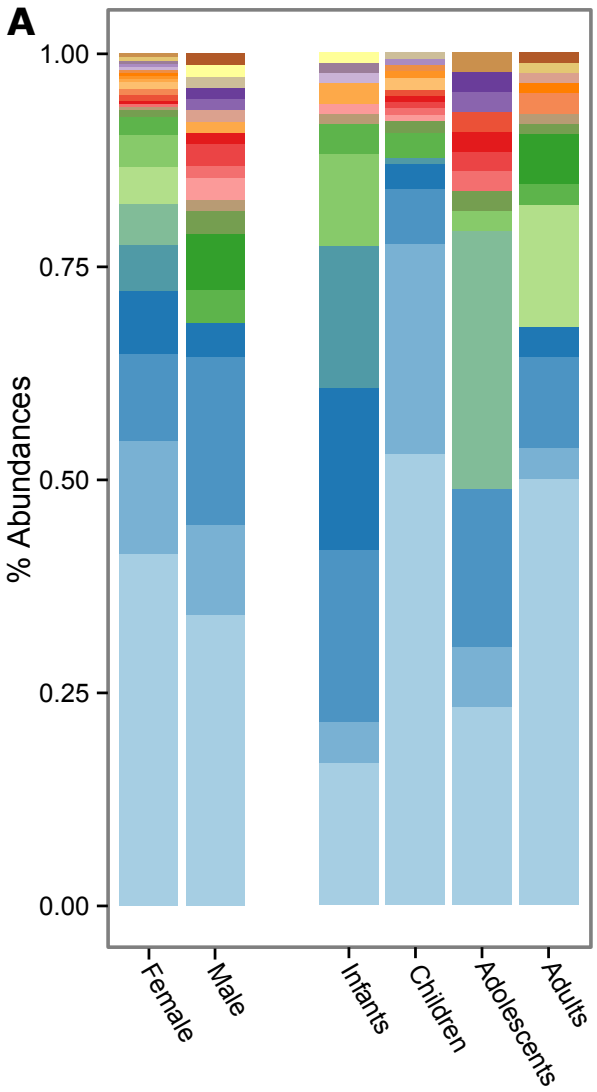
Table S4: Mean relative abundance (%) of OTUs at the genus level of fungal gut microbiota of healthy subjects from metagenomics analysis.

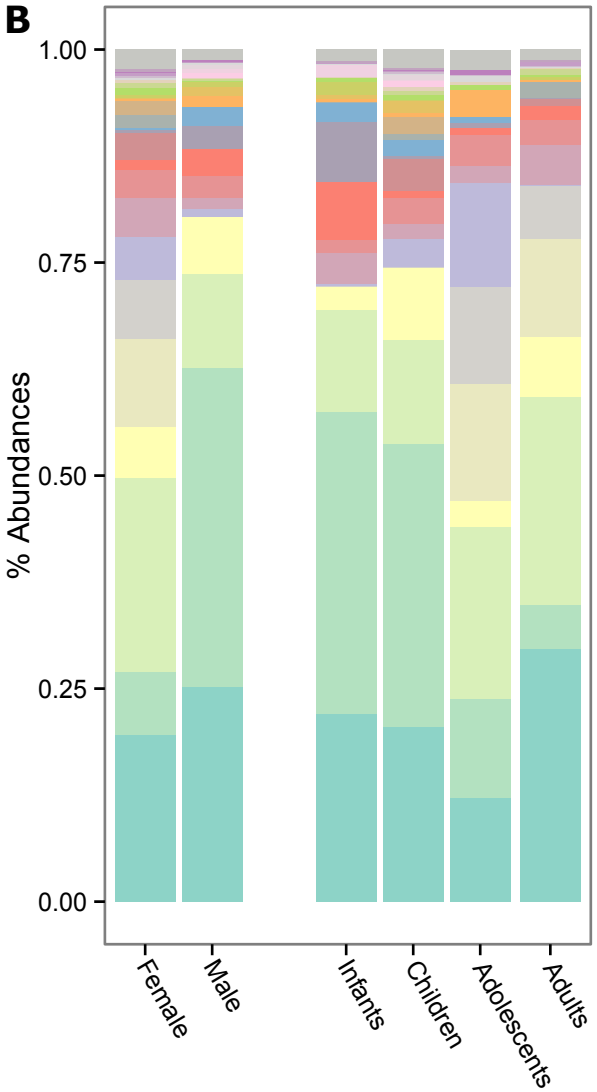
	mean relative abundance (%)						
Taxonomy	**Total**	**Male**	**Female**	**Infants**	**Children**	**Adolescents**	**Adults**
Penicillium	22.362	25.184	19.637	22.041	20.558	12.233	29.652
Aspergillus	22.202	37.509	7.423	35.442	33.227	11.552	5.156
Candida	16.918	10.975	22.657	11.963	12.165	20.226	24.427
Fungi_unidentified_1_1	6.330	6.716	5.958	2.731	8.525	3.110	7.123
Blastocystis	5.311	0.000	10.439	0.000	0.000	13.630	11.393
Pichia	3.485	0.045	6.806	0.126	0.004	11.382	6.248
Mucor	3.057	0.885	5.154	0.233	3.317	12.196	0.080
Debaryomyces	2.979	1.322	4.578	3.644	1.683	2.065	4.693
Malassezia	2.871	2.541	3.190	1.458	3.145	3.608	3.001
Ascomycota_unidentified_1_1	2.133	3.148	1.153	6.803	0.813	0.862	1.692
Eremothecium	1.671	0.010	3.274	0.039	3.671	0.000	0.829
Eurotiomycetes_unidentified_1	1.500	2.690	0.352	7.065	0.459	0.534	0.029
Tremellomycetes_unidentified_1	1.200	2.190	0.245	2.167	1.816	0.680	0.080
Cyberlindnera	0.841	0.108	1.548	0.155	0.724	0.024	1.778
Mucoraceae_unidentified	0.795	0.007	1.555	0.039	1.970	0.024	0.080
Saccharomyces	0.713	1.243	0.201	0.301	0.437	3.122	0.177
Ustilago	0.622	1.062	0.198	0.428	1.391	0.073	0.000
Nectriaceae_unidentified	0.469	0.580	0.362	1.623	0.230	0.000	0.320
Rhodotorula	0.450	0.132	0.757	0.437	0.517	0.595	0.303
Malasseziales_unidentified_1	0.385	0.222	0.543	0.010	0.464	0.012	0.680
Wallemia	0.298	0.132	0.459	0.097	0.552	0.279	0.097
Xeromyces	0.290	0.569	0.020	0.010	0.747	0.000	0.000
Trichosporon	0.247	0.496	0.007	1.390	0.009	0.000	0.000
Preussia	0.237	0.482	0.000	0.000	0.614	0.000	0.000
Pleosporales_unidentified_1	0.188	0.236	0.141	0.068	0.084	0.790	0.109
Phoma	0.174	0.035	0.308	0.097	0.278	0.012	0.160
Trichocomaceae_unidentified	0.174	0.021	0.322	0.049	0.031	0.000	0.514
Aureobasidium	0.167	0.236	0.101	0.000	0.221	0.583	0.000
Botrytis	0.162	0.014	0.305	0.194	0.186	0.061	0.160
Ustilaginomycetes_unidentified_1	0.095	0.010	0.178	0.000	0.000	0.680	0.000
Helotiales_unidentified_1	0.085	0.174	0.000	0.000	0.000	0.607	0.000
Saccharomycetales_unidentified_1	0.085	0.010	0.158	0.117	0.119	0.000	0.063
Talaromyces	0.080	0.017	0.141	0.311	0.022	0.000	0.057
Ustilaginaceae_unidentified	0.080	0.007	0.151	0.000	0.208	0.000	0.000
Rhizopus	0.080	0.024	0.134	0.000	0.053	0.000	0.200
Cryptococcus	0.078	0.017	0.137	0.049	0.000	0.000	0.234
Torulaspora	0.077	0.087	0.067	0.039	0.066	0.255	0.029
Cordyceps	0.075	0.021	0.127	0.000	0.159	0.097	0.000
Sarcosomataceae_unidentified	0.072	0.000	0.141	0.000	0.186	0.000	0.000
Urocystis	0.072	0.146	0.000	0.000	0.186	0.000	0.000
Incertae_sedis_12_unidentified	0.070	0.000	0.137	0.000	0.141	0.000	0.051
Davidiella	0.056	0.021	0.090	0.049	0.062	0.012	0.074
Helminthosporium	0.053	0.000	0.104	0.000	0.137	0.000	0.000

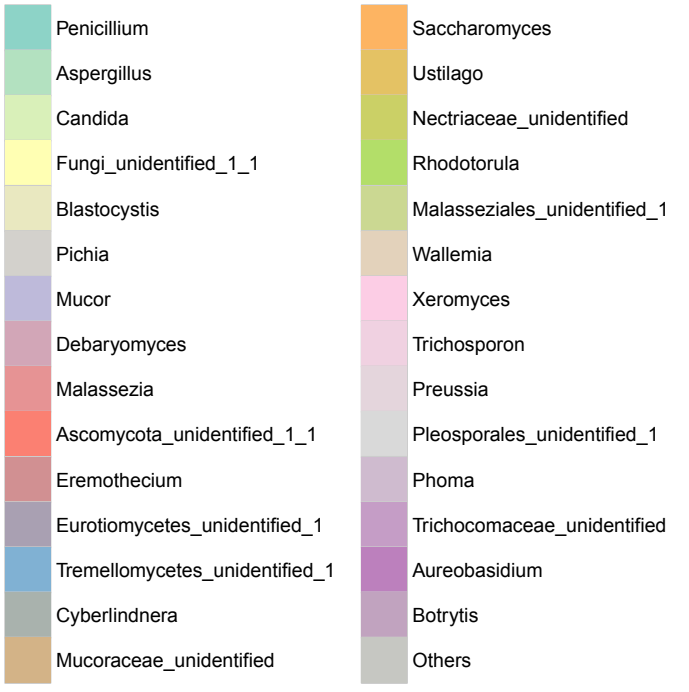
Exophiala	0.051	0.073	0.030	0.204	0.035	0.012	0.000
Wickerhamomyces	0.044	0.017	0.070	0.049	0.000	0.000	0.120
Phaeosphaeriaceae_unidentified	0.039	0.000	0.077	0.000	0.000	0.000	0.131
Basidiomycota_unidentified_1_1	0.038	0.045	0.030	0.068	0.053	0.036	0.000
Tilletia	0.038	0.062	0.013	0.155	0.018	0.024	0.000
Ascosphaera	0.036	0.073	0.000	0.000	0.088	0.000	0.006
Ustilaginaceae_unidentified	0.034	0.000	0.067	0.000	0.088	0.000	0.000
Hypocreales_unidentified_1	0.031	0.042	0.020	0.029	0.009	0.146	0.006
Rasamsonia	0.029	0.031	0.027	0.039	0.027	0.000	0.040
Periconia	0.027	0.056	0.000	0.000	0.071	0.000	0.000
Incertae_sedis_25_unidentified	0.027	0.007	0.047	0.000	0.049	0.024	0.017
Udeniomyces	0.027	0.000	0.054	0.000	0.071	0.000	0.000
Dioszegia	0.026	0.007	0.044	0.019	0.053	0.012	0.000
Thermomyces	0.026	0.031	0.020	0.087	0.000	0.000	0.034
Sporobolomyces	0.020	0.035	0.007	0.058	0.022	0.012	0.000
Lophiostoma	0.017	0.035	0.000	0.000	0.000	0.121	0.000
Embellisia	0.015	0.028	0.003	0.000	0.004	0.097	0.000
Saccharomycetaceae_unidentified	0.015	0.003	0.027	0.010	0.000	0.000	0.046
Guehomyces	0.015	0.007	0.023	0.019	0.031	0.000	0.000
Puccinia	0.012	0.010	0.013	0.010	0.018	0.024	0.000
Pseudozyma	0.012	0.003	0.020	0.010	0.000	0.000	0.034
Ustilaginales_unidentified_1	0.012	0.024	0.000	0.000	0.009	0.061	0.000
Diatrypaceae_unidentified	0.009	0.017	0.000	0.000	0.022	0.000	0.000
Leucosporidiella	0.009	0.017	0.000	0.049	0.000	0.000	0.000
Endoconidioma	0.007	0.000	0.013	0.000	0.018	0.000	0.000
Chaetothyriales_unidentified_1	0.007	0.007	0.007	0.019	0.000	0.024	0.000
Tetracladium	0.007	0.014	0.000	0.000	0.000	0.049	0.000
Kluyveromyces	0.007	0.000	0.013	0.000	0.000	0.000	0.023
Diplodia	0.005	0.000	0.010	0.000	0.013	0.000	0.000
Sarcinomyces	0.005	0.000	0.010	0.000	0.013	0.000	0.000
Gymnoascus	0.005	0.000	0.010	0.000	0.013	0.000	0.000
Kazachstania	0.005	0.000	0.010	0.000	0.000	0.000	0.017
Arthrinium	0.005	0.000	0.010	0.000	0.013	0.000	0.000
Erythrobasidium	0.005	0.000	0.010	0.000	0.013	0.000	0.000
Dothideomycetes_unidentified_1	0.003	0.003	0.003	0.000	0.000	0.012	0.006
Mycocentrospora	0.003	0.000	0.007	0.000	0.009	0.000	0.000
Knufia	0.003	0.007	0.000	0.000	0.009	0.000	0.000
Hanseniaspora	0.003	0.000	0.007	0.000	0.000	0.000	0.011
Myrothecium	0.003	0.000	0.007	0.000	0.009	0.000	0.000
Sordariomycetes_unidentified_1	0.003	0.000	0.007	0.000	0.009	0.000	0.000
Taphrina	0.003	0.003	0.003	0.000	0.009	0.000	0.000
Wojnowicia	0.002	0.003	0.000	0.000	0.004	0.000	0.000
Podosphaera	0.002	0.000	0.003	0.000	0.004	0.000	0.000
Schizosaccharomyces	0.002	0.003	0.000	0.000	0.000	0.012	0.000
Coniochaeta	0.002	0.000	0.003	0.000	0.000	0.000	0.006
Claviceps	0.002	0.000	0.003	0.000	0.000	0.012	0.000
Sordariaceae_unidentified	0.002	0.003	0.000	0.000	0.004	0.000	0.000

Sakaguchia	0.002	0.003	0.000	0.000	0.000	0.012	0.000
Exobasidiales_unidentified_1	0.002	0.003	0.000	0.000	0.004	0.000	0.000
Quambalaria	0.002	0.000	0.003	0.000	0.000	0.000	0.006
Filobasidium	0.002	0.000	0.003	0.000	0.000	0.000	0.006
Hannaella	0.002	0.000	0.003	0.000	0.004	0.000	0.000
Lichtheimia	0.002	0.000	0.003	0.000	0.004	0.000	0.000

Supplementary Figure 1







Chapter 2

New evidences on the altered gut microbiota in autism spectrum disorders

This chapter has been submitted for publication as an original research article:

Strati F, Cavalieri D, Albanese D, De Felice C, Donati C, Hayek J, Jousson O, Leoncini S, Renzi D, Calabrò A and De Filippo C (2016). *New evidences on the altered gut microbiota in autism spectrum disorders.*

Alterations of the human gut microbiota can contribute to gastrointestinal disorders and inflammation [1] affecting also central nervous system activities through different mechanisms of interaction with the host [2]. Dysbiosis of the gut microbiota have been implicated in a wide variety of neurological disorders, including ASDs [3]. Nevertheless the studies that investigated the human gut microbiota in ASDs, although discovered clear imbalances in the intestinal microbial population of autistic subjects, failed to find common microbial features across all these studies, probably due to differences in sampling strategies and techniques applied [4]. The understanding of the bases leading to the alteration of the microbial community structure in ASDs would be of great importance for the design of therapeutic interventions aimed at the relief of gastrointestinal disorders that often affects these subjects.

The aim of the work presented in this chapter focused on the characterization of the intestinal microbial community structure in autistic subjects, both at bacterial and fungal level, in order to elucidate how the gut microbiota in these individuals may impact on ASDs gastrointestinal pathophysiology. We discovered an altered gut microbiota in autistic subjects characterized by an increased *Bacteroidetes/Firmicutes* ratio and imbalances in several bacterial taxa. We observed that putative pro-inflammatory taxa belonging to *Clostridium XIII* and *Escherichia/Shigella* were enriched in the constipated autistic subject. Furthermore we discovered that the autistic subjects harbour an altered fungal gut microbiota, extending our understanding on the intestinal dysbiotic states associated with ASDs.

Apart from the recruitment and diagnosis of autism in the subjects enrolled in this study, I designed and performed all the experiments, ranging from the extraction of total DNA from stool sample to the analysis of sequencing reads and interpretations of these results by applying state-of-the arts bioinformatics and statistical methods for microbiome research. Finally I wrote the manuscript and generated figures and tables.

References

1. Maslowski KM, Mackay CR: **Diet, gut microbiota and immune responses**. *Nature immunology* 2011, **12**(1):5-9.
2. Sampson TR, Mazmanian SK: **Control of brain development, function, and behavior by the microbiome**. *Cell host & microbe* 2015, **17**(5):565-576.
3. Wang Y, Kasper LH: **The role of microbiome in central nervous system disorders**. *Brain, behavior, and immunity* 2014, **38**:1-12.
4. Mayer EA, Padua D, Tillisch K: **Altered brain-gut axis in autism: comorbidity or causative mechanisms?** *BioEssays : news and reviews in molecular, cellular and developmental biology* 2014, **36**(10):933-939.

New evidences on the altered gut microbiota in autism spectrum disorders

Francesco Strati[1,2], Duccio Cavalieri[3], Davide Albanese[1], Claudio De Felice[4], Claudio Donati[1], Joussef Hayek[5], Olivier Jousson[2], Silvia Leoncini[5], Daniela Renzi[6], Antonio Calabrò[6], Carlotta De Filippo[7,*]

[1] Computational Biology Research Unit, Research and Innovation Centre, Fondazione Edmund Mach, Via E. Mach 1, 38010 San Michele all' Adige, Italy

[2] Centre for Integrative Biology, University of Trento, Via Sommarive 9, 38123 Trento, Italy

[3] Department of Biology, University of Florence, Via Madonna del Piano 6, 50019 Sesto Fiorentino, Florence, Italy

[4] Neonatal Intensive Care Unit, Siena University Hospital AOUS, Viale Bracci 16, 53100 Siena, Italy

[5] Child Neuropsychiatry Unit, Siena University Hospital AOUS, Viale Bracci 16, 53100 Siena, Italy

[6] Department of Experimental and Clinical Biomedical Sciences, Gastroenterology Unit, University of Florence, Viale Morgagni 40, 50139 Florence, Italy

[7] Institute of Biometeorology, National Research Council (CNR), Via Giovanni Caproni 8, 50145 Florence, Italy

Correspondence to: Carlotta De Filippo, Institute of Biometeorology, National Research Council (CNR), Via Giovanni Caproni 8 - I-50145 Florence, Italy; e-mail: *c.de.filippo@ibimet.cnr.it*

Abstract

Background

Autism Spectrum Disorders (ASDs) are neurodevelopmental conditions characterized by social and behavioural impairments. In addition to neurological symptoms, ASDs subjects frequently suffer from gastrointestinal abnormalities, thus implying a role of the gut microbiota in ASDs gastrointestinal pathophysiology.

Results

Here we characterized the bacterial and fungal gut microbiota in a cohort of autistic individuals demonstrating the presence of an altered microbial community structure. A fraction of 90% of the autistic subjects were classified as severe ASDs. We observed a reduced relative abundance of *Bacteroidetes* in the ASDs group with *Collinsella, Corynebacterium, Dorea* and *Lactobacillus* being the bacterial taxa characterizing the gut microbiota of the ASDs cohort. On the contrary, the relative abundance of the bacterial taxa *Alistipes*, *Bilophila, Dialister*, *Parabacteroides* and *Veillonella* were significantly reduced in ASDs subjects. Furthermore constipation has been associated with different bacterial patterns in autistic and neurotypical subjects, with constipated autistic individuals characterized by high levels of bacterial taxa belonging to *Escherichia/Shigella* and *Clostridium cluster XVIII*. We also observed that the fungal genus *Candida* was more than two times more abundant, in terms of relative abundance, in autistic than neurotypical subjects.

Conclusions

The finding that, besides the bacterial gut microbiota, also the gut mycobiota contributes to the alteration of the intestinal microbial community structure in autism opens the possibility for new potential intervention strategies aimed at the relief of gastrointestinal symptoms in ASDs.

Keywords: Autism spectrum disorders, gut microbiota, mycobiota, constipation, metataxonomy

Introduction

The term "autism spectrum disorders" (ASDs) refer to a group of neurodevelopmental disorders with an early life stage onset characterized by alterations in social interactions and communication and by restricted and repetitive behaviour [1]. It is now well accepted the contribution of both genetic and environmental factors in the aetiology of ASDs [2, 3]. Among the non-neurological symptoms associated with ASDs, several studies indicate gastrointestinal (GI) symptoms as common comorbidities [4-7]. Alterations in the composition of the gut microbiota have been implicated in a wide variety of human diseases, including ASDs [8]. Since the gut microbiota makes critical contributions to metabolism, maintenance of immune homeostasis and may control central nervous system (CNS) activities through neural, endocrine and immune pathways [9] it has been hypothesized the active role of the gut microbiota in ASDs pathophysiology. There is more than a subtle link between the gut microbiota and the CNS, through the so-called "microbiome-gut-brain axis". Indeed it has been demonstrated a direct interaction between the gut microbiota and enteric neurons [10, 11], its role in the regulation of the HPA axis [12] and the production of many chemicals important in brain functioning (*e.g.* serotonin, dopamine, kynurenine, γ-aminobutyric acid, SCFAs, *p*-cresol) [13, 14]. A dysbiotic microbial community could lead to systemic inflammation due to hyper-activation of T-helper 1 and T-helper 17 cell responses [15] affecting also the reactivity of peripheral immune cells to the CNS [16] and the integrity of blood-brain barrier [17] which is known to be altered in ASDs [18]. Several evidences suggested an early immune activation with chronic inflammation and cytokine dysregulation in ASDs [19, 20] and it has been shown that systemic inflammation induced by LPS provokes behavioural changes and impairs the blood-brain barrier in animal models [17, 21]. Furthermore, fungal infections that may originate from alterations in commensal bacteria population [22], could shift the indoleamine 2,3-dioxygenase's activity [23, 24] reducing the levels of kynurenine [25], a neuroprotective agent. Despite several reports disclosed an aberrant gut microbiota in ASDs, consensus across studies has not yet been established [8]. Here we characterized the bacterial gut microbiota and the less studied gut mycobiota of subjects affected by autism through amplicon-based metataxonomics analysis of the V3-V5 region of the prokaryotic 16S rDNA and of the ITS1 region of the fungal rDNA in order to better understand the microbial community structure associated with ASDs and its impact on GI abnormalities.

Results

Autistic subjects harbour an altered bacterial gut microbiota

For the characterization of the gut microbiota associated with autism we recruited 40 autistic subjects (36 out of 40 autistic subjects were classified as severe ASDs, CARS value >37) and 40 neurotypical controls (Table 1, Supplementary Table 1).

Table 1: Characteristics of study participants

	Autistic	**Neurotypical**
Subjects (n)	40	40
Age (1st - 3rd quartile)	10 (5 - 17)	7 (3.6 - 12)
Gender (n)		
female	22.5% (9)	30% (12)
male	77.5% (31)	70% (28)
Constipation (n)		
constipated	12.5% (5)	27.5% (11)
non-constipated	72.5% (29)	72.5% (29)
NA	15% (6)	0% (0)
Calprotectin (1st - 3rd quartile)	36.9 (17.6 - 76.0) µg/g	40.9 (17 - 74.7) µg/g
constipated	39.1 (22.9 - 70.0) µg/g	27.9 (20.3 - 97.6) µg/g
non-constipated	35.9 (15.0 - 57.8) µg/g	50.5 (15.0 - 73.8) µg/g
CARS (1st - 3rd quartile)	47 (40 - 50.5)	NA
constipated	50 (36 - 52.0)	NA
non-constipated	48 (42 - 50.0)	NA
ESR (1st - 3rd quartile)	7.5 (3.25 - 17.7) mm/h	NA
constipated	22.0 (12.0 - 25.0) mm/h	NA
non-constipated	7.0 (2.7 - 11.2) mm/h	NA
Serum IgA (1st - 3rd quartile)	131.0 (70.0 - 172.2) mg/ml	NA
constipated	97.0 (82.0 - 153.0) mg/ml	NA
non-constipated	133.0 (67.0 - 181.0) mg/ml	NA

Data expressed as medians with interquartile ranges when applicable. AD, autistic subjects; NT, neurotypical subjects; NA, not applicable; CARS, childhood autism rating scale; ESR, erythrocyte sedimentation rate.

Analysis of *alpha*-diversity revealed no significant differences between autistic and neurotypical subjects (hereinafter termed AD and NT respectively). However the analysis of the *beta*-diversity calculated on the unweighted, weighted UniFrac distances and the Bray-Curtis dissimilarity, revealed that the bacterial microbiota of AD clusters apart from that of NT ($p<0.05$, PERMANOVA; Figure 1, Supplementary Table 2). Since we enrolled subjects suffering from constipation among NT and AD subjects, the impact of constipation on the *beta*-diversity of the two groups of study was also analysed. We observed that constipation has a significant effect on the microbial community structure within NT subjects ($p<0.05$, PERMANOVA), as expected [26, 27], but not within AD subjects (Supplementary Table 2).

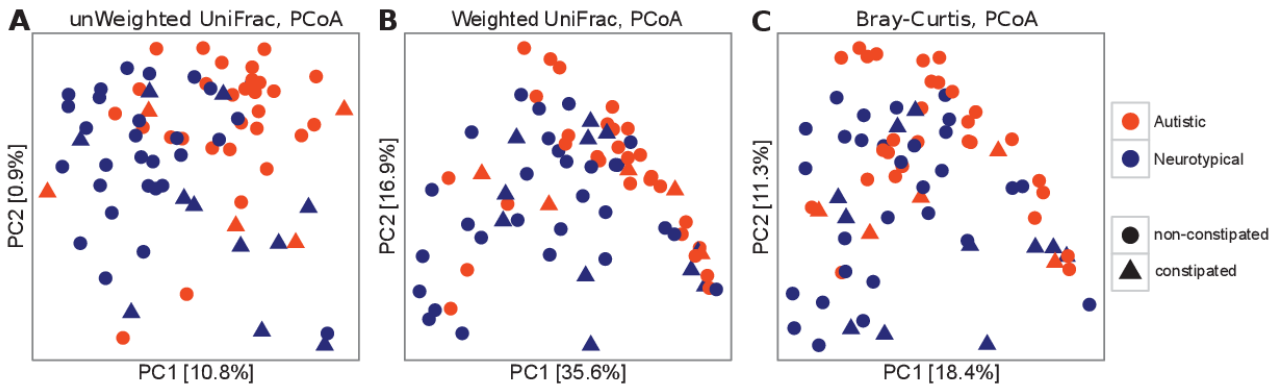


Figure 1: PCoA of bacterial *beta*-diversity based on **A)** the unweighted and **B)** weighted UniFrac distances and **C)** Bray-Curtis dissimilarity. Autistic and neurotypical subjects are coloured in orange and blue, respectively. The constipation status of the subjects is indicated according to different shapes, circles for non-constipated and triangles for constipated individuals.

Furthermore we observed that the severity of the autistic phenotype, as measured by CARS scores, does not affect the bacterial community structure among AD individuals (p>0.05, PERMANOVA; Supplementary Table 3). Phylum level analysis showed a clear alteration of the bacterial gut community in AD characterized by a higher *Firmicutes/Bacteroidetes* ratio (p<0.005, Wilcoxon rank-sum test; Figure 2A) in AD than NT due to a significant reduction of the relative abundance of *Bacteroidetes* (9.2% AD, 19.4% NT) (FDR-corrected p<0.05, Welch t-test; Figure 2B).

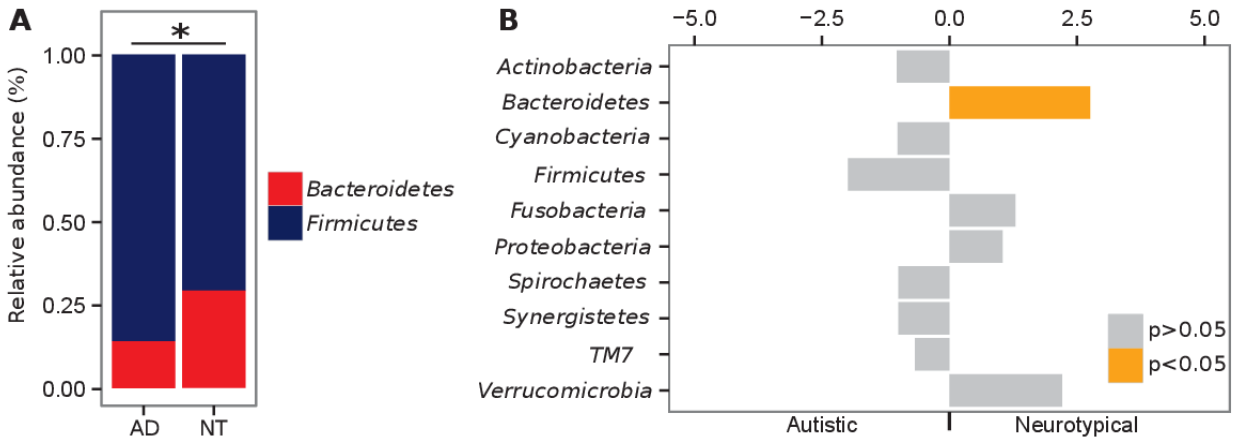


Figure 2: A) Mean relative abundances (%) of *Firmicutes* and *Bacteroidetes* in autistic (AD) and neurotypical (NT) subjects; * p<0.005, Wilcoxon sum-rank test on the *Firmicutes/Bacteroidetes* ratio. **B)** Welch's t-test statistics of the relative abundances of bacterial phyla in autistic and neurotypical subjects. Orange bars indicate significant FDR-corrected p-values adjusted for multiple comparison controlling the family-wise Type I error rate.

Genus level analysis showed that the top ten most abundant genera in both AD and NT subjects were *Bifidobacterium, Bacteroides, Faecalibacterium, Unknown Lachnospiraceae, Blautia, Ruminococcus, Clostridium XI, Streptococcus, Gemmiger* and *Lachnospiracea incertae sedis* (Supplementary Fig. 1, Supplementary Table 4). Interestingly, the genus *Prevotella* was only barely

represented in AD with respect to NT (0.05% AD, 1.5% NT), in agreement with a previous study on the gut microbiota in ASD children [28] although this difference of relative abundance was not supported by the statistical analysis. We further analysed the bacterial community structure associated with AD and NT by using LEfSe, an algorithm for high-dimensional biomarker discovery which uses linear discriminant analysis (LDA) to estimate the effect size of each taxa which are differentially represented in cases and controls [29]. LEfSe analysis revealed a significant increase of the relative abundance of different bacterial taxa in AD than in NT among which *Collinsella*, *Corynebacterium*, *Dorea* and *Lactobacillus* and a significant reduction of the taxa *Alistipes*, *Bilophila, Dialister*, *Parabacteroides* and *Veillonella* in AD than in NT (p<0.01, Wilcoxon rank-sum test; LDA>2.0; Figure 3).

Constipation select different bacterial taxa in autistic subjects and neurotypical healthy controls

Autistic subjects frequently suffer of GI comorbidities [4-7] and constipation is a GI symptom often reported in these subjects, known to alter the physiology of the human GI tract and the gut microbiota itself [27, 30, 31]. Correlation analysis of the bacterial relative abundances between constipated and non-constipated subjects, both autistic and neurotypical, revealed that among the most abundant bacterial genera (with relative abundance >0.5% and detectable in at least the 70% of the investigated subjects) the taxa *Gemmiger* and *Ruminococcus* anti-correlates with the constipation status (Spearman's correlation r= -0.39 and -0.36, respectively; FDR-corrected p<0.05; Supplementary Table 5) while *Escherichia/Shigella* and *Clostridium cluster XVIII* positively correlates with this GI symptom (Spearman's correlation r= 0.31 and 0.38, respectively; FDR-corrected p<0.05; Supplementary Table 5). We further compared the relative abundance of these taxa among constipated and non-constipated subjects within and between groups. We observed that *Escherichia/Shigella* and *Clostridium cluster XVIII* were significantly more abundant in constipated AD compared to the non-constipated ones (FDR-corrected p<0.05, Wilcoxon rank sum test; Figure 4A and 4B) while no differences have been detected between constipated and non-constipated NT for these taxa. On the other hand, the genus *Gemmiger* was significantly less abundant in constipated compared to non-constipated NT (FDR-corrected p<0.05, Wilcoxon rank sum test; Figure 4C). Remarkably, no significant differences have been observed in the levels of faecal calprotectin between AD and NT as well as between constipated and non-constipated subjects in both groups (Table 1 and Supplementary Table 1). Furthermore we analysed the levels of other two biomarkers of inflammations *i.e.* serum IgA and ESR in the autistic subjects and we did not observe significant differences among constipated and non-constipated AD (Table 1 and Supplementary Table 1).

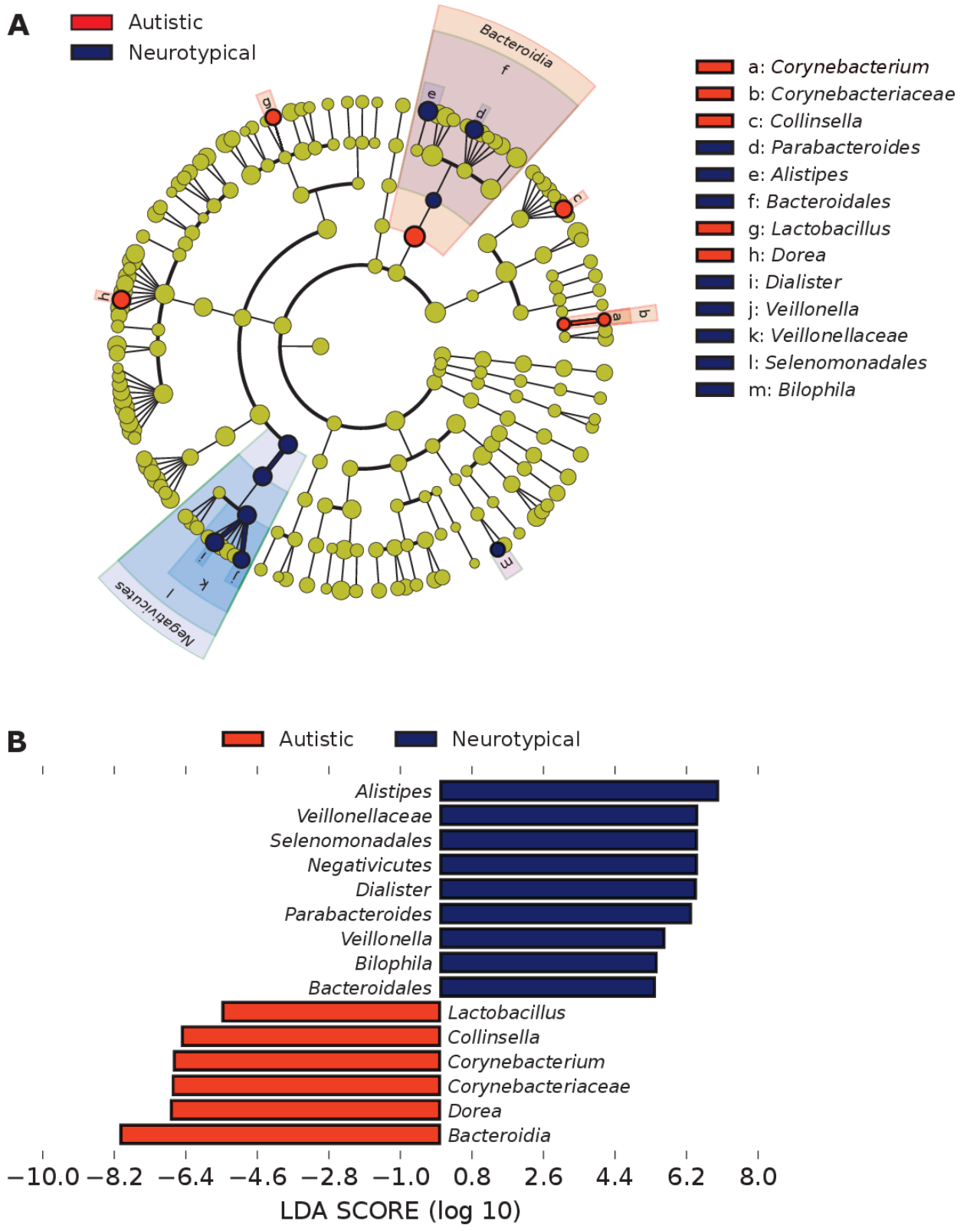


Figure 3: A) Cladograms generated by LEfSe indicating differences in the bacterial taxa between autistic (AD) and neurotypical (NT) subjects. Nodes in orange indicate taxa that were enriched in AD compared to those in NT, while nodes in blue indicate taxa that were enriched in NT compared to those in AD. **B)** LDA scores for the bacterial taxa differentially abundant between autistic (AD) and neurotypical (NT) subjects. Positive and negative LDA scores indicate the bacterial taxa enriched in NT and AD subjects, respectively. Only the taxa having a p<0.01 (Wilcoxon rank-sum test) and LDA> 2.0 are shown in the figure legend

Therefore, while constipation resulted in a significant increase of *Escherichia/Shigella* and *Clostridium cluster XVIII* no differences have been observed in the levels of inflammation between constipated and non-constipated autistic subjects suggesting that constipation and the related alterations of the gut microbiota in autistic subjects as well as in neurotypical individuals is not associated with an increase of intestinal inflammation. It should be noted that the number of enrolled constipated subjects was quite low and therefore these analyses could be underpowered.

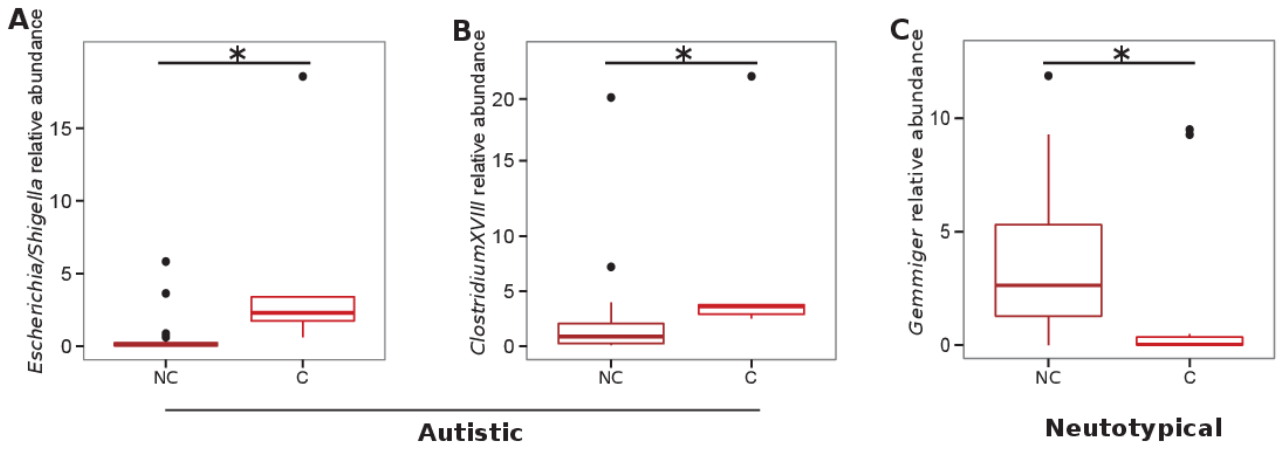


Figure 4: Box plot representation of the relative abundances of bacterial genera correlating with the constipation status of the subjects enrolled in this study. Comparisons between **A, B)** constipated (C) and non-constipated (NC) autistic subjects and **C)** constipated (C) and non-constipated (NC) neurotypical subjects; * FDR-corrected $p<0.05$, Wilcoxon sum-rank test.

Autistic subjects harbour an altered gut mycobiota

We then investigated the gut mycobiota of our study cohort through amplicon-based sequencing of fungal ITS1 region. High quality fungal sequences were detected respectively in 35 out of 40 autistic subjects and 38 out of 40 NT. As occurred for the bacterial gut microbiota we did not observe significant differences in fungal *alpha*-diversity between AD and NT. The analysis of *beta*-diversity revealed that the gut mycobiota of AD was different compared to NT as calculated by PCoA and PERMANOVA on the weighted UniFrac distance and Bray-Curtis dissimilarity ($p<0.05$, Figure 5). As for the bacterial *beta*-diversity, constipation showed a significant effect within NT subjects ($p=0.046$, PERMANOVA on Bray-Curtis dissimilarities) but not within AD subjects (Supplementary Table 6). Furthermore the severity of the autistic phenotype does not affect the gut mycobiota community structure among AD individuals ($p>0.05$, PERMANOVA; Supplementary Table 3). An in-depth analysis of the gut mycobiota lead to the identification of 50 fungal taxa fully classified to the genus level and 30 only partially classified. Genus level analysis showed *Aspergillus* (24.2% AD; 28% NT), *Candida* (37.7% AD; 14.1% NT), *Penicillium* (13.2% AD; 23.5% NT) and *Malassezia* (3.05% AD; 3.3% NT) as the most abundant and widely distributed genera in our study cohort in terms of relative abundance (Supplementary Fig. 2, Supplementary Table 7).

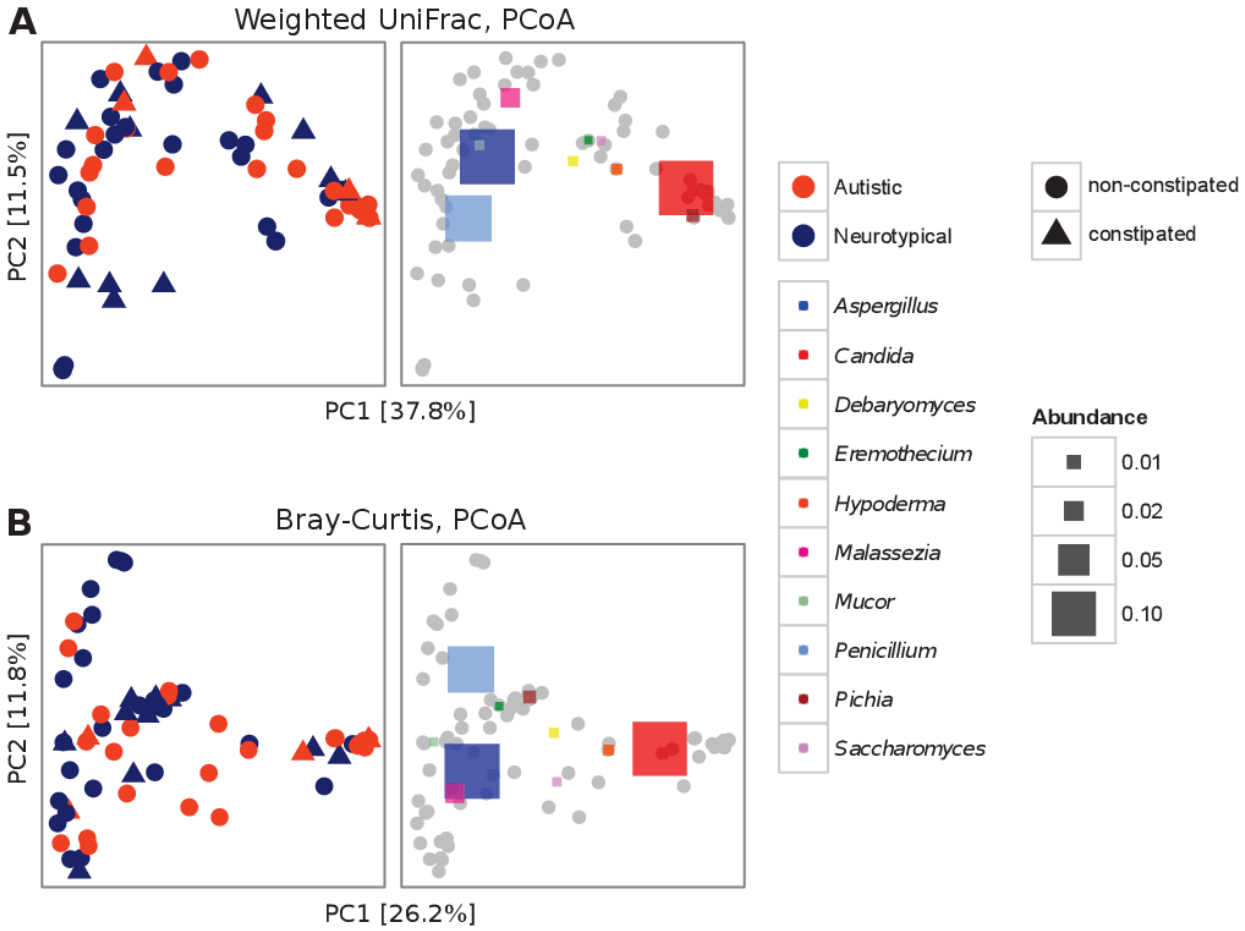


Figure 5: PCoAs of fungal *beta*-diversity based on **A)** weighted UniFrac distance and **B)** Bray-Curtis dissimilarity. The right panel of the graphs **A** and **B** shows the same PCoA coordinates with the most abundant OTUs superimposed as coloured squares, with the size being proportional to the mean relative abundance of the taxon across all samples (grey dots). Autistic and neurotypical subjects are coloured in orange and blue, respectively. The constipation status of the subjects is indicated according to different shapes, circles for non-constipated and triangles for constipated individuals.

Although the relative abundance of the genus *Candida* was more than twice as much in AD than NT we did not identify significant differences among the two study groups (Welch *t*-test, FDR-corrected *p-value*=0.09, raw *p-value*=0.006; Supplementary Fig. 3) probably due to a larger dispersion of values in AD when compared to NT ($p<0.001$; Levene's test). However the superimposition of the most abundant genera over the PCoA plots revealed that high levels of *Candida* abundance was associated with a group of subjects mainly affected by autism (Figure 5) suggesting that *Candida* could play a role in the altered microbial community associated with the autistic subjects. Correlation analyses among the most abundant fungi and bacteria (with relative abundance >0.5% and detectable in at least the 70% of the investigated subjects) revealed no significant correlations among autistic subjects while a significant positive correlation between the genera *Aspergillus* and *Bifidobacterium* was found within NT individuals (Spearman's $r= 0.6$, FDR-corrected $p=0.004$) (Supplementary Table 8).

Discussion

The gut microbiota is a crucial factor for the maintenance of GI tract functions and immune homeostasis. It is well known that dysbiosis of the GI tract could lead to inflammation and immune activation in several pathologies [15]. The frequent occurrence of GI symptoms in autistic subjects imply the possible involvement of the gut microbiota in ASDs gastrointestinal pathophysiology, further supported by the speculations on the increased incidence of ASDs cases due to "Western" habits (*i.e.* diet, medications and excessive overall hygiene) that can affect the composition of the gut microbiota [32]. Several studies demonstrated alterations in the bacterial gut microbiota of ASDs individuals, even if the differences reported in these studies were in some cases discordant, possibly due to variance in sampling strategies and methodologies used [8]. In addition, our recent findings showed an altered gut microbiota in Rett syndrome [33], a genetically determined neurodevelopmental disorder previously categorized in the ASDs group sharing some features of these conditions. We characterized the gut microbiota associated with autism, disclosing an altered microbial community both at bacterial and fungal level. We observed a significant increase in the *Firmicutes*/*Bacteroidetes* ratio in autistic subjects due to a significant reduction of *Bacteroidetes* in these individuals. Several inflammatory conditions have been related to an increase in the *Firmicutes*/*Bacteroidetes* ratio such as IBDs [34] and obesity [35]. Consistently with these observations, an increased *Firmicutes*/*Bacteroidetes* ratio has been reported also in subjects with autism [36, 37]. Furthermore we discovered that the relative abundances of the genera *Collinsella*, *Corynebacterium*, *Dorea* and *Lactobacillus* were significantly increased in the gut microbiota of autistic subjects with respect to that of the neurotypical subjects while the relative abundance of the genera *Alistipes*, *Bilophila*, *Dialister*, *Parabacteroides* and *Veillonella* were significantly reduced in these individuals. A recent study on a mouse model of ASDs demonstrated that treatments with a PSA^+ *Bacteroides fragilis* strain restore autism-related behavioural and GI abnormalities, also reducing the reported high levels of *Lachnospiraceae* and 4-ethylphenylsulfate, a metabolite produced by this bacterial family related to *p*-cresol, a putative metabolic marker for autism [38]. Overall, these data are consistent with our findings and remark the importance of *Bacteroidetes* in ASDs pathophysiology. Moreover *Lactobacillus* resulted to be enriched in the gut microbiota of autistic individuals while *Dialister* and *Veillonella* resulted to be depleted, in line with the results obtained in previous studies [28, 39]. Since constipation is a common gastrointestinal problem in subjects with ASDs [4-7] we compared our data between constipated and non-constipated subjects in order to evaluate the contribution of constipation in shaping the gut microbiota of autistic subjects. Indeed it has been proposed that GI symptoms may be related to ASDs [40]. The evidence that the taxa belonging to the *Clostridium cluster XVIII* and the putative pro-inflammatory

Escherichia/Shigella [41, 42] positively correlated with the constipation status of the subjects as well as their enrichment in constipated autistic subjects supports the hypothesis that GI problems and related alterations of the gut microbiota may contribute to ASDs gastrointestinal symptoms [40]. Because of their ability to produce exotoxins and propionate that may exacerbate autistic symptoms [43], the role of clostridia in ASDs has been extensively explored. The species belonging to the *Clostridium cluster XVIII* have been shown to produce exotoxins [44] and to promote conditions favouring inflammation [45, 46] although other studies observed their potential ability to induce homeostatic T-reg responses [47]. It is also interesting to underline the occurrence of a subclinical acute phase response in ASDs plasma, as evidenced by advanced proteomic analysis [48].

Despite the importance given to the implications of the gut microbiota in health and disease few reports have explored the relevance of the fungal component of the gut microbiota in GI (patho)physiology [49]. Furthermore, none of the published studies on ASDs' gut microbiota have assessed the fungal gut community structure associated with autism. Our dataset of autistic subjects displayed a different fungal community structure compared to neurotypical subjects. In particular the genus *Candida* was one of the most abundant taxa in the gut mycobiota of this study cohort, being two times more abundant in AD than NT. To the best of our knowledge this is the first time that alterations of the intestinal fungal microbiota are associated with ASDs. Although *Candida* is one of the most common and abundant genus of the human gut mycobiota [50, 51], its implication in phenomena of fungal dysbiosis have been reported in several GI and inflammatory conditions [52-54] as well as in Rett syndrome [33]. It is therefore possible that alteration of the intestinal fungal population driven by an expansion of *Candida* in the gut mycobiota of autistic individuals may negatively impact on GI abnormalities through cytokine dysregulation. The gut microbiota, in particular some species of *Lactobacillus*, modulates the immunological responses to *Candida* in the GI tract by providing tryptophan-derived aryl-hydrocarbon receptor ligands that stimulate the immune system, principally ILC3 cells, to produce IL-22 [55]. Together with IL-17, IL-22 avoids the excessive proliferation of *Candida* and other fungal commensals in the gut. It is therefore possible that alterations of the gut microbiota in ASDs could lead to an expansion of the *Candida* population preventing from a full restoration of the bacterial community structure. Indeed it has been observed that alterations of the bacterial gut microbiota due to prolonged antibiotic usage and the subsequent colonization with *C. albicans* prevents from a full restoration of the bacterial community structure resulting in altered abundances of *Bacteriodetes*, *Lactobacillaceae*, *Ruminococcaceae* and *Lachnospiraceae* [56].

Conclusions

Here we observed an altered intestinal microbial community associated with autism, both at bacterial and fungal level. Furthermore such alterations of the gut microbiota did not depend by the constipation status of autistic individuals although a larger cohort would be necessary to validate these findings. Our results encourage an in-depth, extensive study on the impact of the gut microbiota in the GI physiology and neuroplastic changes in ASDs, as well as the integration of such data with immunology and metabolomics to further establish the relevance of the gut microbiota in the ASDs pathophysiology.

Methods

Study participants and samples handling and collection

We recruited 40 subjects with clinical diagnosis of autism (average age 11.1±6.8; sex, male:female, 31:9) and we compared them with 40 age and sex-matched neurotypical healthy subjects (average age 9.2±7.9; sex, male:female, 28:12). Autistic subjects with clinically evident inflammatory conditions were excluded. Constipation and inflammation (*i.e.* serum IgA, erythrocyte sedimentation rate and faecal calprotectin levels) were also assessed. The autistic subjects were consecutively admitted to the Child Neuropsychiatry Unit of the University Hospital of Siena, and ASDs were diagnosed according with the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition [1], and evaluated using Autism Diagnostic Observation Schedule and Autism Behaviour Checklist. Childhood Autism Rating Scale scores (CARS) [57] were calculated by an experienced child neuropsychiatrist. Average CARS values were 46.2 ± 6.8 (value range 32-57); a fraction of 90 % (36/40) were classified as severe ASDs (CARS value >37), with 10 % (4/40) being moderately severe ASDs (CARS values from 30 to 36) (Supplementary Table 9). No specific comorbidities in the autistic cohort were present with the single exception of a coexisting celiac disease in two patients (5 %). Constipation has been defined according to Rome III criteria [58]. Stool samples from enrolled subjects were collected, aliquoted as it is and stored at -80°C until analysis. All subjects of this study were under a Mediterranean-based diet and no antibiotics, probiotics or prebiotics have been taken in the 3 month prior the sample collection. None of the subjects were on anti-inflammatory or antioxidant drugs. The study was conducted after the approval by the Institutional Review Board of the Siena University Hospital (AOUS, Siena, Italy) and all written informed consents were obtained from either the parents or the legal tutors of the enrolled subjects, in compliance with national legislation and the Code of Ethical Principles for Medical Research Involving Human Subjects of the World Medical Association (Declaration of Helsinki).

Faecal Calprotectin Assay

Calprotectin determination was performed by using a polyclonal antibody in an enzyme-linked immunosorbent assay (Calprest, Eurospital, Trieste, Italy) according to the manufacturer's instructions. Calprotectin values < 50 µg/g per stool sample were considered normal.

Pyrosequencing and Data analysis

Total DNA extraction from faecal samples (250mg, wet weight) was performed using the FastDNA™ SPIN Kit for Feces (MP Biomedicals, Santa Ana, CA, USA) following manufacturer's instructions. For each DNA sample, we amplified respectively the bacterial 16S rRNA genes using a primer set specific for V3-V5 hypervariable region (F357: 5'-TCCTACGGGAGGCAGCAG-3' and R937: 5'-TGTGCGGGCCCCCGTCAATT-3') and the Internal Transcribed Spacer (ITS) using a primer set specific for fungal ITS1 rDNA region (18SF: 5'-GTAAAAGTCGTAACAAGGTTTC-3' and 5.8S1R: 5'-GTTCAAAGAYTCGATGATTCAC-3') [59] containing adaptors, key sequence and barcode sequences as described by the 454 Sequencing System Guidelines for Amplicon Experimental Design (Roche, Basel, Switzerland). The PCR products obtained were then purified, quantified and pooled in equimolar way in a final amplicon library. The 454 pyrosequencing was carried out on the GS FLX+ system using the XL+ chemistry following the manufacturer's recommendations (Roche, Basel, Switzerland). Raw 454 data were demultiplexed using the Roche's sff file software and submitted to the European Nucleotide Archive (ENA) with accession numbers PRJEB15418 and PRJEB15420. Sample accessions IDs and metadata are available in Supplementary Table 9. Reads were pre-processed using the MICCA pipeline (v. 0.1) (http://www.micca.org) [60]. operational taxonomic units (OTUs) were assigned by clustering the sequences with a threshold of 97% pairwise identity and their representative sequences were classified using the RDP classifier version 2.7 on 16S data and using the RDP classifier version 2.8 [61] against the UNITE fungal ITS database [62] on ITS1 data. Template-guided multiple sequence alignment (MSA) was performed using PyNAST [63] (v. 0.1) against the multiple alignment of the Greengenes [64] database (release 13_05) filtered at 97% similarity for bacterial sequences and through de novo MSA using T-Coffee [65] for fungal sequences. Fungal taxonomy assignments were then manually curated using BLASTn against the GenBank's database for accuracy. High quality fungal sequences have been also manually filtered out for sequences belonging to *Agaricomycetes* (unlikely to be residents of the human gut due to their ecology [66]). The phylogenetic tree was inferred using micca-phylogeny [67]. Sampling heterogeneity was reduced by rarefaction. *Alpha* (within-sample richness) and *beta*-diversity (between-sample dissimilarity) estimates were computed using the phyloseq R package [68]. Permutational MANOVA (PERMANOVA) test was performed using the adonis()function in the R

package vegan with 999 permutations. Permutations have been constrained within age groups (corresponding to 0-2 y/o, 3-10 y/o, 11-17 y/o and >18 y/o) or gender to evaluate possible biases related to the unequal age and gender distributions among subjects using the "*strata*" argument within the adonis() function. Two-sided, unpaired Welch *t*-statistics were computed using the function mt() in the phyloseq library and the *p*-values were adjusted for multiple comparison controlling the family-wise Type I error rate (minP procedure) [69]. Spearman's correlation tests were computed using the *psych* R package [70]. Linear discriminant effect size analysis (LEfSe) was performed to find features (taxa) differentially represented between autistic and neurotypical subjects. LEfSe combines Kruskal-Wallis test or pairwise Wilcoxon rank-sum test with linear discriminant analysis (LDA). It ranks features by effect size, which put features that explain most of the biological difference at top. LEfSe analysis was performed under the following conditions: α value for the statistical test equal to 0.01 and threshold on the logarithmic LDA score for discriminative features equal to 2.0 [29]. All statistical analyses were performed using R [71] and *p*-values were FDR-corrected [72].

Abbreviations

AD, autistic subjects; ASDs, autism spectrum disorders; CARS, childhood autism rating scale; CNS, central nervous system; ESR, erythrocyte sedimentation rate; FDR, false discovery rate; GI, gastro-intestinal; IBD, inflammatory bowel disease; ITS, Internal Transcribed Spacer; NT, neurotypical subjects; OTU, operational taxonomic unit; PCoA, principal coordinates analysis; PERMANOVA, permutational multivariate analysis of variance; SCFAs, short chain fatty acids.

Acknowledgements

The authors would like to thank Maddalena Sordo, Irene Stefanini, Massimo Pindo and the lab staff of the Sequencing Platform from Fondazione Edmund Mach for technical support and the helpful discussions. We thank Unifarm S.p.A. for financial support to FS. We thank Cesare Furlanello and Paolo Zanini for the insightful discussions on the connection between diet, the globalized world, microbiota and autism. All the authors are grateful to the autistic subjects and their families.

Author contribution

FS, CDF, DC and AC conceived the study and designed the experiments. FS performed the experiments and wrote the manuscript. DR and SL analysed and collected the clinical data. FS analysed the metagenomics data. DA and CD supervised and contributed to data analysis. CDFe, JH, SL, DR and AC recruited subjects and collected specimens. CDFe, JH, SL, DR and AC provided

further clinical patient information and reagents for the study. AC, DC and CDF approved the final manuscript. All the authors critically reviewed and approved the manuscript.

Availability of supporting data

Raw sequences are available in the European Nucleotide Archive (ENA) with accession numbers PRJEB15418 and PRJEB15420.

Conflict of interest

Authors declare no conflict of interests.

Ethics approval and consent to participate

The study was conducted after the approval by the Institutional Review Board of the Siena University Hospital (AOUS, Siena, Italy) and all written informed consents were obtained from either the parents or the legal tutors of the enrolled subjects, in compliance with national legislation and the Code of Ethical Principles for Medical Research Involving Human Subjects of the World Medical Association (Declaration of Helsinki).

Funding

This work was supported from the "Accordo di Programma Integrato MetaFoodLabs" funded by the research office of the Provincia Autonoma di Trento (Italy) (PAT Prot. S116/2012/537723).

References

1. Association AP: **Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)**: American Psychiatric Pub; 2013.
2. Kim YS, Leventhal BL: **Genetic epidemiology and insights into interactive genetic and environmental effects in autism spectrum disorders**. *Biological psychiatry* 2015, **77**(1):66-74.
3. Colvert E, Tick B, McEwen F, Stewart C, Curran SR, Woodhouse E, Gillan N, Hallett V, Lietz S, Garnett T *et al*: **Heritability of Autism Spectrum Disorder in a UK Population-Based Twin Sample**. *JAMA psychiatry* 2015, **72**(5):415-423.
4. McElhanon BO, McCracken C, Karpen S, Sharp WG: **Gastrointestinal symptoms in autism spectrum disorder: a meta-analysis**. *Pediatrics* 2014, **133**(5):872-883.
5. Hsiao EY: **Gastrointestinal issues in autism spectrum disorder**. *Harvard review of psychiatry* 2014, **22**(2):104-111.
6. Furuta GT, Williams K, Kooros K, Kaul A, Panzer R, Coury DL, Fuchs G: **Management of constipation in children and adolescents with autism spectrum disorders**. *Pediatrics* 2012, **130 Suppl 2**:S98-105.
7. Gorrindo P, Williams KC, Lee EB, Walker LS, McGrew SG, Levitt P: **Gastrointestinal dysfunction in autism: parental report, clinical evaluation, and associated factors**.

Autism research : official journal of the International Society for Autism Research 2012, **5**(2):101-108.

8. Mayer EA, Padua D, Tillisch K: **Altered brain-gut axis in autism: comorbidity or causative mechanisms?** *BioEssays : news and reviews in molecular, cellular and developmental biology* 2014, **36**(10):933-939.
9. Sampson TR, Mazmanian SK: **Control of brain development, function, and behavior by the microbiome**. *Cell host & microbe* 2015, **17**(5):565-576.
10. Barajon I, Serrao G, Arnaboldi F, Opizzi E, Ripamonti G, Balsari A, Rumio C: **Toll-like receptors 3, 4, and 7 are expressed in the enteric nervous system and dorsal root ganglia**. *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society* 2009, **57**(11):1013-1023.
11. Brun P, Giron MC, Qesari M, Porzionato A, Caputi V, Zoppellaro C, Banzato S, Grillo AR, Spagnol L, De Caro R *et al*: **Toll-like receptor 2 regulates intestinal inflammation by controlling integrity of the enteric nervous system**. *Gastroenterology* 2013, **145**(6):1323-1333.
12. Sudo N: **Role of microbiome in regulating the HPA axis and its relevance to allergy**. *Chemical immunology and allergy* 2012, **98**:163-175.
13. De Vadder F, Kovatcheva-Datchary P, Goncalves D, Vinera J, Zitoun C, Duchampt A, Backhed F, Mithieux G: **Microbiota-generated metabolites promote metabolic benefits via gut-brain neural circuits**. *Cell* 2014, **156**(1-2):84-96.
14. Ray K: **Gut microbiota: microbial metabolites feed into the gut-brain-gut circuit during host metabolism**. *Nature reviews Gastroenterology & hepatology* 2014, **11**(2):76.
15. Kamada N, Seo SU, Chen GY, Nunez G: **Role of the gut microbiota in immunity and inflammatory disease**. *Nature reviews Immunology* 2013, **13**(5):321-335.
16. Berer K, Krishnamoorthy G: **Commensal gut flora and brain autoimmunity: a love or hate affair?** *Acta neuropathologica* 2012, **123**(5):639-651.
17. Stolp HB, Dziegielewska KM, Ek CJ, Potter AM, Saunders NR: **Long-term changes in blood-brain barrier permeability and white matter following prolonged systemic inflammation in early development in the rat**. *The European journal of neuroscience* 2005, **22**(11):2805-2816.
18. Onore CE, Nordahl CW, Young GS, Van de Water JA, Rogers SJ, Ashwood P: **Levels of soluble platelet endothelial cell adhesion molecule-1 and P-selectin are decreased in children with autism spectrum disorder**. *Biological psychiatry* 2012, **72**(12):1020-1025.
19. Onore C, Careaga M, Ashwood P: **The role of immune dysfunction in the pathophysiology of autism**. *Brain, behavior, and immunity* 2012, **26**(3):383-392.
20. Theoharides TC, Asadi S, Patel AB: **Focal brain inflammation and autism**. *Journal of neuroinflammation* 2013, **10**:46.
21. Stolp HB, Johansson PA, Habgood MD, Dziegielewska KM, Saunders NR, Ek CJ: **Effects of neonatal systemic inflammation on blood-brain barrier permeability and behaviour in juvenile and adult rats**. *Cardiovascular psychiatry and neurology* 2011, **2011**:469046.
22. Oever JT, Netea MG: **The bacteriome-mycobiome interaction and antifungal host defense**. *European journal of immunology* 2014, **44**(11):3182-3191.
23. De Luca A, Montagnoli C, Zelante T, Bonifazi P, Bozza S, Moretti S, D'Angelo C, Vacca C, Boon L, Bistoni F *et al*: **Functional yet balanced reactivity to Candida albicans requires TRIF, MyD88, and IDO-dependent inhibition of Rorc**. *Journal of immunology* 2007, **179**(9):5999-6008.
24. Cheng SC, van de Veerdonk F, Smeekens S, Joosten LA, van der Meer JW, Kullberg BJ, Netea MG: **Candida albicans dampens host defense by downregulating IL-17 production**. *Journal of immunology* 2010, **185**(4):2450-2457.

25. Bozza S, Fallarino F, Pitzurra L, Zelante T, Montagnoli C, Bellocchio S, Mosci P, Vacca C, Puccetti P, Romani L: **A crucial role for tryptophan catabolism at the host/Candida albicans interface**. *Journal of immunology* 2005, **174**(5):2910-2918.
26. Vandeputte D, Falony G, Vieira-Silva S, Tito RY, Joossens M, Raes J: **Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates**. *Gut* 2016, **65**(1):57-62.
27. Zhu L, Liu W, Alkhouri R, Baker RD, Bard JE, Quigley EM, Baker SS: **Structural changes in the gut microbiome of constipated patients**. *Physiological genomics* 2014, **46**(18):679-686.
28. Kang DW, Park JG, Ilhan ZE, Wallstrom G, Labaer J, Adams JB, Krajmalnik-Brown R: **Reduced incidence of Prevotella and other fermenters in intestinal microflora of autistic children**. *PloS one* 2013, **8**(7):e68322.
29. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C: **Metagenomic biomarker discovery and explanation**. *Genome biology* 2011, **12**(6):R60.
30. Zoppi G, Cinquetti M, Luciano A, Benini A, Muner A, Bertazzoni Minelli E: **The intestinal ecosystem in chronic functional constipation**. *Acta paediatrica* 1998, **87**(8):836-841.
31. Son JS, Zheng LJ, Rowehl LM, Tian X, Zhang Y, Zhu W, Litcher-Kelly L, Gadow KD, Gathungu G, Robertson CE *et al*: **Comparison of Fecal Microbiota in Children with Autism Spectrum Disorders and Neurotypical Siblings in the Simons Simplex Collection**. *PloS one* 2015, **10**(10):e0137725.
32. Wang Y, Kasper LH: **The role of microbiome in central nervous system disorders**. *Brain, behavior, and immunity* 2014, **38**:1-12.
33. Strati F, Cavalieri D, Albanese D, De Felice C, Donati C, Hayek J, Jousson O, Leoncini S, Pindo M, Renzi D *et al*: **Altered gut microbiota in Rett syndrome**. *Microbiome* 2016, **4**(1):41.
34. Frank DN, St Amand AL, Feldman RA, Boedeker EC, Harpaz N, Pace NR: **Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases**. *Proceedings of the National Academy of Sciences of the United States of America* 2007, **104**(34):13780-13785.
35. Turnbaugh PJ, Backhed F, Fulton L, Gordon JI: **Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome**. *Cell host & microbe* 2008, **3**(4):213-223.
36. Williams BL, Hornig M, Buie T, Bauman ML, Cho Paik M, Wick I, Bennett A, Jabado O, Hirschberg DL, Lipkin WI: **Impaired carbohydrate digestion and transport and mucosal dysbiosis in the intestines of children with autism and gastrointestinal disturbances**. *PloS one* 2011, **6**(9):e24585.
37. Tomova A, Husarova V, Lakatosova S, Bakos J, Vlkova B, Babinska K, Ostatnikova D: **Gastrointestinal microbiota in children with autism in Slovakia**. *Physiology & behavior* 2015, **138**:179-187.
38. Hsiao EY, McBride SW, Hsien S, Sharon G, Hyde ER, McCue T, Codelli JA, Chow J, Reisman SE, Petrosino JF *et al*: **Microbiota modulate behavioral and physiological abnormalities associated with neurodevelopmental disorders**. *Cell* 2013, **155**(7):1451-1463.
39. Finegold SM, Dowd SE, Gontcharova V, Liu C, Henley KE, Wolcott RD, Youn E, Summanen PH, Granpeesheh D, Dixon D *et al*: **Pyrosequencing study of fecal microflora of autistic and control children**. *Anaerobe* 2010, **16**(4):444-453.
40. Adams JB, Johansen LJ, Powell LD, Quig D, Rubin RA: **Gastrointestinal flora and gastrointestinal status in children with autism--comparisons to typical children and correlation with autism severity**. *BMC gastroenterology* 2011, **11**:22.

41. Morgan XC, Tickle TL, Sokol H, Gevers D, Devaney KL, Ward DV, Reyes JA, Shah SA, LeLeiko N, Snapper SB *et al*: **Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment**. *Genome biology* 2012, **13**(9):R79.
42. Arthur JC, Gharaibeh RZ, Mühlbauer M, Perez-Chanona E, Uronis JM, McCafferty J, Fodor AA, Jobin C: **Microbial genomic analysis reveals the essential role of inflammation in bacteria-induced colorectal cancer**. *Nature communications* 2014, **5**.
43. Frye RE, Rose S, Slattery J, MacFabe DF: **Gastrointestinal dysfunction in autism spectrum disorder: the role of the mitochondria and the enteric microbiome**. *Microbial ecology in health and disease* 2015, **26**:27458.
44. Stiles BG, Pradhan K, Fleming JM, Samy RP, Barth H, Popoff MR: **Clostridium and bacillus binary enterotoxins: bad for the bowels, and eukaryotic being**. *Toxins* 2014, **6**(9):2626-2656.
45. Matsuda H, Fujiyama Y, Andoh A, Ushijima T, Kajinami T, Bamba T: **Characterization of antibody responses against rectal mucosa-associated bacterial flora in patients with ulcerative colitis**. *Journal of gastroenterology and hepatology* 2000, **15**(1):61-68.
46. Woting A, Pfeiffer N, Loh G, Klaus S, Blaut M: **Clostridium ramosum promotes high-fat diet-induced obesity in gnotobiotic mouse models**. *mBio* 2014, **5**(5):e01530-01514.
47. Atarashi K, Tanoue T, Oshima K, Suda W, Nagano Y, Nishikawa H, Fukuda S, Saito T, Narushima S, Hase K: **Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota**. *Nature* 2013, **500**(7461):232-236.
48. Cortelazzo A, De Felice C, Guerranti R, Signorini C, Leoncini S, Zollo G, Leoncini R, Timperio AM, Zolla L, Ciccoli L *et al*: **Expression and oxidative modifications of plasma proteins in autism spectrum disorders: Interplay between inflammatory response and lipid peroxidation**. *Proteomics Clinical applications* 2016.
49. Rizzetto L, De Filippo C, Cavalieri D: **Richness and diversity of mammalian fungal communities shape innate and adaptive immunity in health and disease**. *European journal of immunology* 2014, **44**(11):3166-3181.
50. Scanlan PD, Marchesi JR: **Micro-eukaryotic diversity of the human distal gut microbiota: qualitative assessment using culture-dependent and -independent analysis of faeces**. *The ISME journal* 2008, **2**(12):1183-1193.
51. Strati F, Di Paola M, Stefanini I, Albanese D, Rizzetto L, Lionetti P, Calabro A, Jousson O, Donati C, Cavalieri D *et al*: **Age and Gender Affect the Composition of Fungal Population of the Human Gastrointestinal Tract**. *Frontiers in microbiology* 2016, **7**:1227.
52. Li Q, Wang C, Tang C, He Q, Li N, Li J: **Dysbiosis of gut fungal microbiota is associated with mucosal inflammation in Crohn's disease**. *Journal of clinical gastroenterology* 2014, **48**(6):513.
53. Luan C, Xie L, Yang X, Miao H, Lv N, Zhang R, Xiao X, Hu Y, Liu Y, Wu N *et al*: **Dysbiosis of fungal microbiota in the intestinal mucosa of patients with colorectal adenomas**. *Scientific reports* 2015, **5**:7980.
54. Iliev ID, Funari VA, Taylor KD, Nguyen Q, Reyes CN, Strom SP, Brown J, Becker CA, Fleshner PR, Dubinsky M *et al*: **Interactions between commensal fungi and the C-type lectin receptor Dectin-1 influence colitis**. *Science* 2012, **336**(6086):1314-1317.
55. Zelante T, Iannitti RG, Cunha C, De Luca A, Giovannini G, Pieraccini G, Zecchi R, D'Angelo C, Massi-Benedetti C, Fallarino F *et al*: **Tryptophan catabolites from microbiota engage aryl hydrocarbon receptor and balance mucosal reactivity via interleukin-22**. *Immunity* 2013, **39**(2):372-385.
56. Erb Downward JR, Falkowski NR, Mason KL, Muraglia R, Huffnagle GB: **Modulation of post-antibiotic bacterial community reassembly and host response by Candida albicans**. *Scientific reports* 2013, **3**:2191.

57. Schopler E, Reichler RJ, DeVellis RF, Daly K: **Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS)**. *Journal of autism and developmental disorders* 1980, **10**(1):91-103.
58. Longstreth GF, Thompson WG, Chey WD, Houghton LA, Mearin F, Spiller RC: **Functional bowel disorders**. *Gastroenterology* 2006, **130**(5):1480-1491.
59. Findley K, Oh J, Yang J, Conlan S, Deming C, Meyer JA, Schoenfeld D, Nomicos E, Park M, Program NIHISCCS *et al*: **Topographic diversity of fungal and bacterial communities in human skin**. *Nature* 2013, **498**(7454):367-370.
60. Albanese D, Fontana P, De Filippo C, Cavalieri D, Donati C: **MICCA: a complete and accurate software for taxonomic profiling of metagenomic data**. *Scientific reports* 2015, **5**:9743.
61. Wang Q, Garrity GM, Tiedje JM, Cole JR: **Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy**. *Applied and environmental microbiology* 2007, **73**(16):5261-5267.
62. Koljalg U, Nilsson RH, Abarenkov K, Tedersoo L, Taylor AF, Bahram M, Bates ST, Bruns TD, Bengtsson-Palme J, Callaghan TM *et al*: **Towards a unified paradigm for sequence-based identification of fungi**. *Molecular ecology* 2013, **22**(21):5271-5277.
63. Caporaso JG, Bittinger K, Bushman FD, DeSantis TZ, Andersen GL, Knight R: **PyNAST: a flexible tool for aligning sequences to a template alignment**. *Bioinformatics* 2010, **26**(2):266-267.
64. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL: **Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB**. *Applied and environmental microbiology* 2006, **72**(7):5069-5072.
65. Notredame C, Higgins DG, Heringa J: **T-Coffee: A novel method for fast and accurate multiple sequence alignment**. *Journal of molecular biology* 2000, **302**(1):205-217.
66. Hibbett DS: **A phylogenetic overview of the Agaricomycotina**. *Mycologia* 2006, **98**(6):917-925.
67. Price MN, Dehal PS, Arkin AP: **FastTree 2--approximately maximum-likelihood trees for large alignments**. *PloS one* 2010, **5**(3):e9490.
68. McMurdie PJ, Holmes S: **phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data**. *PloS one* 2013, **8**(4):e61217.
69. Westfall PH, Young SS: **Resampling-based multiple testing: Examples and methods for p-value adjustment**, vol. 279: John Wiley & Sons; 1993.
70. Revelle W: **psych: procedures for psychological, psychometric, and personality research. R package version 1.3. 10. Northwestern University, Evanston, IL**. In.; 2013.
71. Team RC: **R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2012**. In.: ISBN 3-900051-07-0; 2014.
72. Benjamini Y, Hochberg Y: **Controlling the false discovery rate: a practical and powerful approach to multiple testing**. *Journal of the Royal Statistical Society Series B (Methodological)* 1995:289-300.

Supplementary material chapter 2

Supplementary Figure 1: Relative abundances at the genus level of the bacterial gut microbiota of autistic and neurotypical subjects.

Supplementary Figure 2: Relative abundances at the genus level of the fungal gut microbiota of autistic and neurotypical subjects.

Supplementary Figure 3: *Candida* relative abundance in autistic (AD) and neurotypical (NT) subjects. *Candida* relative abundances are reported as mean ± standard error.

Supplementary Table 1: Statistical comparisons (Wilcoxon rank sum test) of clinical data among autistic (AD) and neurotypical (NT) subjects both constipated (C) and non-constipated (NC).

Supplementary Table 2: Permutational multivariate analysis of variance (PERMANOVA) tests of the bacterial gut microbiota on the unweighted and weighted UniFrac distances and the Bray-Curtis dissimilarity according to individuals' health status and constipation.

Supplementary Table 3: Permutational multivariate analysis of variance (PERMANOVA) tests of the bacterial and fungal gut microbiota on the unweighted and weighted UniFrac distances and the Bray-Curtis dissimilarity according to the severity of the autistic phenotype.

Supplementary Table 4: Mean relative abundance (%) ± standard deviation (SD) of bacterial taxa at genus levels in autistic (AD) subjects and neurotypical (NT) controls subjects both constipated (C) and non-constipated (NC).

Supplementary Table 5: Spearman's correlation analysis among the most abundant bacterial genera and the constipation status of the subjects of the study cohort.

Supplementary Table 6: Permutational multivariate analysis of variance (PERMANOVA) tests of the fungal gut microbiota on the unweighted and weighted UniFrac distances and the Bray-Curtis dissimilarity according to individuals' health status and constipation.

Supplementary Table 7: Mean relative abundance (%) ± standard deviation (SD) of fungal taxa at genus levels in autistic (AD) subjects and neurotypical (NT) controls subjects both constipated (C) and non-constipated (NC).

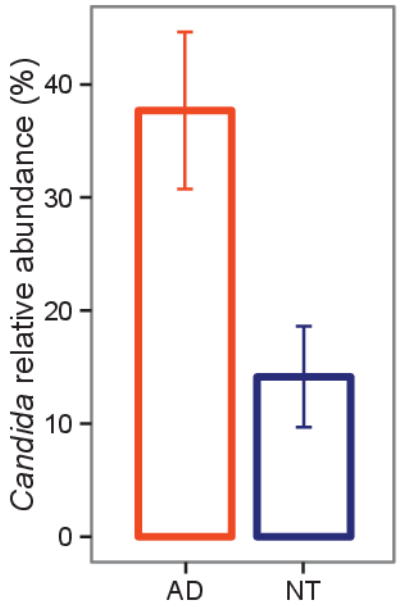
Supplementary Table 8: Spearman's correlation analysis among the most abundant bacterial genera and fungal genera in autistic and neurotypical subjects.

Supplementary Table 9: Correspondences among deposited metagenomics data and samples.

Supplementary Figure 1

Supplementary Figure 2

Supplementary Figure 3



Supplementary Table 1: Statistical comparisons (Wilcoxon rank sum test) of clinical data among autistic (AD) and neurotypical (NT) subjects both constipated (C) and non-constipated (NC).

	AD vs NT	AD-C vs NT-C	AD-NC vs NT-NC	AD-C vs AD-NC	NT-C vs NT-NC
Age	0.14	0.20	0.64	0.88	0.10
Calprotectin	0.75	0.94	0.53	0.56	0.68
CARS	NA	NA	NA	0.98	NA
ESR	NA	NA	NA	0.98	NA
Serum IgA	NA	NA	NA	0.45	NA

Supplementary Table 2: Permutational multivariate analysis of variance (PERMANOVA) tests of the bacterial gut microbiota on the unweighted and weighted UniFrac distances and the Bray-Curtis dissimilarity according to individuals' health status and constipation.

	Metric	F	R^2	*p-value*
NT (n=40) *vs* **AD** (n=40)	Unweighted Unifrac	2.31	0.02	0.001
	Weighted Unifrac	3.29	0.04	0.010
	Bray-Curtis	2.35	0.02	0.005
AD-C (n=5) *vs* **AD-NC** (n=29)	Unweighted Unifrac	1.36	0.04	0.058
	Weighted Unifrac	0.98	0.03	0.397
	Bray-Curtis	1.23	0.03	0.226
NT-C (n=11) *vs* **NT-NC** (n=29)	Unweighted Unifrac	2.21	0.05	0.001
	Weighted Unifrac	2.13	0.05	0.048
	Bray-Curtis	1.81	0.04	0.030
AD-C (n=5) *vs* **NT-C** (n=11)	Unweighted Unifrac	0.44	0.03	0.991
	Weighted Unifrac	0.37	0.02	0.925
	Bray-Curtis	0.44	0.03	0.989
AD-NC (n=29) *vs* **NT-NC** (n=29)	Unweighted Unifrac	3.10	0.05	0.001
	Weighted Unifrac	4.49	0.07	0.002
	Bray-Curtis	2.93	0.04	0.002

NT, neurotypical subjects; AD, autistic subjects; NT-C, constipated neurotypical subjects; NT-NC, non-constipated neurotypical subjects; AD-C, constipated autistic subjects; AD-NC, non-constipated autistic subjects.

Supplementary Table 3: Permutational multivariate analysis of variance (PERMANOVA) tests of the bacterial and fungal gut microbiota on the unweighted and weighted UniFrac distances and the Bray-Curtis dissimilarity according to the severity of the autistic phenotype.

Bacterial microbiota	**Metric**	**F**	R^2	*p-value*
AD severe (n=36) *vs* **AD moderate** (n=4)	Unweighted Unifrac	1.04	0.02	0.43
	Weighted Unifrac	1.19	0.03	0.27
	Bray-Curtis	1.27	0.03	0.24
Fungal microbiota	**Metric**	**F**	R^2	*p-value*
AD severe (n=31) *vs* **AD moderate** (n=4)	Unweighted Unifrac	0.44	0.01	0.92
	Weighted Unifrac	0.77	0.02	0.52
	Bray-Curtis	0.54	0.01	0.69

Supplementary Table 4: Mean relative abundance (%) ± standard deviation (SD) of bacterial taxa at genus levels in autistic (AD) and neurotypical (NT) subjects both constipated (C) and non-constipated (NC).

	AD		NT		AD-NC		AD-C		NT-NC		NT-C	
Genus	**mean**	**SD**	**mean**	**SD**	**mean**	**SD**	**mean**	**SD**	**mean**	**SD**	**mean**	**SD**
Bifidobacterium	24.9000	18.1659	21.7741	20.2050	26.2459	19.3322	22.2564	17.4641	18.7076	16.1969	29.8585	27.5398
Bacteroides	5.6545	9.2705	12.6658	15.8310	4.8177	9.0817	5.4283	6.4654	14.8957	17.2179	6.7872	9.7427
Faecalibacterium	10.2981	11.7366	6.2955	6.1588	11.5902	11.9463	2.2778	2.7370	7.2501	6.6415	3.7789	3.8489
Lachnospiraceae;Unknown	6.3966	4.5323	6.4080	4.8425	6.8836	4.5400	5.4699	4.5756	6.6496	4.4803	5.7713	5.8849
Blautia	6.2533	6.7229	4.0792	4.3493	6.8504	6.8236	6.9322	8.1733	3.1402	2.5915	6.5549	6.7580
Ruminococcus	3.4483	3.9080	3.9006	4.4747	3.8566	4.0712	1.7964	3.8777	4.5480	4.2181	2.1940	4.8837
Clostridium XI	3.5446	3.2881	3.0968	3.9781	3.3952	3.0237	5.0716	5.6175	3.4110	4.4716	2.2685	2.1657
Streptococcus	1.7234	3.3565	4.6497	12.6877	1.7628	3.6635	2.8009	2.9063	4.0161	11.7806	6.3204	15.3292
Gemmiger	3.0106	4.2047	3.0436	3.3370	3.5163	4.6064	1.0224	1.3213	3.7067	3.4961	1.2953	2.1374
Lachnospiracea_incertae_sedis	3.0411	2.7147	2.4181	2.2914	2.7028	1.3934	6.2676	6.3233	2.5923	2.5157	1.9589	1.5609
Ruminococcaceae;Unknown	2.3921	3.6106	2.4617	2.3196	1.6564	1.9501	4.3666	8.1926	2.6479	2.1183	1.9708	2.8379
Escherichia/Shigella	1.4706	3.8944	3.1831	8.7676	0.4731	1.2099	5.3189	7.4712	2.5425	9.3382	4.8721	7.1628
Alistipes	1.2848	3.1149	3.1659	5.9674	0.6116	2.1424	2.8081	5.2410	2.4785	3.0930	4.9780	10.3494
Anaerostipes	2.4452	3.4284	1.9146	1.8001	2.6335	3.7884	3.0161	2.4568	1.4996	1.2319	3.0089	2.5636
Clostridium XVIII	2.4645	4.6033	1.2291	1.6523	1.8462	3.6987	6.6457	8.5055	1.0161	1.5696	1.7908	1.8084
Dorea	2.1164	2.2171	1.1022	1.5706	2.3000	2.4804	1.4492	1.0954	1.0497	1.5281	1.2407	1.7473
Collinsella	2.4212	3.0206	0.7377	1.7124	2.2217	2.8077	2.5441	1.6025	0.7855	1.6358	0.6117	1.9798
Clostridium sensu stricto	1.7014	3.2238	1.1334	1.4337	1.9206	3.6399	1.2483	1.2948	1.0142	1.0385	1.4477	2.2014
Erysipelotrichaceae_incertae_sedis	1.6220	4.7726	0.4596	0.8299	1.7466	5.4441	1.2091	1.8392	0.4071	0.8734	0.5982	0.7218
Clostridiales;Unknown	1.0777	1.3254	0.9136	1.0758	0.9062	1.3068	1.1163	1.0965	0.9408	1.0560	0.8419	1.1760
Dialister	0.5714	1.3153	1.4077	2.7441	0.6744	1.4609	0.0464	0.0773	1.6265	3.0762	0.8311	1.5381
Prevotellaceae;Unknown	1.3324	7.7123	0.4882	2.8855	1.7257	8.9074	0.1938	0.4234	0.6721	3.3867	0.0032	0.0108
Coprococcus	0.8026	0.5336	0.8585	0.6871	0.8014	0.5555	0.9059	0.5469	0.8047	0.6772	1.0003	0.7257
Unknown	0.8041	0.8209	0.8570	1.1623	0.9125	0.8936	0.3329	0.2022	1.0150	1.3098	0.4404	0.4499
Prevotella	0.0501	0.2717	1.5614	9.0676	0.0652	0.3137	0.0000	0.0000	2.0918	10.6478	0.1632	0.4924
Coriobacteriaceae;Unknown	0.5268	0.9162	0.8099	1.3768	0.3230	0.3489	0.1153	0.1998	0.9298	1.5587	0.4939	0.6622

Lactobacillus	1.0018	3.5838	0.2511	1.0308	1.1847	4.1198	0.4446	0.9513	0.2577	1.1633	0.2340	0.5951
Clostridium IV	0.4069	0.6500	0.8271	1.3817	0.3021	0.5829	0.5243	0.9899	0.8672	1.5054	0.7214	1.0408
Turicibacter	0.7895	1.5079	0.3824	0.7666	0.9308	1.7086	0.0725	0.0489	0.4725	0.8798	0.1448	0.1982
Barnesiella	0.2219	0.6002	0.8845	1.9255	0.0535	0.1953	0.5041	0.9901	1.1470	2.1907	0.1924	0.5430
Oscillibacter	0.4162	0.9883	0.5668	0.7707	0.3132	1.0566	0.3828	0.6825	0.6461	0.8050	0.3577	0.6602
Clostridium XlVa	0.4605	0.4844	0.4947	0.6779	0.4421	0.4928	0.5552	0.4492	0.5299	0.7682	0.4020	0.3561
Firmicutes;Unknown	0.3529	0.4567	0.5375	0.7557	0.3396	0.4758	0.2306	0.1516	0.6582	0.8487	0.2194	0.2379
Parabacteroides	0.1896	0.3689	0.6876	1.1413	0.0765	0.1286	0.4530	0.6515	0.9002	1.2781	0.1270	0.1711
Bifidobacteriaceae;Unknown	0.4625	0.3040	0.3696	0.3211	0.4831	0.3105	0.4006	0.2835	0.3687	0.3218	0.3718	0.3348
Enterococcus	0.1428	0.4268	0.4562	1.8010	0.1008	0.2299	0.5445	1.0733	0.1857	0.5507	1.1694	3.3210
Sarcina	0.3079	0.7661	0.2204	0.4950	0.3380	0.8600	0.1379	0.1350	0.1847	0.3732	0.3145	0.7432
Butyricicoccus	0.3326	0.5226	0.1754	0.2275	0.3974	0.5849	0.1379	0.1575	0.1945	0.2308	0.1248	0.2209
Peptostreptococcaceae;Unknown	0.2614	0.2492	0.1574	0.1747	0.2525	0.1991	0.4363	0.4903	0.1572	0.1919	0.1578	0.1257
Megasphaera	0.3021	1.3036	0.0961	0.4098	0.3196	1.4333	0.0202	0.0419	0.1238	0.4801	0.0232	0.0382
Veillonella	0.0627	0.1681	0.3180	1.0394	0.0336	0.0701	0.2972	0.4005	0.0924	0.1538	0.9127	1.9001
Enterobacteriaceae;Unknown	0.0852	0.1755	0.2826	0.8837	0.0375	0.0637	0.4113	0.3301	0.0781	0.1533	0.8219	1.5934
Eggerthella	0.1005	0.1384	0.2366	0.2953	0.0912	0.1315	0.1890	0.1960	0.1781	0.2389	0.3907	0.3800
Bacteroidetes;Unknown	0.2511	1.5845	0.0639	0.3946	0.0004	0.0022	0.0000	0.0000	0.0881	0.4633	0.0000	0.0000
Haemophilus	0.2372	1.1607	0.0727	0.1800	0.3259	1.3354	0.0048	0.0106	0.0676	0.1746	0.0859	0.2017
Roseburia	0.1624	0.1926	0.1434	0.1685	0.1970	0.2090	0.0690	0.0729	0.1640	0.1818	0.0892	0.1175
Erysipelotrichaceae;Unknown	0.1806	0.2920	0.0993	0.1429	0.1660	0.2107	0.3864	0.6496	0.0818	0.0820	0.1454	0.2399
Megamonas	0.0000	0.0000	0.2794	1.7487	0.0000	0.0000	0.0000	0.0000	0.3854	2.0536	0.0000	0.0000
Parasutterella	0.1560	0.6878	0.0898	0.1907	0.0232	0.0775	0.0321	0.0623	0.1084	0.2089	0.0405	0.1266
Lactobacillales;Unknown	0.1061	0.1861	0.1259	0.2441	0.1068	0.1882	0.2128	0.2320	0.0990	0.1875	0.1967	0.3554
Clostridiaceae 1;Unknown	0.1388	0.2252	0.0877	0.1618	0.1523	0.2481	0.1415	0.1486	0.0615	0.0766	0.1567	0.2800
Clostridium XlVb	0.1277	0.3296	0.0901	0.2042	0.1201	0.3331	0.2627	0.4710	0.0717	0.1485	0.1383	0.3120
Odoribacter	0.0878	0.2304	0.1266	0.1962	0.0264	0.0839	0.2794	0.5542	0.1658	0.2159	0.0232	0.0588
Flavonifractor	0.0740	0.1251	0.1367	0.1392	0.0510	0.0818	0.0820	0.0971	0.1261	0.1407	0.1648	0.1376
Lactococcus	0.1953	0.6485	0.0140	0.0338	0.1287	0.4673	0.8156	1.3953	0.0100	0.0320	0.0243	0.0378
Eubacterium	0.0939	0.1480	0.0942	0.1122	0.0691	0.1275	0.0832	0.1264	0.1082	0.1164	0.0573	0.0955
Clostridia;Unknown	0.0892	0.1915	0.0811	0.1913	0.0533	0.1709	0.0713	0.1466	0.0570	0.1236	0.1448	0.3061
Catenibacterium	0.1268	0.6975	0.0198	0.1240	0.0240	0.1258	0.0000	0.0000	0.0273	0.1457	0.0000	0.0000

Paraprevotella	0.1097	0.6417	0.0366	0.1304	0.1416	0.7405	0.0464	0.1037	0.0502	0.1516	0.0005	0.0018
Akkermansia	0.0380	0.0833	0.1031	0.1674	0.0277	0.0748	0.0380	0.0547	0.0816	0.1679	0.1600	0.1597
Phascolarctobacterium	0.0724	0.2345	0.0525	0.1542	0.0092	0.0241	0.1201	0.2138	0.0717	0.1781	0.0016	0.0054
Porphyromonadaceae;Unknown	0.0196	0.1054	0.1015	0.3790	0.0020	0.0078	0.0119	0.0266	0.1386	0.4415	0.0038	0.0107
Peptostreptococcus	0.0004	0.0021	0.1089	0.6841	0.0004	0.0022	0.0012	0.0027	0.0006	0.0033	0.3945	1.3044
Butyricimonas	0.0259	0.0718	0.0367	0.1108	0.0018	0.0063	0.1153	0.1596	0.0504	0.1280	0.0005	0.0018
Bilophila	0.0108	0.0247	0.0477	0.1039	0.0039	0.0101	0.0095	0.0108	0.0640	0.1184	0.0049	0.0074
Actinomyces	0.0260	0.0252	0.0294	0.0404	0.0258	0.0200	0.0262	0.0486	0.0219	0.0254	0.0492	0.0629
Sutterella	0.0022	0.0084	0.0508	0.1827	0.0016	0.0062	0.0083	0.0186	0.0701	0.2123	0.0000	0.0000
Gordonibacter	0.0199	0.0295	0.0305	0.0690	0.0164	0.0183	0.0297	0.0371	0.0305	0.0781	0.0303	0.0387
Coprobacillus	0.0104	0.0285	0.0397	0.1089	0.0055	0.0183	0.0440	0.0611	0.0375	0.1016	0.0454	0.1314
Deltaproteobacteria;Unknown	0.0486	0.2397	0.0006	0.0029	0.0000	0.0000	0.2889	0.6460	0.0002	0.0011	0.0016	0.0054
Desulfovibrio	0.0205	0.0637	0.0260	0.0609	0.0006	0.0024	0.1011	0.1392	0.0246	0.0559	0.0297	0.0752
Bacteroidales;Unknown	0.0123	0.0375	0.0278	0.0465	0.0053	0.0131	0.0499	0.0989	0.0367	0.0519	0.0043	0.0080
Mitsuokella	0.0385	0.1761	0.0000	0.0000	0.0336	0.1779	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Allisonella	0.0324	0.1082	0.0042	0.0194	0.0131	0.0557	0.0464	0.1037	0.0057	0.0227	0.0000	0.0000
Gracilibacteraceae;Unknown	0.0250	0.0837	0.0089	0.0409	0.0010	0.0044	0.0440	0.0984	0.0123	0.0478	0.0000	0.0000
Acidaminococcus	0.0091	0.0333	0.0232	0.1286	0.0057	0.0230	0.0000	0.0000	0.0281	0.1501	0.0103	0.0341
Slackia	0.0265	0.1181	0.0052	0.0204	0.0115	0.0607	0.0095	0.0213	0.0072	0.0238	0.0000	0.0000
Anaerotruncus	0.0111	0.0272	0.0146	0.0306	0.0025	0.0057	0.0428	0.0638	0.0139	0.0334	0.0162	0.0229
Eubacteriaceae;Unknown	0.0177	0.0812	0.0071	0.0282	0.0055	0.0211	0.0036	0.0080	0.0033	0.0127	0.0173	0.0500
Alphaproteobacteria;Unknown	0.0059	0.0229	0.0175	0.0635	0.0027	0.0141	0.0000	0.0000	0.0240	0.0738	0.0005	0.0018
Pasteurellaceae;Unknown	0.0031	0.0188	0.0181	0.0840	0.0043	0.0217	0.0000	0.0000	0.0068	0.0221	0.0481	0.1575
Olsenella	0.0166	0.0677	0.0025	0.0099	0.0014	0.0049	0.0273	0.0611	0.0035	0.0115	0.0000	0.0000
Desulfovibrionaceae;Unknown	0.0120	0.0618	0.0064	0.0329	0.0133	0.0705	0.0143	0.0319	0.0084	0.0386	0.0011	0.0024
Holdemania	0.0076	0.0167	0.0088	0.0138	0.0039	0.0077	0.0071	0.0106	0.0107	0.0149	0.0038	0.0089
Streptophyta	0.0103	0.0247	0.0052	0.0194	0.0096	0.0217	0.0226	0.0473	0.0031	0.0134	0.0108	0.0304
Burkholderiales;Unknown	0.0046	0.0264	0.0098	0.0311	0.0057	0.0304	0.0036	0.0080	0.0129	0.0360	0.0016	0.0054
Enterorhabdus	0.0065	0.0256	0.0049	0.0236	0.0029	0.0152	0.0000	0.0000	0.0068	0.0276	0.0000	0.0000
Alloscardovia	0.0007	0.0028	0.0097	0.0611	0.0010	0.0031	0.0000	0.0000	0.0000	0.0000	0.0351	0.1165
Cellulosilyticum	0.0091	0.0573	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Varibaculum	0.0006	0.0038	0.0079	0.0379	0.0000	0.0000	0.0000	0.0000	0.0094	0.0444	0.0038	0.0067

Hallella	0.0076	0.0479	0.0000	0.0000	0.0105	0.0553	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Anaerofustis	0.0042	0.0085	0.0030	0.0084	0.0043	0.0084	0.0024	0.0033	0.0029	0.0092	0.0032	0.0062
Peptoniphilus	0.0045	0.0160	0.0027	0.0075	0.0031	0.0093	0.0000	0.0000	0.0027	0.0083	0.0027	0.0049
Anaerofilum	0.0027	0.0075	0.0043	0.0103	0.0008	0.0043	0.0071	0.0160	0.0035	0.0085	0.0065	0.0144
Succiniclasticum	0.0000	0.0000	0.0070	0.0442	0.0000	0.0000	0.0000	0.0000	0.0096	0.0519	0.0000	0.0000
Corynebacterium	0.0058	0.0113	0.0010	0.0035	0.0033	0.0058	0.0143	0.0254	0.0012	0.0040	0.0005	0.0018
Gemella	0.0037	0.0089	0.0021	0.0055	0.0031	0.0067	0.0036	0.0080	0.0018	0.0055	0.0027	0.0056
Acidaminococcaceae;Unknown	0.0006	0.0038	0.0051	0.0320	0.0000	0.0000	0.0000	0.0000	0.0070	0.0375	0.0000	0.0000
Pseudomonas	0.0027	0.0150	0.0022	0.0104	0.0033	0.0174	0.0000	0.0000	0.0020	0.0110	0.0027	0.0090
Solobacterium	0.0015	0.0042	0.0030	0.0065	0.0016	0.0044	0.0000	0.0000	0.0029	0.0069	0.0032	0.0056
Lactobacillaceae;Unknown	0.0043	0.0113	0.0000	0.0000	0.0039	0.0111	0.0059	0.0133	0.0000	0.0000	0.0000	0.0000
Sporobacter	0.0006	0.0023	0.0037	0.0088	0.0000	0.0000	0.0012	0.0027	0.0039	0.0089	0.0032	0.0090
Carnobacteriaceae;Unknown	0.0000	0.0000	0.0040	0.0132	0.0000	0.0000	0.0000	0.0000	0.0025	0.0093	0.0081	0.0203
Anaerococcus	0.0009	0.0039	0.0031	0.0126	0.0010	0.0044	0.0000	0.0000	0.0010	0.0039	0.0086	0.0231
Granulicatella	0.0004	0.0021	0.0030	0.0143	0.0006	0.0024	0.0000	0.0000	0.0008	0.0035	0.0086	0.0268
Oxalobacter	0.0019	0.0078	0.0015	0.0077	0.0002	0.0011	0.0036	0.0053	0.0020	0.0090	0.0000	0.0000
Raoultella	0.0012	0.0059	0.0022	0.0066	0.0004	0.0022	0.0071	0.0160	0.0006	0.0033	0.0065	0.0105
Rothia	0.0021	0.0061	0.0012	0.0043	0.0025	0.0067	0.0024	0.0053	0.0006	0.0024	0.0027	0.0072
Rikenellaceae;Unknown	0.0012	0.0051	0.0019	0.0074	0.0000	0.0000	0.0071	0.0129	0.0027	0.0086	0.0000	0.0000
Proteus	0.0013	0.0068	0.0016	0.0086	0.0000	0.0000	0.0083	0.0186	0.0000	0.0000	0.0059	0.0162
Delftia	0.0027	0.0169	0.0000	0.0000	0.0037	0.0195	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Atopobium	0.0004	0.0016	0.0022	0.0093	0.0006	0.0018	0.0000	0.0000	0.0012	0.0066	0.0049	0.0143
Actinomycetales;Unknown	0.0016	0.0050	0.0007	0.0028	0.0020	0.0057	0.0012	0.0027	0.0008	0.0031	0.0005	0.0018
Cloacibacillus	0.0015	0.0094	0.0009	0.0042	0.0000	0.0000	0.0000	0.0000	0.0008	0.0044	0.0011	0.0036
Fusobacterium	0.0001	0.0009	0.0021	0.0096	0.0002	0.0011	0.0000	0.0000	0.0010	0.0055	0.0049	0.0161
Staphylococcus	0.0012	0.0034	0.0006	0.0023	0.0014	0.0037	0.0000	0.0000	0.0006	0.0024	0.0005	0.0018
Clostridiales_Incertae Sedis XI;Unknown	0.0018	0.0095	0.0000	0.0000	0.0004	0.0022	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Porphyromonas	0.0012	0.0051	0.0004	0.0021	0.0006	0.0024	0.0000	0.0000	0.0004	0.0022	0.0005	0.0018
TM7_genera_incertae_sedis	0.0012	0.0066	0.0004	0.0021	0.0014	0.0076	0.0000	0.0000	0.0004	0.0022	0.0005	0.0018
Peptococcus	0.0001	0.0009	0.0013	0.0085	0.0000	0.0000	0.0000	0.0000	0.0018	0.0099	0.0000	0.0000
Parvimonas	0.0007	0.0024	0.0007	0.0031	0.0008	0.0026	0.0000	0.0000	0.0006	0.0033	0.0011	0.0024
Actinomycetaceae;Unknown	0.0006	0.0029	0.0007	0.0039	0.0008	0.0034	0.0000	0.0000	0.0010	0.0045	0.0000	0.0000

Brevundimonas	0.0010	0.0066	0.0003	0.0019	0.0014	0.0076	0.0000	0.0000	0.0004	0.0022	0.0000	0.0000
Ralstonia	0.0012	0.0052	0.0000	0.0000	0.0016	0.0060	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Mesorhizobium	0.0009	0.0034	0.0001	0.0009	0.0012	0.0039	0.0000	0.0000	0.0000	0.0000	0.0005	0.0018
Bacillales;Unknown	0.0009	0.0056	0.0000	0.0000	0.0012	0.0065	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Abiotrophia	0.0003	0.0019	0.0006	0.0023	0.0004	0.0022	0.0000	0.0000	0.0004	0.0015	0.0011	0.0036
Leptospira	0.0009	0.0056	0.0000	0.0000	0.0012	0.0065	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Actinobacteria;Unknown	0.0003	0.0019	0.0004	0.0028	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0016	0.0054
Leptotrichia	0.0000	0.0000	0.0007	0.0047	0.0000	0.0000	0.0000	0.0000	0.0010	0.0055	0.0000	0.0000
Neisseriaceae;Unknown	0.0000	0.0000	0.0007	0.0047	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0027	0.0090
Desulfovibrionales;Unknown	0.0001	0.0009	0.0006	0.0038	0.0000	0.0000	0.0012	0.0027	0.0000	0.0000	0.0022	0.0072
Leuconostoc	0.0001	0.0009	0.0004	0.0021	0.0002	0.0011	0.0000	0.0000	0.0004	0.0022	0.0005	0.0018
Paraeggerthella	0.0000	0.0000	0.0004	0.0028	0.0000	0.0000	0.0000	0.0000	0.0006	0.0033	0.0000	0.0000
Weissella	0.0000	0.0000	0.0004	0.0021	0.0000	0.0000	0.0000	0.0000	0.0006	0.0024	0.0000	0.0000
Mogibacterium	0.0000	0.0000	0.0004	0.0028	0.0000	0.0000	0.0000	0.0000	0.0006	0.0033	0.0000	0.0000
Sneathia	0.0004	0.0028	0.0000	0.0000	0.0006	0.0033	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Ochrobactrum	0.0004	0.0028	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Cardiobacterium	0.0000	0.0000	0.0004	0.0028	0.0000	0.0000	0.0000	0.0000	0.0006	0.0033	0.0000	0.0000
Microbacterium	0.0000	0.0000	0.0003	0.0013	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0011	0.0024
Propionibacteriaceae;Unknown	0.0000	0.0000	0.0003	0.0019	0.0000	0.0000	0.0000	0.0000	0.0004	0.0022	0.0000	0.0000
Dysgonomonas	0.0001	0.0009	0.0001	0.0009	0.0000	0.0000	0.0000	0.0000	0.0002	0.0011	0.0000	0.0000
Pediococcus	0.0000	0.0000	0.0003	0.0019	0.0000	0.0000	0.0000	0.0000	0.0004	0.0022	0.0000	0.0000
Murdochiella	0.0001	0.0009	0.0001	0.0009	0.0002	0.0011	0.0000	0.0000	0.0000	0.0000	0.0005	0.0018
Curvibacter	0.0003	0.0019	0.0000	0.0000	0.0004	0.0022	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Pyramidobacter	0.0003	0.0019	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Rhodococcus	0.0001	0.0009	0.0000	0.0000	0.0002	0.0011	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Propionibacterium	0.0001	0.0009	0.0000	0.0000	0.0002	0.0011	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Bacillus	0.0000	0.0000	0.0001	0.0009	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0005	0.0018
Dolosigranulum	0.0000	0.0000	0.0001	0.0009	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0005	0.0018
Acetivibrio	0.0000	0.0000	0.0001	0.0009	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0005	0.0018
Shinella	0.0000	0.0000	0.0001	0.0009	0.0000	0.0000	0.0000	0.0000	0.0002	0.0011	0.0000	0.0000

Supplementary Table 5: Spearman's correlation analysis among the most abundant bacterial genera and the constipation status of the subjects of the study cohort.

Genus	**Spearman's *r***	***p*-values**	**FDR-corrected *p***
Gemmiger	-0.3950	0.0005	0.0097
Clostridium XVIII	0.3811	0.0008	0.0097
Ruminococcus	-0.3606	0.0016	0.0128
Escherichia/Shigella	0.3061	0.0080	0.0480
Faecalibacterium	-0.2843	0.0141	0.0676
Dialister	-0.2258	0.0531	0.2123
Anaerostipes	0.1952	0.0956	0.2550
Erysipelotrichaceae incertae sedis	0.1953	0.0955	0.2550
Turicibacter	-0.1960	0.0941	0.2550
Collinsella	-0.1654	0.1590	0.3816
Oscillibacter	-0.1440	0.2210	0.4822
Dorea	-0.1360	0.2479	0.4958
Streptococcus	0.1245	0.2906	0.5083
Lactobacillus	0.1230	0.2965	0.5083
Alistipes	0.1079	0.3603	0.5686
Coprococcus	0.1037	0.3791	0.5686
Bacteroides	-0.0922	0.4346	0.6135
Blautia	0.0645	0.5848	0.7387
Clostridium IV	-0.0692	0.5581	0.7387
Clostridium XlVa	0.0553	0.6396	0.7675
Bifidobacterium	0.0430	0.7158	0.8181
Lachnospiracea incertae sedis	0.0300	0.7999	0.8726
Clostridium sensu stricto	0.0108	0.9275	0.9678
Clostridium XI	0.0008	0.9948	0.9948

Supplementary Table 6: Permutational multivariate analysis of variance (PERMANOVA) tests of the fungal gut microbiota on the unweighted and weighted UniFrac distances and the Bray-Curtis dissimilarity according to individuals' health status and constipation.

	Metric	F	R^2	*p-value*
NT (n=38) *vs* **AD** (n=35)	Unweighted Unifrac	1.55	0.02	0.095
	Weighted Unifrac	2.59	0.02	0.040
	Bray-Curtis	2.91	0.03	0.029
AD-C (n=5) *vs* **AD-NC** (n=24)	Unweighted Unifrac	0.61	0.02	0.777
	Weighted Unifrac	0.63	0.02	0.588
	Bray-Curtis	0.57	0.02	0.623
NT-C (n=11) *vs* **NT-NC** (n=27)	Unweighted Unifrac	1.06	0.02	0.336
	Weighted Unifrac	1.26	0.03	0.240
	Bray-Curtis	2.36	0.06	0.046
AD-C (n=5) *vs* **NT-C** (n=11)	Unweighted Unifrac	0.82	0.05	0.570
	Weighted Unifrac	1.60	0.10	0.169
	Bray-Curtis	1.48	0.09	0.204
AD-NC (n=24) *vs* **NT-NC** (n=27)	Unweighted Unifrac	1.27	0.02	0.199
	Weighted Unifrac	2.47	0.04	0.049
	Bray-Curtis	3.01	0.05	0.020

NT, neurotypical subjects; AD, autistic subjects; NT-C, constipated neurotypical subjects; NT-NC, non-constipated neurotypical subjects; AD-C, constipated autistic subjects; AD-NC, non-constipated autistic subjects.

Supplementary Table 7: Mean relative abundance (%) ± standard deviation (SD) of fungal taxa at genus levels in autistic (AD) and neurotypical (NT) subjects both constipated (C) and non-constipated (NC).

	AD		NT		AD-NC		AD-C		NT-NC		NT-C	
Genus	**mean**	**SD**	**mean**	**SD**	**mean**	**SD**	**mean**	**SD**	**mean**	**SD**	**mean**	**SD**
Aspergillus	24.225	23.867	28.029	30.298	23.468	23.512	12.378	18.558	22.299	24.064	42.094	39.808
Candida	37.701	41.069	14.164	27.493	37.431	40.429	56.324	50.747	12.053	26.803	19.343	29.788
Penicillium	13.241	19.683	23.482	30.851	15.399	22.740	6.127	8.946	30.810	33.553	5.494	9.671
Fungi_unidentified_1_1	4.887	7.175	4.350	10.143	5.075	8.112	3.863	5.164	5.081	11.816	2.558	3.569
Blastocystis	0.000	0.000	7.742	26.718	0.000	0.000	0.000	0.000	10.896	31.308	0.000	0.000
Malassezia	3.047	5.128	3.293	7.523	2.661	4.739	4.818	8.844	3.827	8.750	1.982	2.769
Tremellomycetes_unidentified_1	3.103	7.905	2.998	4.488	3.405	9.445	1.620	2.523	2.811	3.411	3.455	6.631
Eurotiomycetes;Unknown	2.038	9.419	1.634	9.987	0.017	0.085	0.353	0.789	2.296	11.848	0.009	0.031
Pichia	2.762	16.126	0.044	0.270	4.028	19.470	0.000	0.000	0.000	0.000	0.151	0.501
Basidiomycota_unidentified_1_1	0.050	0.298	2.618	15.795	0.074	0.360	0.000	0.000	0.008	0.028	9.025	29.317
Ascomycota_unidentified_1_1	0.365	1.381	2.011	8.528	0.402	1.603	0.602	1.080	0.796	2.327	4.994	15.533
Hypoderma	1.917	11.285	0.000	0.000	2.795	13.627	0.000	0.000	0.000	0.000	0.000	0.000
Debaryomyces	0.288	0.622	1.418	3.261	0.260	0.595	0.208	0.220	0.769	1.690	3.011	5.290
Saccharomyces	1.270	3.333	0.273	0.437	1.246	3.032	2.721	6.084	0.265	0.378	0.293	0.578
Mucor	0.344	1.623	1.123	3.739	0.498	1.953	0.021	0.046	1.558	4.383	0.057	0.097
Dothideomycetes_unidentified_1	0.688	2.088	0.697	1.304	0.394	0.652	2.679	5.363	0.635	1.047	0.850	1.845
Eremothecium	0.000	0.000	1.112	6.804	0.000	0.000	0.000	0.000	1.561	8.072	0.009	0.031
Xeromyces	0.095	0.544	0.618	2.401	0.138	0.657	0.000	0.000	0.385	1.421	1.189	3.945
Aureobasidium	0.504	1.722	0.199	0.767	0.199	0.889	0.852	1.904	0.235	0.898	0.113	0.269
Davidiella	0.368	1.166	0.246	0.567	0.169	0.358	0.395	0.768	0.188	0.476	0.387	0.755
Cyberlindnera	0.312	0.918	0.213	0.728	0.234	0.503	0.997	2.229	0.285	0.854	0.038	0.125
Trichosporon	0.006	0.035	0.473	2.914	0.009	0.042	0.000	0.000	0.000	0.000	1.633	5.417
Podosphaera	0.475	2.754	0.000	0.000	0.684	3.327	0.000	0.000	0.000	0.000	0.000	0.000
Mucoraceae_unidentified	0.131	0.772	0.314	1.903	0.190	0.933	0.000	0.000	0.442	2.257	0.000	0.000
Thermomyces	0.418	2.369	0.019	0.118	0.000	0.000	0.125	0.279	0.000	0.000	0.066	0.219
Malasseziales;Unknown	0.027	0.158	0.336	1.456	0.039	0.191	0.000	0.000	0.208	1.079	0.651	2.160
Fusarium	0.003	0.018	0.353	2.156	0.000	0.000	0.000	0.000	0.000	0.000	1.218	4.005

Rhodotorula	0.148	0.663	0.158	0.692	0.056	0.196	0.768	1.718	0.188	0.818	0.085	0.160
Preussia	0.000	0.000	0.279	1.718	0.000	0.000	0.000	0.000	0.392	2.038	0.000	0.000
Urocystidales;Unknown	0.000	0.000	0.246	1.448	0.000	0.000	0.000	0.000	0.346	1.717	0.000	0.000
Pleosporales_unidentified_1	0.083	0.411	0.156	0.827	0.022	0.106	0.478	1.068	0.188	0.979	0.076	0.174
Helminthosporium	0.000	0.000	0.230	1.415	0.000	0.000	0.000	0.000	0.323	1.679	0.000	0.000
Herpotrichiellaceae_unidentified	0.228	1.334	0.000	0.000	0.333	1.610	0.000	0.000	0.000	0.000	0.000	0.000
Torulaspora	0.045	0.143	0.183	0.554	0.022	0.075	0.000	0.000	0.135	0.369	0.302	0.873
Exophiala	0.128	0.719	0.071	0.438	0.009	0.042	0.852	1.904	0.100	0.520	0.000	0.000
Cryptococcus	0.131	0.643	0.049	0.209	0.000	0.000	0.748	1.672	0.027	0.121	0.104	0.344
Pseudeurotium	0.160	0.948	0.000	0.000	0.000	0.000	1.121	2.508	0.000	0.000	0.000	0.000
Helotiales;Unknown	0.036	0.211	0.107	0.657	0.052	0.254	0.000	0.000	0.150	0.779	0.000	0.000
Rhizopus	0.074	0.439	0.044	0.253	0.108	0.530	0.000	0.000	0.058	0.300	0.009	0.031
Urocystis	0.000	0.000	0.107	0.514	0.000	0.000	0.000	0.000	0.012	0.060	0.340	0.939
Alternaria	0.083	0.457	0.022	0.135	0.009	0.042	0.540	1.207	0.031	0.160	0.000	0.000
Phoma	0.030	0.132	0.071	0.438	0.043	0.159	0.000	0.000	0.100	0.520	0.000	0.000
Tilletia	0.039	0.165	0.060	0.371	0.017	0.066	0.187	0.418	0.085	0.440	0.000	0.000
Pseudogymnoascus	0.086	0.372	0.000	0.000	0.125	0.447	0.000	0.000	0.000	0.000	0.000	0.000
Talaromyces	0.047	0.281	0.033	0.186	0.069	0.339	0.000	0.000	0.046	0.220	0.000	0.000
Dothioraceae_unidentified	0.080	0.474	0.000	0.000	0.000	0.000	0.561	1.254	0.000	0.000	0.000	0.000
Unknown	0.071	0.303	0.003	0.017	0.104	0.364	0.000	0.000	0.004	0.020	0.000	0.000
Wallemia	0.009	0.039	0.057	0.223	0.004	0.021	0.042	0.093	0.081	0.263	0.000	0.000
Sordariomycetes_unidentified_1	0.062	0.369	0.000	0.000	0.000	0.000	0.436	0.975	0.000	0.000	0.000	0.000
Fungi;Unknown	0.000	0.000	0.060	0.240	0.000	0.000	0.000	0.000	0.085	0.282	0.000	0.000
Botrytis	0.006	0.024	0.044	0.175	0.009	0.029	0.000	0.000	0.062	0.206	0.000	0.000
Guehomyces	0.000	0.000	0.046	0.194	0.000	0.000	0.000	0.000	0.042	0.200	0.057	0.188
Periconia	0.033	0.176	0.014	0.084	0.048	0.212	0.000	0.000	0.000	0.000	0.047	0.157
Ascosphaera	0.000	0.000	0.041	0.253	0.000	0.000	0.000	0.000	0.000	0.000	0.142	0.470
Lophiostoma	0.000	0.000	0.036	0.219	0.000	0.000	0.000	0.000	0.050	0.260	0.000	0.000
Trichocomaceae_unidentified	0.015	0.088	0.019	0.087	0.000	0.000	0.000	0.000	0.008	0.028	0.047	0.157
Chaetothyriales_unidentified_1	0.030	0.176	0.000	0.000	0.043	0.212	0.000	0.000	0.000	0.000	0.000	0.000
Paraconiothyrium	0.024	0.140	0.000	0.000	0.035	0.170	0.000	0.000	0.000	0.000	0.000	0.000
Lalaria	0.021	0.123	0.000	0.000	0.030	0.148	0.000	0.000	0.000	0.000	0.000	0.000

Diatrypaceae_unidentified	0.006	0.035	0.014	0.084	0.009	0.042	0.000	0.000	0.019	0.100	0.000	0.000
Leptosphaeriaceae_unidentified	0.015	0.088	0.003	0.017	0.000	0.000	0.104	0.232	0.004	0.020	0.000	0.000
Hypocreaceae;Unknown	0.003	0.018	0.011	0.040	0.004	0.021	0.000	0.000	0.004	0.020	0.028	0.067
Tetracladium	0.000	0.000	0.014	0.084	0.000	0.000	0.000	0.000	0.019	0.100	0.000	0.000
Cordyceps	0.000	0.000	0.014	0.084	0.000	0.000	0.000	0.000	0.000	0.000	0.047	0.157
Incertae_sedis_3_unidentified	0.012	0.070	0.000	0.000	0.000	0.000	0.083	0.186	0.000	0.000	0.000	0.000
Sporobolomyces	0.009	0.053	0.003	0.017	0.000	0.000	0.000	0.000	0.004	0.020	0.000	0.000
Eurotiales;Unknown	0.003	0.018	0.008	0.051	0.004	0.021	0.000	0.000	0.012	0.060	0.000	0.000
Puccinia	0.000	0.000	0.011	0.067	0.000	0.000	0.000	0.000	0.000	0.000	0.038	0.125
Amphisphaeriaceae_unidentified	0.009	0.053	0.000	0.000	0.013	0.064	0.000	0.000	0.000	0.000	0.000	0.000
Pseudozyma	0.000	0.000	0.008	0.051	0.000	0.000	0.000	0.000	0.000	0.000	0.028	0.094
Kluyveromyces	0.006	0.035	0.000	0.000	0.009	0.042	0.000	0.000	0.000	0.000	0.000	0.000
Ustilaginales_unidentified_1	0.000	0.000	0.005	0.034	0.000	0.000	0.000	0.000	0.008	0.040	0.000	0.000
Golovinomyces	0.003	0.018	0.000	0.000	0.004	0.021	0.000	0.000	0.000	0.000	0.000	0.000
Pezizomycetes_unidentified_1	0.003	0.018	0.000	0.000	0.004	0.021	0.000	0.000	0.000	0.000	0.000	0.000
Knufia	0.000	0.000	0.003	0.017	0.000	0.000	0.000	0.000	0.000	0.000	0.009	0.031
Dipodascaceae_unidentified	0.000	0.000	0.003	0.017	0.000	0.000	0.000	0.000	0.000	0.000	0.009	0.031
Wickerhamomyces	0.000	0.000	0.003	0.017	0.000	0.000	0.000	0.000	0.000	0.000	0.009	0.031
Schizosaccharomyces	0.000	0.000	0.003	0.017	0.000	0.000	0.000	0.000	0.004	0.020	0.000	0.000
Coniochaeta	0.000	0.000	0.003	0.017	0.000	0.000	0.000	0.000	0.004	0.020	0.000	0.000
Sordariaceae_unidentified	0.000	0.000	0.003	0.017	0.000	0.000	0.000	0.000	0.004	0.020	0.000	0.000

Supplementary Table 8: Spearman's correlation analysis among the most abundant bacterial genera and fungal genera in autistic and neurotypical subjects.

Autistic subjects	**Spearman's *r* correlations**				**FDR-corrected *p*-values**			
Taxa	*Aspergillus*	*Candida*	*Penicillium*	*Malassezia*	*Aspergillus*	*Candida*	*Penicillium*	*Malassezia*
Bifidobacterium	-0.145	-0.021	0.047	0.029	0.945	0.996	0.996	0.996
Bacteroides	0.213	-0.109	-0.143	0.048	0.945	0.945	0.945	0.996
Faecalibacterium	-0.037	0.143	-0.079	-0.140	0.996	0.945	0.996	0.945
Blautia	-0.110	0.121	0.241	-0.011	0.945	0.945	0.945	0.996
Ruminococcus	0.072	-0.167	0.174	0.044	0.996	0.945	0.945	0.996
Clostridium XI	0.098	-0.129	0.230	-0.001	0.970	0.945	0.945	0.996
Streptococcus	-0.050	0.144	0.178	-0.109	0.996	0.945	0.945	0.945
Gemmiger	0.007	0.174	-0.026	-0.391	0.996	0.945	0.996	0.645
Lachnospiracea incertae sedis	-0.165	0.301	0.013	-0.235	0.945	0.941	0.996	0.945
Escherichia/Shigella	-0.094	0.110	-0.008	0.136	0.976	0.945	0.996	0.945
Alistipes	0.147	-0.205	-0.111	0.130	0.945	0.945	0.945	0.945
Anaerostipes	-0.041	0.277	-0.211	-0.162	0.996	0.941	0.945	0.945
Clostridium XVIII	-0.139	0.167	-0.145	-0.108	0.945	0.945	0.945	0.945
Dorea	0.089	0.009	-0.024	0.023	0.996	0.996	0.996	0.996
Collinsella	0.059	-0.054	0.301	0.193	0.996	0.996	0.941	0.945
Clostridium sensu stricto	0.111	0.100	0.187	-0.139	0.945	0.970	0.945	0.945
Dialister	0.156	-0.107	-0.050	0.058	0.945	0.945	0.996	0.996
Erysipelotrichaceae incertae sedis	0.009	-0.290	0.015	0.018	0.996	0.941	0.996	0.996
Coprococcus	-0.149	-0.004	-0.171	-0.250	0.945	0.996	0.945	0.945
Clostridium IV	0.302	-0.231	-0.003	0.020	0.941	0.945	0.996	0.996
Lactobacillus	0.062	0.045	-0.002	0.009	0.996	0.996	0.996	0.996
Turicibacter	0.334	-0.334	0.425	0.417	0.941	0.941	0.606	0.606
Oscillibacter	0.225	-0.281	-0.065	0.246	0.945	0.941	0.996	0.945
Clostridium XlVa	0.154	-0.018	-0.038	-0.002	0.945	0.996	0.996	0.996

Neutotypical subjects	**Spearman's *r* correlations**				**FDR-corrected *p*-values**			
Taxa	*Aspergillus*	*Candida*	*Penicillium*	*Malassezia*	*Aspergillus*	*Candida*	*Penicillium*	*Malassezia*
Bifidobacterium	0.617	-0.097	-0.060	0.170	0.004	0.821	0.909	0.716
Bacteroides	-0.231	-0.049	0.229	-0.307	0.690	0.947	0.690	0.592
Faecalibacterium	-0.324	0.014	0.238	-0.254	0.592	0.978	0.690	0.690
Blautia	-0.111	0.207	-0.181	0.163	0.783	0.690	0.690	0.716
Ruminococcus	-0.321	-0.127	0.112	-0.069	0.592	0.774	0.783	0.903
Clostridium XI	0.154	0.055	0.185	0.004	0.725	0.927	0.690	0.983
Streptococcus	-0.010	-0.072	0.037	0.198	0.982	0.903	0.956	0.690
Gemmiger	-0.196	0.091	0.247	0.005	0.690	0.843	0.690	0.983
Lachnospiracea incertae sedis	-0.236	-0.024	-0.065	-0.166	0.690	0.978	0.903	0.716
Escherichia/Shigella	0.141	0.389	-0.085	0.143	0.747	0.378	0.865	0.747
Alistipes	-0.209	-0.004	-0.306	-0.152	0.690	0.983	0.592	0.725
Anaerostipes	-0.041	0.180	-0.273	0.016	0.956	0.690	0.690	0.978
Clostridium XVIII	0.111	0.163	-0.158	-0.206	0.783	0.716	0.716	0.690
Dorea	0.071	0.071	0.101	0.037	0.903	0.903	0.817	0.956
Collinsella	-0.116	0.210	0.019	0.123	0.783	0.690	0.978	0.774
Clostridium sensu stricto	0.099	0.066	0.216	0.127	0.820	0.903	0.690	0.774
Dialister	-0.255	-0.158	0.123	0.030	0.690	0.716	0.774	0.967
Erysipelotrichaceae incertae sedis	0.264	0.043	-0.205	0.208	0.690	0.956	0.690	0.690
Coprococcus	0.128	0.106	-0.184	0.188	0.774	0.802	0.690	0.690
Clostridium IV	-0.189	-0.014	-0.060	-0.188	0.690	0.978	0.909	0.690
Lactobacillus	0.142	0.121	-0.397	0.299	0.747	0.774	0.378	0.596
Turicibacter	0.037	0.021	0.353	-0.215	0.956	0.978	0.573	0.690
Oscillibacter	-0.317	-0.159	-0.127	-0.208	0.592	0.716	0.774	0.690
Clostridium XlVa	-0.270	-0.032	0.013	-0.415	0.690	0.967	0.978	0.378

Supplementary Table 9: Correspondences among deposited metagenomics data and samples.

16S dataset

Sample_unique_name	Health_status	Constipation	Calprotectin (ug/g)	CARS	ERS (mm/h)	serum IgA (mg/ml)	File_name
AD109	Autism	no	51.3	46	3	138	AUTISMO1_16S.MID44.fastq
AD121	Autism	NA	40.1	48	NA	71	AUTISMO1_16S.MID45.fastq
AD122	Autism	yes	114.6	50	2	192	AUTISMO1_16S.MID46.fastq
AD128	Autism	no	36.9	39	4	207	AUTISMO1_16S.MID50.fastq
AD14	Autism	no	12.4	40	21	137	AUTISMO1_16S.MID1.fastq
AD152	Autism	no	33.5	42	60	88	AUTISMO2_16S.MID1.fastq
AD156	Autism	no	35.9	48	22	147	AUTISMO2_16S.MID7.fastq
AD169	Autism	NA	135.7	40	18	63	AUTISMO2_16S.MID13.fastq
AD170	Autism	NA	101.8	56	NA	NA	AUTISMO2_16S.MID14.fastq
AD18	Autism	no	150	42	2	89	AUTISMO1_16S.MID7.fastq
AD185	Autism	no	57.8	50	2	34	AUTISMO2_16S.MID16.fastq
AD186	Autism	NA	45.5	46	17	156	AUTISMO2_16S.MID17.fastq
AD194	Autism	no	41	50	22	296	AUTISMO2_16S.MID19.fastq
AD196	Autism	no	36	48	2	54	AUTISMO2_16S.MID20.fastq
AD197	Autism	no	15	42	8	281	AUTISMO2_16S.MID22.fastq
AD198	Autism	no	74.4	48	25	70	AUTISMO2_16S.MID24.fastq
AD202	Autism	yes	23.1	57	NA	153	AUTISMO2_16S.MID28.fastq
AD203	Autism	no	15	35	8	156	AUTISMO2_16S.MID29.fastq
AD204	Autism	no	403.9	40	12	66	AUTISMO2_16S.MID33.fastq
AD208	Autism	no	15	52	5	188	AUTISMO2_16S.MID36.fastq
AD216	Autism	no	15	48	4	54	AUTISMO2_16S.MID41.fastq
AD258	Autism	yes	NA	32	28	82	AUTISMO2_16S.MID44.fastq
AD259	Autism	no	83.4	42	NA	75	AUTISMO2_16S.MID45.fastq
AD26	Autism	no	8	39	8	158	AUTISMO1_16S.MID13.fastq
AD260	Autism	yes	22.4	36	NA	70	AUTISMO2_16S.MID46.fastq
AD262	Autism	NA	80.7	40	NA	NA	AUTISMO2_16S.MID50.fastq
AD27	Autism	no	14.6	45	2	26	AUTISMO1_16S.MID14.fastq
AD28	Autism	NA	28.1	45	6	161	AUTISMO1_16S.MID16.fastq

AD29	Autism	yes	55.2	52	22	97	AUTISMO1_16S.MID17.fastq
AD30	Autism	no	10.9	54	7	181	AUTISMO1_16S.MID19.fastq
AD31	Autism	no	20.3	40	4	67	AUTISMO1_16S.MID20.fastq
AD32	Autism	no	23.7	46	2	75	AUTISMO2_16S.MID42.fastq
AD40	Autism	no	77.6	57	7	234	AUTISMO1_16S.MID22.fastq
AD48	Autism	no	91.2	56	30	129	AUTISMO1_16S.MID24.fastq
AD49	Autism	no	9.8	50	2	55	AUTISMO1_16S.MID28.fastq
AD61	Autism	no	11.2	56	8	280	AUTISMO1_16S.MID29.fastq
AD62	Autism	no	44.1	50	2	225	AUTISMO1_16S.MID33.fastq
AD67	Autism	no	123.7	55	4	176	AUTISMO1_16S.MID36.fastq
AD89	Autism	no	33.2	52	8	60	AUTISMO1_16S.MID41.fastq
AD95	Autism	no	47.5	33	11	133	AUTISMO1_16S.MID42.fastq
HC13	Neurotypical	no	NA	NA	NA	NA	RETT1_16S.MID94.fastq
HC20	Neurotypical	no	NA	NA	NA	NA	RETT1_16S.MID106.fastq
HC21	Neurotypical	no	NA	NA	NA	NA	RETT1_16S.MID121.fastq
HC22	Neurotypical	no	NA	NA	NA	NA	RETT1_16S.MID122.fastq
HC23	Neurotypical	no	NA	NA	NA	NA	RETT1_16S.MID125.fastq
HC30	Neurotypical	no	NA	NA	NA	NA	RETT1_16S.MID91.fastq
HC45	Neurotypical	no	NA	NA	NA	NA	RETT1_16S.MID71.fastq
HC47	Neurotypical	no	NA	NA	NA	NA	RETT1_16S.MID79.fastq
HC48	Neurotypical	no	NA	NA	NA	NA	RETT1_16S.MID82.fastq
NT1	Neurotypical	no	NA	NA	NA	NA	AUTISMO1_16S.MID54.fastq
NT10	Neurotypical	no	15.00	NA	NA	NA	AUTISMO1_16S.MID77.fastq
NT11	Neurotypical	no	73.8	NA	NA	NA	AUTISMO1_16S.MID79.fastq
NT12	Neurotypical	no	15.00	NA	NA	NA	AUTISMO1_16S_MFB035.MID82.fastq
NT13	Neurotypical	yes	119.4	NA	NA	NA	AUTISMO1_16S_MFB037.MID82.fastq
NT14	Neurotypical	no	15.00	NA	NA	NA	AUTISMO1_16S.MID84.fastq
NT15	Neurotypical	yes	15.00	NA	NA	NA	AUTISMO1_16S.MID85.fastq
NT16	Neurotypical	yes	97.00	NA	NA	NA	AUTISMO1_16S.MID92.fastq
NT17	Neurotypical	no	31.3	NA	NA	NA	AUTISMO1_16S.MID94.fastq
NT18	Neurotypical	no	73.2	NA	NA	NA	AUTISMO2_16S.MID61.fastq
NT19	Neurotypical	no	20.2	NA	NA	NA	AUTISMO2_16S.MID67.fastq

NT2	Neurotypical	no	NA	NA	NA	NA	AUTISMO1_16S.MID105.fastq
NT20	Neurotypical	no	62.5	NA	NA	NA	AUTISMO2_16S.MID71.fastq
NT21	Neurotypical	yes	15	NA	NA	NA	AUTISMO2_16S.MID77.fastq
NT22	Neurotypical	no	50.5	NA	NA	NA	AUTISMO2_16S.MID79.fastq
NT23	Neurotypical	no	15	NA	NA	NA	AUTISMO2_16S.MID82.fastq
NT24	Neurotypical	no	74.4	NA	NA	NA	AUTISMO2_16S.MID84.fastq
NT25	Neurotypical	no	99.1	NA	NA	NA	AUTISMO2_16S.MID85.fastq
NT26	Neurotypical	yes	26.8	NA	NA	NA	AUTISMO2_16S.MID91.fastq
NT27	Neurotypical	yes	15	NA	NA	NA	AUTISMO2_16S.MID92.fastq
NT28	Neurotypical	no	75.8	NA	NA	NA	AUTISMO2_16S.MID94.fastq
NT29	Neurotypical	yes	25.7	NA	NA	NA	AUTISMO2_16S.MID105.fastq
NT3	Neurotypical	no	NA	NA	NA	NA	AUTISMO1_16S.MID55.fastq
NT30	Neurotypical	yes	69.5	NA	NA	NA	AUTISMO2_16S.MID106.fastq
NT31	Neurotypical	yes	27.2	NA	NA	NA	AUTISMO2_16S.MID121.fastq
NT4	Neurotypical	no	NA	NA	NA	NA	AUTISMO1_16S.MID106.fastq
NT5	Neurotypical	no	NA	NA	NA	NA	AUTISMO1_16S.MID58.fastq
NT6	Neurotypical	no	NA	NA	NA	NA	AUTISMO1_16S.MID121.fastq
NT7	Neurotypical	no	NA	NA	NA	NA	AUTISMO1_16S.MID61.fastq
NT8	Neurotypical	yes	500.00	NA	NA	NA	AUTISMO1_16S.MID67.fastq
NT9	Neurotypical	yes	98.2	NA	NA	NA	AUTISMO1_16S.MID71.fastq

ITS1 dataset

Sample_unique_name	Health_status	Constipation	Calprotectin (ug/g)	CARS	ERS (mm/h)	serum IgA (mg/ml)	File_name
AD109	Autism	no	51.3	46	3	138	AUTISMO_POOL1_ITS.MID83.fastq
AD121	Autism	NA	40.1	48	NA	71	AUTISMO_POOL1_ITS.MID86.fastq
AD122	Autism	yes	114.6	50	2	192	AUTISMO_POOL1_ITS.MID88.fastq
AD128	Autism	no	36.9	39	4	207	AUTISMO_POOL2_ITS.MID4.fastq
AD14	Autism	no	12.4	40	21	137	AUTISMO_POOL1_ITS.MID4.fastq
AD152	Autism	no	33.5	42	60	88	AUTISMO_POOL2_ITS.MID6.fastq
AD156	Autism	no	35.9	48	22	147	AUTISMO_POOL2_ITS.MID10.fastq
AD169	Autism	NA	135.7	40	18	63	AUTISMO_POOL2_ITS.MID11.fastq
AD170	Autism	NA	101.8	56	NA	NA	AUTISMO_POOL2_ITS.MID21.fastq

AD18	Autism	no	150	42	2	89	AUTISMO_POOL1_ITS.MID6.fastq
AD185	Autism	no	57.8	50	2	34	AUTISMO_POOL2_ITS.MID23.fastq
AD186	Autism	NA	45.5	46	17	156	AUTISMO_POOL2_ITS.MID35.fastq
AD194	Autism	no	41	50	22	296	AUTISMO_POOL2_ITS.MID37.fastq
AD196	Autism	no	36	48	2	54	AUTISMO_POOL2_ITS.MID39.fastq
AD197	Autism	no	15	42	8	281	AUTISMO_POOL2_ITS.MID47.fastq
AD198	Autism	no	74.4	48	25	70	AUTISMO_POOL2_ITS.MID52.fastq
AD202	Autism	yes	23.1	57	NA	153	AUTISMO_POOL2_ITS.MID59.fastq
AD203	Autism	no	15	35	8	156	AUTISMO_POOL2_ITS.MID62.fastq
AD204	Autism	no	403.9	40	12	66	AUTISMO_POOL2_ITS.MID65.fastq
AD208	Autism	no	15	52	5	188	AUTISMO_POOL2_ITS.MID68.fastq
AD216	Autism	no	15	48	4	54	AUTISMO_POOL2_ITS.MID70.fastq
AD258	Autism	yes	NA	32	28	82	AUTISMO_POOL2_ITS.MID81.fastq
AD259	Autism	no	83.4	42	NA	75	AUTISMO_POOL2_ITS.MID83.fastq
AD26	Autism	no	8	39	8	158	AUTISMO_POOL1_ITS.MID10.fastq
AD260	Autism	yes	22.4	36	NA	70	AUTISMO_POOL2_ITS.MID86.fastq
AD262	Autism	NA	80.7	40	NA	NA	AUTISMO_POOL2_ITS.MID88.fastq
AD27	Autism	no	14.6	45	2	26	AUTISMO_POOL1_ITS.MID11.fastq
AD28	Autism	NA	28.1	45	6	161	AUTISMO_POOL1_ITS.MID21.fastq
AD29	Autism	yes	55.2	52	22	97	AUTISMO_POOL1_ITS.MID23.fastq
AD30	Autism	no	10.9	54	7	181	AUTISMO_POOL1_ITS.MID35.fastq
AD31	Autism	no	20.3	40	4	67	AUTISMO_POOL1_ITS.MID37.fastq
AD32	Autism	no	23.7	46	2	75	AUTISMO_POOL1_ITS.MID39.fastq
AD40	Autism	no	77.6	57	7	234	AUTISMO_POOL1_ITS.MID47.fastq
AD48	Autism	no	91.2	56	30	129	AUTISMO_POOL1_ITS.MID52.fastq
AD49	Autism	no	9.8	50	2	55	AUTISMO_POOL1_ITS.MID59.fastq
AD61	Autism	no	11.2	56	8	280	AUTISMO_POOL1_ITS.MID62.fastq
AD62	Autism	no	44.1	50	2	225	AUTISMO_POOL1_ITS.MID65.fastq
AD67	Autism	no	123.7	55	4	176	AUTISMO_POOL1_ITS.MID68.fastq
AD89	Autism	no	33.2	52	8	60	AUTISMO_POOL1_ITS.MID70.fastq
AD95	Autism	no	47.5	33	11	133	AUTISMO_POOL1_ITS.MID81.fastq
HC13	Neurotypical	no	NA	NA	NA	NA	AUTISMO_POOL2_ITS.MID142.fastq

HC20	Neurotypical	no	NA	NA	NA	NA	AUTISMO_POOL2_ITS.MID144.fastq
HC21	Neurotypical	no	NA	NA	NA	NA	AUTISMO_POOL1_ITS.MID144.fastq
HC22	Neurotypical	no	NA	NA	NA	NA	AUTISMO_POOL2_ITS.MID141.fastq
HC23	Neurotypical	no	NA	NA	NA	NA	AUTISMO_POOL2_ITS.MID143.fastq
HC30	Neurotypical	no	NA	NA	NA	NA	AUTISMO_POOL1_ITS.MID146.fastq
HC45	Neurotypical	no	NA	NA	NA	NA	AUTISMO_POOL2_ITS.MID145.fastq
HC47	Neurotypical	no	NA	NA	NA	NA	AUTISMO_POOL2_ITS.MID146.fastq
HC48	Neurotypical	no	NA	NA	NA	NA	AUTISMO_POOL1_ITS.MID145.fastq
NT1	Neurotypical	no	NA	NA	NA	NA	AUTISMO_POOL1_ITS.MID97.fastq
NT10	Neurotypical	no	15.00	NA	NA	NA	AUTISMO_POOL1_ITS.MID117.fastq
NT11	Neurotypical	no	73.8	NA	NA	NA	AUTISMO_POOL1_ITS.MID119.fastq
NT12	Neurotypical	no	15.00	NA	NA	NA	AUTISMO_POOL1_ITS.MID126.fastq
NT13	Neurotypical	yes	119.4	NA	NA	NA	AUTISMO_POOL1_ITS.MID124.fastq
NT14	Neurotypical	no	15.00	NA	NA	NA	AUTISMO_POOL1_ITS.MID131.fastq
NT15	Neurotypical	yes	15.00	NA	NA	NA	AUTISMO_POOL1_ITS.MID135.fastq
NT16	Neurotypical	yes	97.00	NA	NA	NA	AUTISMO_POOL1_ITS.MID136.fastq
NT17	Neurotypical	no	31.3	NA	NA	NA	AUTISMO_POOL1_ITS.MID137.fastq
NT18	Neurotypical	no	73.2	NA	NA	NA	AUTISMO_POOL2_ITS.MID97.fastq
NT19	Neurotypical	no	20.2	NA	NA	NA	AUTISMO_POOL2_ITS.MID102.fastq
NT2	Neurotypical	no	NA	NA	NA	NA	AUTISMO_POOL1_ITS.MID141.fastq
NT20	Neurotypical	no	62.5	NA	NA	NA	AUTISMO_POOL2_ITS.MID103.fastq
NT21	Neurotypical	yes	15	NA	NA	NA	AUTISMO_POOL2_ITS.MID112.fastq
NT22	Neurotypical	no	50.5	NA	NA	NA	AUTISMO_POOL2_ITS.MID114.fastq
NT23	Neurotypical	no	15	NA	NA	NA	AUTISMO_POOL2_ITS.MID115.fastq
NT24	Neurotypical	no	74.4	NA	NA	NA	AUTISMO_POOL2_ITS.MID117.fastq
NT25	Neurotypical	no	99.1	NA	NA	NA	AUTISMO_POOL2_ITS.MID119.fastq
NT26	Neurotypical	yes	26.8	NA	NA	NA	AUTISMO_POOL2_ITS.MID126.fastq
NT27	Neurotypical	yes	15	NA	NA	NA	AUTISMO_POOL2_ITS.MID124.fastq
NT28	Neurotypical	no	75.8	NA	NA	NA	AUTISMO_POOL2_ITS.MID131.fastq
NT29	Neurotypical	yes	25.7	NA	NA	NA	AUTISMO_POOL2_ITS.MID135.fastq
NT3	Neurotypical	no	NA	NA	NA	NA	AUTISMO_POOL1_ITS.MID102.fastq
NT30	Neurotypical	yes	69.5	NA	NA	NA	AUTISMO_POOL2_ITS.MID136.fastq

NT31	Neurotypical	yes	27.2	NA	NA	NA	AUTISMO_POOL2_ITS.MID137.fastq
NT4	Neurotypical	no	NA	NA	NA	NA	AUTISMO_POOL1_ITS.MID142.fastq
NT5	Neurotypical	no	NA	NA	NA	NA	AUTISMO_POOL1_ITS.MID103.fastq
NT6	Neurotypical	no	NA	NA	NA	NA	AUTISMO_POOL1_ITS.MID143.fastq
NT7	Neurotypical	no	NA	NA	NA	NA	AUTISMO_POOL1_ITS.MID112.fastq
NT8	Neurotypical	yes	500.00	NA	NA	NA	AUTISMO_POOL1_ITS.MID114.fastq
NT9	Neurotypical	yes	98.2	NA	NA	NA	AUTISMO_POOL1_ITS.MID115.fastq

Chapter 3

Altered gut microbiota in Rett syndrome

This chapter has been reprinted* from:

Strati F, Cavalieri D, Albanese D, De Felice C, Donati C, Hayek J, Jousson O, Leoncini S, Pindo M, Renzi D, Rizzetto L, Stefanini I, Calabrò A and De Filippo C (2016). *Altered gut microbiota in Rett syndrome*. Microbiome 4:41, doi: 10.1186/s40168-016-0185-y.

Supported by the increasing appreciation on the important role of the microbiome-gut-brain axis we studied the gut microbiota in Rett syndrome, a genetic, neurological disorder characterized also by gastrointestinal dysfunctions [1]. Since Rett syndrome is a genetic-determined disease due to impairments of the epigenetic regulator MeCP2 [2], this represented a unique occasion to better understand how the composition of the gut microbiota in RTT subjects may be affected by mutations in a gene that appears to be essential for the normal function of nerve cells but apparently is unrelated to gastrointestinal dysfunctions. The results of this study were fascinating. We discovered for the first time and to the best of our knowledge that RTT is characterized by an altered microbial community structure, both at bacterial and fungal level. Furthermore we observed that the gut microbiota of RTT subjects was characterized by high levels of SCFAs-producing bacteria, with *Bifidobacterium* being the most abundant bacterial taxa. As expected, but in contrast with the common perception on the health-promoting role of SCFAs in intestinal physiology, we discovered high levels of these metabolites in RTT suggesting the strengthening role of SCFAs in the establishment of constipation in RTT subjects. This project, more than any other in this Ph.D. thesis, required my total involvement and dedication. Thanks to the expertise of the physicians of the University Hospital of Siena and Florence we enrolled 50 subjects diagnosed with RTT, collecting also several clinical data that I integrated with the amplicon-sequencing data. During this work I applied several analytical pipelines and bioinformatics tool for microbiome research greatly improving my knowledge of multivariate analyses in microbial ecology. Furthermore I learned how to prepare and analyse faecal samples for targeted metabolomics analysis *i.e.* SCFAs analysis, by mean of GC-MS. As in the previous manuscripts, I designed and performed all the experiments related to the analysis of the gut microbiota, analysed the data, wrote the manuscript, and generated figures and tables.

References

1. Chahrour M, Zoghbi HY: **The story of Rett syndrome: from clinic to neurobiology**. *Neuron* 2007, **56**(3):422-437.
2. Liyanage VR, Rastegar M: **Rett syndrome and MeCP2**. *Neuromolecular medicine* 2014, **16**(2):231-264.

Strati et al. Microbiome (2016) 4:41
DOI 10.1186/s40168-016-0185-y

Microbiome

RESEARCH Open Access

Altered gut microbiota in Rett syndrome

Francesco Strati[1,2], Duccio Cavalieri[1,3], Davide Albanese[1], Claudio De Felice[4], Claudio Donati[1], Joussef Hayek[5], Olivier Jousson[2], Silvia Leoncini[5], Massimo Pindo[6], Daniela Renzi[7], Lisa Rizzetto[1], Irene Stefanini[1], Antonio Calabrò[7] and Carlotta De Filippo[8*]

Abstract

Background: The human gut microbiota directly affects human health, and its alteration can lead to gastrointestinal abnormalities and inflammation. Rett syndrome (RTT), a progressive neurological disorder mainly caused by mutations in *MeCP2* gene, is commonly associated with gastrointestinal dysfunctions and constipation, suggesting a link between RTT's gastrointestinal abnormalities and the gut microbiota. The aim of this study was to evaluate the bacterial and fungal gut microbiota in a cohort of RTT subjects integrating clinical, metabolomics and metagenomics data to understand if changes in the gut microbiota of RTT subjects could be associated with gastrointestinal abnormalities and inflammatory status.

Results: Our findings revealed the occurrence of an intestinal sub-inflammatory status in RTT subjects as measured by the elevated values of faecal calprotectin and erythrocyte sedimentation rate. We showed that, overall, RTT subjects harbour bacterial and fungal microbiota altered in terms of relative abundances from those of healthy controls, with a reduced microbial richness and dominated by microbial taxa belonging to *Bifidobacterium*, several Clostridia (among which *Anaerostipes*, *Clostridium XlVa*, *Clostridium XlVb*) as well as *Erysipelotrichaceae*, *Actinomyces*, *Lactobacillus*, *Enterococcus*, *Eggerthella*, *Escherichia/Shigella* and the fungal genus *Candida*.
We further observed that alterations of the gut microbiota do not depend on the constipation status of RTT subjects and that this dysbiotic microbiota produced altered short chain fatty acids profiles.

Conclusions: We demonstrated for the first time that RTT is associated with a dysbiosis of both the bacterial and fungal component of the gut microbiota, suggesting that impairments of MeCP2 functioning favour the establishment of a microbial community adapted to the costive gastrointestinal niche of RTT subjects. The altered production of short chain fatty acids associated with this microbiota might reinforce the constipation status of RTT subjects and contribute to RTT gastrointestinal physiopathology.

Keywords: Gut microbiota, Mycobiota, Rett syndrome, SCFAs, Metataxonomics, Intestinal dysbiosis, Constipation

Background

Rett syndrome (RTT; OMIM #312750) is a severe and progressive neurological disorder that almost exclusively affects females with an incidence of ~1:10,000 live births [1]. Loss-of-function mutations of the X-linked methyl-CpG binding protein 2 (*MeCP2*) gene is the major cause (approximately 90 %) of classical cases of RTT while cyclin-dependent kinase-like 5 (*CDKL5*) and forkhead box protein G1 (*FOXG1*) gene mutations represent the remaining 10 % of the cases [2]. MeCP2 is a fundamental mediator of synaptic development and plasticity, and its function is critical in the regulation of synaptic activities during early postnatal development [1]. Different *MeCP2* mutations are also known to correlate with the clinical severity of RTT [3, 4] and the role of MeCP2 in other neurodevelopmental disorders, such as autism, has been demonstrated [5]. RTT subjects develop normally up to 6–18 months of age after which they undergo a period of neurological regression [1]. Microcephaly, dyspraxia, stereotyped hand movements, transient autistic features, respiratory abnormalities, bruxism, seizures and gastrointestinal (GI) dysfunctions are symptoms commonly reported in RTT indicating it as a multisystemic disorder [1]. Among the above-mentioned comorbidities, several epidemiological studies indicated that GI dysfunctions are prevalent through the entire life of RTT

* Correspondence: c.de.filippo@ibimet.cnr.it
[8]Institute of Biometeorology (IBIMET), National Research Council (CNR), Via Giovanni Caproni 8, I-50145 Florence, Italy
Full list of author information is available at the end of the article

subjects [6, 7] with constipation as one of the most frequently reported GI symptoms [6, 7]. Recently, it has been shown that the phenomenon of gut dismotility observed in RTT may arise from impairments in the function of MeCP2 in the enteric nervous system (ENS) [8]. Nevertheless, the direct causes of these GI dysfunctions are still unclear and the role of gut microbiota in host physiology should not be neglected. Indeed, the human gut microbiota plays a crucial role in the function and integrity of the GI tract, maintenance of immune homeostasis and host energy metabolism [9]. Alterations in the composition of commensal bacterial population can lead to chronic inflammation encompassing hyperactivation of T-helper 1 and T-helper 17 immune responses [10], also predisposing individuals to fungal infections [11]. Abnormal immunological response to fungi can in turn contribute to systemic responses including chronic inflammation as observed in inflammatory bowel diseases [12]. Dysbioses of the gut microbiota have been associated with an increasing number of health conditions [13]. A strict relationship between the gut microbiota and the central nervous system (CNS) has been observed, and numerous studies have shown alterations of the gut microbiota in the heterogeneous group of neurological disorders belonging to the autism spectrum disorders (ASDs) [14]. In addition, the gut microbiota may modulate CNS activities through neural, endocrine, metabolic and immune pathways [15] affecting complex physiological and behavioural states of the host [15, 16] so it is possible to hypothesize gut microbiota alterations in RTT as occur in ASDs. Supported by the increasing appreciation of the gut-microbiome-brain axis, we asked whether MeCP2 impairments in RTT might affect also the composition of the gut microbiota resulting in an eventual intestinal dysbiosis in RTT subjects. Indeed, in the case of RTT, it is possible that alterations in the composition of gut microbiota triggered by the neurophysiological changes typical of the disease could contribute to GI abnormalities and be an additional factor relevant to previously observed cytokine dysregulation and systemic inflammation [17, 18]. Here, we characterized for the first time the intestinal microbiota, both bacterial and fungal, in subjects affected by RTT in order to investigate the implication of gut microorganisms and their metabolism on RTT gastrointestinal physiology evaluating also how the constipation status may affect the composition of the gut microbiota in RTT subjects.

Results

RTT is associated with mild intestinal inflammation

We analysed the inflammatory status and GI abnormalities in a cohort of 50 RTT subjects by measuring the erythrocyte sedimentation rate (ESR), C-reactive protein (CRP), serum IgA and faecal calprotectin (Additional file 1: Table S1). We found that RTT subjects presented elevated values of ESR (median value 22 mm/h; interquartile range 10–36.5 mm/h; average 26.8 ± 22.4 mm/h) and faecal calprotectin (median value 63.45 μg/g; interquartile range 44–123 μg/g; average 104 ± 97.8 μg/g) compared to a healthy population [19, 20]. Constipation, one of the most common GI symptoms of RTT, was present in 70.8 % of the examined RTT cohort, and it correlated with the titre of serum IgA antibodies (Spearman's $r = 0.43$, $p = 0.011$) while the levels of serum IgA positively correlated with the ESR values (Spearman's $r = 0.462$, $p = 0.011$). Constipation, ESR and serum IgA titre correlated with the age of RTT subjects (Spearman's $r = 0.35$, 0.409 and 0.596, respectively, $p < 0.05$) (Additional file 2: Figure S1, Additional file 3: Table S2). Altogether, these parameters confirmed the presence of a mild GI inflammation in RTT subjects.

RTT subjects harbour an altered and less diverse gut microbiota

We characterized the bacterial gut microbiota associated with RTT by means of high-throughput sequencing of the V3-V5 region of the 16S rDNA gene. We quantified the bacterial richness within each sample (*alpha*-diversity) of the two groups, RTT subjects and healthy controls (hereinafter termed HC). Three different *alpha*-diversity estimators were used, namely the observed number of OTUs, the Chao1 index and the Shannon entropy index. The bacterial gut microbiota of RTT subjects was significantly less diverse compared to that of HC ($p < 0.005$, Wilcoxon rank-sum test) with all the three estimators used.

Since constipation affects more than the 70 % of the RTT study cohort, we asked whether the constipation status might be responsible for the differences observed between RTT subjects and HC. The analysis of *alpha*-diversity revealed that, even when analysed separately, both constipated and non-constipated RTT subjects (respectively RTT-C and RTT-NC) harbour a less diverse gut microbiota with respect to HC ($p < 0.05$, Wilcoxon rank-sum test; Fig. 1a and Additional file 3: Figure S2a) while there was no significant difference between RTT-C and RTT-NC ($p > 0.05$, Wilcoxon rank-sum test; Fig. 1a and Additional file 4: Figure S2a). To assess the robustness of these results, we repeated the computation of the *alpha*-diversity using 100 independent rarefactions and for different values of the rarefaction depth. *Alpha* rarefaction curves are reported in the Additional file 5: Figure S3a.

To identify possible differences between the bacterial components of the gut microbiota of RTT subjects and HC, we calculated the *beta*-diversity of the samples using

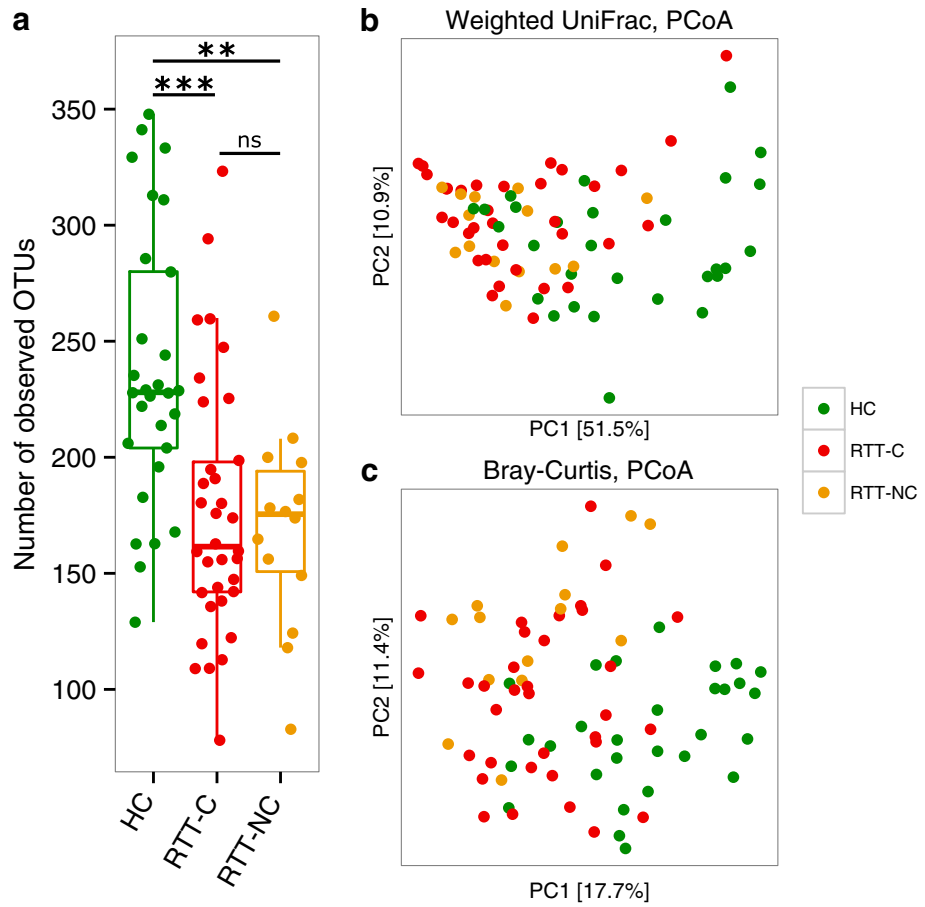


Fig. 1 Measures of bacterial diversity. **a** *Alpha*-diversity calculated on the number of observed OTUs; ***$p < 0.0001$; **$p < 0.001$; *ns*, not significant; Wilcoxon rank-sum test. **b**, **c** PCoA plots of bacterial *beta*-diversity based on **b** the Weighted UniFrac distance and **c** the Bray-Curtis dissimilarity analysed according to individuals' health status. Constipated Rett syndrome subjects (RTT-C), non-constipated Rett syndrome subjects (RTT-NC) and healthy controls (HC) are coloured in *red*, *orange* or *green*, respectively

the unweighted and weighted UniFrac distances and the Bray-Curtis dissimilarity. The Principal Coordinates Analysis (PCoA) based on these measures (Fig. 1b and Additional file 4: Figure S2b) revealed that the gut microbiota of RTT subjects was distinct from those of the HC ($p \leq 0.003$, PERMANOVA) (Table 1). The analysis of *beta*-diversity among HC, RTT-C and RTT-NC revealed significant differences between HC and both RTT-C and RTT-NC ($p \leq 0.003$, PERMANOVA), but no significant difference was detected when comparing the gut microbiota of RTT-C and RTT-NC (Table 1). Multiple-rarefaction PCoA plots ("jackknifed" PCoA plots, [21]) (Additional file 6: Figure S4a and Additional file 7: Figure S5a) were computed to assess the robustness of bacterial *beta*-diversity analyses, showing that these results hold independently from rarefaction.

Bifidobacterium is the hallmark of intestinal dysbiosis in RTT

Since we did not observe population differences in the gut microbiota of RTT subjects in function of the constipation status, only the health condition of the subjects (i.e., healthy or RTT) was considered in the following analyses. To identify the taxa that were differentially represented in HC and RTT subjects, we compared the relative abundances between these two groups at different taxonomic levels.

Phylum level analysis showed that *Actinobacteria* was the most abundant phylum in RTT with a significant increase of its relative abundance in RTT subjects compared to HC ($p = 0.0017$, Wilcoxon rank-sum test; Additional file 8: Table S3 and Additional file 9: Table S4). Furthermore, while *Firmicutes* was the most abundant phylum in HC, we observed a significant decrease of the relative abundance of *Bacteroidetes* in RTT subjects ($p = 0.002$, Wilcoxon rank-sum test; Additional file 8: Table S3 and Additional file 9: Table S4). Indeed, the significant increase of the *Firmicutes*/*Bacteroidetes* ratio, a rough estimator of intestinal dysbiosis, in RTT subjects (median value 3.95) compared to HC (median value 1.64) ($p = 0.003$, Wilcoxon rank-sum test) indicated the presence of an intestinal dysbiosis associated with RTT.

Table 1 Permutational multivariate analysis of variance (PERMANOVA) tests of the bacterial gut microbiota on the unweighted and weighted UniFrac distances and the Bray-Curtis dissimilarity according to individuals' health status and constipation

	Metric	*F*	R^2	*p* value*
HC (*n* = 29) vs RTT (*n* = 50)	Unweighted UniFrac	6.84	0.08	≤0.003
	Weighted UniFrac	15.6	0.16	≤0.003
	Bray-Curtis	6.94	0.08	≤0.003
HC (*n* = 29) vs RTT-C (*n* = 34)	Unweighted UniFrac	5.59	0.08	≤0.003
	Weighted UniFrac	12.1	0.16	≤0.003
	Bray-Curtis	5.39	0.08	≤0.003
HC (*n* = 29) vs RTT-NC (*n* = 14)	Unweighted UniFrac	5.27	0.11	≤0.003
	Weighted UniFrac	8.75	0.17	≤0.003
	Bray-Curtis	5.47	0.11	≤0.003
RTT-C (*n* = 34) vs RTT-NC (*n* = 14)	Unweighted UniFrac	1.31	0.02	0.180
	Weighted UniFrac	1.10	0.02	0.687
	Bray-Curtis	1.76	0.03	0.180

HC healthy controls, *RTT-C* constipated Rett syndrome subjects, *RTT-NC* non-constipated Rett syndrome subjects
*Bonferroni corrected *p* values

Analysis of the relative abundance of bacterial taxonomic groups at the genus level showed that the ten most abundant genera were *Bifidobacterium* (mean relative abundance, RTT, 36.7 %; HC, 17.2 %), *Bacteroides* (RTT, 12.3 %; HC, 18.3 %), *Faecalibacterium* (RTT, 3.6 %; HC, 9.2 %), *Lachnospiracea incertae sedis* (RTT, 4.6 %; HC, 3.9 %), *Blautia* (RTT, 4.7 %; HC, 3.7 %), *Escherichia/Shigella* (RTT, 5.2 %; HC, 2.4 %), *Alistipes* (RTT, 1.3 %; HC, 4.7 %), *Streptococcus* (RTT, 2.3 %; HC, 2.9 %), *Gemmiger* (RTT, 1.1 %; HC, 3.1 %) and *Ruminococcus* (RTT, 1.5 %; HC, 2.2 %) (Additional file 10: Figure S6 and Additional file 8: Table S3). Comparing the relative abundance of all the genera among the two groups of study, we discovered *Actinomyces*, *Bifidobacterium*, *Clostridium XIVa Eggerthella*, *Enterococcus*, *Erysipelotrichaceae incertae sedis*, *Escherichia/Shigella* and *Megasphaera*, as significantly more abundant in RTT subjects compared to HC while several other bacterial genera usually associated with a healthy human gut were less abundant in RTT subjects compared to HC ($p < 0.05$, Wilcoxon rank-sum test; Additional file 11: Figure S7, Additional file 8: Table S3 and Additional file 9: Table S4).

To define more precisely the taxa that were driving the differentiation of the microbiota of the groups, we performed an analysis based on PhyloRelief [22], a recent phylogenetic-based feature weighting algorithm for metagenomics data. Being independent from a precompiled taxonomy, PhyloRelief includes in the analysis unclassified taxa that would be otherwise ignored by other methods. The PhyloRelief analysis confirmed that the *Bifidobacterium* clade was significantly more represented in RTT subjects with respect to HC (Fig. 2). In addition, several OTUs classified as belonging to different members of Clostridia (e.g., *Anaerostipes*, *Clostridium XIVb*, *unknown-Lachnospiraceae*) as well as *Erysipelotrichaceae* (*Clostridium XVIII* and *Erysipelotrichaceae incertae sedis*), *Actinomyces*, *Lactobacillus*, *Eggerthella*, *Enterococcus* and *Enterobacteriaceae* (in particular *Escherichia/Shigella*) were significantly more abundant in the gut microbiota of RTT subjects compared to HC (Fig. 2). Remarkably, the PhyloRelief analysis identified significant differences in the taxa *Anaerostipes*, *Clostridium XIVb*, *Clostridium XVIII*, *Lactobacillus* and *Clostridium IV* that went undetected by using the Wilcoxon rank-sum test (Additional file 9: Table S4).

Differentially abundant taxa were further confirmed by LEfSe [23], an algorithm for high-dimensional biomarker discovery which exploits linear discriminant analysis (LDA) to robustly identify features statistically different among classes. Figure 3 shows the most relevant clades identified by LEfSe (logarithmic LDA score > 2.0; see also Additional file 12: Figure S8).

To evaluate the absolute amount of *Bifidobacterium* and validate their increase in RTT-associated dysbiosis, we performed quantitative PCR analysis (qPCR). We observed that *Bifidobacterium* was twofold more abundant in RTT subjects than in HC (median values $4.08*10^8$ CFU/g and $1.99*10^8$ CFU/g, respectively; $p = 0.009$, Wilcoxon rank-sum test; Fig. 4a). We also noticed that among the *Bifidobacterium* species investigated, *Bifidobacterium longum ssp. longum* was significantly more abundant in RTT subjects compared to HC (median values $4.94*10^8$ CFU/g and $2.44*10^8$ CFU/g respectively; $p = 0.019$, Wilcoxon rank-sum test; Fig. 4b).

High levels of faecal short chain fatty acids in RTT subjects

PICRUSt was used for inference of microbial metabolic pathways [24] in the gut microbiota of RTT subjects and HC (Additional file 13: Table S5). The analysis predicted the enrichment, among others, of carbohydrate and propanoate metabolism in the gut microbiota of RTT subjects, which are metabolic pathways related also to the metabolism of short chain fatty Acids (SCFAs). Since SCFAs are important for colonic health and may act on neuronal physiology [9, 25, 26], we measured the faecal content of SCFAs in our samples by means of GC-MS.

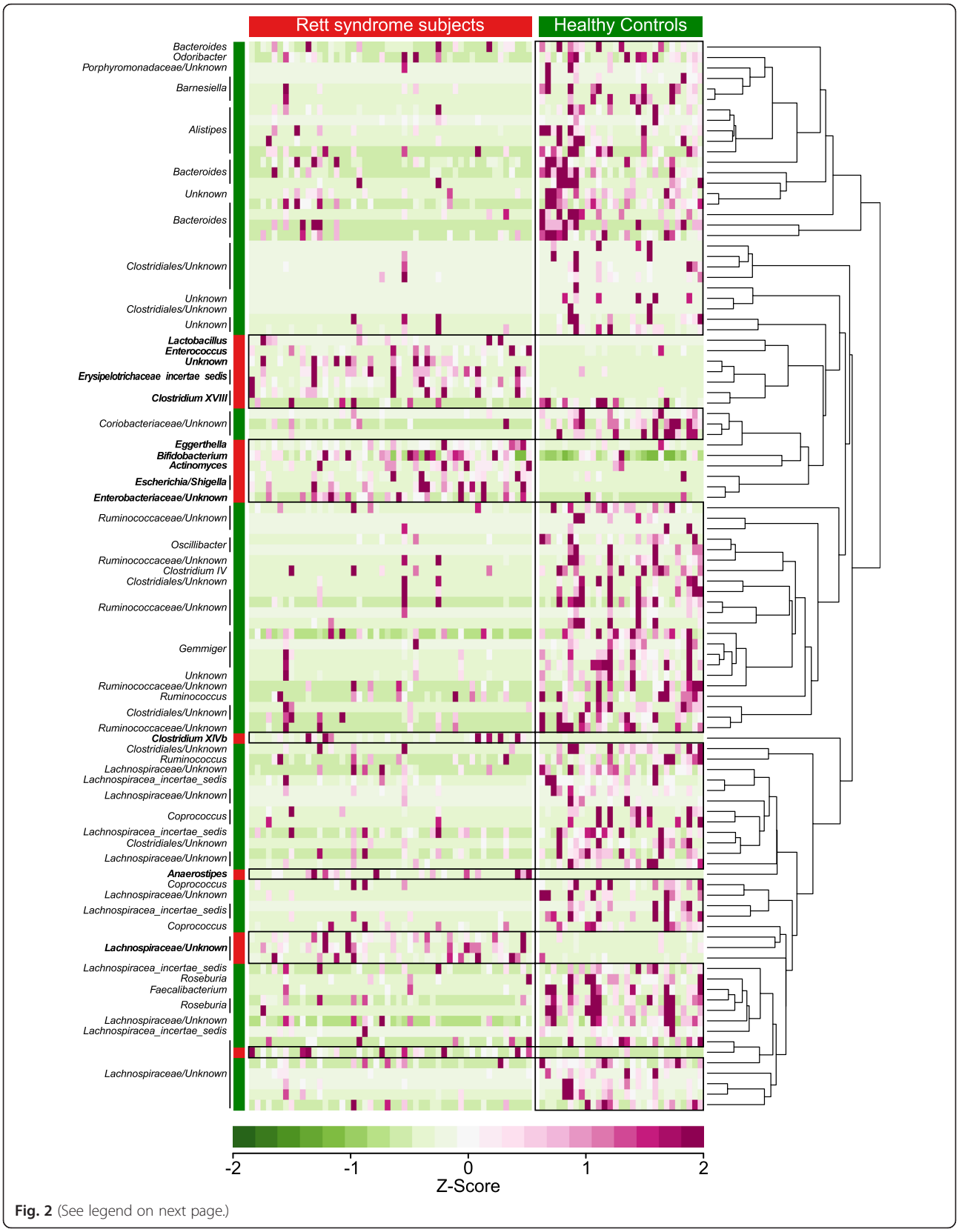


Fig. 2 (See legend on next page.)

(See figure on previous page.)
Fig. 2 PhyloRelief analysis (RTT vs HC) of bacterial OTUs using the unweighted UniFrac distance. The heat-map shows the relative abundances of the OTUs that are differentially represented in Rett syndrome (RTT) subjects and healthy controls (HC) (PhyloRelief selected clades with FDR-corrected $p < 0.01$, Kruskal-Wallis test). OTUs are classified according to their genus on the left side of the figure. The OTUs more represented in RTT subjects than HC are highlighted in *bold characters*. The ultrametric pruned phylogenetic tree of the OTUs is shown on the right side of the figure. RTT subjects and healthy controls are coloured in *red* and *green*, respectively. Abundances are expressed in terms of their z-score

We observed that the overall content of SCFAs in RTT subjects' faeces was higher than in HC (median values RTT, 191.5 μmol/g; HC 156.6 μmol/g). Particularly, we observed that propionate (median values RTT, 20.4 μmol/g; HC 13.2 μmol/g), isovalerate\2-methylbutyrate (median values RTT, 4.7 μmol/g; HC 2.2 μmol/g) and isobutyrate (median values RTT, 3.1 μmol/g; HC 1.6 μmol/g) were significantly more abundant in RTT subjects than HC ($p < 0.05$, Wilcoxon rank-sum test; Fig. 5).

RTT-associated gut mycobiota shows clear population composition differences compared to HC

The human gut mycobiota has been poorly explored so far, although there is an increasing awareness of its importance in human (patho)physiology [27]. We characterized the gut mycobiota of the study cohort by means of high-throughput sequencing of the ITS1 region of the ribosomal Internal Transcribed Spacer (ITS).

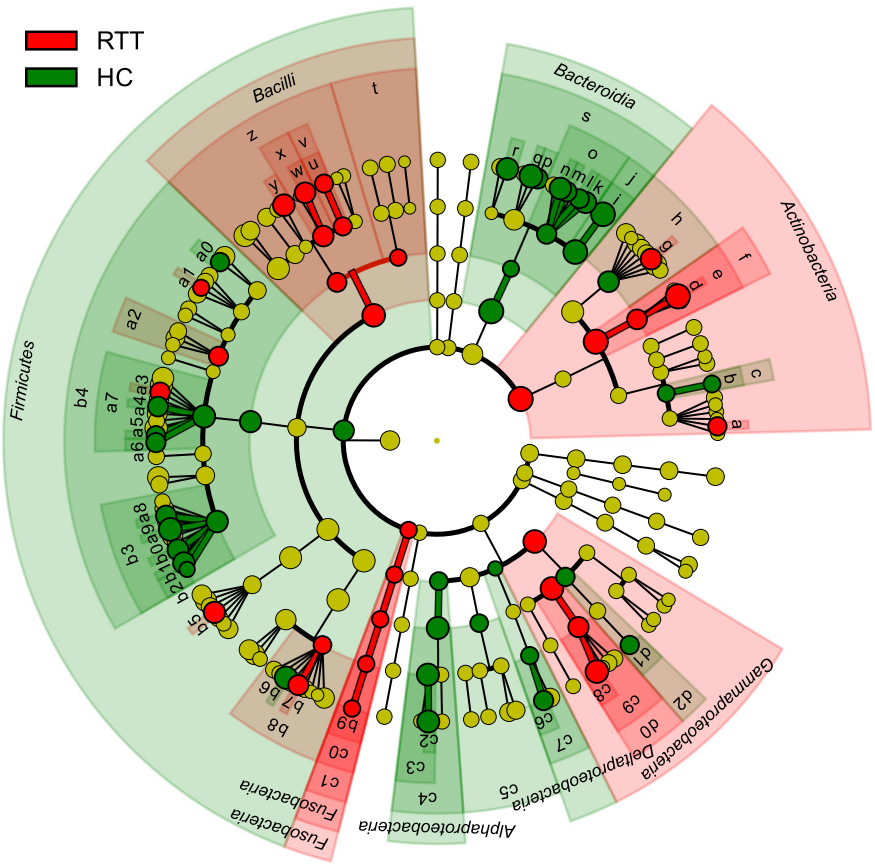
High-quality fungal sequences were detected respectively in 49 out of 50 RTT subjects and 28 out of 29 HC. The analysis of the *alpha*-diversity revealed that the gut mycobiota of RTT subjects, both constipated and non-constipated, was slightly less diverse compared to HC even if no significant differences were observed ($p > 0.05$, Wilcoxon rank-sum test; Additional file 14: Figure S9a). *Alpha* rarefaction curves (using 100 independent rarefactions) are reported in the Additional file 5: Figure S3b. As for the bacterial microbiota, a PERMANOVA analysis on the unweighted, weighted UniFrac distances and Bray-Curtis dissimilarity revealed that the gut mycobiota of RTT subjects was significantly different from that of HC ($p < 0.05$, PERMANOVA; Table 2, Fig. 6 and Additional file 14: Figure S9b), while no significant difference was detected between the gut mycobiota of constipated and non-constipated RTT subjects (Table 2). Multiple-rarefaction PCoA plots ("jackknifed" PCoA plots, [21]) (Additional file 6: Figure S4b and Additional file 7: Figure S5b) computed to assess the robustness of the fungal *beta*-diversity analyses showed that the unweighted UniFrac measure on the gut mycobiota was sensitive to rarefaction (although PERMANOVA *p* values were significant, see Additional file 7: Figure S5b), while the other *beta*-diversity measures support a differentiation between the gut mycobiota of HC and RTT subjects.

The genus *Candida* predominates in the altered gut mycobiota of RTT subjects

Metataxonomics analysis of the gut mycobiota led to the identification of 77 fungal taxa unambiguously classified to the genus level and 19 taxa only partially classified. The ten most abundant annotated fungal genera were *Candida* (RTT 61.3 %; HC 25.5 %), *Penicillium* (RTT 13.5 %; HC 19.4 %), *Aspergillus* (RTT 7.3 %; HC 6.5 %), *Malassezia* (RTT 3.5 %; HC 4.5 %), *Debaryomyces* (RTT 1.7 %; HC 5.5 %), *Mucor* (RTT 1.1 %; HC 4.6 %), *Eremothecium* (RTT 0.07 %; HC 3.7 %), *Pichia* (RTT 0.1 %; HC 3.5 %), *Cyberlindnera* (RTT 0.5 %; HC 1.7 %), and *Trichosporon* (RTT 1.3 %; HC 0.007 %) (Additional file 15: Figure S10). The relative abundance of the genus *Candida* was significantly higher in RTT subjects than HC ($p = 0.002$, Wilcoxon rank-sum test; Additional file 16: Figure S11). Remarkably, we detected sequences belonging to the single-cell protozoa *Blastocystis* in different healthy controls (in 24.1 % of the inspected healthy individuals) while this genus was present only in one RTT subject. *Blastocystis* is an important eukaryote of the GI tract of healthy individuals [28] being less common in subjects affected by irritable bowel syndrome and inflammatory bowel diseases [29]. The reported high relative abundance of the genus *Candida* in the gut mycobiota of RTT subjects was further confirmed by LEfSe analysis (Additional file 17: Figure S12).

Discussion

Our study identified a clear dysbiosis of the fungal and bacterial gut microbiota in individuals affected by RTT, a neurological disorder also associated with gastrointestinal symptoms and systemic inflammation. The elevated values of calprotectin and ESR measured in RTT subjects correlating also with the titre of serum IgA antibodies indicated the occurrence of an intestinal sub-inflammatory status, in line with previous indication of a pro-inflammatory status in *MeCP2*-related RTT [18]. A state of intestinal inflammation is also related to loss of intestinal barrier function and the subsequent translocation of pathobionts that may induce systemic responses [10]. RTT subjects displayed on average a lower bacterial gut microbiota richness and diversity compared to HC. We observed a significant increase in the *Firmicutes*/*Bacteroidetes* ratio in RTT subjects due to a reduction of the relative abundance of *Bacteroidetes* in these subjects. An increased *Firmicutes*/*Bacteroidetes* ratio has recently been reported also



a: *Actinomyces*
b: *Corynebacterium*
c: *Corynebacteriaceae*
d: *Bifidobacterium*
e: *Bifidobacteriaceae*
f: *Bifidobacteriales*
g: *Eggerthella*
h: *Coriobacteriaceae*
i: *Bacteroides*
j: *Bacteroidaceae*
k: *Barnesiella*
l: *Butyricimonas*
m: *Odoribacter*
n: *Parabacteroides*
o: *Porphyromonadaceae*
p: *Paraprevotella*
q: *Prevotella*
r: Alistipes
s: *Bacteroidales*
t: *Bacillales*
u: *Granulicatella*
v: *Carnobacteriaceae*
w: *Enterococcus*
x: *Enterococcaceae*
y: *Lactobacillus*
z: *Lactobacillales*
a0: *Sarcina*
a1: *Parvimonas*
a2: *Eubacteriaceae*
a3: *ClostridiumXIVa*
a4: *Coprococcus*
a5: *Lactonifactor*
a6: *Roseburia*
a7: *Lachnospiraceae*
a8: *Butyricicoccus*
a9: *Faecalibacterium*
b0: *Oscillibacter*
b1: *Ruminococcus*
b2: *Sporobacter*
b3: *Ruminococcaceae*
b4: *Clostridiales*
b5: *Erysipelotrichaceae incertae sedis*
b6: *Dialister*
b7: *Megasphaera*
b8: *Veillonellaceae*
b9: *Fusobacterium*
c0: *Fusobacteriaceae*
c1: *Fusobacteriales*
c2: *Gemmiger*
c3: *Hyphomicrobiaceae*
c4: *Rhizobiales*
c5: *Burkholderiales*
c6: *Bilophila*
c7: *Desulfovibrionaceae*
c8: *Escherichia/Shigella*
c9: *Enterobacteriaceae*
d0: *Enterobacteriales*
d1: *Haemophilus*
d2: *Pasteurellales*

Fig. 3 Cladogram showing the most discriminative bacterial clades identified by LEfSe. Coloured regions/branches indicate differences in the bacterial population structure between Rett syndrome (RTT) subjects and healthy controls (HC). Regions in *red* indicate clades that were enriched in RTT subjects compared to those in HC, while regions in *green* indicate clades that were enriched in HC compared to those in RTT subjects

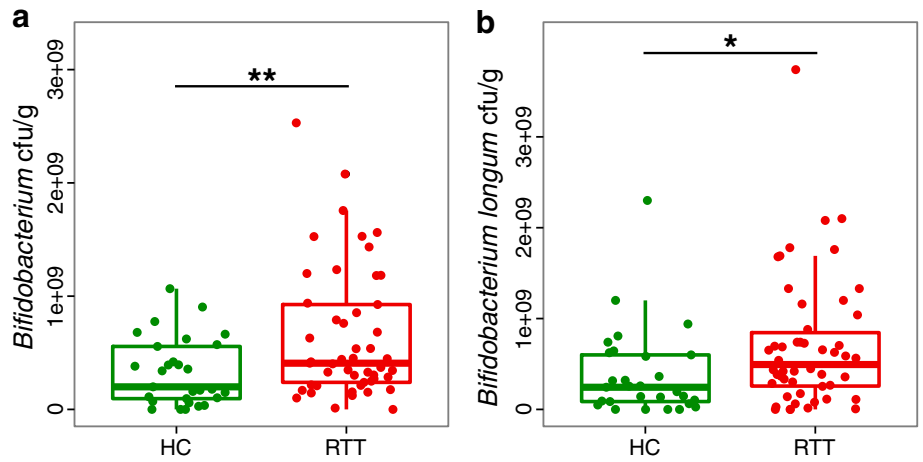


Fig. 4 Absolute quantification of *Bifidobacterium*. qPCR analysis of **a** the genus *Bifidobacterium* and **b** the species *Bifidobacterium longum ssp. longum* in Rett syndrome (RTT) subjects versus healthy controls (HC); $^{**}p < 0.01$, $^{*}p < 0.05$, Wilcoxon rank-sum test

in children affected by autism [30], and treatment with *Bacteroides fragilis* has been shown to restore autism-related behavioural and GI abnormalities in a mouse model of neurodevelopmental disorders [31].

An in-depth analysis of bacterial taxa revealed that the relative abundances of *Bifidobacterium* and several Clostridia, i.e. *Anaerostipes*, *Clostridium XIVa*, and *Clostridium XIVb*, as well as *Erysipelotrichaceae* (*Clostridium XVIII* and *Erysipelotrichaceae incertae sedis)*, *Actinomyces*, *Eggerthella*, *Enterococcus*, *Escherichia/Shigella* and *Lactobacillus*, were significantly higher in RTT subjects than in HC. Bifidobacteria are well recognized as health-promoting bacteria [32], with potential probiotic properties [33], and have been rarely associated with pathological states [34, 35]. Measuring the absolute abundances of the most common intestinal *Bifidobacterium* species by qPCR, we found that *Bifidobacterium longum ssp. longum* was the most abundant in RTT subjects, with absolute abundance twofold higher than in HC. The reported high abundance of *Bifidobacterium* in RTT subjects could indicate *Bifidobacterium* adaptation to the GI niche associated with RTT.

In line with our results, various studies on the gut microbiota of ASDs subjects reported also the increased incidence of Clostridia [36, 37], one of the most abundant Gram-positive bacteria known to reside in the human gut. Moreover, *Erysipelotrichaceae*, *Lactobacillus* and *Escherichia/Shigella* resulted to be enriched in ASDs [38, 39] in concomitance with a reduction of *Prevotella* [39], consistently with our observations. *Prevotella* is an important member of the human gut microbiota involved in the maintenance of the microbial community structure [40] while *Escherichia/Shigella* genera may exert pro-inflammatory activities and are abundant in

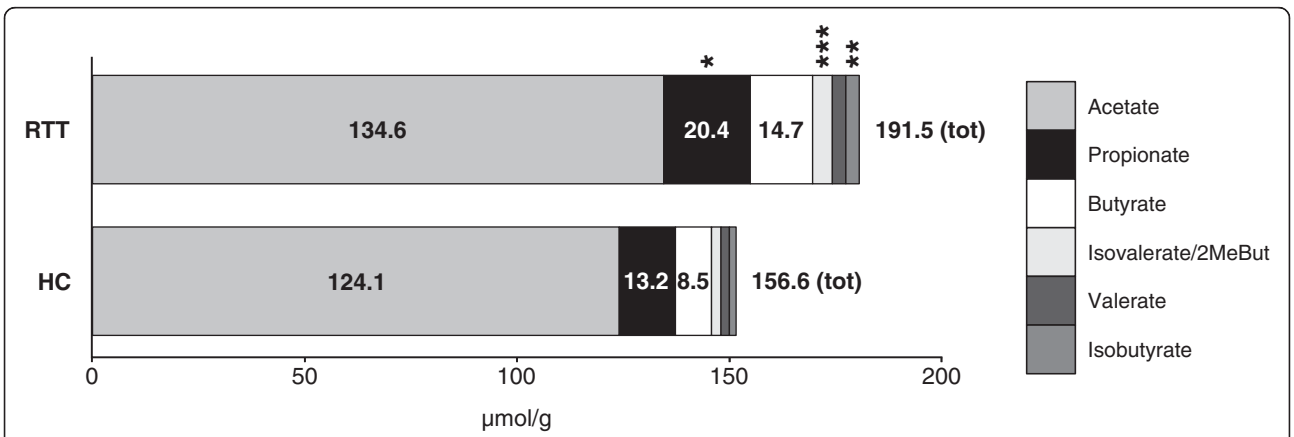


Fig. 5 SCFAs faecal content. Bar-plot representation of the median values of faecal SCFAs in Rett syndrome (RTT) subjects and healthy controls (HC); $^{*}p < 0.05$, $^{**}p < 0.005$, $^{***}p < 0.0005$, Wilcoxon rank-sum test

Table 2 Permutational multivariate analysis of variance (PERMANOVA) tests of the gut mycobiota on the unweighted and weighted UniFrac distances and the Bray-Curtis dissimilarity according to individuals' health status and constipation

	Metric	*F*	R^2	*p* value*
HC (*n* = 28) vs RTT (*n* = 49)	Unweighted UniFrac	2.76	0.03	0.006
	Weighted UniFrac	7.45	0.09	≤0.003
	Bray-Curtis	6.84	0.08	0.006
HC (*n* = 28) vs RTT-C (*n* = 33)	Unweighted UniFrac	2.23	0.03	0.018
	Weighted UniFrac	5.60	0.08	0.009
	Bray-Curtis	5.82	0.08	≤0.003
HC (*n* = 28) vs RTT-NC (*n* = 14)	Unweighted UniFrac	2.00	0.04	0.036
	Weighted UniFrac	5.95	0.12	≤0.003
	Bray-Curtis	4.59	0.10	0.006
RTT-C (*n* = 33) vs RTT-NC (*n* = 14)	Unweighted UniFrac	1.01	0.02	1.000
	Weighted UniFrac	1.27	0.02	0.837
	Bray-Curtis	1.00	0.02	1.000

HC healthy controls, *RTT-C* constipated Rett syndrome subjects, *RTT-NC* non-constipated Rett syndrome subjects
*Bonferroni corrected *p* values

subjects with active states of intestinal inflammation [41, 42]. However, given the lack of a consensus [14] and the methodological differences, it is difficult to draw general conclusions by directly comparing the results of the different studies on ASDs' gut microbiota.

It is well known that perturbations in the composition of commensal bacteria can predispose individuals to fungal infections [43]. We observed a dysbiotic gut mycobiota associated with RTT characterized by an altered community structure dominated by the genus *Candida*. *Candida* is one of the most common fungal commensals of the GI tract [28] but bacterial dysbiosis can shift *Candida* commensalism to pathogenesis, leading to extended infections and candidiasis [44]. It was also observed that the proportion of opportunistic pathogenic fungi, including *Candida*, increases in a mouse model of intestinal inflammation [12].

The herein described intestinal dysbiosis could be associated with changes in gut metabolite profiles consequently affecting the RTT gastrointestinal physiology. *Bifidobacterium* as well as *Anaerostipes*, *Clostridium XIVa*, *Clostridium XIVb* and *Clostridium XVIII* are known producer of SCFAs as fermentation end-products of carbohydrates and proteins [45]. Also, the lactic-acid bacteria *Lactobacillus* and *Enterococcus* can sustain the production of SCFAs through cross-feeding mechanisms that involves lactate-utilizing gut bacteria [46]. We observed that the faecal content of SCFAs in RTT subjects was significantly enriched in propionate, isobutyrate and isovalerate\2-methylbutyrate. Non-physiological high levels of SCFAs in the gut could contribute to GI symptoms, including the constipation status (which affect more of the 70 % of this RTT study cohort), through the alteration of goblet cell mucin discharge [47] and the inhibition of smooth muscle contraction in the colon [48] mediated by the release of the peptide YY from enteroendocrine cells [49]. Also, prolonged exposure to protein-derived SCFAs (among which isobutyrate, 2-methylbutyrate and isovalerate) and other protein fermentation products, such as ammonia, phenolic compounds or *p*-cresol, may affect the metabolism and the physiology of colonocytes [50]. Remarkably, enteric SCFAs, principally propionate and butyrate, can modulate gene expression, brain function and behaviour, affecting neurotransmitter systems, neuronal cell adhesion, inflammation, oxidative stress, lipid metabolism and

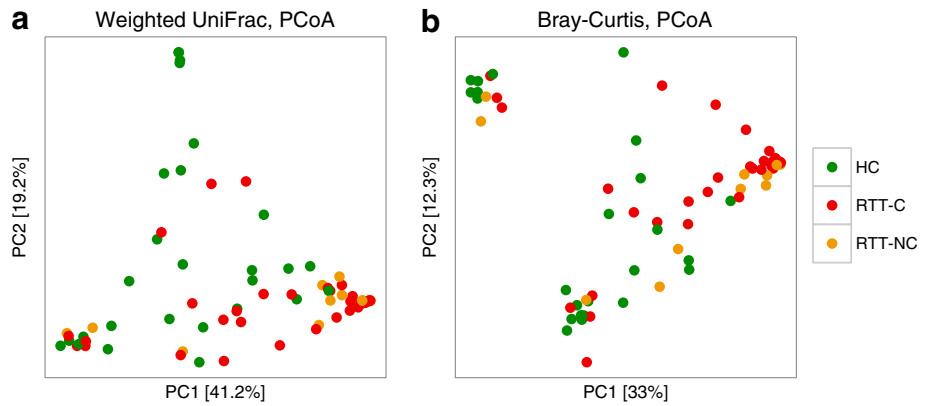


Fig. 6 Measures of fungal *beta*-diversity. PCoA plots of fungal *beta*-diversity based on **a** the Weighted UniFrac distance and **b** the Bray-Curtis dissimilarity analysed according to individuals' health status. Constipated Rett syndrome subjects (RTT-C), non-constipated Rett syndrome subjects (RTT-NC) and healthy controls (HC) are coloured in *red*, *orange* or *green*, respectively

mitochondrial function in rat and in vitro cell models of ASDs [51, 52].

Taken together, these observations suggest that functional impairments of MeCP2 favour the establishment of both constipation and adaptation of an intestinal dysbiotic microbial community that may reinforce the constipation status through non-physiological levels of SCFAs. Nevertheless, we are not able to infer the consequentiality of the two phenomenons, i.e. constipation or dysbiosis. The SCFAs produced by the gut microbiota are thus potentially implicated in the chronic constipation often associated with RTT, yet their role in the pathophysiology of RTT remains elusive. Indeed, the high faecal content of SCFAs in RTT subjects could be also related to a reduced intestinal absorption of these and other metabolites in the gut [53], or to increased liberation of SCFAs due to fibre retention in a costive gut. The establishment of intestinal dysbiosis, both at bacterial and fungal level, may reinforce, rather than determine, constipation, which is one of the most common GI problems in RTT. It is worth noting that the constipation status correlates with age suggesting that constipation and intestinal dysbiosis could be temporally connected with the progression of the disease. Furthermore, the reduction of the mucin layer due to the inhibition of goblet cells induced by the high levels of SCFAs might trigger an immunological response that might boost ERS values, serum IgA and the levels of calprotectin simplifying the putative translocation through the intestinal barrier of pathobionts equipped with proper virulence factors, such as *Candida* and *Escherichia/Shigella*, and overall contributing to systemic inflammation and cytokine dysregulation.

Conclusions

Here, we demonstrated for the first time that RTT is characterized by a dysbiotic bacterial and fungal microbiota showing an overall reduction of the microbial richness and diversity as well as an altered composition of the microbial community structure in RTT subjects. In particular, the increase in the relative abundance of *Bifidobacterium*, Clostridia and *Candida* drives the dysbiotic state associated with RTT. We hypothesize that impairments of MeCP2 functioning promote the establishment of a dysbiotic intestinal microbial community that, in turn, could affect RTT gastrointestinal physiopathology through altered SCFAs production, reinforcing the constipation status itself and favouring inflammation and cytokine dysregulation. Due to the importance that our findings might have in the design of potential therapeutic interventions aimed at gastrointestinal relief in RTT, we are planning to further investigate the gut microbiota dynamics during the progression of the disease in a *MeCP2*-null mouse model applying specific probiotics and prebiotics treatments.

Methods

Study participants and sample handling and collection

We recruited 50 female subjects with clinical diagnosis of RTT (average age 12 ± 7.3), genotyped for *MeCP2* and *CDKL5* gene mutations (Additional file 1: Table S1) and 29 age-matched healthy subjects as controls (average age 17 ± 9.6) (Additional file 18: Table S6). RTT subjects with clinically evident inflammatory conditions (i.e. upper respiratory tract infection, pneumonia, urinary infection, stomatitis and periodontal inflammation), either acute or chronic, were excluded. A "compressed" clinical severity score (CSS) was attributed to RTT subjects following thirteen criteria: regression, microcephaly, somatic growth, independent sitting, ambulation, hand use, scoliosis, language, nonverbal communication, respiratory dysfunction, autonomic symptoms, stereotypies and seizures [4]. ESR, CRP and serum IgA levels were assessed as markers of inflammation or GI abnormalities. Gastrointestinal symptoms (i.e. constipation) and intestinal inflammation (i.e. faecal calprotectin levels) [54] were also assessed (Additional file 1: Table S1). The diagnosis of constipation was defined according to Rome III criteria [55]. Stool samples from enrolled subjects were collected, aliquoted as it is and stored at −80 °C until analysis. All subjects of this study were under a Mediterranean-based diet and no antibiotics, probiotics or prebiotics have been taken in the 3 months prior the sample collection. The study was approved by the institutional review board of the Siena University Hospital (AOUS, Siena, Italy), and all enrolled subjects or tutors gave written informed consent in accordance with the sampling protocol approved by the local Ethical Committee (No. 2012-005021-76).

Faecal calprotectin assay

Calprotectin determination was performed by using a polyclonal antibody in an enzyme-linked immunosorbent assay (Calprest, Eurospital, Trieste, Italy).

Briefly, frozen stool samples were thawed at room temperature; 100 mg of faeces (wet weight) were weighed and placed in a disposable screw cap-tube containing the extraction buffer (weight/volume 1:50). Samples were then mixed vigorously for 30 s, homogenized for 25 min on a shaker, and centrifuged for 20 min at 10,000*g* at room temperature; 0.5 ml of clear extract supernatant was transferred to new tubes and stored at −80 °C. Finally, samples were diluted 1:50, and absorbance was measured at 405 nm. According to the manufacturer's instructions, normal values were considered <50 μg/g of calprotectin per faecal sample.

DNA extraction, PCR amplification of the V3-V5 region of bacterial 16S rDNA and of the ITS1 region of fungal rDNA

Total DNA extraction from faecal samples (250 mg, wet weight) was performed using the FastDNA™ SPIN Kit for Feces (MP Biomedicals, Santa Ana, CA, USA) following manufacturer's instructions. DNA integrity and quality were checked on 1 % agarose gel TAE 1X and quantified with a NanoDrop® spectrophotometer. For each DNA sample, 16S rRNA gene was amplified using fusion primer set specific for V3-V5 hypervariable regions (357F: 5′-TCCTACGGGAGGCAGCAG-3′ and 937R: 5′-TGTGCG GGCCCCCGTCAATT-3′) containing adaptors, key sequence and barcode (Multiple IDentifier) sequences as described by the 454 Sequencing System Guidelines for Amplicon Experimental Design (Roche, Basel, Switzerland).

PCR reactions were performed using the FastStart High Fidelity PCR system (Roche, Basel, Switzerland) according to the following protocol: 5 min at 95 °C, 25 cycles of 30 s at 95 °C, 30 s at 58 °C and 1 min at 72 °C, followed by a final extension of 8 min at 72 °C. The PCR reaction mix contained 1X FastStart High Fidelity PCR buffer 1.8 mM $MgCl_2$, 200 μM of dNTPs, 0.4 μM of each primer (Eurofins, PRIMM, Milano, Italy), 2.5 U of FastStart High Fidelity Polymerase Blend and 10 ng of gDNA as template. For ITS1 amplicon sequencing, fusion primer sets were designed as described above coupled with forward primer 18SF (5′-GTAAAAGTCG-TAACAAGGTTTC-3′) and reverse primer 5.8S1R (5′-GTTCAAAGAYTCGATGATTCAC-3′) [56] specific for fungal ITS1 rDNA region. The PCR reaction mix contained 1X FastStart High Fidelity PCR buffer, 2 mM $MgCl_2$, 200 μM of dNTPs, 0.4 μM of each primer (PRIMM, Milano, Italy), 2.5 U of FastStart High Fidelity Polymerase Blend and 100 ng of gDNA as template. Thermal cycling conditions used were 5 min at 95 °C, 35 cycles of 45 s at 95 °C, 45 s at 56 °C and 1.30 min at 72 °C followed by a final extension of 10 min at 72 °C. All PCR experiments were carried out in triplicates using a Veriti® Thermal Cycler (Applied Biosystems, Foster City, CA, USA).

Library construction and pyrosequencing

The PCR products obtained were analysed by gel electrophoresis and cleaned using the AMPure XP beads kit (Beckman Coulter, Brea, CA, USA) following the manufacturer's instructions, quantified via quantitative PCR using the Library quantification kit—Roche 454 titanium (KAPA Biosystems, Boston, MA) and pooled in equimolar way in a final amplicon library. The 454 pyrosequencing was carried out on the GS FLX+ system using the XL+ chemistry following the manufacturer's recommendations (Roche, Basel, Switzerland).

qPCR analysis

Amplifications of *Bifidobacterium sp.* 16S rDNA gene were performed in triplicate for each sample using the KAPA SYBR® FAST qPCR Kit Optimized for LightCycler® 480 (Kapa Biosystems, Inc., Wilmington, MA, USA) and the LightCycler® 480 II instrument (Roche, Basel, Switzerland) with primers and protocols described previously [57, 58]. The PCR reaction mix contained 1X KAPA SYBR FAST qPCR Master Mix, 0.2 μM of each *Bifidobacterium*-specific primer and 10 ng of gDNA as template. For quantification of each *Bifidobacterium* species, we constructed a seven-point standard curve consisting in tenfold serial dilutions of gDNA extracted from a pure culture at known concentration. For quantification of the genus *Bifidobacterium*, we used the strain *Bifidobacterium animalis ssp. lactis* BB12. The following *Bifidobacterium* type strains were used for the species-specific quantification: *B. angulatum* ATCC27535, *B. adolescentis* ATCC15703, *B. animalis ssp. lactis* ATCC15705, *B. bifidum* DSM20456, *B. breve* ATCC15700, *B. dentium* ATCC27534, *B. longum ssp. infantis* ATCC15697, *B. longum ssp. longum* ATCC15707 and *B. pseudocatenulatum* ATCC27917. Amplification specificity of target gene was checked by melting curve analysis. Efficiency and reliability of PCR amplifications were calculated.

SCFAs analysis

Frozen faecal samples (~150 mg, wet weight) were diluted 1:10 in sterile PBS 1 M (pH 7.2) and centrifuged at 13.000*g* for 5 min. Supernatants were then filtered using a 0.2-μm polycarbonate syringe filter and acidified by the addition of one volume of HCl 6 M to three volumes of sample. After 10-min incubation at room temperature, samples were centrifuged at 13.000*g* for 5 min. One volume of 10 mM 2-ethylbutyric acid was added to four volumes of sample as internal standard. Calibration was done using standard solutions of acetate, propionate, butyrate, isobutyrate, 2-methyl-butyrate (2-MeBut), valerate and isovalerate in acidified water (pH 2). Standard solutions containing 50, 20, 10, 5, 1 and 0.5 mM of each external standard were used.

Analysis was performed using a TRACE™ Ultra Gas Chromatograph (Thermo Scientific, Waltham, MA, USA) coupled to a TSQ Quantum GC mass spectrometric detector (Thermo Scientific, Waltham, MA, USA). SCFAs were separated using a Restek Stabilwax-DA (30 m × 0.25 mm; 0.25-μm film thickness) (Restek corp., Bellafonte, PA, USA). The injected sample volume was 1 μl in split mode with a ratio of 10:1. The initial oven temperature was at 90 °C and maintained for 0.5 min and then increased 20 °C/min to 240 °C. The carrier gas helium was delivered at a flow rate of 1 ml min^{-1}. The temperatures of the inlet,

transfer line and electron impact (EI) ion source were set at 280, 250 and 250 °C, respectively. The electron energy was 70 eV, and the mass spectral data was collected in a full scan mode (m/z 30–200).

Data analysis

Pyrosequencing resulted in a total of 2,227,864 reads for 16S rDNA with a mean of 27,987 ± 5782 sequences per sample and 1,678,227 reads for ITS1 region with a mean of 21,118 ± 13,270 sequences per sample. Raw 454 files were demultiplexed using the Roche's sff file software and submitted to the European Nucleotide Archive with accession number PRJEB12607. Sample accession IDs and metadata are available in Additional file 18: Table S6. Reads were pre-processed using the MICCA pipeline (http://www.micca.org) [59]. Forward and reverse primer trimming and quality filtering were performed using micca-preproc. De novo sequence clustering, chimera filtering and taxonomy assignment were performed by micca-otu-denovo: operational taxonomic units (OTUs) were assigned by clustering the sequences with a threshold of 97 % pairwise identity, and their representative sequences were classified using the RDP classifier version 2.7 on 16S rDNA data and using the RDP classifier version 2.8 [60] against the UNITE fungal ITS database [61] on ITS1 data. Template-guided multiple sequence alignment (MSA) was performed using PyNAST [62] (v. 0.1) against the multiple alignment of the Greengenes [63] database (release 13_05) filtered at 97 % similarity for bacterial sequences and through de novo MSA using T-Coffee [64] for fungal sequences. Fungal taxonomy assignments were also manually curated using BLASTn against the GenBank's database for accuracy. High-quality fungal sequences have been also manually filtered out for sequences belonging to *Agaricomycetes* (unlikely to be residents of the human gut due to their ecology) [65]. Samples with less than 1000 reads have been excluded from the analysis. The phylogenetic trees were inferred using micca-phylogeny [66]. Sampling heterogeneity was reduced rarefying samples at 90 % of the less abundant sample (16S data) and at the depth of the less abundant sample (ITS1 data). *Alpha-* (within-sample richness) and *beta*-diversity (between-sample dissimilarity) estimates were computed using the phyloseq R package [67]. Multiple-rarefaction PCoA plots ("jackknifed" PCoA plots, [21]) were computed to assess the robustness of the *beta*-diversity analyses. Permutational MANOVA (PERMANOVA) was performed on the UniFrac distances and Bray-Curtis dissimilarity using the adonis() function of the *vegan* R package with 999 permutations, and p values were corrected using the Bonferroni correction [68]. The non-parametric Wilcoxon rank-sum test was used for the comparison of relative abundances of microbial taxa between groups, and the resulting p values were corrected for multiple testing controlling the false discovery rate [69] at all taxonomic levels taken into account. Further identification of taxa differentially distributed in case/control groups was obtained by PhyloRelief, a phylogenetic-based feature weighting algorithm for metagenomics data. This method unambiguously groups taxa into clades without relying on a precompiled taxonomy and accomplishes a ranking of the clades according to their contribution to the sample differentiation [22]. Linear discriminant effect size analysis (LEfSe) [23] with default parameters was performed to find taxonomic clades differentially represented between RTT subjects and HC. LEfSe combines Kruskal-Wallis test and Wilcoxon rank-sum tests with linear discriminant analysis (LDA). LEfSe ranks features by effect size, putting at the top features that explain most of the biological difference. In order to investigate the microbial metabolic potential of the gut microbiota in healthy controls and RTT subjects, we applied PICRUSt (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States) [24], a computational approach used to predict the functional composition of a metagenome using marker gene data and a database of reference genomes. PICRUSt uses an extended ancestral-state reconstruction algorithm to predict which gene families are present and then combines gene families to estimate the composite metagenome starting from the taxonomic composition estimated from 16S rDNA data. Starting from a table of OTUs with associated Greengenes identifiers, we obtained the final output from metagenome prediction as an annotated table of predicted gene family counts for each sample, where the encoded function of each gene family be orthologous groups or other identifiers such as KEGG orthologues. Spearman's correlation tests for each correlation were computed using the *psych* R package [70]. All statistical analyses were performed using R [71].

Additional files

Additional file 1: Table S1. Characteristics of study participants affected by Rett syndrome. (DOCX 24 kb)

Additional file 2: Figure S1. Correlation plots of clinical data from RTT subjects. Significant positive correlations were observed among age, IgA and ESR. (PDF 37 kb)

Additional file 3: Table S2. Spearman's correlation analysis among RTT clinical data. (DOCX 14 kb)

Additional file 4: Figure S2. Measures of bacterial diversity. a) *Alpha*-diversity estimated on the Chao1 estimator and the Shannon entropy; ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$; Wilcoxon rank-sum test. b) PCoA plot based on the unweighted UniFrac distance among samples analysed according to individuals' health status. Constipated Rett syndrome subjects (RTT-C), non-constipated Rett syndrome subjects (RTT-NC) and

healthy controls (HC) are coloured in red, orange or green, respectively. (PDF 58 kb)

Additional file 5: Figure S3. *Alpha*-diversity rarefaction curves. The plot shows the *alpha*-diversity for HC, RTT-C and RTT-NC averaged over 100 independent rarefactions as a function of the rarefaction depth for a) the bacterial gut microbiota and b) the fungal gut microbiota. The points in the curves are the averages, while the whiskers represent the standard deviations. (PDF 47 kb)

Additional file 6: Figure S4. Multiple-rarefaction PCoA plots. Each PCoA replicate was optimally superimposed by Procrustes analysis on the master PCoA scatter plot (used in the main text). Points represent the average location of 100 rarefaction replicates. Ellipses show the 95 % confidence region assuming a multivariate normal distribution. a) PCoA plots of bacterial *beta*-diversity based on the unweighted and weighted UniFrac distances and the Bray-Curtis dissimilarity analysed according to individuals' health status; b) PCoA plots of fungal *beta*-diversity based on the unweighted and weighted UniFrac distances and the Bray-Curtis dissimilarity analysed according to individuals' health status. Constipated Rett syndrome subjects (RTT-C), non-constipated Rett syndrome subjects (RTT-NC) and healthy controls (HC) are coloured in red, orange or green, respectively. (PDF 185 kb)

Additional file 7: Figure S5. PERMANOVA *p* values distribution. For each rarefaction replicate (n = 100), a PERMANOVA test was conducted to assess the robustness of the results over the rarefaction. a) PERMANOVA *p* values distributions of bacterial *beta*-diversity based on the unweighted and weighted UniFrac distances and the Bray-Curtis dissimilarity analysed according to individuals' health status; b) PERMANOVA *p* values distributions of fungal *beta*-diversity based on the unweighted and weighted UniFrac distances and the Bray-Curtis dissimilarity analysed according to individuals' health status. (PDF 27 kb)

Additional file 8: Table S3. Mean relative abundance (%) ± standard deviation (SD) of bacterial OTUs at the phylum and genus levels in Rett syndrome (RTT) subjects and healthy controls (HC). (DOCX 32 kb)

Additional file 9: Table S4. Wilcoxon rank-sum test comparison of bacterial relative abundances at phylum and genus levels. (DOCX 22 kb)

Additional file 10: Figure S6. Genus level relative abundances of the bacterial gut microbiota of healthy controls (HC) and Rett syndrome (RTT) subjects. (PDF 5470 kb)

Additional file 11: Figure S7. Bacterial taxa which relative abundances were significantly different ($p < 0.05$; Wilcoxon rank-sum test) between healthy controls (HC) and Rett syndrome (RTT) subjects. (PDF 319 kb)

Additional file 12: Figure S8. LDA scores of the most discriminant bacterial taxa identified by LEfSe. Positive and negative LDA scores indicate the taxa enriched in healthy controls (HC) and Rett syndrome (RTT) subjects, respectively. (PDF 62 kb)

Additional file 13: Table S5. Statistics of the significantly different metabolic pathways (KEGG categories) inferred with PICRUSt in the gut microbiota of healthy controls (HC) and Rett syndrome (RTT) subjects (Welch's *t* test, $p < 0.05$ FDR-corrected) from 16S rDNA data. (DOCX 26 kb)

Additional file 14: Figure S9. Measures of fungal diversity. a) Three estimators of *alpha*-diversity have been calculated: the number of observed OTUs, the Chao1 estimator and the Shannon entropy; b) PCoA plot based on the unweighted UniFrac distance among samples analysed according to individuals' health status. Constipated Rett syndrome subjects (RTT-C), non-constipated Rett syndrome subjects (RTT-NC) and healthy controls (HC) are coloured in red, orange or green, respectively. (PDF 63 kb)

Additional file 15: Figure S10. Genus level relative abundances of the fungal gut microbiota of healthy controls (HC) and Rett syndrome (RTT) subjects. (PDF 2981 kb)

Additional file 16: Figure S11. *Candida* relative abundance in the gut microbiota of healthy controls (HC) and Rett syndrome (RTT) subjects ($p < 0.05$; Wilcoxon rank-sum test). (PDF 20 kb)

Additional file 17: Figure S12. a) LDA scores of the most discriminant fungal taxa identified by LEfSe. Positive and negative LDA scores indicate the taxa enriched in healthy controls (HC) and Rett syndrome (RTT) subjects, respectively. b) Cladogram showing the most discriminative fungal clades identified by LEfSe. Coloured regions/branches indicate differences in the fungal population structure between Rett syndrome (RTT) subjects and healthy controls (HC). Regions in red indicate clades that were enriched in RTT subjects compared to those in HC, while regions in green indicate clades that were enriched in HC compared to those in RTT subjects. (PDF 116 kb)

Additional file 18: Table S6. Correspondences among deposited metagenomics data and samples, unrarefied OTU tables and taxonomic classifications of the 16S and ITS1 datasets. (XLSX 449 kb)

Additional file 19. MICCA pipelines used for the analysis of the 16S and ITS1 datasets. (PDF 7 kb)

Abbreviations

ASDs, autism spectrum disorders; CDKL5, cyclin-dependent kinase-like 5; CNS, central nervous system; CRP, C-reactive protein; CSS, clinical severity score; ENS, enteric nervous system; ESR, erythrocyte sedimentation rate; FDR, false discovery rate; GC-MS, gas chromatography–mass spectrometry; GI, gastrointestinal; HC, healthy controls; ITS, internal transcribed spacer; LDA, linear discriminant analysis; LEfSe, linear discriminant effect size analysis; MeCP2, methyl-CpG binding protein 2; OTU, operational taxonomic unit; PCoA, principal coordinates analysis; PERMANOVA, permutational multivariate analysis of variance; PICRUSt, phylogenetic investigation of communities by reconstruction of unobserved states; qPCR, quantitative PCR; RTT, Rett syndrome; RTT-C, constipated Rett syndrome subjects, RTT-NC, non-constipated Rett syndrome subjects; SCFAs, short chain fatty acids.

Acknowledgements

The authors would like to thank Prof. Paola Mattarelli from the University of Bologna for providing the *Bifidobacterium* strains used in this work; Maddalena Sordo, Francesca Fava, Urska Vrhovsek, Silvia Carlin and the lab staff of the Sequencing Platform from Fondazione Edmund Mach for their technical support and Kieran Tuohy and Nicola Segata for their helpful discussion. We thank also Unifarm S.p.A. for the financial support to FS.

Funding

This work was financially supported from the "Accordo di Programma Integrato MetaFoodLabs" and funded by the research office of the Provincia Autonoma di Trento (Italy) (PAT Prot. S116/2012/537723).

Availability of supporting data

Raw sequences are available in the European Nucleotide Archive (ENA) with accession number PRJEB12607 (http://www.ebi.ac.uk/ena/data/view/PRJEB12607). Sample metadata, unrarefied OTU tables and taxonomic classifications are available in the Additional file 18: Table S6. Furthermore, the pipelines used for the processing of the raw data are available as Additional file 19.

Authors' contributions

FS designed and performed the experiments, analysed the data and wrote the manuscript. DR and SL analysed and collected the clinical data. IS, DA, MP and CD supervised and contributed to the data analysis. CDFe, JH and AC recruited the subjects and collected the specimens. IS, LR, OJ, AC and CD critically reviewed the manuscript. DC, AC and CDF conceived the study and approved the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Written informed consent was obtained from the enrolled subjects or tutors.

Ethics approval and consent to participate

The study was approved by the institutional review board of the Siena University Hospital (AOUS, Siena, Italy), and all enrolled subjects or tutors gave written informed consent in accordance with the sampling protocol approved by the local Ethical Committee (No. 2012-005021-76).

Author details
[1]Department of Computational Biology, Research and Innovation Centre, Fondazione Edmund Mach, Via E. Mach 1, 38010 San Michele all' Adige, Italy. [2]Centre for Integrative Biology, University of Trento, Via Sommarive 9, 38123 Trento, Italy. [3]Department of Biology, University of Florence, Via Madonna del Piano 6, 50019 Sesto Fiorentino, Florence, Italy. [4]Neonatal Intensive Care Unit, University Hospital AOUS, Viale Bracci 16, 53100 Siena, Italy. [5]Child Neuropsychiatry Unit, University Hospital AOUS, Viale Bracci 16, 53100 Siena, Italy. [6]Department of Genomics and Biology of Fruit Crop, Research and Innovation Centre, Fondazione Edmund Mach, Via E. Mach 1, 38010 San Michele all' Adige, Italy. [7]Department of Experimental and Clinical Biomedical Sciences, Gastroenterology Unit, University of Florence, Viale Morgagni 40, 50139 Florence, Italy. [8]Institute of Biometeorology (IBIMET), National Research Council (CNR), Via Giovanni Caproni 8, I-50145 Florence, Italy.

Received: 18 April 2016 Accepted: 21 July 2016
Published online: 30 July 2016

References

1. Chahrour M, Zoghbi HY. The story of Rett syndrome: from clinic to neurobiology. Neuron. 2007;56(3):422–37.
2. Liyanage VR, Rastegar M. Rett syndrome and MeCP2. Neruomol Med. 2014; 16(2):231–64.
3. Cuddapah VA, Pillai RB, Shekar KV, Lane JB, Motil KJ, Skinner SA, Tarquinio DC, Glaze DG, McGwin G, Kaufmann WE, et al. Methyl-CpG-binding protein 2 (MECP2) mutation type is associated with disease severity in Rett syndrome. J Med Genet. 2014;51(3):152–8.
4. Neul JL, Fang P, Barrish J, Lane J, Caeg EB, Smith EO, Zoghbi H, Percy A, Glaze DG. Specific mutations in methyl-CpG-binding protein 2 confer different severity in Rett syndrome. Neurology. 2008;70(16):1313–21.
5. Gonzales ML, LaSalle JM. The role of MeCP2 in brain development and neurodevelopmental disorders. Curr Psychiatry Rep. 2010;12(2):127–34.
6. Motil KJ, Caeg E, Barrish JO, Geerts S, Lane JB, Percy AK, Annese F, McNair L, Skinner SA, Lee HS, et al. Gastrointestinal and nutritional problems occur frequently throughout life in girls and women with Rett syndrome. J Pediatr Gastroenterol Nutr. 2012;55(3):292–8.
7. Leonard H, Ravikumara M, Baikie G, Naseem N, Ellaway C, Percy A, Abraham S, Geerts S, Lane J, Jones M, et al. Assessment and management of nutrition and growth in Rett syndrome. J Pediatr Gastroenterol Nutr. 2013;57(4):451–60.
8. Wahba G, Schock SC, Claridge E, Bettolli M, Grynspan D, Humphreys P, Staines WA. MeCP2 in the enteric nervous system. Neurogastroenterol Motil. 2015;27(8):1156–61.
9. Maslowski KM, Mackay CR. Diet, gut microbiota and immune responses. Nat Immunol. 2011;12(1):5–9.
10. Kamada N, Seo SU, Chen GY, Nunez G. Role of the gut microbiota in immunity and inflammatory disease. Nat Rev Immunol. 2013;13(5):321–35.
11. Underhill DM, Iliev ID. The mycobiota: interactions between commensal fungi and the host immune system. Nat Rev Immunol. 2014;14(6):405–16.
12. Iliev ID, Funari VA, Taylor KD, Nguyen Q, Reyes CN, Strom SP, Brown J, Becker CA, Fleshner PR, Dubinsky M, et al. Interactions between commensal fungi and the C-type lectin receptor Dectin-1 influence colitis. Science. 2012;336(6086):1314–7.
13. Clemente JC, Ursell LK, Parfrey LW, Knight R. The impact of the gut microbiota on human health: an integrative view. Cell. 2012;148(6):1258–70.
14. Mayer EA, Padua D, Tillisch K. Altered brain-gut axis in autism: comorbidity or causative mechanisms? Bioessays. 2014;36(10):933–9.
15. Sampson TR, Mazmanian SK. Control of brain development, function, and behavior by the microbiome. Cell Host Microbe. 2015;17(5):565–76.
16. Wang Y, Kasper LH. The role of microbiome in central nervous system disorders. Brain Behav Immun. 2014;38:1–12.
17. Leoncini S, De Felice C, Signorini C, Zollo G, Cortelazzo A, Durand T, Galano JM, Guerranti R, Rossi M, Ciccoli L, et al. Cytokine dysregulation in MECP2- and CDKL5-related Rett syndrome: relationships with aberrant redox homeostasis, inflammation, and omega-3 PUFAs. Oxidative Med Cell Longev. 2015;2015:421624.
18. Cortelazzo A, De Felice C, Guerranti R, Signorini C, Leoncini S, Pecorelli A, Zollo G, Landi C, Valacchi G, Ciccoli L, et al. Subclinical inflammatory status in Rett syndrome. Mediat Inflamm. 2014;2014:480980.
19. Bochen K, Krasowska A, Milaniuk S, Kulczynska M, Prystupa A, Dzida G. Erythrocyte sedimentation rate–an old marker with new applications. Journal of Pre-clinical and Clinical Research 2011;5(2):50–55.
20. Joshi S, Lewis SJ, Creanor S, Ayling RM. Age-related faecal calprotectin, lactoferrin and tumour M2-PK concentrations in healthy volunteers. Ann Clin Biochem. 2010;47(Pt 3):259–63.
21. Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R. UniFrac: an effective distance metric for microbial community comparison. ISME J. 2011;5(2):169–72.
22. Albanese D, De Filippo C, Cavalieri D, Donati C. Explaining diversity in metagenomic datasets by phylogenetic-based feature weighting. PLoS Comput Biol. 2015;11(3):e1004186.
23. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C. Metagenomic biomarker discovery and explanation. Genome Biol. 2011;12(6):R60.
24. Langille MG, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, Clemente JC, Burkepile DE, Vega Thurber RL, Knight R, et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat Biotechnol. 2013;31(9):814–21.
25. Mitchell RW, On NH, Del Bigio MR, Miller DW, Hatch GM. Fatty acid transport protein expression in human brain and potential role in fatty acid transport across human brain microvessel endothelial cells. J Neurochem. 2011;117(4):735–46.
26. Frost G, Sleeth ML, Sahuri-Arisoylu M, Lizarbe B, Cerdan S, Brody L, Anastasovska J, Ghourab S, Hankir M, Zhang S, et al. The short-chain fatty acid acetate reduces appetite via a central homeostatic mechanism. Nat Commun. 2014;5:3611.
27. Rizzetto L, De Filippo C, Cavalieri D. Richness and diversity of mammalian fungal communities shape innate and adaptive immunity in health and disease. Eur J Immunol. 2014;44(11):3166–81.
28. Scanlan PD, Marchesi JR. Micro-eukaryotic diversity of the human distal gut microbiota: qualitative assessment using culture-dependent and -independent analysis of faeces. ISME J. 2008;2(12):1183–93.
29. Andersen LO, Stensvold CR. Blastocystis in Health and Disease: Are We Moving from a Clinical to a Public Health Perspective? J Clin Microbiol. 2016;54(3):524–528.
30. Tomova A, Husarova V, Lakatosova S, Bakos J, Vlkova B, Babinska K, Ostatnikova D. Gastrointestinal microbiota in children with autism in Slovakia. Physiol Behav. 2015;138:179–87.
31. Hsiao EY, McBride SW, Hsien S, Sharon G, Hyde ER, McCue T, Codelli JA, Chow J, Reisman SE, Petrosino JF, et al. Microbiota modulate behavioral and physiological abnormalities associated with neurodevelopmental disorders. Cell. 2013;155(7):1451–63.
32. Ventura M, Turroni F, Motherway MO, MacSharry J, van Sinderen D. Host-microbe interactions that facilitate gut colonization by commensal bifidobacteria. Trends Microbiol. 2012;20(10):467–76.
33. Fanning S, Hall LJ, Cronin M, Zomer A, MacSharry J, Goulding D, Motherway MO, Shanahan F, Nally K, Dougan G, et al. Bifidobacterial surface-exopolysaccharide facilitates commensal-host interaction through immune modulation and pathogen protection. Proc Natl Acad Sci U S A. 2012;109(6):2108–13.
34. Pathak P, Trilligan C, Rapose A. Bifidobacterium–friend or foe? A case of urinary tract infection with Bifidobacterium species. BMJ case reports. 2014; doi:10.1136/bcr-2014-205122.
35. Tena D, Losa C, Medina MJ, Saez-Nieto JA. Peritonitis caused by Bifidobacterium longum: case report and literature review. Anaerobe. 2014;27:27–30.
36. Song Y, Liu C, Finegold SM. Real-time PCR quantitation of clostridia in feces of autistic children. Appl Environ Microbiol. 2004;70(11):6459–65.
37. Parracho HM, Bingham MO, Gibson GR, McCartney AL. Differences between the gut microflora of children with autistic spectrum disorders and that of healthy children. J Med Microbiol. 2005;54(Pt 10):987–91.
38. de Theije CG, Wopereis H, Ramadan M, van Eijndthoven T, Lambert J, Knol J, Garssen J, Kraneveld AD, Oozeer R. Altered gut microbiota and activity in a murine model of autism spectrum disorders. Brain Behav Immun. 2014;37:197–206.
39. Kang DW, Park JG, Ilhan ZE, Wallstrom G, Labaer J, Adams JB, Krajmalnik-Brown R. Reduced incidence of Prevotella and other fermenters in intestinal microflora of autistic children. PLoS One. 2013;8(7):e68322.
40. Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, Fernandes GR, Tap J, Bruls T, Batto JM, et al. Enterotypes of the human gut microbiome. Nature. 2011;473(7346):174–80.
41. Morgan XC, Tickle TL, Sokol H, Gevers D, Devaney KL, Ward DV, Reyes JA, Shah SA, LeLeiko N, Snapper SB, et al. Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. Genome Biol. 2012;13(9):R79.
42. Arthur JC, Gharaibeh RZ, Mühlbauer M, Perez-Chanona E, Uronis JM, McCafferty J, Fodor AA, Jobin C. Microbial genomic analysis reveals the

essential role of inflammation in bacteria-induced colorectal cancer. Nat Commun. 2014;5:4724. doi:10.1038/ncomms5724.
43. Oever JT, Netea MG. The bacteriome-mycobiome interaction and antifungal host defense. Eur J Immunol. 2014;44(11):3182–91.
44. Erb Downward JR, Falkowski NR, Mason KL, Muraglia R, Huffnagle GB. Modulation of post-antibiotic bacterial community reassembly and host response by Candida albicans. Sci Rep. 2013;3:2191.
45. Cummings J, Macfarlane G. The control and consequences of bacterial fermentation in the human colon. J Appl Bacteriol. 1991;70(6):443–59.
46. Duncan SH, Louis P, Flint HJ. Lactate-utilizing bacteria, isolated from human feces, that produce butyrate as a major fermentation product. Appl Environ Microbiol. 2004;70(10):5810–7.
47. Barcelo A, Claustre J, Moro F, Chayvialle JA, Cuber JC, Plaisancie P. Mucin secretion is modulated by luminal factors in the isolated vascularly perfused rat colon. Gut. 2000;46(2):218–24.
48. Squires PE, Rumsey RD, Edwards CA, Read NW. Effect of short-chain fatty acids on contractile activity and fluid flow in rat colon in vitro. Am J Phys. 1992;262(5 Pt 1):G813–817.
49. Cherbut C, Ferrier L, Roze C, Anini Y, Blottiere H, Lecannu G, Galmiche JP. Short-chain fatty acids modify colonic motility through nerves and polypeptide YY release in the rat. Am J Phys. 1998;275(6 Pt 1):G1415–1422.
50. Blachier F, Mariotti F, Huneau JF, Tome D. Effects of amino acid-derived luminal metabolites on the colonic epithelium and physiopathological consequences. Amino Acids. 2007;33(4):547–62.
51. Nankova BB, Agarwal R, MacFabe DF, La Gamma EF. Enteric bacterial metabolites propionic and butyric acid modulate gene expression, including CREB-dependent catecholaminergic neurotransmission, in PC12 cells—possible relevance to autism spectrum disorders. PLoS One. 2014;9(8):e103740.
52. MacFabe DF. Enteric short-chain fatty acids: microbial messengers of metabolism, mitochondria, and mind: implications in autism spectrum disorders. Microb Ecol Health Dis. 2015;26:28177.
53. Manokas T, Fromkes JJ, Sundaram U. Effect of chronic inflammation on ileal short-chain fatty acid/bicarbonate exchange. Am J Physiol Gastrointest Liver Physiol. 2000;278(4):G585–590.
54. Aomatsu T, Yoden A, Matsumoto K, Kimura E, Inoue K, Andoh A, Tamai H. Fecal calprotectin is a useful marker for disease activity in pediatric patients with inflammatory bowel disease. Dig Dis Sci. 2011;56(8):2372–7.
55. Longstreth GF, Thompson WG, Chey WD, Houghton LA, Mearin F, Spiller RC. Functional bowel disorders. Gastroenterology. 2006;130(5):1480–91.
56. Findley K, Oh J, Yang J, Conlan S, Deming C, Meyer JA, Schoenfeld D, Nomicos E, Park M, Program NIHISCCS, et al. Topographic diversity of fungal and bacterial communities in human skin. Nature. 2013;498(7454):367–70.
57. Rinttila T, Kassinen A, Malinen E, Krogius L, Palva A. Development of an extensive set of 16S rDNA-targeted primers for quantification of pathogenic and indigenous bacteria in faecal samples by real-time PCR. J Appl Microbiol. 2004;97(6):1166–77.
58. Matsuki T, Watanabe K, Fujimoto J, Kado Y, Takada T, Matsumoto K, Tanaka R. Quantitative PCR with 16S rRNA-gene-targeted species-specific primers for analysis of human intestinal bifidobacteria. Appl Environ Microbiol. 2004;70(1):167–73.
59. Albanese D, Fontana P, De Filippo C, Cavalieri D, Donati C. MICCA: a complete and accurate software for taxonomic profiling of metagenomic data. Sci Rep. 2015;5:9743.
60. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol. 2007;73(16):5261–7.
61. Koljalg U, Nilsson RH, Abarenkov K, Tedersoo L, Taylor AF, Bahram M, Bates ST, Bruns TD, Bengtsson-Palme J, Callaghan TM, et al. Towards a unified paradigm for sequence-based identification of fungi. Mol Ecol. 2013;22(21):5271–7.
62. Caporaso JG, Bittinger K, Bushman FD, DeSantis TZ, Andersen GL, Knight R. PyNAST: a flexible tool for aligning sequences to a template alignment. Bioinformatics. 2010;26(2):266–7.
63. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol. 2006;72(7):5069–72.
64. Notredame C, Higgins DG, Heringa J. T-Coffee: a novel method for fast and accurate multiple sequence alignment. J Mol Biol. 2000;302(1):205–17.
65. Hibbett DS. A phylogenetic overview of the Agaricomycotina. Mycologia. 2006;98(6):917–25.
66. Price MN, Dehal PS, Arkin AP. FastTree 2—approximately maximum-likelihood trees for large alignments. PLoS One. 2010;5(3):e9490.
67. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One. 2013;8(4):e61217.
68. Bonferroni CE. Teoria statistica delle classi e calcolo delle probabilità. 8th ed Florence R. Istituto superiore di scienze economiche e commerciali, Libreria internazionale Seeber; 1936.
69. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B (Methodological). 1995;57(1):289–300.
70. Revelle W. psych: procedures for psychological, psychometric, and personality research. R package version 1.3. 10. Evanston, IL: Northwestern University; 2013.
71. Team RC. R: A language and environment for statistical computing. 2014th ed. Vienna, Austria: R Foundation for Statistical Computing; 2012. ISBN 3-900051-07-0.

Supplementary material chapter 3

Supplementary Table 1: Characteristics of study participants affected by Rett syndrome.

Supplementary Figure 1: Correlation plots of clinical data from RTT subjects. Significant positive correlations were observed among age, IgA and ESR.

Supplementary Table 2: Spearman's correlation analysis among RTT clinical data.

Supplementary Figure 2: Measures of bacterial diversity. **a)** *Alpha*-diversity estimated on the Chao1 estimator and the Shannon entropy; ***, $p<0.001$; **, $p<0.01$; *, $p<0.05$; Wilcoxon rank-sum test. **b)** PCoA plot based on the unweighted UniFrac distance among samples analysed according to individuals' health status. Constipated Rett syndrome subjects (RTT-C), non-constipated Rett syndrome subjects (RTT-NC) and healthy controls (HC) are coloured in red, orange or green, respectively.

Supplementary Figure 3: *Alpha*-diversity rarefaction curves. The plot shows the *alpha*-diversity for HC, RTT-C and RTT-NC averaged over 100 independent rarefactions as a function of the rarefaction depth for **a)** the bacterial gut microbiota and **b)** the fungal gut microbiota. The points in the curves are the averages, while the whiskers represent the standard deviations.

Supplementary Figure 4: Multiple-rarefaction PCoA plots. Each PCoA replicate was optimally superimposed by Procrustes analysis on the master PCoA scatter plot (used in the main text). Points represent the average location of 100 rarefaction replicates. Ellipses show the 95% confidence region assuming a multivariate normal distribution. **a)** PCoA plots of bacterial *beta*-diversity based on the unweighted and weighted UniFrac distances and the Bray-Curtis dissimilarity analysed according to individuals' health status; **b)** PCoA plots of fungal *beta*-diversity based on the unweighted and weighted UniFrac distances and the Bray-Curtis dissimilarity analysed according to individuals' health status. Constipated Rett syndrome subjects (RTT-C), non-constipated Rett syndrome subjects (RTT-NC) and healthy controls (HC) are coloured in red, orange or green, respectively.

Supplementary Figure 5: PERMANOVA *p-values* distribution. For each rarefaction replicate (n=100), a PERMANOVA test was conducted to assess the robustness of the results over the rarefaction. **a)** PERMANOVA *p-values* distributions of bacterial *beta*-diversity based on the

unweighted and weighted UniFrac distances and the Bray-Curtis dissimilarity analysed according to individuals' health status; **b)** PERMANOVA *p-values* distributions of fungal *beta*-diversity based on the unweighted and weighted UniFrac distances and the Bray-Curtis dissimilarity analysed according to individuals' health status.

Supplementary Table 3: Mean relative abundance (%) ± standard deviation (SD) of bacterial OTUs at phylum and genus levels in Rett syndrome (RTT) subjects and healthy controls (HC).

Supplementary Table 4: Wilcoxon rank-sum test comparison of bacterial relative abundances at phylum and genus levels.

Supplementary Figure 6: Genus level relative abundances of the bacterial gut microbiota of healthy controls (HC) and Rett syndrome (RTT) subjects.

Supplementary Figure 7: Bacterial taxa which relative abundances were significantly different ($p<0.05$; Wilcoxon rank-sum test) between healthy controls (HC) and Rett syndrome (RTT) subjects.

Supplementary Figure 8: LDA scores of the most discriminant bacterial taxa identified by LEfSe. Positive and negative LDA scores indicate the taxa enriched in healthy controls (HC) and Rett syndrome (RTT) subjects, respectively

Supplementary Table 5: Statistics of the significantly different metabolic pathways (KEGG categories) inferred with PICRUSt in the gut microbiota of healthy controls (HC) and Rett syndrome (RTT) subjects (Welch's *t*-test, $p<0.05$ FDR-corrected) from 16S rDNA data.

Supplementary Figure 9: Measures of fungal diversity. **a)** Three estimators of *alpha*-diversity have been calculated: the number of observed OTUs, the Chao1 estimator and the Shannon entropy; **b)** PCoA plot based on the unweighted UniFrac distance among samples analysed according to individuals' health status. Constipated Rett syndrome subjects (RTT-C), non-constipated Rett syndrome subjects (RTT-NC) and healthy controls (HC) are coloured in red, orange or green, respectively.

Supplementary Figure 10: Genus level relative abundances of the fungal gut microbiota of healthy controls (HC) and Rett syndrome (RTT) subjects.

Supplementary Figure 11: *Candida* relative abundance in the gut microbiota of healthy controls (HC) and Rett syndrome (RTT) subjects ($p<0.05$; Wilcoxon rank-sum test).

Supplementary Figure 12: a) LDA scores of the most discriminant fungal taxa identified by LEfSe. Positive and negative LDA scores indicate the taxa enriched in healthy controls (HC) and Rett syndrome (RTT) subjects, respectively; **b)** cladogram showing the most discriminative fungal clades identified by LEfSe. Coloured regions/branches indicate differences in the fungal population structure between Rett syndrome (RTT) subjects and healthy controls (HC). Regions in red indicate clades that were enriched in RTT subjects compared to those in HC, while regions in green indicate clades that were enriched in HC compared to those in RTT subjects.

Supplementary Table 6: Correspondences among deposited metagenomics data and samples, unrarefied OTU tables and taxonomic classifications of the 16S and ITS1 datasets.

- ✓ This table can be found within this paper online at: http://microbiomejournal.biomedcentral.com/articles/10.1186/s40168-016-0185-y

Supplementary methods: MICCA pipelines used for the analysis of the 16S and ITS1 datasets MICCA (v 0.1) (http://www.micca.org)

```
# 16S dataset pipeline
micca-preproc -f TCCTACGGGAGGCAGCAG -r TGTGCGGGCCCCCGTCAATT -O 16 -q 20 -l
400 fastq/*.fastq -o preprocessed

micca-otu-denovo preprocessed/*.fastq -s 0.97 -d -l 200 -c -o otus

micca-phylogeny otus/representatives.fasta -a template -o otus/phylo --
template-min-perc 75 --template-file
~/greengenes_2013_05/rep_set_aligned/97_otus.fasta

micca-midpoint-root otus/phylo/tree.tre otus/phylo/tree_rooted.tre


# ITS1 dataset pipeline
micca-preproc -f GTAAAAGTCGTAACAAGGTTTC -r GTTCAAAGAYTCGATGATTCAC -O 18 -q 15
-l 150 fastq/*.fastq -o preprocessed
```

```
micca-otu-denovo preprocessed/*.fastq -s 0.97 -c -d -l 140 -o otus -t rdp --rdp-gene fungalits_unite --rdp max-memory 2000

micca-phylogeny otus/representatives.fasta -a denovo_tcoffe --tcoffe-num-threads 1 -o otus/phylo

micca-midpoint-root otus/phylo/tree.tre otus/phylo/tree_rooted.tre
```

Supplementary Table 1: Characteristics of study participants affected by Rett syndrome.

Sample ID	Age	MECp2 Nt mutation*	MECp2 aa mutation*	type of MECp2 mutation*	Domain change location*	Clinic phenotype#	CSS	Constipation	ESR (mm/h)	IgA (mg/ml)	CRP (mg/100ml)	Calprotectin (µg/g)
RTT4	25	c.747_751dup5	p.P251fs	Frameshift in/del	TRD	intermediate	7	yes	55	128	<0.35	116.9
RTT9	3	deletion exon 3. 4a. 4b	deletion exon 3. 4a. 4b	Large deletion	Large deletion	intermediate	6	no	NA	52	<0.35	57.3
RTT10	21	c.1163_1197del35	p.P388fs	Frameshift in/del	C-term	intermediate	7	yes	74	372	0.69	171.4
RTT34	9	c.502C>T	p.R168X	Nonsense	Inter-domain region	intermediate	8	yes	49	171	0.92	110.1
RTT35	7	c.1072_1186del113	Large deletion	Large deletion	C-term	intermediate	7	no	13	9.9	0.85	64.3
RTT38	7	deletion exon 4	deletion exon 4	Large deletion	Large deletion	mild	4	yes	21	122	<0.35	21.9
RTT39	17	Large deletion	Large deletion	Large deletion	Large deletion	intermediate	8	yes	9	253	<0.35	296.1
RTT43	13	c.473C>T	p.T158M	Missense	MBD	severe	10	no	39	215	2.39	50.5
RTT51	7	c.397C>T	p.R133C	Missense	MBD	intermediate	7	no	12	171	<0.35	72.1
RTT53	15	del.MECp2	del.MECp2	Large deletion	Large deletion	severe	9	yes	40	381	<0.35	28.8
RTT72	31	c.880C>T	p.R294X	Nonsense	TRD	severe	9	yes	4	22	<0.35	379.8

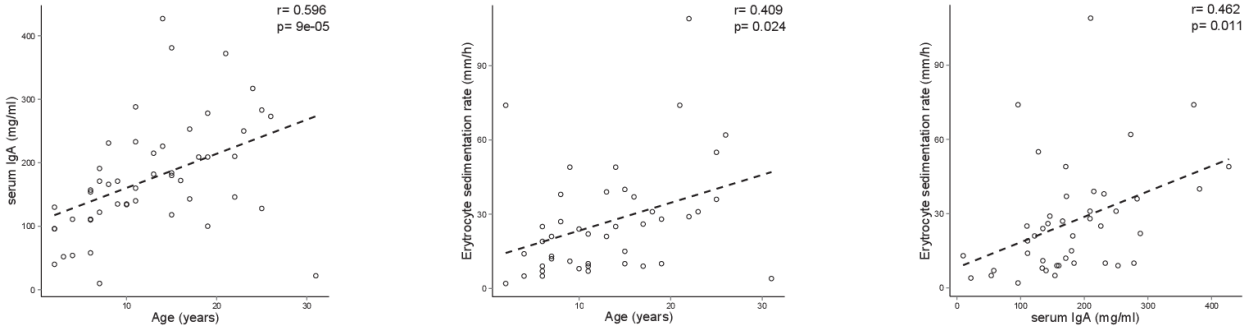
RTT80	4	c.473C>T	p.T158M	Missense	MBD	NA	NA	no	5	54	<0.35	44.9
RTT91	24	c.916C>T	p.R306C	Missense	TRD	severe	11	yes	NA	317	0.55	41.1
RTT97	19	c.547G>C	p.G183R	Missense	Inter-domain region	intermediate	5	yes	10	278	<0.35	92.2
RTT99	19	c.763C>T	p.R255X	Nonsense	TRD	severe	9	yes	28	209	<0.35	91.4
RTT100	6	c.808C>T	p.R270X	Nonsense	TRD-NLS	intermediate	8	no	7	58	<0.35	76.2
RTT101	23	c.455C>G	p.P152R	Missense	MBD	intermediate	8	yes	31	250	6.91	191.2
RTT111	17	c.763C>T	p.R255X	Nonsense	TRD	intermediate	7	yes	26	143	<0.35	28.3
RTT112	14	c.808C>T	p.R270X	Nonsense	TRD-NLS	intermediate	8	yes	25	226	<0.35	73.9
RTT114	15	c.880C>T	p.R294X	Nonsense	TRD	intermediate	6	yes	10	184	<0.35	26.6
RTT116	11	c.808C>T	p.R270X	Nonsense	TRD-NLS	severe	9	yes	10	233	<0.35	45.7
RTT119	2	c.808C>T	p.R270X	Nonsense	TRD-NLS	severe	9	no	74	96	2.54	71.2
RTT123	22	c.808C>T	p.R270X	Nonsense	TRD-NLS	severe	11	yes	29	146	0.44	206.4
RTT129	4	c.502C>T	p.R168X	Nonsense	Inter-domain region	intermediate	6	yes	14	111	<0.35	41.5
RTT132	8	deletion exon 3. 4a. 4b	deletion exon 3. 4a. 4b	Large deletion	Large deletion	intermediate	8	yes	38	231	1	55.5
RTT134	11	deletion exon 1. 2	deletion exon 1. 2	Large deletion	Large deletion	severe	10	yes	22	288	<0.35	32.2
RTT135	18	c.1152_1192del41	p.P385fs	Frameshift in/del	C-term	intermediate	6	no	31	209	2.01	143.4
RTT136	9	c.1157_1200del44	p.L386fs	Frameshift in/del	C-term	intermediate	8	yes	11	135	<0.35	57.5

RTT137	16	c.[1111_1115del; 1116_1137inv]	p.H371LfsX34	Frameshift in/del	C-term	severe	10	no	37	172	<0.35	122.3
RTT138	2	c.473C>T	p.T158M	Missense	MBD	severe	11	yes	2	96	<0.35	41.4
RTT140	11	c.398G>T	p.R133L	Missense	MBD	severe	9	yes	9	160	<0.35	48.2
RTT142	10	CDKL5 mutation	CDKL5 mutation	CDKL5 mutation	NA	intermediate	8	no	24	135	<0.35	500
RTT147	2	c.431delA	p.K144fs	Frameshift in/del	MBD	intermediate	7	no	NA	130	<0.35	56.8
RTT151	6	c.763C>T	p.R255X	Nonsense	TRD	intermediate	6	yes	5	154	0.46	96.2
RTT154	6	c.905C>T	p.P302L	Missense	TRD	severe	11	no	25	110	1	240.9
RTT155	25	CDKL5 mutation	CDKL5 mutation	CDKL5 mutation	NA	severe	9	yes	36	283	<0.35	124.4
RTT158	6	c.1164_1172del14	p.P388fs	Frameshift in/del	C-term	intermediate	6	yes	9	157	<0.35	49.3
RTT159	7	c.502C>T	p.R168X	Nonsense	Inter-domain region	severe	10	yes	NA	191	<0.35	62.6
RTT163	10	c.808C>T	p.R270X	Nonsense	TRD-NLS	intermediate	6	no	8	134	<0.35	NA
RTT164	15	c.808C>T	p.R270X	Nonsense	TRD-NLS	severe	11	yes	15	180	<0.35	NA
RTT165	15	c.1157_1197del41	p.L386fs	Frameshift in/del	C-term	intermediate	7	no	NA	118	NA	31.9
RTT167	8	c.455C>G	p.P152R	Missense	MBD	severe	10	yes	27	166	<0.35	121.1
RTT171	6	c.916C>T	p.R306C	Missense	TRD	intermediate	7	yes	19	111	<0.35	59.3
RTT172	13	c.431delA	p.K144fs	Frameshift in/del	MBD	severe	11	yes	21	182	<0.35	19.7

RTT177	11	c.1162_1179del18	p.P388_P393del	In-frame in/del	C-term	intermediate	5	NA	7	140	<0.35	61.5
RTT190	14	c.1165_1233del69ins21	p.P389_P411del23ins7	In-frame in/del	C-term	severe	11	yes	49	427	0.75	206.6
RTT199	2	c.808C>T	p.R270X	Nonsense	TRD-NLS	intermediate	8	yes	NA	40	NA	213.8
RTT200	26	c.1157_1197del41	p.L386fs	Frameshift in/del	C-term	mild	3	yes	62	273	1.51	204.3
RTT215	19	c.547G>C	p.G183R	Missense	Inter-domain region	intermediate	5	NA	NA	NA	NA	15
RTT250	22	Large deletion	Large deletion	Large deletion	Large deletion	severe	11	yes	109	210	<0.35	29.8

*****According to RettBASE (http://mecp2.chw.edu.au/). CSS, Clinical Severity Score; ESR, Erythrocyte Sedimentation Rate; CRP, C-reactive protein; NA, not applicable. [#]Clinic phenotype attributed according to the CSS (mild, 1-4; intermediate, 5-8; severe, 9-13).

Supplementary Figure 1

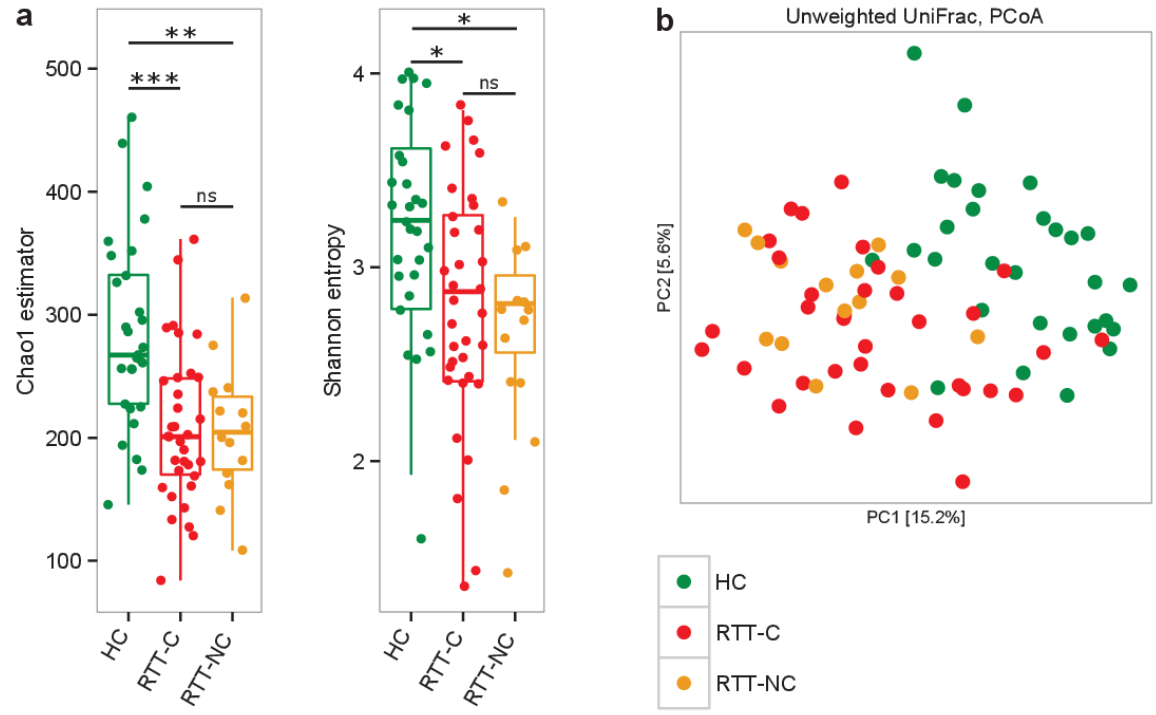


Supplementary Table 2: Spearman's correlation analysis among RTT clinical data.

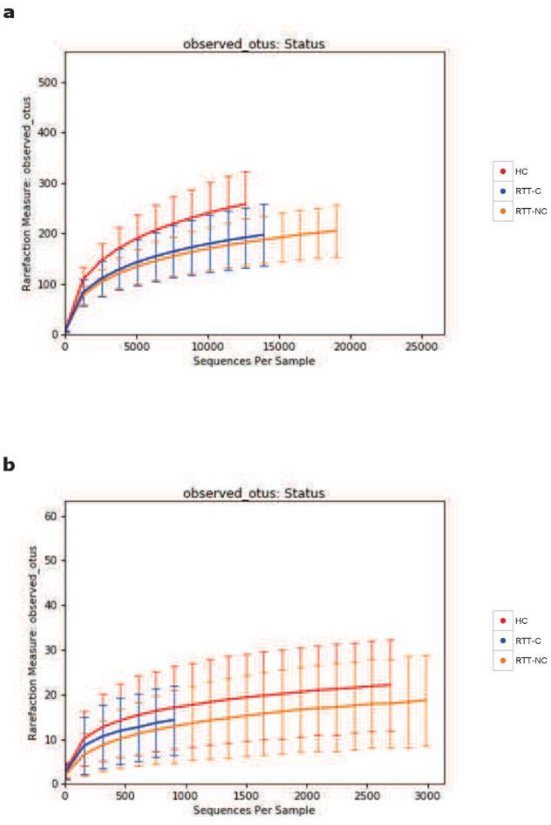
Spearman's r	**Age**	**CSS**	**Constipation**	**IgA**	**ESR**	**Calprotectin**
Age	1.000	0.090	0.350	0.596	0.409	0.177
	CSS	1.000	0.137	0.211	0.234	0.092
		Constipation	1.000	0.430	0.051	-0.089
			IgA	1.000	0.462	0.017
				ESR	1.000	0.205
					Calprotectin	1.000

p-value	**Age**	**CSS**	**Constipation**	**IgA**	**ESR**	**Calprotectin**
Age	0.000	0.642	0.045	0.00009	0.024	0.379
	CSS	0.000	0.539	0.323	0.323	0.642
		Constipation	0.000	0.011	0.800	0.642
			IgA	0.000	0.011	0.907
				ESR	0.000	0.374
					Calprotectin	0.000

Supplementary Figure 2

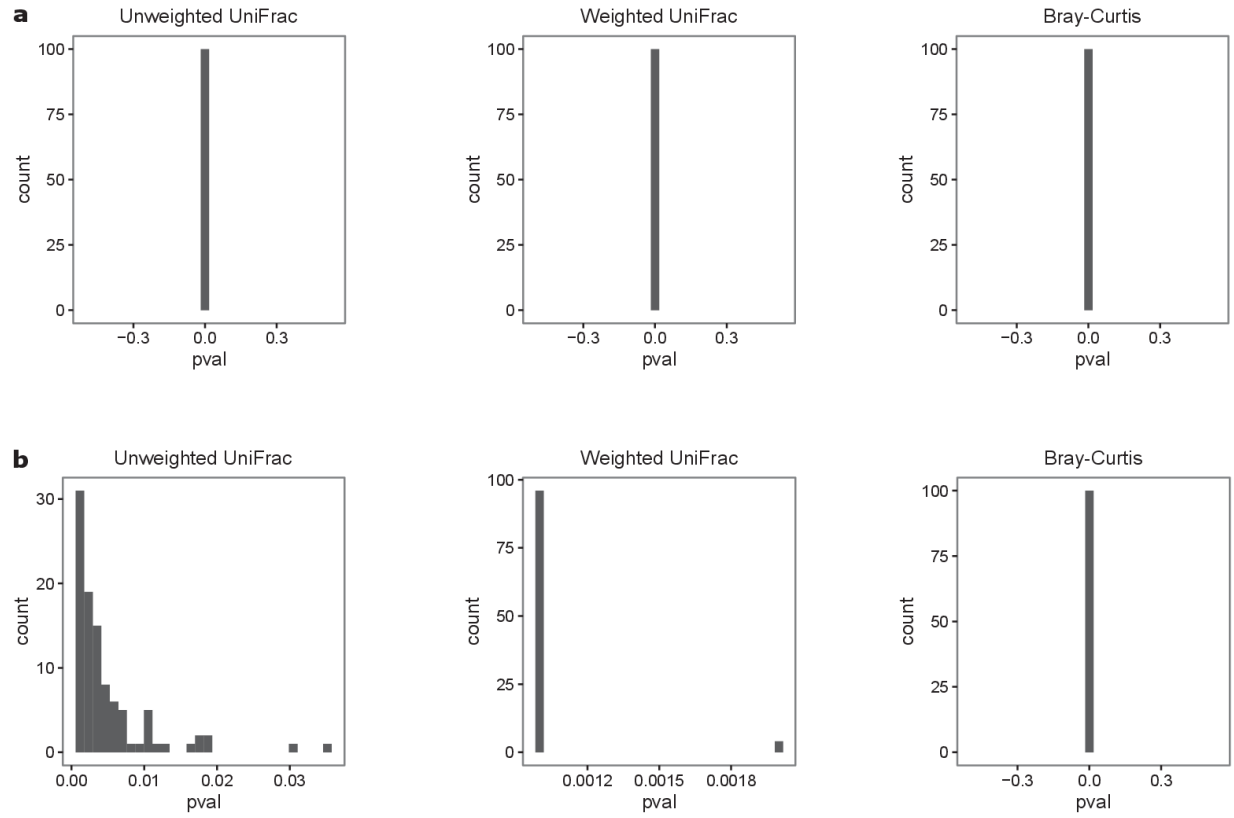


Supplementary Figure 3



Supplementary Figure 4

Supplementary Figure 5



Supplementary Table 3: Mean relative abundance (%) ± standard deviation (SD) of bacterial OTUs at phylum and genus levels in Rett syndrome (RTT) subjects and healthy controls (HC).

	RTT		HC	
Phylum	**mean**	**SD**	**mean**	**SD**
Firmicutes	36.6560	16.3355	43.6145	14.4643
Actinobacteria	40.5576	21.7507	19.5674	17.4821
Bacteroidetes	15.1573	15.9453	30.2396	17.9768
Proteobacteria	7.0502	11.3086	6.2200	9.7279
Unknown	0.4947	0.8526	0.3044	0.2567
Verrucomicrobia	0.0589	0.1550	0.0450	0.0804
Synergistetes	0.0110	0.0591	0.0028	0.0120
Fusobacteria	0.0080	0.0210	0.0003	0.0015
TM7	0.0027	0.0064	0.0019	0.0104
Lentisphaerae	0.0022	0.0158	0.0022	0.0091
Elusimicrobia	0.0000	0.0000	0.0017	0.0089
Cyanobacteria	0.0013	0.0061	0.0003	0.0015

	RTT		HC	
Genus	**mean**	**SD**	**mean**	**SD**
Bifidobacterium	36.7042	21.9180	17.2935	16.4407
Bacteroides	12.3152	13.8314	18.3145	15.2315
Faecalibacterium	3.6678	4.9853	9.2303	9.4653
Lachnospiracea incertae sedis	4.6103	4.9332	3.9995	2.5715
Blautia	4.7655	7.0012	3.7541	5.6668
Escherichia/Shigella	5.2332	11.0378	2.4657	9.0906
Lachnospiraceae;Unknown	2.5140	2.2392	4.3575	2.9834
Alistipes	1.3665	2.0459	4.7119	6.2269
Streptococcus	2.3251	4.1986	2.9697	11.3418
Gemmiger	1.1117	2.0365	3.1772	3.1064
Ruminococcaceae;Unknown	1.2905	1.7899	2.9735	2.8355
Ruminococcus	1.5287	2.5080	2.2076	1.9194
Clostridium XI	1.5268	1.8281	1.2691	1.3289
Collinsella	1.9867	3.3628	0.7221	1.0007
Clostridiales;Unknown	0.6234	1.3096	2.0696	2.5707

Parabacteroides	1.0523	1.9446	1.4590	1.5945
Clostridium XVIII	1.3167	1.8063	1.1703	1.4937
Anaerostipes	1.2669	1.5410	1.2065	1.0610
Prevotella	0.0285	0.0968	2.2137	10.4531
Erysipelotrichaceae incertae sedis	1.3897	2.0083	0.7190	1.6414
Dialister	0.3793	1.3711	1.5186	3.5555
Barnesiella	0.1787	0.4704	1.5752	2.8878
Oscillibacter	0.5675	1.3085	1.1060	1.3587
Enterococcus	1.2751	3.2763	0.2677	0.9199
Coriobacteriaceae;Unknown	0.5619	1.3252	0.9567	1.6284
Lactobacillus	1.2687	3.8816	0.0447	0.0689
Clostridium sensu stricto	0.6903	1.5974	0.6045	0.7939
Prevotellaceae;Unknown	0.0507	0.1724	1.1107	3.8195
Firmicutes;Unknown	0.2289	0.7742	0.7980	1.2740
Eggerthella	0.7638	1.8753	0.2498	0.5586
Clostridium IV	0.4129	0.6337	0.4747	0.5825
Veillonella	0.8229	2.1017	0.0547	0.1136
Unknown	0.4947	0.8526	0.3044	0.2522
Megamonas	0.6157	2.7667	0.0803	0.4250
Phascolarctobacterium	0.1319	0.2905	0.4954	0.8446
Flavonifractor	0.2888	0.3677	0.2713	0.3345
Enterobacteriaceae;Unknown	0.4359	0.8197	0.1041	0.2101
Butyricicoccus	0.2009	0.5159	0.3265	0.4766
Catenibacterium	0.4675	2.6156	0.0276	0.1461
Bifidobacteriaceae;Unknown	0.2497	0.1642	0.1615	0.1800
Dorea	0.1787	0.2538	0.2219	0.2247
Megasphaera	0.3791	1.4683	0.0000	0.0000
Clostridium XlVa	0.2955	0.4094	0.0704	0.0715
Odoribacter	0.1001	0.2170	0.2448	0.3198
Coprococcus	0.0842	0.1999	0.2357	0.2602
Acidaminococcus	0.1823	0.5856	0.1223	0.4516
Turicibacter	0.2044	0.6219	0.0696	0.1336
Clostridium XlVb	0.1561	0.2811	0.1063	0.1835
Bacteroidetes;Unknown	0.0050	0.0194	0.2186	0.7603

Olsenella	0.1324	0.6047	0.0618	0.2847
Eubacterium	0.0751	0.1190	0.1143	0.1453
Bilophila	0.0355	0.0899	0.1460	0.1932
Paraprevotella	0.0125	0.0557	0.1568	0.3086
Roseburia	0.0403	0.0522	0.1231	0.1111
Parasutterella	0.0953	0.3672	0.0588	0.1199
Butyricimonas	0.0331	0.0896	0.1140	0.1654
Eubacteriaceae;Unknown	0.1372	0.4310	0.0019	0.0078
Desulfovibrio	0.0482	0.1399	0.0789	0.1998
Lactobacillales;Unknown	0.0874	0.1231	0.0334	0.0749
Lactococcus	0.1087	0.4339	0.0097	0.0279
Lactonifactor	0.0413	0.0465	0.0734	0.0600
Porphyromonadaceae;Unknown	0.0040	0.0211	0.1021	0.3562
Akkermansia	0.0584	0.1546	0.0450	0.0790
Sarcina	0.0018	0.0075	0.0961	0.3427
Granulicatella	0.0652	0.1346	0.0315	0.1098
Peptostreptococcaceae;Unknown	0.0551	0.0655	0.0386	0.0532
Coprobacillus	0.0453	0.1166	0.0362	0.0851
Actinomyces	0.0631	0.1206	0.0155	0.0213
Anaerotruncus	0.0519	0.1073	0.0248	0.0678
Clostridia;Unknown	0.0115	0.0386	0.0638	0.2839
Erysipelotrichaceae;Unknown	0.0499	0.1590	0.0229	0.0395
Haemophilus	0.0207	0.0938	0.0491	0.1256
Gordonibacter	0.0303	0.0561	0.0337	0.1103
Slackia	0.0171	0.0546	0.0425	0.1459
Clostridiaceae 1;Unknown	0.0157	0.0334	0.0389	0.0719
Mitsuokella	0.0311	0.2185	0.0213	0.1125
Sutterella	0.0154	0.0480	0.0356	0.1047
Holdemania	0.0239	0.0409	0.0232	0.0282
Alphaproteobacteria;Unknown	0.0019	0.0082	0.0439	0.1366
Proteobacteria;Unknown	0.0232	0.1641	0.0039	0.0165
Veillonellaceae;Unknown	0.0261	0.0729	0.0000	0.0000
Enterorhabdus	0.0080	0.0371	0.0177	0.0822
Burkholderiales;Unknown	0.0005	0.0034	0.0248	0.1123

Allisonella	0.0216	0.0912	0.0006	0.0029
Peptoniphilus	0.0118	0.0479	0.0066	0.0236
Varibaculum	0.0155	0.0898	0.0022	0.0103
Peptostreptococcus	0.0163	0.0536	0.0008	0.0024
Desulfovibrionaceae;Unknown	0.0002	0.0011	0.0132	0.0470
Pyramidobacter	0.0106	0.0590	0.0025	0.0117
Sutterellaceae;Unknown	0.0062	0.0305	0.0063	0.0270
Anaerofustis	0.0072	0.0131	0.0044	0.0112
Porphyromonas	0.0061	0.0353	0.0052	0.0187
Anaerococcus	0.0088	0.0231	0.0019	0.0058
Pediococcus	0.0093	0.0432	0.0011	0.0046
Morganella	0.0077	0.0236	0.0011	0.0058
Corynebacterium	0.0056	0.0100	0.0030	0.0118
Fusobacterium	0.0080	0.0210	0.0003	0.0015
Bacteroidales;Unknown	0.0011	0.0040	0.0069	0.0160
Clostridiales Incertae Sedis XI;Unknown	0.0027	0.0181	0.0052	0.0263
Parvimonas	0.0074	0.0259	0.0003	0.0015
Finegoldia	0.0056	0.0173	0.0019	0.0062
Weissella	0.0067	0.0453	0.0003	0.0015
Gemella	0.0050	0.0110	0.0014	0.0042
Actinomycetales;Unknown	0.0051	0.0135	0.0008	0.0032
Staphylococcus	0.0059	0.0198	0.0000	0.0000
Proteus	0.0058	0.0206	0.0000	0.0000
Oxalobacter	0.0011	0.0061	0.0044	0.0134
Anaerofilum	0.0029	0.0107	0.0025	0.0060
Pasteurellaceae;Unknown	0.0030	0.0136	0.0022	0.0091
TM7 genera incertae sedis	0.0027	0.0064	0.0019	0.0102
Peptococcus	0.0043	0.0153	0.0003	0.0015
Victivallis	0.0022	0.0158	0.0022	0.0089
Actinobacteria;Unknown	0.0011	0.0051	0.0033	0.0080
Howardella	0.0043	0.0306	0.0000	0.0000
Anaerovorax	0.0003	0.0023	0.0039	0.0148
Rikenella	0.0011	0.0079	0.0030	0.0161
Sporobacter	0.0008	0.0037	0.0030	0.0057

Anaeroglobus	0.0037	0.0186	0.0000	0.0000
Atopobium	0.0034	0.0111	0.0003	0.0015
Kocuria	0.0030	0.0113	0.0000	0.0000
Hydrogenoanaerobacterium	0.0000	0.0000	0.0028	0.0146
Dysgonomonas	0.0021	0.0136	0.0006	0.0029
Xylanibacter	0.0003	0.0023	0.0022	0.0078
Actinomycetaceae;Unknown	0.0021	0.0082	0.0003	0.0015
Propionibacterium	0.0018	0.0073	0.0006	0.0029
Solobacterium	0.0021	0.0094	0.0000	0.0000
Mogibacterium	0.0005	0.0034	0.0014	0.0052
Deltaproteobacteria;Unknown	0.0002	0.0011	0.0017	0.0049
Elusimicrobium	0.0000	0.0000	0.0017	0.0088
Succiniclasticum	0.0016	0.0113	0.0000	0.0000
Streptophyta	0.0013	0.0061	0.0003	0.0015
Desulfovibrionales;Unknown	0.0006	0.0027	0.0008	0.0044
Comamonas	0.0011	0.0079	0.0003	0.0015
Acinetobacter	0.0008	0.0046	0.0006	0.0020
Mobiluncus	0.0002	0.0011	0.0008	0.0032
Cardiobacterium	0.0000	0.0000	0.0008	0.0044
Scardovia	0.0008	0.0037	0.0000	0.0000
Pseudomonas	0.0008	0.0040	0.0000	0.0000
Cloacibacillus	0.0005	0.0034	0.0003	0.0015
Pseudoramibacter	0.0006	0.0036	0.0000	0.0000
Clostridiales Incertae Sedis XIII;Unknown	0.0006	0.0045	0.0000	0.0000
Rothia	0.0003	0.0023	0.0003	0.0015
Gordonia	0.0000	0.0000	0.0006	0.0029
Paraeggerthella	0.0000	0.0000	0.0006	0.0029
Actinobaculum	0.0005	0.0025	0.0000	0.0000
Trueperella	0.0005	0.0034	0.0000	0.0000
Abiotrophia	0.0005	0.0034	0.0000	0.0000
Murdochiella	0.0005	0.0025	0.0000	0.0000
Schwartzia	0.0005	0.0034	0.0000	0.0000
Rhizobacter	0.0005	0.0034	0.0000	0.0000
Puniceicoccaceae;Unknown	0.0005	0.0034	0.0000	0.0000

Leclercia	0.0002	0.0011	0.0003	0.0015
Arcanobacterium	0.0003	0.0023	0.0000	0.0000
Devosia	0.0003	0.0023	0.0000	0.0000
Enhydrobacter	0.0003	0.0023	0.0000	0.0000
Rikenellaceae;Unknown	0.0000	0.0000	0.0003	0.0015
Pseudoflavonifractor	0.0000	0.0000	0.0003	0.0015
Mesorhizobium	0.0000	0.0000	0.0003	0.0015
Alloscardovia	0.0002	0.0011	0.0000	0.0000
Bacillus	0.0002	0.0011	0.0000	0.0000
Facklamia	0.0002	0.0011	0.0000	0.0000
Tetragenococcus	0.0002	0.0011	0.0000	0.0000
Lactobacillaceae;Unknown	0.0002	0.0011	0.0000	0.0000
Selenomonas	0.0002	0.0011	0.0000	0.0000

Supplementary Table 4: Wilcoxon rank-sum test comparison of bacterial relative abundances at phylum and genus levels.

Phylum	***p-value***	**FDR**
Actinobacteria	0.00005	0.00170
Bacteroidetes	0.00008	0.00213
Fusobacteria	0.02407	0.08647
Firmicutes	0.07023	0.20600
TM7	0.07675	0.21441
Lentisphaerae	0.11732	0.27065
Elusimicrobia	0.19816	0.36712
Cyanobacteria/Chloroplast	0.61100	0.73654
Unknown	0.72561	0.78848
Verrucomicrobia	0.82300	0.87787
Proteobacteria	1.00000	1.00000
Synergistetes	1.00000	1.00000

Genus	***p-value***	**FDR**
Clostridiales;Unknown	0.00003	0.00152
Firmicutes;Unknown	0.00002	0.00152
Gemmiger	0.00001	0.00152
Barnesiella	0.00004	0.00170
Coprococcus	0.00008	0.00213
Bifidobacterium	0.00011	0.00222
Roseburia	0.00011	0.00222
Eggerthella	0.00016	0.00286
Ruminococcaceae;Unknown	0.00027	0.00434
Escherichia/Shigella	0.00033	0.00446
Veillonellaceae;Unknown	0.00031	0.00446
Enterococcus	0.00044	0.00518
Porphyromonadaceae;Unknown	0.00042	0.00518
Lactobacillales;Unknown	0.00048	0.00533
Erysipelotrichaceae incertae sedis	0.00082	0.00847
Alistipes	0.00104	0.01014

Odoribacter	0.00116	0.01076
Lachnospiraceae;Unknown	0.00129	0.01131
Butyricimonas	0.00137	0.01151
Bilophila	0.00156	0.01251
Enterobacteriaceae;Unknown	0.00169	0.01280
Haemophilus	0.00175	0.01280
Megasphaera	0.00197	0.01387
Ruminococcus	0.00273	0.01846
Clostridium XlVa	0.00309	0.02012
Lactonifactor	0.00322	0.02025
Oscillibacter	0.00382	0.02317
Eubacteriaceae;Unknown	0.00429	0.02515
Dialister	0.00451	0.02562
Coriobacteriaceae;Unknown	0.00547	0.03006
Faecalibacterium	0.00592	0.03159
Actinomyces	0.00658	0.03311
Bifidobacteriaceae;Unknown	0.00659	0.03311
Butyricicoccus	0.00700	0.03423
Sporobacter	0.00894	0.04251
Desulfovibrionaceae;Unknown	0.01314	0.05928
Parabacteroides	0.01291	0.05928
Granulicatella	0.01377	0.06060
Alphaproteobacteria;Unknown	0.01456	0.06102
Burkholderiales;Unknown	0.01441	0.06102
Paraprevotella	0.01553	0.06358
Bacteroidales;Unknown	0.01689	0.06756
Bacteroides	0.01804	0.07056
Sarcina	0.02012	0.07696
Lactobacillus	0.02140	0.08013
Fusobacterium	0.02407	0.08647
Parvimonas	0.02457	0.08648
Prevotella	0.02635	0.09094
Corynebacterium	0.03468	0.11737
Deltaproteobacteria;Unknown	0.03929	0.13047

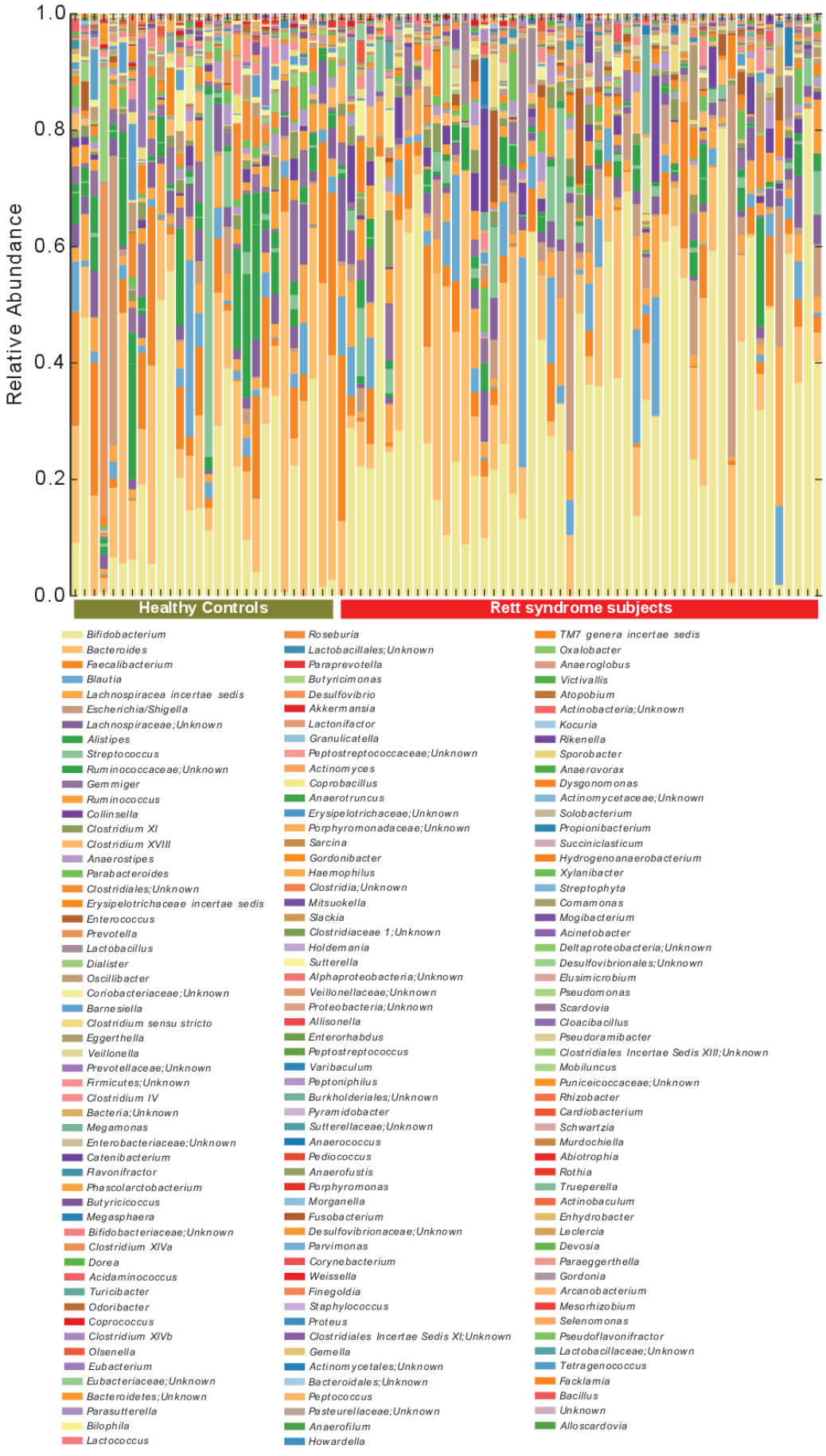
Phascolarctobacterium	0.04454	0.14517
Anaerotruncus	0.05115	0.16368
Kocuria	0.05544	0.16824
Proteus	0.05543	0.16824
Staphylococcus	0.05544	0.16824
Allisonella	0.06077	0.18127
TM7 genera incertae sedis	0.07675	0.21441
Veillonella	0.07562	0.21441
Solobacterium	0.08268	0.22737
Peptostreptococcus	0.09151	0.24777
Bacteroidetes;Unknown	0.09792	0.25344
Dorea	0.09768	0.25344
Morganella	0.09518	0.25344
Anaerococcus	0.09997	0.25499
Actinobacteria;Unknown	0.11330	0.27065
Clostridium IV	0.11594	0.27065
Eubacterium	0.11708	0.27065
Gemella	0.11841	0.27065
Oxalobacter	0.11592	0.27065
Victivallis	0.11732	0.27065
Xylanibacter	0.10821	0.27065
Lactococcus	0.12012	0.27103
Anaeroglobus	0.12372	0.27563
Streptococcus	0.14862	0.32697
Actinomycetales;Unknown	0.15086	0.32779
Clostridiaceae 1;Unknown	0.16208	0.34631
Sutterella	0.16331	0.34631
Atopobium	0.19007	0.36712
Cardiobacterium	0.19816	0.36712
Elusimicrobium	0.19816	0.36712
Gordonia	0.19816	0.36712
Hydrogenoanaerobacterium	0.19816	0.36712
Mesorhizobium	0.19816	0.36712
Paraeggerthella	0.19816	0.36712

Parasutterella	0.18578	0.36712
Pseudoflavonifractor	0.19816	0.36712
Rikenellaceae;Unknown	0.19816	0.36712
Scardovia	0.18658	0.36712
Clostridium XlVb	0.21116	0.38713
Gordonibacter	0.24322	0.44130
Actinobaculum	0.28677	0.48033
Anaerovorax	0.26886	0.48033
Mobiluncus	0.27554	0.48033
Mogibacterium	0.28929	0.48033
Murdochiella	0.28677	0.48033
Peptococcus	0.27305	0.48033
Peptostreptococcaceae;Unknown	0.27728	0.48033
Pseudomonas	0.28677	0.48033
Pseudoramibacter	0.28677	0.48033
Megamonas	0.29292	0.48181
Proteobacteria;Unknown	0.29637	0.48297
Acidaminococcus	0.30026	0.48482
Olsenella	0.30607	0.48971
Varibaculum	0.33763	0.53535
Anaerofustis	0.35423	0.55665
Clostridium sensu stricto	0.36173	0.56340
Finegoldia	0.37028	0.57166
Erysipelotrichaceae;Unknown	0.37391	0.57224
Holdemania	0.38664	0.58662
Abiotrophia	0.46215	0.59371
Alloscardovia	0.46215	0.59371
Anaerofilum	0.40288	0.59371
Anaerostipes	0.40712	0.59371
Arcanobacterium	0.46215	0.59371
Bacillus	0.46215	0.59371
Clostridiales Incertae Sedis XIII;Unknown	0.46215	0.59371
Devosia	0.46215	0.59371
Enhydrobacter	0.46215	0.59371

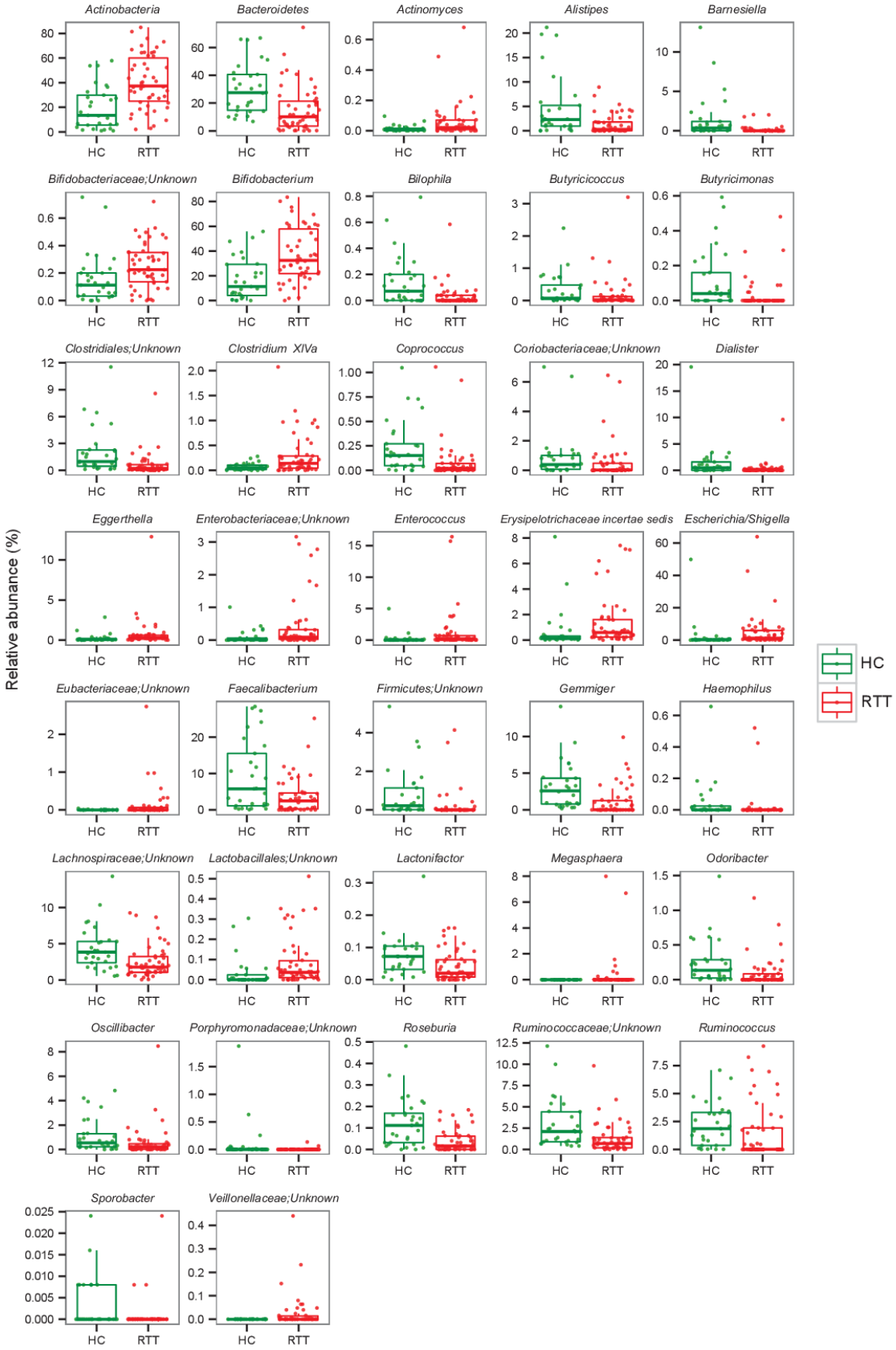
Facklamia	0.46215	0.59371
Howardella	0.46215	0.59371
Lactobacillaceae;Unknown	0.46215	0.59371
Propionibacterium	0.43349	0.59371
Puniceicoccaceae;Unknown	0.46215	0.59371
Rhizobacter	0.46215	0.59371
Schwartzia	0.46215	0.59371
Selenomonas	0.46215	0.59371
Succiniclasticum	0.46215	0.59371
Tetragenococcus	0.46215	0.59371
Trueperella	0.46215	0.59371
Unknown	0.41255	0.59371
Prevotellaceae;Unknown	0.48211	0.61486
Turicibacter	0.48796	0.61785
Collinsella	0.52711	0.66265
Coprobacillus	0.53092	0.66271
Unknown	0.57398	0.71142
Flavonifractor	0.58619	0.72147
Acinetobacter	0.60160	0.73529
Streptophyta	0.61100	0.73654
Catenibacterium	0.62043	0.74282
Blautia	0.64352	0.75506
Pediococcus	0.63506	0.75506
Pyramidobacter	0.64239	0.75506
Desulfovibrionales;Unknown	0.65869	0.76774
Slackia	0.67791	0.78495
Cloacibacillus	0.72253	0.78848
Clostridia;Unknown	0.69455	0.78848
Comamonas	0.72253	0.78848
Lachnospiracea incertae sedis	0.72185	0.78848
Leclercia	0.70856	0.78848
Porphyromonas	0.72576	0.78848
Rikenella	0.69472	0.78848
Rothia	0.72253	0.78848

Unknown	0.72561	0.78848
Desulfovibrio	0.74359	0.80289
Akkermansia	0.80929	0.86850
Enterorhabdus	0.88503	0.93246
Sutterellaceae;Unknown	0.88503	0.93246
Weissella	0.89007	0.93246
Clostridium XVIII	0.90285	0.93566
Dysgonomonas	0.91439	0.93566
Mitsuokella	0.91439	0.93566
Pasteurellaceae;Unknown	0.91364	0.93566
Peptoniphilus	0.92302	0.93902
Clostridium XI	0.96755	0.97867

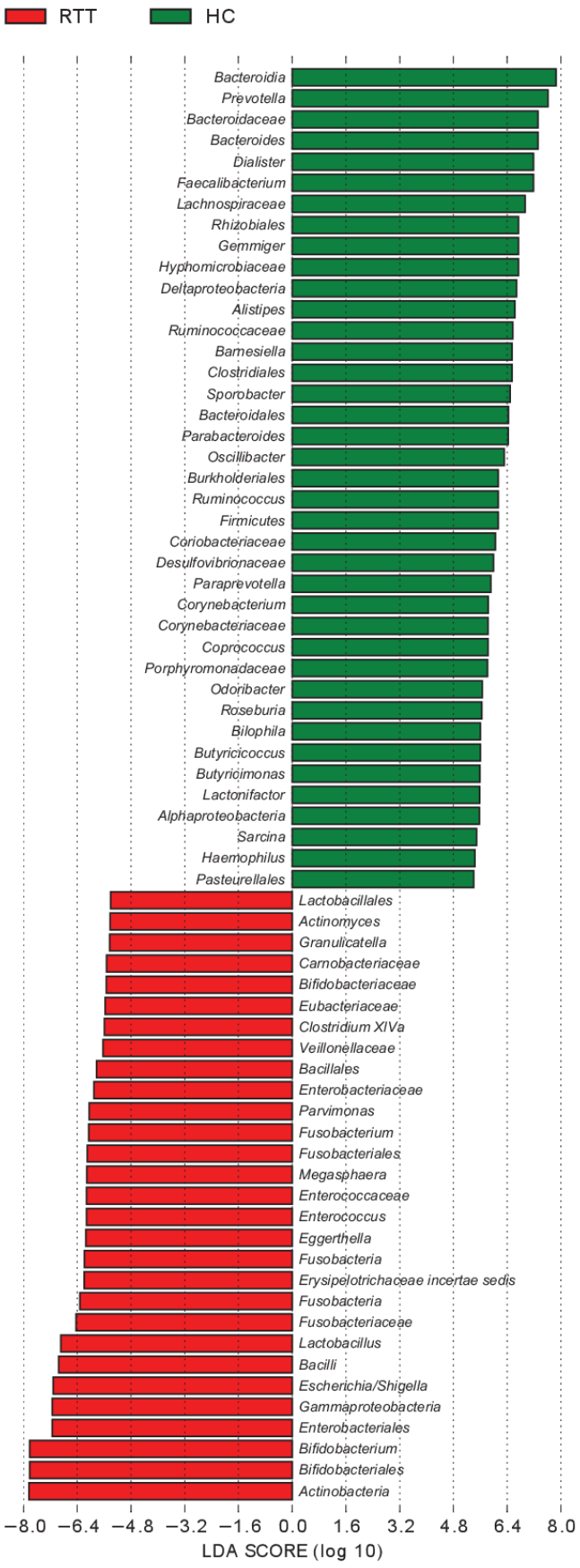
Supplementary Figure 6



Supplementary Figure 7



Supplementary Figure 8



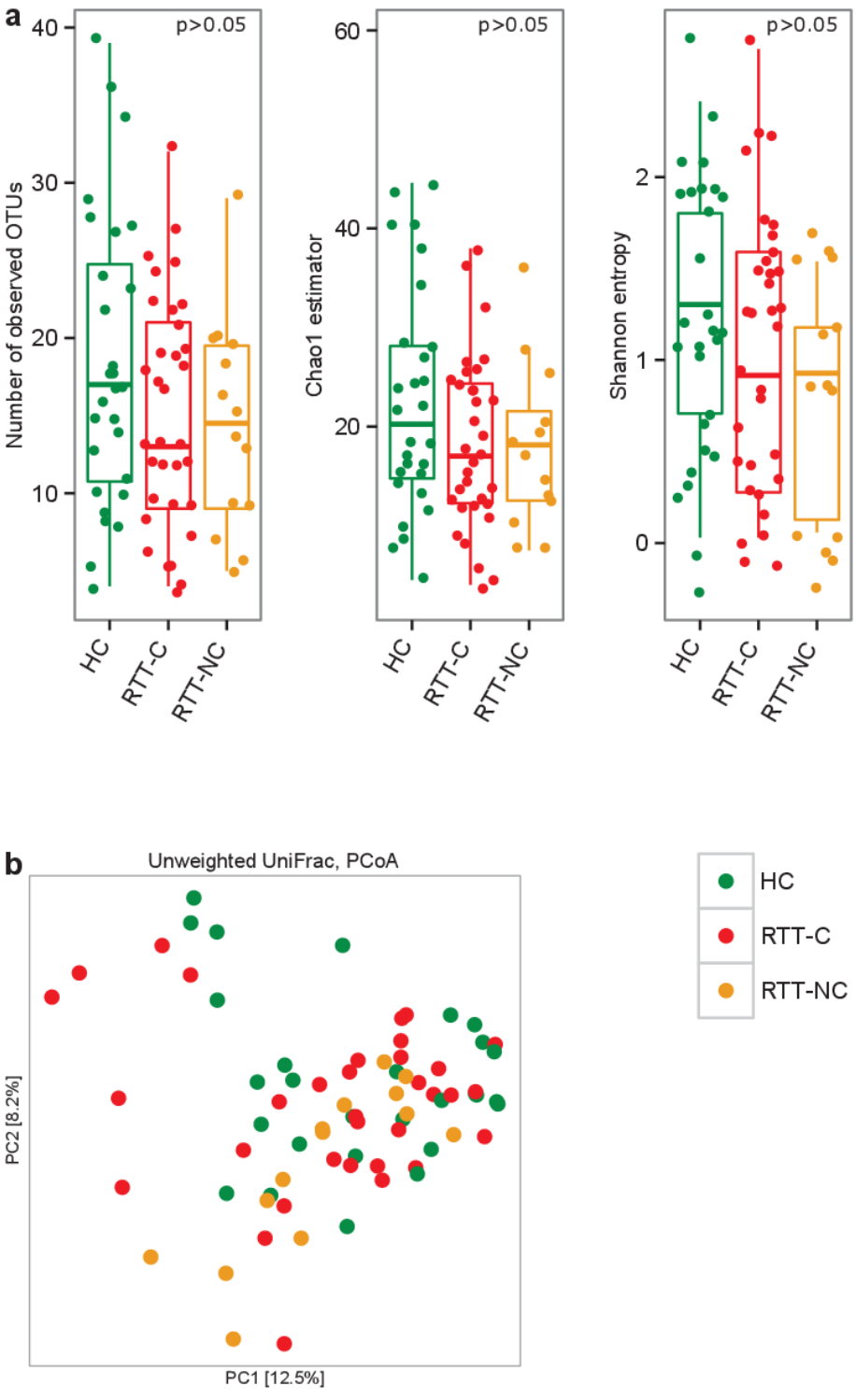
Supplementary Table 5: Statistics of the significantly different metabolic pathways (KEGG categories) inferred with PICRUSt in the gut microbiota of healthy controls (HC) and Rett syndrome (RTT) subjects (Welch's *t*-test, p<0.05 FDR-corrected) from 16S rDNA data.

KEGG categories	Healthy controls		RTT patients		p-values	p-adj	95.0% lower CI	95.0% upper CI
	mean rel. freq. (%)	std. dev. (%)	mean rel. freq. (%)	std. dev. (%)				
Phosphotransferase system (PTS)	0.371433	0.123433	0.598653	0.258281	0.000002	0.000536	-0.314180	-0.140261
Transporters	6.808851	0.965836	7.889148	0.943105	0.000014	0.000887	-1.534563	-0.626031
Chromosome	1.575911	0.091076	1.479764	0.062469	0.000011	0.000936	0.057055	0.135238
Restriction enzyme	0.216111	0.043790	0.166141	0.042151	0.000009	0.000982	0.029472	0.070469
Chaperones and folding catalysts	1.008852	0.092890	0.911429	0.077359	0.000021	0.001145	0.055760	0.139086
Transcription machinery	0.976725	0.100412	0.867090	0.102835	0.000025	0.001186	0.061622	0.157648
Lipid biosynthesis proteins	0.579641	0.036582	0.534886	0.042453	0.000007	0.001227	0.026391	0.063119
Ascorbate and aldarate metabolism	0.098602	0.022726	0.128234	0.036042	0.000032	0.001324	-0.042985	-0.016279
Plant-pathogen interaction	0.140824	0.018886	0.120213	0.021061	0.000040	0.001456	0.011284	0.029937
General function prediction only	3.608275	0.137215	3.458786	0.155383	0.000045	0.001468	0.081304	0.217673
Xylene degradation	0.058345	0.019100	0.076697	0.017200	0.000101	0.001944	-0.027109	-0.009596
ABC transporters	3.178954	0.422589	3.605578	0.441775	0.000092	0.002010	-0.630201	-0.223046
Protein processing in endoplasmic reticulum	0.064067	0.015210	0.048903	0.015541	0.000099	0.002028	0.007897	0.022431
Chagas disease (American trypanosomiasis)	0.008754	0.008443	0.018124	0.011049	0.000084	0.002111	-0.013845	-0.004894
Protein export	0.603850	0.032683	0.571587	0.031658	0.000091	0.002125	0.016933	0.047595
Transcription related proteins	0.003437	0.002811	0.007184	0.005401	0.000145	0.002163	-0.005612	-0.001881
Epithelial cell signaling in Helicobacter pylori infection	0.094189	0.011868	0.082544	0.012551	0.000144	0.002253	0.005904	0.017387
Chlorocyclohexane and chlorobenzene degradation	0.011154	0.005722	0.017594	0.007702	0.000083	0.002257	-0.009515	-0.003365
Energy metabolism	0.869032	0.121532	0.750978	0.127371	0.000159	0.002265	0.059453	0.176656
Drug metabolism - other enzymes	0.335110	0.037774	0.302137	0.020652	0.000127	0.002322	0.017333	0.048613
Prostate cancer	0.040448	0.009196	0.031359	0.009916	0.000143	0.002348	0.004607	0.013571
Dioxin degradation	0.060495	0.020803	0.079556	0.016884	0.000139	0.002406	-0.028330	-0.009793
African trypanosomiasis	0.009944	0.008824	0.019585	0.011132	0.000083	0.002462	-0.014238	-0.005045

Chloroalkane and chloroalkene degradation	0.197584	0.035617	0.230084	0.031046	0.000182	0.002483	-0.048675	-0.016324
Glycerophospholipid metabolism	0.556524	0.045365	0.514331	0.043530	0.000197	0.002484	0.020979	0.063406
Riboflavin metabolism	0.211331	0.039616	0.176516	0.029317	0.000191	0.002507	0.017544	0.052085
Ion channels	0.017188	0.010138	0.026582	0.010465	0.000276	0.003348	-0.014256	-0.004533
Drug metabolism - cytochrome P450	0.032544	0.015327	0.047231	0.017704	0.000300	0.003519	-0.022366	-0.007008
Metabolism of xenobiotics by cytochrome P450	0.032463	0.015297	0.047071	0.017762	0.000322	0.003640	-0.022289	-0.006927
Cytoskeleton proteins	0.392817	0.047263	0.350158	0.048095	0.000361	0.003816	0.020111	0.065207
Ubiquitin system	0.013709	0.007738	0.021551	0.010444	0.000350	0.003828	-0.012007	-0.003677
Progesterone-mediated oocyte maturation	0.036074	0.007734	0.028931	0.008823	0.000446	0.004307	0.003287	0.010997
Synthesis and degradation of ketone bodies	0.021521	0.006037	0.016095	0.006313	0.000423	0.004333	0.002518	0.008335
Prion diseases	0.003041	0.001880	0.005617	0.004260	0.000479	0.004365	-0.003981	-0.001172
Antigen processing and presentation	0.036074	0.007734	0.028931	0.008823	0.000446	0.004437	0.003287	0.010997
NOD-like receptor signaling pathway	0.038380	0.009217	0.030182	0.009629	0.000478	0.004476	0.003759	0.012637
Butirosin and neomycin biosynthesis	0.076643	0.012314	0.065715	0.013611	0.000619	0.005491	0.004868	0.016987
One carbon pool by folate	0.639248	0.058686	0.593137	0.042400	0.000683	0.005897	0.020657	0.071563
Flagellar assembly	0.207395	0.095427	0.126996	0.094264	0.000723	0.006078	0.035341	0.125457
Carbon fixation in photosynthetic organisms	0.673432	0.036854	0.643658	0.032598	0.000811	0.006654	0.012967	0.046582
Function unknown	1.221045	0.106717	1.314148	0.124039	0.000931	0.007448	-0.146709	-0.039496
Proximal tubule bicarbonate reclamation	0.020030	0.008701	0.012757	0.009208	0.001010	0.007886	0.003063	0.011484
Others	0.927654	0.072299	0.983807	0.060567	0.001070	0.008159	-0.088631	-0.023676
Polyketide sugar unit biosynthesis	0.237731	0.030644	0.213450	0.030187	0.001377	0.010264	0.009825	0.038737
Carbohydrate metabolism	0.195214	0.030237	0.220358	0.034839	0.001486	0.010832	-0.040277	-0.010010
Histidine metabolism	0.690466	0.074480	0.637199	0.052650	0.001721	0.012269	0.021108	0.085426
Inositol phosphate metabolism	0.101559	0.012481	0.112350	0.016401	0.001789	0.012485	-0.017420	-0.004162
Fatty acid metabolism	0.258928	0.041966	0.294572	0.053866	0.001912	0.012801	-0.057683	-0.013604
Proteasome	0.052386	0.012189	0.062559	0.014996	0.001883	0.012868	-0.016450	-0.003896
Ribosome biogenesis in eukaryotes	0.055503	0.013070	0.066227	0.015763	0.002045	0.013417	-0.017394	-0.004053
Translation factors	0.552452	0.040026	0.524852	0.028837	0.002494	0.016043	0.010250	0.044950

Base excision repair	0.440707	0.021299	0.424864	0.021213	0.002622	0.016537	0.005758	0.025928
Tyrosine metabolism	0.360993	0.032997	0.388238	0.043420	0.002793	0.016967	-0.044782	-0.009707
Polycyclic aromatic hydrocarbon degradation	0.117622	0.012024	0.108404	0.013412	0.002870	0.017114	0.003280	0.015157
Bacterial chemotaxis	0.288587	0.078897	0.232275	0.071952	0.002990	0.017207	0.020004	0.092621
Primary immunodeficiency	0.048920	0.007695	0.055219	0.010062	0.002940	0.017222	-0.010377	-0.002222
Transcription factors	1.690215	0.177764	1.823132	0.183453	0.002786	0.017244	-0.218151	-0.047684
Shigellosis	0.000000	0.000000	0.000004	0.000010	0.003219	0.018207	-0.000007	-0.000002
Pathways in cancer	0.039025	0.008329	0.032571	0.010061	0.003486	0.019379	0.002200	0.010708
Carbon fixation pathways in prokaryotes	0.949869	0.081717	0.892225	0.077806	0.003703	0.020242	0.019527	0.095761
Lipid metabolism	0.123550	0.023902	0.104193	0.033699	0.004484	0.024113	0.006202	0.032512
Bacterial motility proteins	0.556595	0.189692	0.425357	0.184722	0.004626	0.024474	0.042100	0.220374
Ribosome Biogenesis	1.429153	0.078921	1.378516	0.061728	0.005296	0.027571	0.015793	0.085481
Selenocompound metabolism	0.387517	0.028658	0.409540	0.038965	0.005928	0.030379	-0.037503	-0.006542
Glutathione metabolism	0.194576	0.037021	0.220928	0.042683	0.006022	0.030386	-0.044885	-0.007818
Membrane and intracellular structural molecules	0.489043	0.137976	0.392971	0.152215	0.006243	0.031026	0.028225	0.163918
Other transporters	0.256346	0.029236	0.232426	0.045811	0.006605	0.031862	0.006862	0.040978
Zeatin biosynthesis	0.056130	0.008212	0.051131	0.005685	0.006582	0.032220	0.001468	0.008530
Ethylbenzene degradation	0.040695	0.014248	0.031044	0.015234	0.007055	0.033537	0.002729	0.016573
Propanoate metabolism	0.461379	0.036262	0.485831	0.040058	0.008005	0.035966	-0.042292	-0.006612
Styrene degradation	0.011358	0.005122	0.015654	0.008668	0.007786	0.035968	-0.007425	-0.001166
Arginine and proline metabolism	1.282244	0.086147	1.228878	0.071400	0.007691	0.036038	0.014776	0.091956
Retinol metabolism	0.035106	0.012620	0.044640	0.017861	0.007927	0.036114	-0.016493	-0.002574
Glycosphingolipid biosynthesis - ganglio series	0.069709	0.042315	0.042320	0.042644	0.008474	0.037561	0.007270	0.047508
Insulin signaling pathway	0.089897	0.011488	0.082731	0.011126	0.010051	0.042816	0.001778	0.012555
Tuberculosis	0.155036	0.012125	0.147335	0.012348	0.009963	0.043000	0.001914	0.013487
Basal transcription factors	0.002927	0.003096	0.001199	0.001756	0.009916	0.043365	0.000439	0.003016
Glutamatergic synapse	0.116836	0.011225	0.109912	0.010650	0.010393	0.043703	0.001694	0.012154
Biosynthesis of ansamycins	0.112961	0.011872	0.104325	0.017073	0.011067	0.045948	0.002033	0.015239

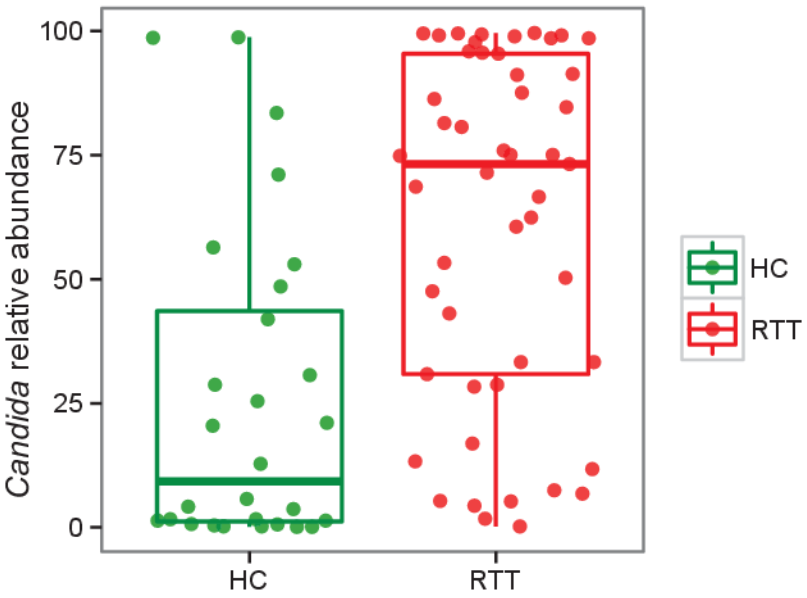
Supplementary Figure 9



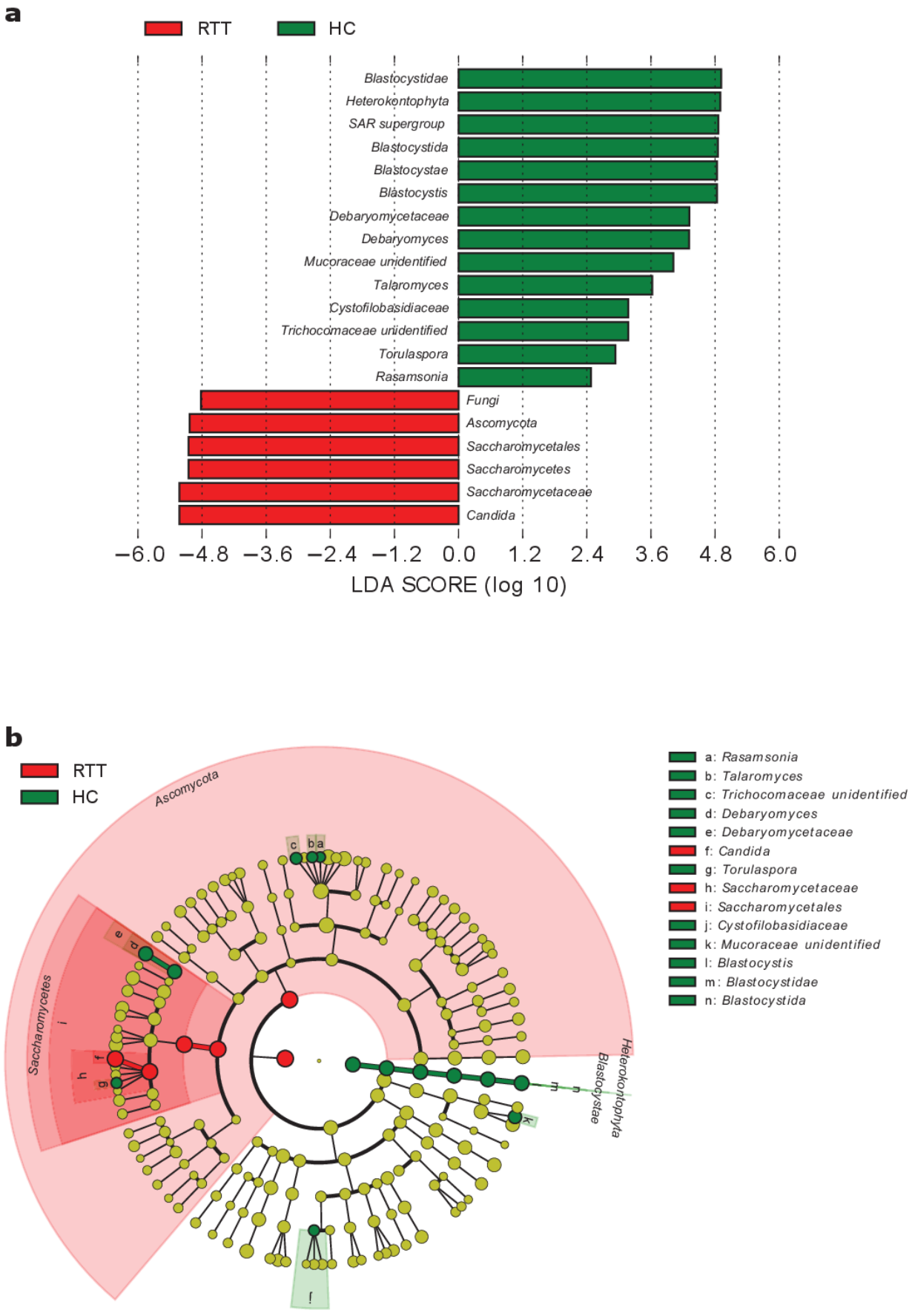
Supplementary Figure 10

Candida
Penicillium
Aspergillus
Blastocystis
Malassezia
Ascomycota unidentified
Debaryomyces
Mucor
Fungi unidentified
Eremothecium
Pichia
Cyberlindnera
Trichosporon
Mucoraceae unidentified
Wallemia
Eurotiomycetes;Unknown
Rhodotorula
Davidiella
Phoma
Botrytis
Clavispora
Ustilaginomycetes;Unknown
Issatchenkia
Nectriaceae;Unknown
Wickerhamomyces
Rhizopus
Aureobasidium
Trichocomaceae unidentified
Podosphaera
Tremellales;Incertae sedis 12
Saccharomycetales unidentified
Cryptococcus
Blumeria
Pleosporales unidentified
Cordyceps
Dioszegia
Sporidiobolales;Incertae sedis 25
Sarcosomataceae unidentified
Torulaspora
Basidiomycota unidentified
Rasamsonia
Eupenicillium
Lachnum
Talaromyces
Helotiales unidentified 1
Exophiala
Phaeosphaeriaceae unidentified
Sporobolomyces
Urocystidales;Unknown
Guehomyces
Hypocreales;Unknown
Helotiales;Incertae sedis 2
Tremellomycetes unidentified
Saccharomyces
Monascus
Udeniomyces
Mycocentrospora
Ustilaginaceae;Unknown
Dothideomycetes unidentified
Kazachstania
Hanseniaspora
Thermomyces
Filobasidium
Dothideaceae;Unknown
Tilletia
Ustilago
Alternaria
Trichosphaeriales;Incertae sedis 26
Syzygospora
Leptosphaerulina
Bulleromyces
Pseudozyma
Fusarium
Diplodia
Physcia
Arthrinium
Coniothyrium
Claviceps
Endoconidioma
Xeromyces
Gymnoascus
Peziza
Sordariomycetes;Unknown
Letendraea
Zygosaccharomyces
Pestalotiopsis
Ustilaginales unidentified
Golovinomyces
Quambalaria
Ascosphaera
Cystofilobasidium
Chaetothyriales unidentified
Dipodascaceae unidentified
Trichoderma
Capnodiales;Unknown
Mortierella
Dekkera

Supplementary Figure 11



Supplementary Figure 12



Chapter 4

Intestinal Candida isolates from Rett syndrome subjects bear potential virulent traits and capacity to persist within the host

This chapter will be shortly submitted for publication as an original research article:

Strati F, De Filippo C, Rizzetto L, De Felice C, Hayek J, Jousson O, Leoncini S, Renzi D, Calabrò A, Donati C and Cavalieri D (2016). *Intestinal Candida isolates from Rett syndrome subjects bear potential virulent traits and capacity to persist within the host.*

The gut mycobiota, together with its bacterial counterpart, exerts key roles in maintaining microbial community structure, metabolic functions and has strong immunomodulatory properties, being a main actor in host physiopathology [1, 2]. Since previous observations indicated the presence of a subclinical inflammatory status [3], cytokine dysregulation in both Th1 and Th17 responses [4, 5] and intestinal dysbiosis in RTT subjects, we asked whether fungal isolates from stool samples of RTT subjects may hold virulence traits and immunomodulatory properties favouring the not yet well understood sub-inflammatory status triggered by MeCP2 deficiency.

We discovered that *C. parapsilosis* was the most abundant species retrieved in RTT's faecal samples. These isolates were characterized by high resistance to azoles antifungals and using RAPD genotyping we found them genetically unrelated to the *C. parapsilosis* isolates from healthy controls. Furthermore, RTT *C. parapsilosis* isolates exerted strong immunological responses in human PBMCs inducing high levels of IL-10 and a mixed population of Th1/Th17 cells suggesting the capacity of these fungi to persist within the host being potentially involved in chronic, pro-inflammatory responses.

This research project gave me the possibility to gain experience in immunology and cellular microbiology being specifically trained in isolation, stimulation of human peripheral blood mononuclear cells, intracellular staining of transcription factors and measurements of cytokines by using flow cytometry and immunoassays. As first author I designed and performed all the experiments, analysed the data, wrote the manuscript and generated figures and tables.

References

1. Wheeler ML, Limon JJ, Bar AS, Leal CA, Gargus M, Tang J, Brown J, Funari VA, Wang HL, Crother TR *et al*: **Immunological Consequences of Intestinal Fungal Dysbiosis**. *Cell host & microbe* 2016, **19**(6):865-873.
2. Underhill DM, Iliev ID: **The mycobiota: interactions between commensal fungi and the host immune system**. *Nature reviews Immunology* 2014, **14**(6):405-416.
3. Cortelazzo A, De Felice C, Guerranti R, Signorini C, Leoncini S, Pecorelli A, Zollo G, Landi C, Valacchi G, Ciccoli L *et al*: **Subclinical inflammatory status in Rett syndrome**. *Mediators of inflammation* 2014, **2014**:480980.
4. Leoncini S, De Felice C, Signorini C, Zollo G, Cortelazzo A, Durand T, Galano JM, Guerranti R, Rossi M, Ciccoli L *et al*: **Cytokine Dysregulation in MECP2- and CDKL5-Related Rett Syndrome: Relationships with Aberrant Redox Homeostasis, Inflammation, and omega-3 PUFAs**. *Oxidative medicine and cellular longevity* 2015, **2015**:421624.
5. Li C, Jiang S, Liu S-Q, Lykken E, Zhao L-t, Sevilla J, Zhu B, Li Q-J: **MeCP2 enforces Foxp3 expression to promote regulatory T cells' resilience to inflammation**. *Proceedings of the National Academy of Sciences* 2014, **111**(27):E2807-E2816.

Intestinal *Candida* isolates from Rett syndrome subjects bear potential virulent traits and capacity to persist within the host

Francesco Strati[1,2], Carlotta De Filippo[3], Lisa Rizzetto[4], Claudio De Felice[5], Joussef Hayek[6], Olivier Jousson[2], Silvia Leoncini[6], Daniela Renzi[7], Antonio Calabrò[7], Claudio Donati[1], Duccio Cavalieri[8,*]

[1]Computational Biology Research Unit, Research and Innovation Centre, Fondazione Edmund Mach, Via E. Mach 1, 38010 San Michele all' Adige, Italy
[2]Centre for Integrative Biology, University of Trento, Via Sommarive 9, 38123 Trento, Italy
[3]Institute of Biometeorology (IBIMET), National Research Council (CNR), Via Giovanni Caproni 8, 50145 Florence, Italy
[4]Nutrition and Nutrigenomics unit, Research and Innovation Centre, Fondazione Edmund Mach, Via E. Mach 1, 38010 San Michele all' Adige, Italy
[5]Neonatal Intensive Care Unit, University Hospital AOUS, Viale Bracci 16, 53100 Siena, Italy
[6]Child Neuropsychiatry Unit, University Hospital AOUS, Viale Bracci 16, 53100 Siena, Italy
[7]Department of Experimental and Clinical Biomedical Sciences, Gastroenterology Unit, University of Florence, Viale Morgagni 40, 50139 Florence, Italy
[8]Department of Biology, University of Florence, Via Madonna del Piano 6, 50019 Sesto Fiorentino, Florence, Italy

*__Correspondence to__: Duccio Cavalieri, Department of Biology, University of Florence, via Madonna del Piano 6, I-50019 Sesto Fiorentino, Firenze, Italy; e-mail: *duccio.cavalieri@unifi.it*

Abstract

Rett syndrome (RTT) is a neurological disorder mainly caused by mutations in *MeCP2* gene. It has been shown that MeCP2 impairments can lead to cytokine dysregulation due to the MeCP2 regulatory role in T-helper and T-reg mediated responses, therefore contributing to the pro-inflammatory status associated with RTT. Furthermore the intestinal dysbiosis recently demonstrated in RTT is a known factor responsible for hyper activation of pro-inflammatory immune responses in different intestinal and extra-intestinal pathologies. Here, we studied the cultivable gut mycobiota of RTT subjects characterizing the faecal fungal isolates for their virulence-related traits, antifungal resistance and immunomodulatory properties in order to elucidate the role of fungi in RTT's intestinal dysbiosis and gastrointestinal physiology. Phenotypical analysis of the gut mycobiota in RTT revealed that RTT's isolates produced more biofilm and were significantly more resistant to fluconazole compared to the isolates from the healthy subjects. Furthermore *C. parapsilosis* was the most abundant yeast species in RTT subjects, showing distinct genotypic profiles if compared to healthy controls' *C. parapsilosis* isolates as measured by hierarchical clustering analysis from RAPD genotyping. Such *C. parapsilosis* isolates were also characterized by high levels of resistance to azoles antifungals, making them a difficult target in case of fungal infections. In addition, the high levels of IL-10 produced by PBMCs and the mixed Th1/Th17 cells population induced by RTT *C. parapsilosis* isolates, suggest the capacity of these intestinal fungi to persist within the host, being potentially involved in chronic, pro-inflammatory responses which mechanisms are still to be elucidated in RTT.

Introduction

Rett syndrome (RTT) is a neurological disorder that almost exclusively affects females with an incidence of 1:10,000 live births [1] due to a loss-of-function mutations of the X-linked methyl-CpG binding protein 2 (*MeCP2*) gene in approximately the 90% of classical cases of RTT [2]. RTT subjects develop normally up to 18 months of age after which they undergo a period of neurological regression [1]. RTT affects several organs and system among which the autonomic nervous system [1], the gastrointestinal tract [3] and the immune system [4] making it eligible as a multisystemic disease [1]. Indeed, it has been shown that MeCP2 deficiency is able to lead to cytokine dysregulation [4, 5], to influence the expression of Foxp3 [6], important transcription factor involved in the generation of regulatory T (T-reg) cells, and to determine the significant increase of secreted IL-17A [6]. Since the T-helper (Th)17/T-reg balance is crucial for the development of autoimmune/inflammatory disorders it is possible to hypothesize the presence of an autoimmune component in RTT [7]. To this regard, intestinal dysbiosis may cause chronic intestinal inflammation and autoimmunity as occur in Inflammatory Bowel Diseases (IBDs) [8]. The disruption of the

microbial community structure may lead to inflammation involving hyper activation of Th1 and Th17 immune responses [9]. Altered immunological response to fungi can in turn contribute to systemic inflammatory responses. Remarkably, fungal infections shift indoleamine 2,3-dioxygenase's activity (IDO1) [10, 11], reducing the levels of kynurenine [12], a neuroprotective agent. The levels of kynurenine are moreover affected by the immune system, dysfunction of which has been implicated in several pathologies, including autism spectrum disorders [13]. The host response to fungi is mediated at first by innate immunity and then by adaptive immune responses, especially by Th1 and Th17 responses [14]. Previous studies indicated the presence of a subclinical inflammatory status in subjects affected by RTT [15] remarked by cytokine dysregulation in both Th1 and Th17 responses [5, 6, 16, 17] and an intestinal dysbiosis characterized by high relative abundance of the genus *Candida* [18]. Therefore we asked whether the intestinal fungal population of RTT subjects may be involved in the sub-inflammatory status triggered by MeCP2 deficiency that still remains elusive. Here we studied the cultivable gut mycobiota of RTT subjects characterizing the isolated fungi for their virulence-related traits and antifungal resistance. Moreover we characterized the genetic diversity of *C. albicans* and *C. parapsilosis* isolates and their ability to induce innate and adaptive immunological responses in human PBMCs in order to elucidate the role of fungi in RTT gastrointestinal physiology.

Results

RTT gut mycobiota shows a reduction of *C. albicans* and an expansion of *C. parapsilosis* population

We identified 122 fungal isolates belonging to different species (Supplementary Table 1). Twenty-four of such isolates were obtained from stool samples of RTT subjects (Supplementary Table 1). We discovered a significant reduction of fungal species richness in RTT subjects compared to HC (p=3.9e-05, Wilcoxon rank-sum test) in agreement with the results obtained in our previous work [18]. *Candida* was the most abundant genus present in both RTT subjects (91.7%) and HC (71.4%) with *C. albicans* and *C. parapsilosis* as the two most abundant species in both RTT subjects and HC. Interestingly we observed a particular trend in which the higher was the number of isolated *C. albicans* the lower was the number of isolated *C. parapsilosis* and *vice versa i.e.* in RTT subjects 4 out of 24 fungal isolates belonged to *C. albicans* (16.7%) while 14 out of 24 belonged to *C. parapsilosis* (58.3%); in the HC, 49 out of 98 fungal isolates belonged to *C. albicans* (50%) while 15 out of 98 belonged to *C. parapsilosis* (15.3%) (Supplementary Fig. 1). Although the cultivable gut mycobiota of the inspected individuals was often characterized by a predominant fungal species, the concomitant presence of different *Candida* species in different subjects, as well as other fungi, was

also observed (Supplementary Table 1) making possible to hypothesize the presence of phenomena of competitive exclusion.

We then characterized the fungal isolates for putative virulence-associated traits and resistance to antifungals (Supplementary Table 1). The 50% and the 63.8% of fungal isolates from RTT subjects and HC respectively were able to form hyphae or pseudohyphae (Supplementary Table 1). In addition, we observed that the morphotype switch to hyphae or pseudohyphae was related to the isolates' invasiveness, with hyphae- and pseudohyphae-forming isolates being the most invasive (Supplementary Fig. 2a). We also observed that RTT isolates produced more biofilm (p=1.3e-05, Wilcoxon rank-sum test; Supplementary Fig. 2b) and were significantly more resistant to fluconazole compared to HC isolates (45.8% of RTT isolates were resistant vs 18.1% of HC isolates; p=5.1e-06, Wilcoxon rank-sum test). As previously observed, we found the co-occurrence of azole cross-resistance between fluconazole and itraconazole (Spearman's correlation r=0.57; p=2.2e-10) [19]. Almost the totality of the isolates were susceptible to 5-flucytosine with MIC≤0.125µg/ml (Supplementary Table 1). *C. parapsilosis* isolates from HC were sensible to fluconazole (MIC90=2µg/ml; R=7.7%) and itraconazole (MIC90=0.0156µg/ml; R=0%) while *C. parapsilosis* isolates from RTT subjects exhibited a high resistance to these antifungals (fluconazole, MIC90>64µg/ml, R=35.7%, p=0.003, Wilcoxon rank-sum test, Figure 1a; itraconazole, MIC90>8µg/ml, R=35.7%; Table 1). On the contrary, *C. albicans* isolates from HC exhibited low susceptibility to fluconazole (MIC90>64µg/ml, R=23.8%) and itraconazole (MIC90>8µg/ml, R=47.6%; p=0.03, Wilcoxon rank-sum test; Figure 1b) while the totality of RTT *C. albicans* isolates were susceptible to such azoles (Table 1).

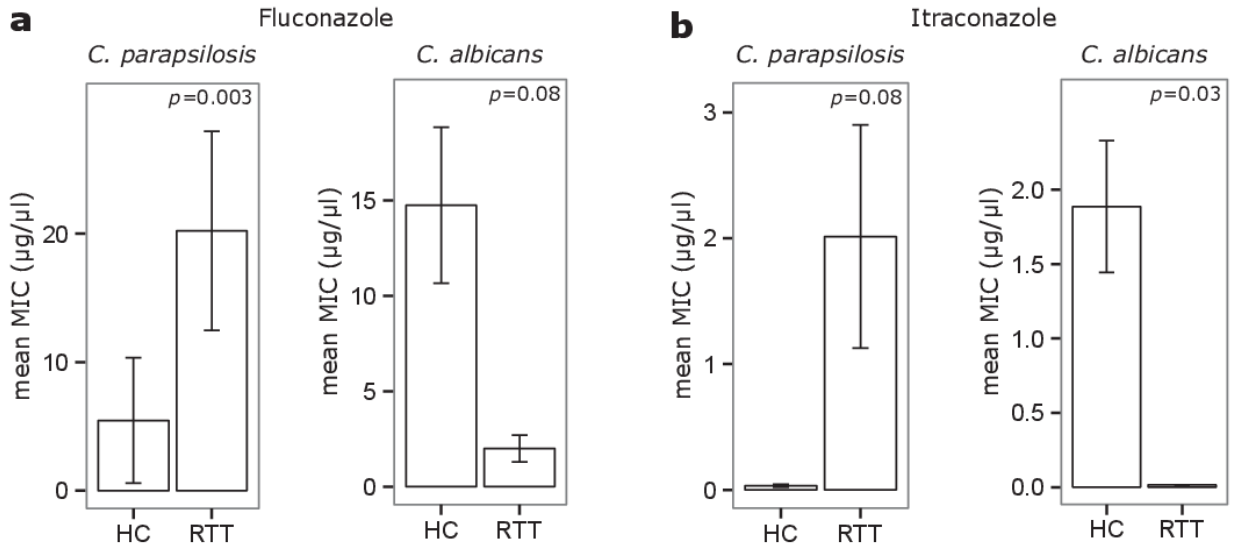


Figure 1: a) Fluconazole and **b)** itraconazole resistance as measured by MIC values in *C. albicans* and *C. parapsilosis* isolates from HC and RTT subjects. MIC values are reported as means ± standard errors. Exact *p-values* are reported and considered significant if < 0.05.

Noteworthy, all the others *Candida* species isolated from RTT subjects (*i.e. C. glabrata, C. pararugosa* and *C. tropicalis*) were resistant to fluconazole (MIC90>64μg/ml; R=100%) and itraconazole (MIC90=8μg/ml; R=100%) while the *Candida* spp. isolated from HC (*i.e. C. deformans, C. intermedia* and *C. lusitaniae*) were completely susceptible to these azoles (Table 1).

Table 1. Antifungals resistance of *Candida* isolates from HC or RTT subjects

Species	Antifungals	**Healthy controls (HC)**					**Rett syndrome (RTT) subjects**				
		MIC (μg/ml)		*[#]Clinical breakpoints*			*MIC (μg/ml)*		*[#]Clinical breakpoints*		
		MIC_{50}	**MIC_{90}**	**%S**	**%SDD**	**%R**	**MIC_{50}**	**MIC_{90}**	**%S**	**%SDD**	**%R**
C. albicans	Fluconazole	0.5	> 64	76.2	0	23.8	1	2	75	25	0
	Itraconazole	0.25	> 8	47.6	4.8	47.6	0.0156	0.0156	100	0	0
	5-Flucytosine	0.125	0.5	95.2	2.4	2.4	0.125	0.125	100	0	0
C. parapsilosis	Fluconazole	0.5	2	92.3	0	7.7	2	> 64	64.3	0	35.7
	Itraconazole	0.0156	0.125	100	0	0	0.0156	> 8	64.3	0	35.7
	5-Flucytosine	0.125	0.5	100	0	0	0.125	0.125	100	0	0
**Candida spp.*	Fluconazole	0.125	0.25	100	0	0	> 64	> 64	0	0	100
	Itraconazole	0.0156	0.0156	100	0	0	8	8	0	0	100
	5-Flucytosine	0.125	0.125	100	0	0	0.125	0.125	100	0	0

[#]According to Pfaller *et al.* 2012; S = sensible; SDD = Sensibility Dose-Dependent or Intermediate; R = Resistant; MIC ranges: Fluconazole 0.125-64 μg/ml; Itraconazole 0.0156-8 μg/ml; 5-Flucytosine 0.125-64 μg/ml. **Candida* spp. isolated from RTT subjects (*i.e. C. glabrata, C. pararugosa* and *C. tropicalis*); *Candida* spp. isolated from HC (*i.e. C. deformans, C. intermedia* and *C. lusitaniae*).

Taken together these results suggest that RTT isolates may be more virulent and difficult to treat in case of infection than HC isolates. Finally, we have been able to isolate *Trichosporon asteroides* and *Saccharomyces cerevisiae* only from RTT subjects. These isolates were both resistant to fluconazole (with MIC>64μg/ml and MIC=8μg/ml respectively) while *Trichosporon asteroides* was also resistant to itraconazole (MIC>8μg/ml). Such species are recognized as potential new emerging fungal pathogens [20] representing a potential treat in RTT subjects.

C. parapsilosis isolates from RTT subjects and HC are genetically distinct

The genetic diversity among *Candida* isolates from HC and RTT subjects was determined by UPGMA hierarchical clustering analysis of Jaccard distances calculated from RAPD genotyping. We observed that *C. parapsilosis* isolates from RTT samples were genetically unrelated to those from HC, with most of RTT *C. parapsilosis* isolates clustering in a single group (Figure 2 and Supplementary Fig. 3; p=0.002, PERMANOVA). On the contrary *C. albicans* isolates from RTT subjects were genetically more variable among them, clustering in different clade of the tree

(Supplementary Fig. 4; p=0.779, PERMANOVA). It is worth to note that we only obtained 4 *C. albicans* isolates from RTT samples.

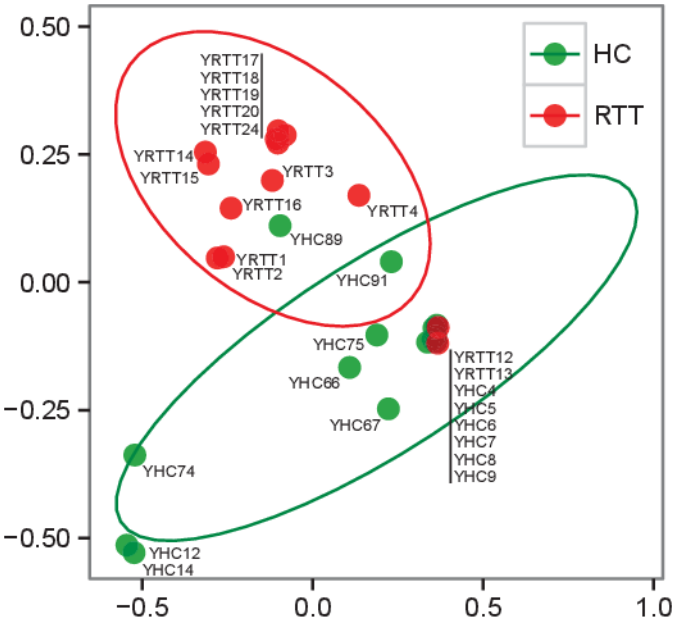


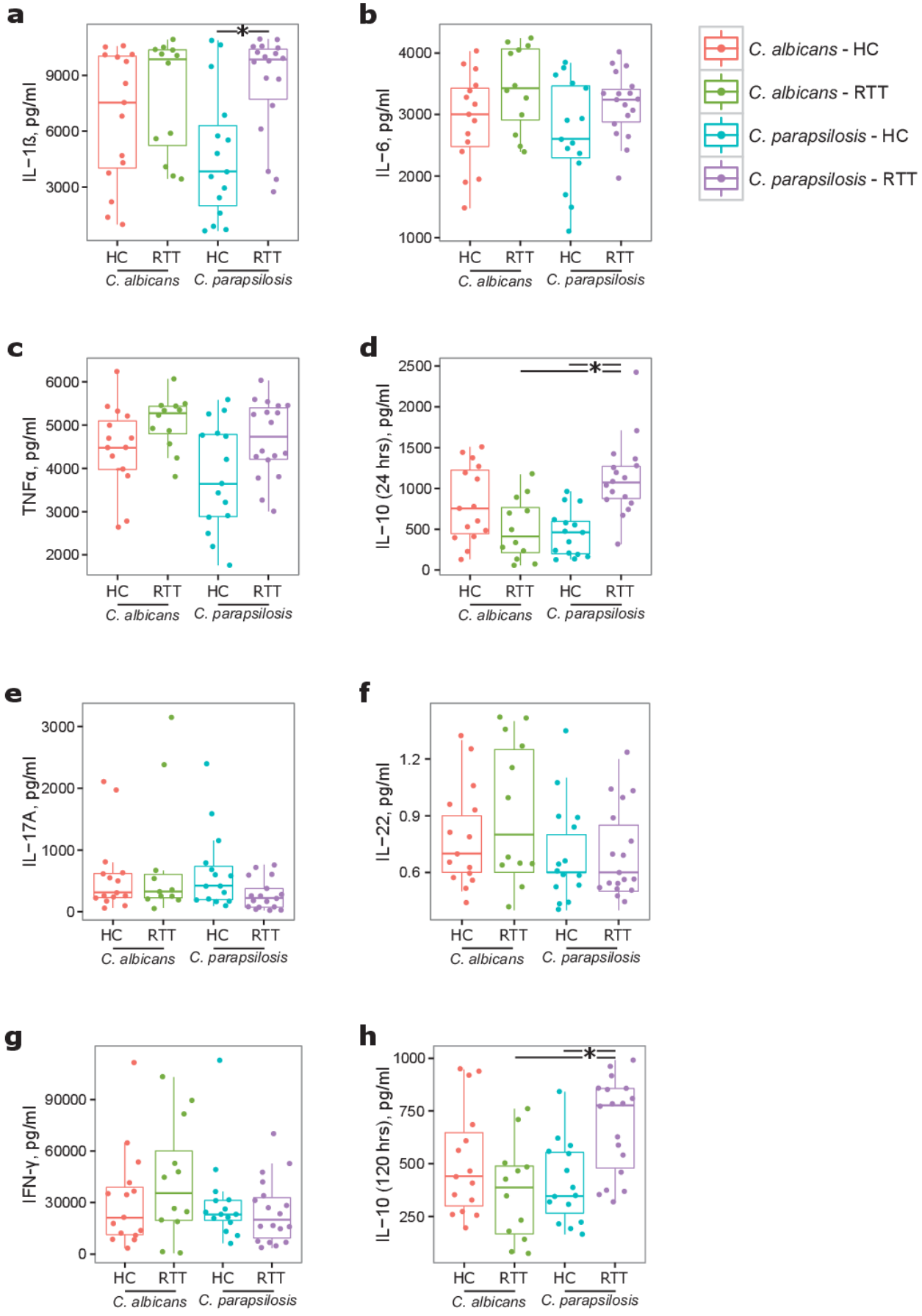
Figure 2: Multidimensional scaling analysis of *C. parapsilosis* genetic diversity calculated by UPGMA hierarchical clustering analysis of samples' distance similarities (Jaccard index) from RAPD genotyping. *C. parapsilosis* isolates from HC and RTT subjects in green and red, respectively.

Candida parapsilosis from RTT subjects induces high levels of IL-10 in PBMCs

The first step in the immunological response against *Candida* is the production of pro-inflammatory cytokines, such as IL-1β, IL-6 and TNFα, by innate immune cells. Later, these cytokines will promote adaptive immunity mediated by Th1 or Th17 responses [14]. Stimulation of PBMCs with *C. albicans* and *C. parapsilosis* isolates from HC and RTT subjects revealed that RTT *Candida* isolates induces higher levels of IL-1β, IL-6 and TNFα with respect to HC *Candida* isolates (Figure 3). On the contrary, any remarkable difference was observed in the levels of IL-17A, IL-22 and INFγ (Figure 3). Furthermore we observed that *C. parapsilosis* isolates from RTT subjects induced highly significant levels of IL-10 compared to HC *C. parapsilosis* isolates and RTT *C. albicans* isolates (p<0.003, Wilcoxon rank-sum test; Figure 3) suggesting an increased fungal tolerance towards these *C. parapsilosis* isolates, potentially favouring fungal persistence within the host.

Figure 3: Cytokines production by peripheral blood mononuclear cells (PBMCs; $5x10^5$ cell) after stimulation with $5x10^6$ heat-killed *C. albicans* or *C. parapsilosis* isolates from HC and RTT subjects. In panels **a-d)** are reported the values for the cytokines produced by innate immune cells (IL-1β, IL-6, TNFα and IL-10 after 24h of PBMCs stimulation) while in **e-h)** the values for the cytokines produced following adaptive immune responses (IL-17A, IL-22, IFNγ and IL-10 after 5 days of PBMCs stimulation). The dots represent each of the three replicates *per* isolate tested; *p<0.05, Wilcoxon rank-sum test. (See figure 3 on next page).

Figure 3


a
IL−1ß, pg/ml
9000
6000
3000
*
HC
RTT
HC
RTT
C. albicans
C. parapsilosis
b
IL−6, pg/ml
4000
3000
2000
1000
HC
RTT
HC
RTT
C. albicans
C. parapsilosis
C. albicans - HC
C. albicans - RTT
C. parapsilosis - HC
C. parapsilosis - RTT
c
TNFα, pg/ml
6000
5000
4000
3000
2000
HC
RTT
HC
RTT
C. albicans
C. parapsilosis
d
IL−10 (24 hrs), pg/ml
2500
2000
1500
1000
500
0
*
HC
RTT
HC
RTT
C. albicans
C. parapsilosis
e
IL−17A, pg/ml
3000
2000
1000
0
HC
RTT
HC
RTT
C. albicans
C. parapsilosis
f
IL−22, pg/ml
1.2
0.9
0.6
HC
RTT
HC
RTT
C. albicans
C. parapsilosis
g
IFN-γ, pg/ml
90000
60000
30000
0
HC
RTT
HC
RTT
C. albicans
C. parapsilosis
h
IL−10 (120 hrs), pg/ml
1000
750
500
250
*
HC
RTT
HC
RTT
C. albicans
C. parapsilosis

Nevertheless we did not observe significant differences in the expression of IDO1 in PBMCs stimulated by *C. parapsilosis* isolates. Since we observed variable levels of Th-driving cytokines, we asked whether *Candida* isolates were able to induce a different Th1/Th17 polarization. We therefore measured the intracellular levels of the key transcription factors T-bet and RORγt (Supplementary Fig. 5), involved in the differentiation of $CD4^+$ naïve cells in Th1 and Th17 cells, respectively [21]. Reflecting the potential of different strains to elicit the immune response at different extents, as previously observed in culture supernatants, we measured variable, but not significant, levels of T-bet and RORγt in response to the different *C. parapsilosis* and *C. albicans* isolates (Supplementary Fig. 5). This could be due to the diverse immune reactivity shown by the different isolates of the same species, as previously observed [22, 23]. However, *C. parapsilosis* isolates from RTT subjects induced more cells co-expressing both RORγt and T-bet compared to HC *C. parapsilosis* isolates (raw *p-value*=0.04, FDR-corrected *p-value*=0.12; Wilcoxon rank-sum test; Supplementary Fig. 5c) suggesting the potentiality of RTT *C. parapsilosis* isolates in promoting pro-inflammatory responses.

Discussion

In this study we showed that in RTT, a multisystemic neurological disorder, faecal *C. parapsilosis* isolates hold potential phenotypic traits favouring the previously observed intestinal sub-inflammatory status [18]. Species level analysis of the cultivable gut mycobiota revealed *C. parapsilosis* as the most abundant yeast species in RTT subjects, genetically unrelated to HC *C. parapsilosis* isolates as measured by hierarchical clustering analysis from RAPD genotyping. Interestingly RTT *C. parapsilosis* isolates were characterized by high levels of resistance to azoles antifungals, making them a difficult target in case of fungal infections in these subjects. Furthermore the high levels of IL-10 produced by PBMCs suggest the RTT *C. parapsilosis*' capacity to persist within the host and to be tolerated by the immune system. IL-10 usually exert a homeostatic control to keep inflammation under control although high levels of IL-10 characterize the cases of chronic fungal infections dominated by non-resolving inflammation [14]. It has been observed that *C. albicans* induces host's immunosuppression through increased IL-10 production by immune cells representing an important mechanisms in *Candida* pathogenesis [24]. Similarly, RTT *C. parapsilosis* isolates could escape immune clearance through a similar mechanism mediated by high levels of IL-10 that, in turn, could impair antifungal Th1 immunity and thus favouring persistent infection.

C. parapsilosis has been described as one of the leading causes of invasive candidal disease [25], to be responsible of macrophage activation and allergic airways inflammation [26] and to be one of the dominant *Candida* species leading IBDs' dysbiosis [27]. Fungal opportunistic infections are generally ascribed to defective host immunity but may require specific microbial population dysbiosis [28] as

recently observed in RTT [18]. Recent studies indicated that fungal infections may originate from individual's own commensal strains and that the ability of a commensal organism to produce disease is not merely a consequence of impaired host immunity [29] as in the case of *C. albicans* GUT (gastrointestinal induced transition) in which virulence-associated genes are down-regulated, enabling fungal adaptation for long-term survival in the large intestine [30]. Therefore *C. parapsilosis* isolates from RTT subjects may be commensals potentially dangerous for the host due to RTT's altered immunological status and the presence of a dysbiotic gut microbiota. Moreover, RTT *C. parapsilosis* isolates induced a higher proportion of a mixed Th1/Th17 cells population compared to HC *C. parapsilosis* isolates. Although Th1 and Th17 responses usually counter-regulate each other, there are increasing evidences of co-operation and dependency between these two immunological responses [31] which are involved in chronic, pro-inflammatory responses as observed in IBDs [32] potentially resulting in adaptive immunity against the commensal microbiota [33]. Interestingly, it has been previously shown that MeCP2 could actually play a regulatory role in T-cells' resilience to inflammation [6]. In particular, emerging evidence indicates that MeCP2 deficiency is able to lead to cytokine dysregulation including macrophage-related cytokines in *Mecp2*-null mice and RTT girls [4, 5], although the understanding of the possible molecular mechanisms underlying this pro-inflammatory status remains elusive. To this regard, the dysbiosis demonstrated in our previous study [18] and the presence of putative virulent, pro-inflammatory intestinal *C. parapsilosis* strains could represent an additional factor in RTT's gastrointestinal pathophysiology.

Methods

Isolation and identification of cultivable fungal species from faeces

Stool samples from a cohort of 50 RTT patients and 29 Healthy Controls (HC) [18] were homogenized in sterile Ringer's solution and plated on solid YPD medium (1% Yeast extract, 2% Bacto-peptone, 2% D-glucose, 2% agar) supplemented with 25U/ml of penicillin, 25µg/ml of streptomycin (Sigma-Aldrich) and incubated aerobically at 27°C for 3-5 days. All fungal isolates grown on the selective medium were further isolated to obtain single-cell pure colonies. Genomic DNA was extracted from pure cultures of the isolated colonies as previously described [34]. Fungal isolates were identified by amplification and sequencing of the ribosomal Internal Transcribed Spacer (ITS) region, using ITS1 (5'-GTTTCCGTAGGTGAACTTGC-3') and ITS4 (5'-TCCTCCGCTTATTGATATGC-3') primers [35]. ITS1-4 sequences were then classified by using the BLAST algorithm in the NCBI database (minimum 97% sequence similarity and 95% coverage with a described species).

Invasive growth

The ability of fungal strains to penetrate the YPD solid medium was tested as previously described [36]. M28-4D and BY4742 *S. cerevisiae* strains, known to be invasive and non-invasive respectively, have been used as controls. The strain invasiveness was assigned with scores from 3 (highly invasive) to 0 (non-invasive).

Hyphal formation

Fungal cells (~10^5 cells/ml) were grown for 7 days in liquid YPD and YNB media (0.67% Yeast Nitrogen Base w/o aminoacids and $(NH_4)_2SO_4$ (Sigma-Aldrich), 2% glucose), both at 27°C and 37°C in order to evaluate hyphae or pseudohyphae formation. Formation of hyphae was inspected by optical microscope observation with a Leica DM1000 led instrument (magnification 40x and 100x).

Biofilm formation

Fungal cells (~10^5 cells/ml) were grown in liquid YPD at 37°C for 48 hours in flat-bottom 96-well microtiter plates. After the incubation period cell suspensions were aspirated and each well with the adhered fungal cells was washed three times with deionized H_2O and one time with PBS 1X. Biofilm-coated wells were then incubated with 0.01% of crystal violet (Sigma) for 30 minutes and washed as above. Finally, each well of the dried microtiter plate was incubated with 100μl of 100% EtOH for 10 minutes and biofilms quantified by optical density measurement at 570nm with a microplate reader (Synergy2, BioTek, USA).

Antifungal susceptibility testing. All fungal isolates were tested for susceptibility to fluconazole, itraconazole and 5-flucytosine (Sigma-Aldrich) by Minimum Inhibitory Concentration (MIC) assays according to the European Committee on Antimicrobial Susceptibility Testing (EUCAST) recommendations [37, 38]. Clinical and Laboratory Standards Institute (CLSI) clinical breakpoints (CBPs) were used to evaluate the antifungal resistance [19, 39]. CBPs have not been established for non-*Candida* yeasts and the non-*Aspergillus* moulds, however have been used as a proxy for the evaluation of antifungals susceptibility in such isolates.

RAPD genotyping and clustering analysis

C. albicans and *C. parapsilosis* isolates were genotyped by Random Amplification of Polymorphic DNA (RAPD) using the primer Oligo 2 (5'-TCACGATGCA-3') as described previously [40]. Amplifications were performed according to the following protocol: 5 min at 94°C, 40 cycles of 30sec at 94°C, 30sec at 36°C and 2min at 72°C, followed by a final extension of 10min at 72°C. The PCR

reaction mix contained 1X PCR buffer 2mM $MgCl_2$, 200µM of dNTPs, 0.4µM of the primer, 2.5U of Taq Polymerase and 10ng of gDNA as template. PCR amplicons were separated using a 1.5% agarose gel in 1x TAE buffer at 90V for 2 hours and visualized with 0.5µg/ml ethidium bromide staining. The presence or absence of an amplicon at any position of the gel was used for the construction of a binary matrix that has been used for the calculation of samples' distance similarity according to the Jaccard index [41] by mean of the "vegdist" function within the *vegan* R package and then clustered hierarchically according to the UPGMA method by using the "hclust" function within the *stats* R package.

Isolation and stimulation of PBMCs

Human peripheral blood mononuclear cells (PBMCs) were isolated by Ficoll-Hypaque density gradient centrifugation (Biochrom, Berlin, Germany) from buffy coats provided by the Transfusion Unit of Ospedale Santa Chiara in Trento. The experimental plan was approved by the local hospital ethical committee, and informed consent was obtained from all the healthy donors (protocol No: 54896583). *Candida* isolates were cultured in YPD medium for 18 h at 37°C, then collected. Fungal cells were harvested by centrifugation, washed twice with PBS, heat-killed for 3h at 65°C and resuspended in culture medium (RPMI1640; Sigma Aldrich). For stimulation experiments, 5 x 10^5 PBMCs in RPMI1640 were incubated with 5 x 10^6 heat-killed *C. albicans*, *C. parapsilosis* or RPMI1640 medium alone (negative control). After the incubation periods (24 hours for IL-1β, IL-6, TNF-α, IL-10 production and 120 hours for IL-17A, INFγ, IL-22, IL-10 production) cell suspensions were centrifuged and supernatants were collected and stored at -20°C until assayed. Each experiment was performed in triplicate.

Cytokine Assays

Cytokine detection *i.e.* IL-17A, INFγ, IL-22, IL-1β, IL-6, TNF-α, IL-10 production, were assayed using the MAP human cytokine/chemokine kit (Merck Millipore) according to the manufacturer's instructions (MagPix technology).

Flow cytometry

PBMCs were collected after stimulation with *Candida* isolates in a ratio of 10:1 (stimuli:cells) and washed with PBS. Intracellular staining for IDO1 (after 24 hours of stimulation), T-bet and RORγt (after 5 day of stimulation) were performed using the fixation/permeabilization buffer kit (Life Technologies) following the manufacturing recommendations. Cells were then stained with adequate concentrations of labelled antibodies diluted in PBS+10% heat-inactivated foetal bovine serum (FBS)

for 20 min at room temperature, A minimum of ten thousand events for each sample were acquired using a Guava easyCyte 8T flow cytometer (Merck Millipore) and using the inCyte software (Merck Millipore). Cells were gated first based on forward and side scatter to exclude dead cells and cell debris. The area of positivity was determined by using an isotype-matched control MAb. Antibodies used: Fluorescein isothiocyanate (FITC)-IDO1 (BD Biosciences, Prodotti Gianni, Italy), FITC-Tbet (Millipore), allophycocyanin (APC)-RORγt (BD Biosciences, Prodotti Gianni, Italy).

Statistical analysis

Wilcoxon rank-sum tests and Spearman's correlations were performed using the R software [42] through the *stats* R package (version 3.1.2) and the *psych* R package, respectively. Permutational MANOVA (PERMANOVA) test was performed by using the adonis()function in the R package "vegan" with 999 permutations. All *p*-values have been corrected for multiple hypothesis testing controlling the false discovery rate [43].

Abbreviations

CBP, clinical breakpoint; HC, healthy controls; IBDs, Inflammatory Bowel Diseases; IDO1, indoleamine 2,3-dioxygenase 1; IL, interleukin; INF, interferon, MeCP2, methyl-CpG binding protein 2; MIC, minimum inhibitory concentration; PBMC, peripheral blood mononuclear cells; RAPD, random amplification of polymorphic DNA; RTT, Rett syndrome; Th, T-helper; TNF, tumor necrosis factor; T-reg cell, regulatory T cell; UPGMA, unweighted pair group method with arithmetic mean.

Acknowledgements

The authors would like to thank Irene Stefanini and Davide Albanese for the helpful discussion. We thank also Unifarm S.p.A. for the financial support to FS.

Author contributions

FS designed and performed the experiments, analysed the data and wrote the manuscript. LR supervised and contributed to the immunological assays. CD supervised and contributed to data analysis. DR, SL, CDFe, JH and AC recruited subjects and collected specimens. DC and CDF conceived the study and approved the final manuscript. All the authors critically reviewed and approved the manuscript.

Conflict of interest

Authors declare no conflict of interests.

Ethics approval and consent to participate

The study was approved by the institutional review board of the Siena University Hospital (AOUS, Siena, Italy) and all enrolled subjects or tutors gave written informed consent in accordance with the sampling protocol approved by the local Ethical Committee (No: 2012-005021-76).

Funding

This work was financially supported from the "Accordo di Programma Integrato MetaFoodLabs" funded by the research office of the Provincia Autonoma di Trento (Italy) (PAT Prot. S116/2012/537723).

References

1. Chahrour M, Zoghbi HY: **The story of Rett syndrome: from clinic to neurobiology**. *Neuron* 2007, **56**(3):422-437.
2. Liyanage VR, Rastegar M: **Rett syndrome and MeCP2**. *Neuromolecular medicine* 2014, **16**(2):231-264.
3. Motil KJ, Caeg E, Barrish JO, Geerts S, Lane JB, Percy AK, Annese F, McNair L, Skinner SA, Lee HS *et al*: **Gastrointestinal and nutritional problems occur frequently throughout life in girls and women with Rett syndrome**. *Journal of pediatric gastroenterology and nutrition* 2012, **55**(3):292-298.
4. Cronk JC, Derecki NC, Litvak V, Kipnis J: **Unexpected cellular players in Rett syndrome pathology**. *Neurobiology of disease* 2016, **92**(Pt A):64-71.
5. Leoncini S, De Felice C, Signorini C, Zollo G, Cortelazzo A, Durand T, Galano JM, Guerranti R, Rossi M, Ciccoli L *et al*: **Cytokine Dysregulation in MECP2- and CDKL5-Related Rett Syndrome: Relationships with Aberrant Redox Homeostasis, Inflammation, and omega-3 PUFAs**. *Oxidative medicine and cellular longevity* 2015, **2015**:421624.
6. Li C, Jiang S, Liu S-Q, Lykken E, Zhao L-t, Sevilla J, Zhu B, Li Q-J: **MeCP2 enforces Foxp3 expression to promote regulatory T cells' resilience to inflammation**. *Proceedings of the National Academy of Sciences* 2014, **111**(27):E2807-E2816.
7. De Felice C, Leoncini S, Signorini C, Cortelazzo A, Rovero P, Durand T, Ciccoli L, Papini AM, Hayek J: **Rett syndrome: An autoimmune disease?** *Autoimmunity reviews* 2016, **15**(4):411-416.
8. Round JL, Mazmanian SK: **The gut microbiota shapes intestinal immune responses during health and disease**. *Nature Reviews Immunology* 2009, **9**(5):313-323.
9. Kamada N, Seo SU, Chen GY, Nunez G: **Role of the gut microbiota in immunity and inflammatory disease**. *Nature reviews Immunology* 2013, **13**(5):321-335.
10. De Luca A, Montagnoli C, Zelante T, Bonifazi P, Bozza S, Moretti S, D'Angelo C, Vacca C, Boon L, Bistoni F *et al*: **Functional yet balanced reactivity to Candida albicans requires TRIF, MyD88, and IDO-dependent inhibition of Rorc**. *Journal of immunology* 2007, **179**(9):5999-6008.
11. Cheng SC, van de Veerdonk F, Smeekens S, Joosten LA, van der Meer JW, Kullberg BJ, Netea MG: **Candida albicans dampens host defense by downregulating IL-17 production**. *Journal of immunology* 2010, **185**(4):2450-2457.

12. Bozza S, Fallarino F, Pitzurra L, Zelante T, Montagnoli C, Bellocchio S, Mosci P, Vacca C, Puccetti P, Romani L: **A crucial role for tryptophan catabolism at the host/Candida albicans interface**. *Journal of immunology* 2005, **174**(5):2910-2918.
13. Schwartz CE: **Aberrant tryptophan metabolism: the unifying biochemical basis for autism spectrum disorders?** *Biomarkers in medicine* 2014, **8**(3):313-315.
14. Romani L: **Immunity to fungal infections**. *Nature reviews Immunology* 2011, **11**(4):275-288.
15. Cortelazzo A, De Felice C, Guerranti R, Signorini C, Leoncini S, Pecorelli A, Zollo G, Landi C, Valacchi G, Ciccoli L *et al*: **Subclinical inflammatory status in Rett syndrome**. *Mediators of inflammation* 2014, **2014**:480980.
16. Jiang S, Li C, McRae G, Lykken E, Sevilla J, Liu SQ, Wan Y, Li QJ: **MeCP2 reinforces STAT3 signaling and the generation of effector CD4+ T cells by promoting miR-124-mediated suppression of SOCS5**. *Science signaling* 2014, **7**(316):ra25.
17. O'Driscoll CM, Lima MP, Kaufmann WE, Bressler JP: **Methyl CpG binding protein 2 deficiency enhances expression of inflammatory cytokines by sustaining NF-kappaB signaling in myeloid derived cells**. *Journal of neuroimmunology* 2015, **283**:23-29.
18. Strati F, Cavalieri D, Albanese D, De Felice C, Donati C, Hayek J, Jousson O, Leoncini S, Pindo M, Renzi D *et al*: **Altered gut microbiota in Rett syndrome**. *Microbiome* 2016, **4**(1):41.
19. Pfaller MA, Diekema DJ: **Progress in antifungal susceptibility testing of Candida spp. by use of Clinical and Laboratory Standards Institute broth microdilution methods, 2010 to 2012**. *Journal of clinical microbiology* 2012, **50**(9):2846-2856.
20. Cornely OA, Cuenca-Estrella M, Meis JF, Ullmann AJ: **European Society of Clinical Microbiology and Infectious Diseases (ESCMID) Fungal Infection Study Group (EFISG) and European Confederation of Medical Mycology (ECMM) 2013 joint guidelines on diagnosis and management of rare and emerging fungal diseases**. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* 2014, **20 Suppl 3**:1-4.
21. Zhu J, Yamane H, Paul WE: **Differentiation of effector CD4 T cell populations (*)**. *Annual review of immunology* 2010, **28**:445-489.
22. Rizzetto L, Giovannini G, Bromley M, Bowyer P, Romani L, Cavalieri D: **Strain dependent variation of immune responses to A. fumigatus: definition of pathogenic species**. *PloS one* 2013, **8**(2):e56651.
23. Marakalala MJ, Vautier S, Potrykus J, Walker LA, Shepardson KM, Hopke A, Mora-Montes HM, Kerrigan A, Netea MG, Murray GI *et al*: **Differential adaptation of Candida albicans in vivo modulates immune recognition by dectin-1**. *PLoS pathogens* 2013, **9**(4):e1003315.
24. Netea MG, Sutmuller R, Hermann C, Van der Graaf CA, Van der Meer JW, van Krieken JH, Hartung T, Adema G, Kullberg BJ: **Toll-like receptor 2 suppresses immunity against Candida albicans through induction of IL-10 and regulatory T cells**. *Journal of immunology* 2004, **172**(6):3712-3718.
25. Trofa D, Gacser A, Nosanchuk JD: **Candida parapsilosis, an emerging fungal pathogen**. *Clinical microbiology reviews* 2008, **21**(4):606-625.
26. Kim YG, Udayanga KG, Totsuka N, Weinberg JB, Nunez G, Shibuya A: **Gut dysbiosis promotes M2 macrophage polarization and allergic airway inflammation via fungi-induced PGE(2)**. *Cell host & microbe* 2014, **15**(1):95-102.
27. Ott SJ, Kuhbacher T, Musfeldt M, Rosenstiel P, Hellmig S, Rehman A, Drews O, Weichert W, Timmis KN, Schreiber S: **Fungi and inflammatory bowel diseases: Alterations of composition and diversity**. *Scandinavian journal of gastroenterology* 2008, **43**(7):831-841.
28. Casadevall A, Pirofski LA: **Accidental virulence, cryptic pathogenesis, martians, lost hosts, and the pathogenicity of environmental microbes**. *Eukaryotic cell* 2007, **6**(12):2169-2174.

29. Odds FC, Davidson AD, Jacobsen MD, Tavanti A, Whyte JA, Kibbler CC, Ellis DH, Maiden MC, Shaw DJ, Gow NA: **Candida albicans strain maintenance, replacement, and microvariation demonstrated by multilocus sequence typing**. *Journal of clinical microbiology* 2006, **44**(10):3647-3658.
30. Pande K, Chen C, Noble SM: **Passage through the mammalian gut triggers a phenotypic switch that promotes Candida albicans commensalism**. *Nature genetics* 2013, **45**(9):1088-1091.
31. Damsker JM, Hansen AM, Caspi RR: **Th1 and Th17 cells: adversaries and collaborators**. *Annals of the New York Academy of Sciences* 2010, **1183**:211-221.
32. Strober W, Fuss IJ: **Proinflammatory cytokines in the pathogenesis of inflammatory bowel diseases**. *Gastroenterology* 2011, **140**(6):1756-1767.
33. Jyonouchi H, Geng L, Cushing-Ruby A, Monteiro IM: **Aberrant responses to TLR agonists in pediatric IBD patients; the possible association with increased production of Th1/Th17 cytokines in response to candida, a luminal antigen**. *Pediatric allergy and immunology : official publication of the European Society of Pediatric Allergy and Immunology* 2010, **21**(4 Pt 2):e747-755.
34. Hoffman CS, Winston F: **A 10-Minute DNA Preparation from Yeast Efficiently Releases Autonomous Plasmids for Transformation of Escherichia-Coli**. *Gene* 1987, **57**(2-3):267-272.
35. Sebastiani F, Barberio C, Casalone E, Cavalieri D, Polsinelli M: **Crosses between Saccharomyces cerevisiae and Saccharomyces bayanus generate fertile hybrids**. *Research in microbiology* 2002, **153**(1):53-58.
36. Vopalenska I, Hulkova M, Janderova B, Palkova Z: **The morphology of Saccharomyces cerevisiae colonies is affected by cell adhesion and the budding pattern**. *Research in microbiology* 2005, **156**(9):921-931.
37. Rodriguez-Tudela J, Arendrup M, Barchiesi F, Bille J, Chryssanthou E, Cuenca-Estrella M, Dannaoui E, Denning D, Donnelly J, Dromer F: **EUCAST Definitive Document EDef 7.1: method for the determination of broth dilution MICs of antifungal agents for fermentative yeasts: Subcommittee on Antifungal Susceptibility Testing (AFST) of the ESCMID European Committee for Antimicrobial Susceptibility Testing (EUCAST)∗**. *Clinical Microbiology and Infection* 2008, **14**(4):398-405.
38. Rodriguez-Tudela J, Arendrup M, Arikan S, Barchiesi F, Bille J, Chryssanthou E, Cuenca-Estrella M, Dannaoui E, Denning D, Donnelly J: **EUCAST DEFINITIVE DOCUMENT E. DEF 9.1: Method for the determination of broth dilution minimum inhibitory concentrations of antifungal agents for conidia forming moulds**. *Def* 2008, **9**:1-13.
39. Castanheira M, Messer SA, Jones RN, Farrell DJ, Pfaller MA: **Activity of echinocandins and triazoles against a contemporary (2012) worldwide collection of yeast and moulds collected from invasive infections**. *International journal of antimicrobial agents* 2014, **44**(4):320-326.
40. Binelli CA, Moretti ML, Assis RS, Sauaia N, Menezes PR, Ribeiro E, Geiger DC, Mikami Y, Miyaji M, Oliveira MS *et al*: **Investigation of the possible association between nosocomial candiduria and candidaemia**. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* 2006, **12**(6):538-543.
41. Levandowsky M, Winter D: **Distance between sets**. *Nature* 1971, **234**(5323):34-35.
42. Team RC: **R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2012**. In.: ISBN 3-900051-07-0; 2014.
43. Benjamini Y, Hochberg Y: **Controlling the false discovery rate: a practical and powerful approach to multiple testing**. *Journal of the Royal Statistical Society Series B (Methodological)* 1995:289-300.

Supplementary material chapter 4

Supplementary Figure 1: Relative abundances of *Candida albicans* and *Candida parapsilosis* isolates in Rett syndrome subjects (RTT) and healthy controls (HC). The total abundance of all the other fungal isolates is also reported as "other species".

Supplementary Figure 2: a) Intestinal fungal isolates ability (or not) to produce hyphae or pseudo-hyphae in relationship with their ability to be invasive on YPD solid medium; **b)** biofilm production by intestinal fungal isolates from HC and RTT subjects; ***$p<0.0001$, Wilcoxon rank-sum test.

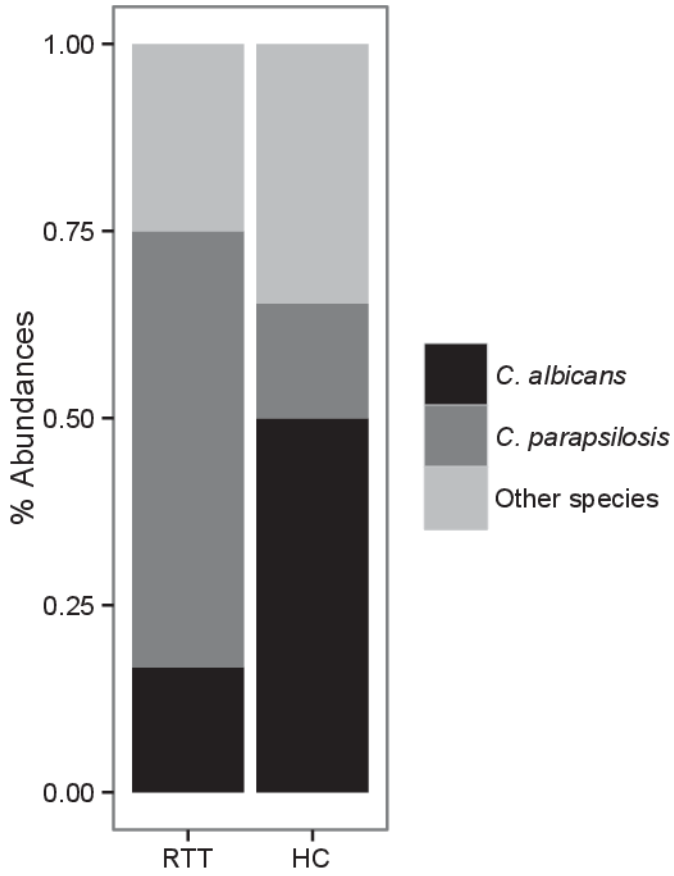
Supplementary Figure 3: UPGMA hierarchical clustering of *C. parapsilosis* genetic diversity calculated by using samples' distance similarities (Jaccard index) from RAPD genotyping. *C. parapsilosis* isolates from HC and RTT subjects in green and red, respectively.

Supplementary Figure 4: UPGMA hierarchical clustering of *C. albicans* genetic diversity calculated by using samples' distance similarities (Jaccard index) from RAPD genotyping. *C. albicans* isolates from HC and RTT subjects in green and red, respectively; in gray the lab strain SC5314.

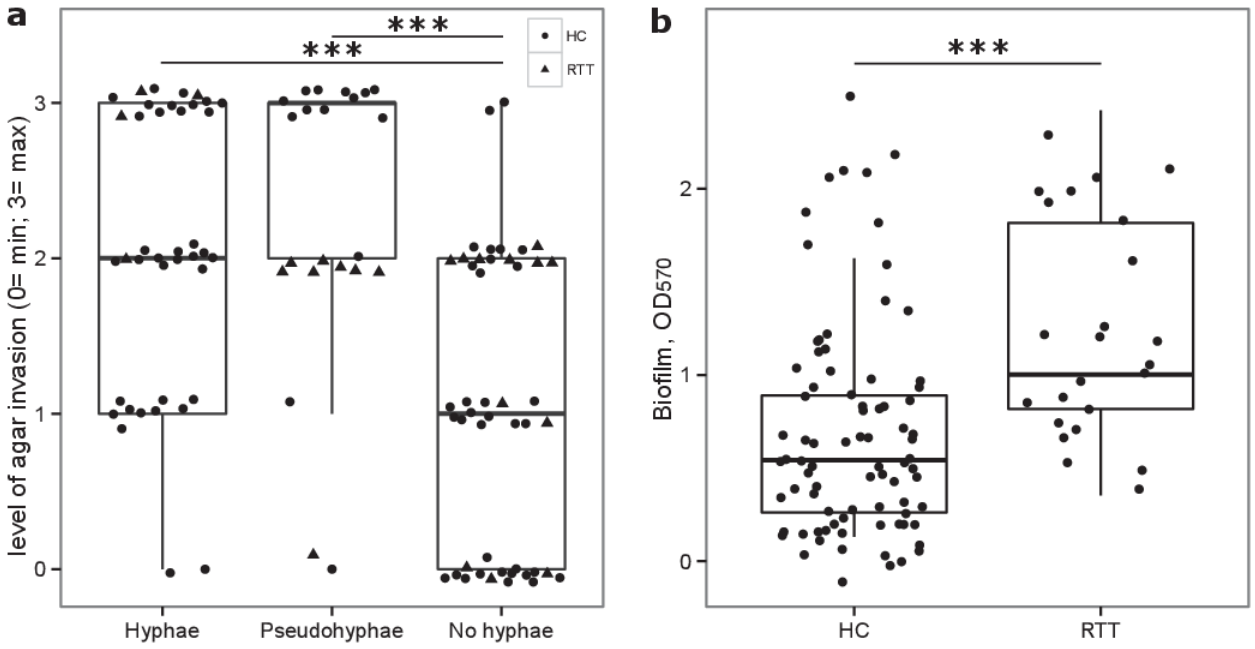
Supplementary Figure 5: Percentage of positive T-cells to T-bet and as measured by intracellular staining and flow cytometry of PBMCs stimulated with **a, b, c)** *C. parapsilosis* isolates and **d, e, f)** *C. albicans* isolates from HC and RTT subjects. Cells were gated for CD4 and data are given as percentage of total gated $CD4^+$ cells.

Supplementary Table 1: Phenotypic characteristics and antifungals susceptibility of fungal isolates.

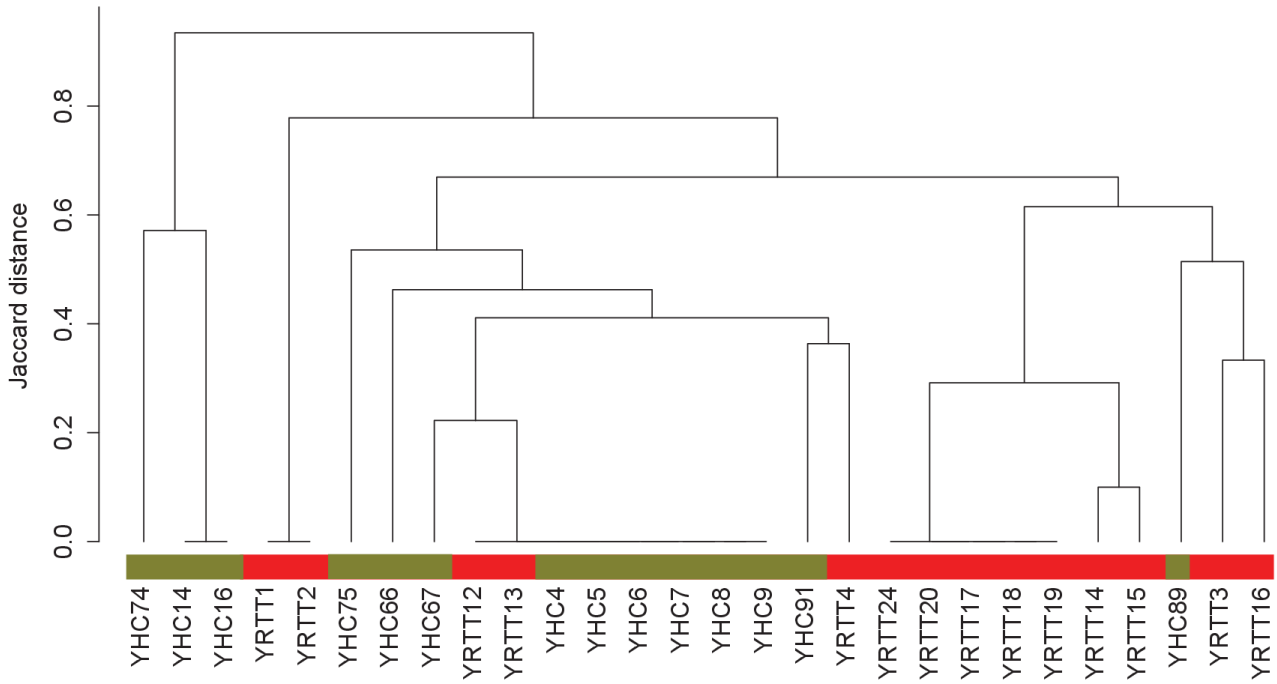
Supplementary Figure 1



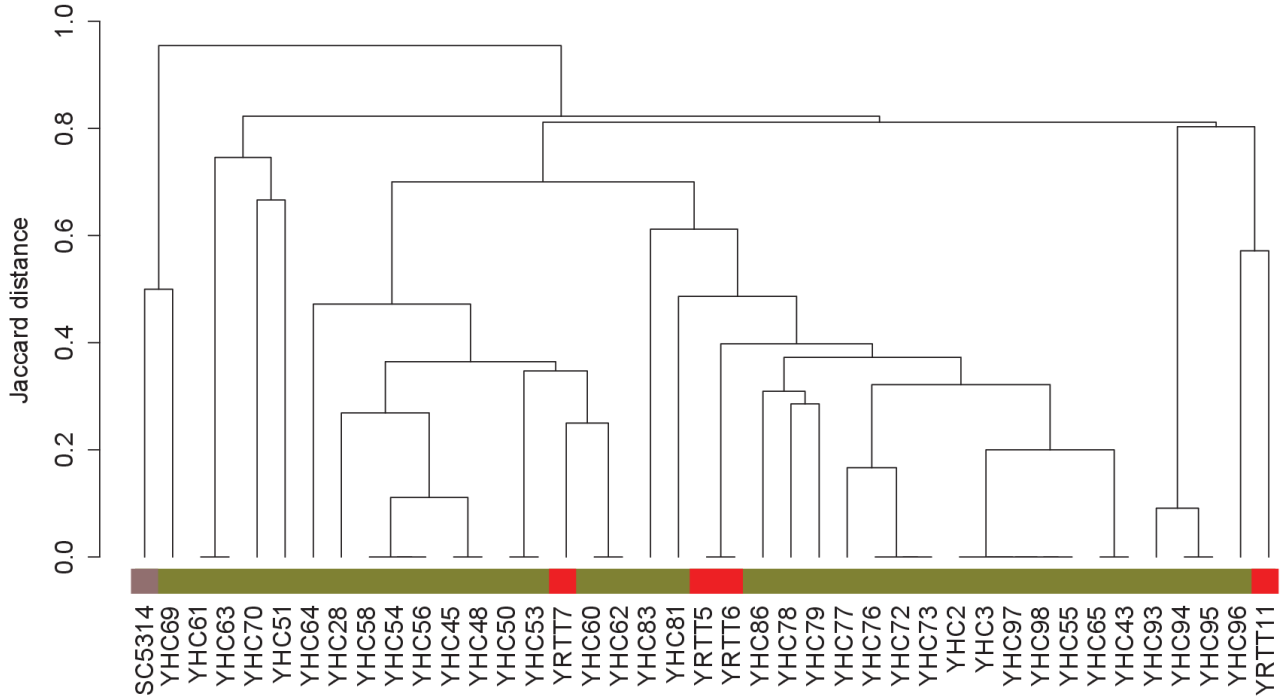
Supplementary Figure 2



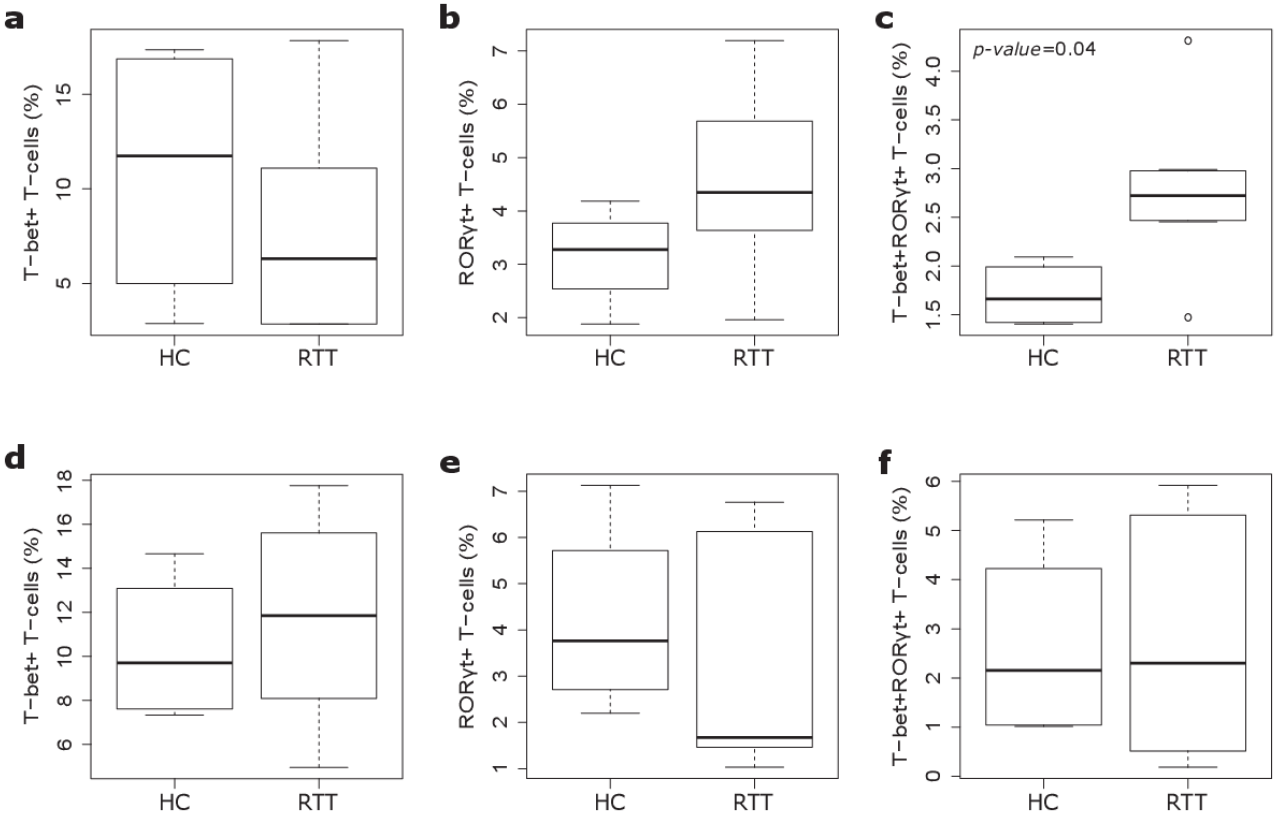
Supplementary Figure 3



Supplementary Figure 4



Supplementary Figure 5



Supplementary Table 1: Phenotypic characteristics and antifungals susceptibility of fungal isolates.

Subject	Status	isolate ID	Fluconazole[§]	Itraconazole[§]	5-flucytosine[§]	Biofilm*	Agar invasion[#]	Hyphae formation	Species
HC21	HC	YHC1	0.5	1	0.125	0.82	2	no hyphae	*Torulaspora delbrueckii*
HC22	HC	YHC2†‡	1	2	0.125	0.66	2	no hyphae	*Candida albicans*
HC22	HC	YHC3	0.5	0.125	0.125	1.18	2	Hyphae	*Candida albicans*
HC29	HC	YHC4	0.5	0.0156	0.125	2.04	3	Pseudohyphae	*Candida parapsilosis*
HC29	HC	YHC5	0.125	0.0156	0.125	1.99	3	Pseudohyphae	*Candida parapsilosis*
HC29	HC	YHC6	0.5	0.0156	0.125	1.59	3	Pseudohyphae	*Candida parapsilosis*
HC29	HC	YHC7†‡	0.5	0.0156	0.125	1.57	3	Pseudohyphae	*Candida parapsilosis*
HC29	HC	YHC8	0.5	0.0156	0.125	1.96	3	Pseudohyphae	*Candida parapsilosis*
HC29	HC	YHC9	0.5	0.0156	0.25	2.03	3	Pseudohyphae	*Candida parapsilosis*
HC29	HC	YHC10	0.5	1	0.125	1.44	3	Pseudohyphae	*Rhodotorula mucilaginosa*
HC29	HC	YHC11	0.25	0.0156	0.125	1.97	3	Pseudohyphae	*Rhodotorula mucilaginosa*
HC31	HC	YHC12†	2	0.0156	1	0.93	3	Pseudohyphae	*Candida parapsilosis*
HC31	HC	YHC13	NA	NA	NA	NA	NA	NA	*Candida albicans*
HC31	HC	YHC14	0.5	0.0156	4	0.94	3	Pseudohyphae	*Candida parapsilosis*
HC31	HC	YHC15	NA	NA	NA	NA	NA	NA	*Candida albicans*
HC31	HC	YHC16	0.25	0.0156	64	1.09	0	no hyphae	*Candida albicans*
HC31	HC	YHC17	NA	NA	NA	NA	0	no hyphae	*Candida albicans*
HC31	HC	YHC18	1	0.031	0.125	NA	0	no hyphae	*Rhodotorula mucilaginosa*
HC31	HC	YHC19	0.25	0.0156	0.5	1.63	0	no hyphae	*Candida lusitaniae*
HC31	HC	YHC20	0.5	0.0156	0.5	NA	0	no hyphae	*Rhodotorula mucilaginosa*
HC31	HC	YHC21	0.5	0.0156	0.125	1.24	0	no hyphae	*Rhodosporidium kratochvilovae*
HC31	HC	YHC22	0.125	0.0156	0.125	1.57	0	no hyphae	*Candida albicans*
HC31	HC	YHC23	NA	NA	NA	NA	NA	NA	*Candida albicans*
HC31	HC	YHC24	NA	NA	NA	NA	NA	NA	*Candida albicans*
HC31	HC	YHC25	NA	NA	NA	NA	NA	NA	*Candida albicans*
HC31	HC	YHC26	0.25	0.0156	0.5	0.20	0	no hyphae	*Candida lusitaniae*
HC31	HC	YHC27	0.125	0.0156	0.125	0.23	1	no hyphae	*Candida albicans*

HC31	HC	YHC28	NA	NA	NA	0.34	1	no hyphae	*Candida albicans*
HC31	HC	YHC29	NA	NA	NA	NA	3	Pseudohyphae	*Candida parapsilosis*
HC31	HC	YHC30	NA	NA	NA	NA	1	no hyphae	*Candida albicans*
HC32	HC	YHC31	0.125	0.0156	0.5	0.25	0	no hyphae	*Pichia manshurica*
HC32	HC	YHC32	0.25	0.0156	> 64	0.47	2	no hyphae	*Pichia manshurica*
HC32	HC	YHC33	NA	NA	NA	NA	2	Hyphae	*Pichia manshurica*
HC32	HC	YHC34	0.25	0.0156	8	0.62	0	no hyphae	*Pichia manshurica*
HC32	HC	YHC35	> 64	> 8	0.125	0.54	0	no hyphae	*Pichia manshurica*
HC32	HC	YHC36	NA	NA	NA	NA	NA	NA	*Pichia manshurica*
HC32	HC	YHC37	0.5	0.0156	0.5	0.30	2	no hyphae	*Pichia manshurica*
HC32	HC	YHC38	0.25	0.0156	0.125	0.30	1	no hyphae	*Pichia manshurica*
HC32	HC	YHC39	0.25	0.0156	0.125	0.67	2	Pseudohyphae	*Pichia manshurica*
HC32	HC	YHC40	> 64	> 8	0.125	0.18	1	no hyphae	*Pichia manshurica*
HC32	HC	YHC41	NA	NA	NA	NA	2	no hyphae	*Pichia manshurica*
HC33	HC	YHC42	NA	NA	NA	NA	0	no hyphae	*Rhodotorula mucilaginosa*
HC33	HC	YHC43	> 64	> 8	0.125	0.24	1	no hyphae	*Candida albicans*
HC33	HC	YHC44	> 64	> 8	0.125	NA	3	Hyphae	*Candida albicans*
HC33	HC	YHC45	0.25	0.0156	0.125	0.14	3	Hyphae	*Candida albicans*
HC33	HC	YHC46	0.5	0.0156	0.125	NA	3	Hyphae	*Candida albicans*
HC33	HC	YHC47	0.25	0.0156	0.125	0.24	2	no hyphae	*Pichia manshurica*
HC33	HC	YHC48	> 64	> 8	0.125	0.57	1	no hyphae	*Candida albicans*
HC33	HC	YHC49	0.25	0.0156	0.125	0.61	3	no hyphae	*Candida albicans*
HC33	HC	YHC50	0.25	0.0156	0.125	0.26	3	Hyphae	*Candida albicans*
HC33	HC	YHC51	0.25	0.0156	0.125	0.41	3	Hyphae	*Candida albicans*
HC33	HC	YHC52	> 64	> 8	0.125	0.17	0	Hyphae	*Rhodosporidium kratochvilovae*
HC34	HC	YHC53	> 64	> 8	0.125	0.18	3	no hyphae	*Candida albicans*
HC37	HC	YHC54	0.25	1	0.5	0.41	3	Hyphae	*Candida albicans*
HC37	HC	YHC55	0.25	2	0.5	0.71	1	Hyphae	*Candida albicans*
HC37	HC	YHC56†ǂ	0.25	2	0.5	0.42	1	Hyphae	*Candida albicans*
HC37	HC	YHC57	0.5	2	0.5	0.75	1	Hyphae	*Torulaspora delbrueckii*
HC37	HC	YHC58	0.5	2	0.5	0.35	2	Hyphae	*Candida albicans*

HC38	HC	YHC59	0.125	0.0156	0.125	0.30	1	no hyphae	*Candida deformans*
HC39	HC	YHC60†	> 64	> 8	0.125	0.39	0	no hyphae	*Candida albicans*
HC39	HC	YHC61	> 64	0.125	0.125	0.16	3	Hyphae	*Candida albicans*
HC39	HC	YHC62	> 64	> 8	0.125	0.38	0	no hyphae	*Candida albicans*
HC39	HC	YHC63	> 64	0.125	0.125	0.19	3	Hyphae	*Candida albicans*
HC39	HC	YHC64	> 64	> 8	0.125	0.36	0	no hyphae	*Candida albicans*
HC39	HC	YHC65	32	1	0.125	0.33	1	no hyphae	*Candida albicans*
HC41	HC	YHC66†	0.25	0.0156	0.125	0.31	2	Hyphae	*Candida parapsilosis*
HC41	HC	YHC67	1	0.0156	0.125	0.90	2	Hyphae	*Candida parapsilosis*
HC41	HC	YHC68	> 64	0.0156	0.125	NA	3	Hyphae	*Aspergillus glaucus*
HC44	HC	YHC69	0.125	0.125	0.125	0.27	1	Hyphae	*Candida albicans*
HC44	HC	YHC70	0.125	0.0156	0.125	0.13	2	no hyphae	*Candida albicans*
HC44	HC	YHC71	0.125	0.25	0.125	0.20	0	Hyphae	*Torulaspora delbrueckii*
HC44	HC	YHC72	0.5	0.25	0.125	0.18	0	no hyphae	*Candida albicans*
HC44	HC	YHC73	0.5	0.25	0.125	0.21	2	Hyphae	*Candida albicans*
HC44	HC	YHC74	0.125	0.0156	0.125	0.63	3	Pseudohyphae	*Candida parapsilosis*
HC44	HC	YHC75†‡	> 64	0.125	0.5	0.17	2	no hyphae	*Candida parapsilosis*
HC46	HC	YHC76	0.5	2	0.125	0.40	3	Hyphae	*Candida albicans*
HC46	HC	YHC77	0.5	2	0.125	0.76	2	Hyphae	*Candida albicans*
HC46	HC	YHC78	0.5	2	0.125	1.42	2	Hyphae	*Candida albicans*
HC46	HC	YHC79†	0.5	2	0.125	0.29	2	Hyphae	*Candida albicans*
HC47	HC	YHC80	0.125	0.0156	0.125	0.20	2	Hyphae	*Torulaspora delbrueckii*
HC47	HC	YHC81	0.25	1	0.125	0.29	2	Hyphae	*Candida albicans*
HC47	HC	YHC82	0.25	0.125	0.125	0.55	2	Hyphae	*Torulaspora delbrueckii*
HC47	HC	YHC83	0.25	1	0.125	0.19	3	Hyphae	*Candida albicans*
HC50	HC	YHC84	0.125	0.0156	0.125	2.54	2	no hyphae	*Candida intermedia*
HC50	HC	YHC85	0.125	0.0156	0.125	0.69	1	no hyphae	*Candida lusitaniae*
HC50	HC	YHC86	0.5	0.0156	0.125	0.86	1	Hyphae	*Candida albicans*
HC50	HC	YHC87	0.25	0.0156	0.125	0.44	1	Pseudohyphae	*Rhodotorula mucilaginosa*
HC50	HC	YHC88	0.25	0.0156	0.125	1.05	1	no hyphae	*Candida lusitaniae*
HC50	HC	YHC89	0.5	0.125	0.125	0.67	1	no hyphae	*Candida parapsilosis*

HC50	HC	YHC90	0.5	0.125	0.125	0.56	1	no hyphae	*Candida lusitaniae*
HC50	HC	YHC91†	0.5	0.0156	0.125	0.65	3	Pseudohyphae	*Candida parapsilosis*
HC50	HC	YHC92	NA	NA	NA	NA	NA	NA	*Yarrowia lipolytica*
HC52	HC	YHC93	0.25	0.0156	0.125	0.63	1	Hyphae	*Candida albicans*
HC52	HC	YHC94	0.25	0.0156	0.125	0.70	1	Hyphae	*Candida albicans*
HC52	HC	YHC95	0.125	0.0156	0.125	0.90	1	Hyphae	*Candida albicans*
HC52	HC	YHC96†	0.5	0.0156	0.125	0.54	3	Hyphae	*Candida albicans*
HC53	HC	YHC97	0.5	0.0156	0.125	1.07	2	Hyphae	*Candida albicans*
HC53	HC	YHC98	0.5	2	0.125	0.81	1	Hyphae	*Candida albicans*
RTT9	RTT	YRTT1	> 64	2	0.125	2.42	2	Pseudohyphae	*Candida parapsilosis*
RTT9	RTT	YRTT2	> 64	2	0.125	2.20	2	Pseudohyphae	*Candida parapsilosis*
RTT9	RTT	YRTT3	64	> 8	0.125	0.60	2	Pseudohyphae	*Candida parapsilosis*
RTT9	RTT	YRTT4†‡	1	0.0625	0.125	1.49	2	no hyphae	*Candida parapsilosis*
RTT9	RTT	YRTT5†‡	1	0.0156	0.125	0.81	0	no hyphae	*Candida albicans*
RTT9	RTT	YRTT6†	1	0.0156	0.125	0.35	0	no hyphae	*Candida albicans*
RTT10	RTT	YRTT7†‡	4	0.0156	0.125	0.77	1	no hyphae	*Candida albicans*
RTT35	RTT	YRTT8	> 64	2	0.125	2.35	2	Hyphae	*Candida tropicalis*
RTT35	RTT	YRTT9	8	0.0156	0.125	1.33	3	Hyphae	*Saccharomyces cerevisiae*
RTT35	RTT	YRTT10	> 64	8	0.125	1.79	3	Hyphae	*Candida tropicalis*
RTT73	RTT	YRTT11†‡	2	0.0156	0.125	1.77	2	no hyphae	*Candida albicans*
RTT73	RTT	YRTT12†	2	0.0156	0.125	2.04	2	no hyphae	*Candida parapsilosis*
RTT73	RTT	YRTT13	2	0.0156	0.125	2.10	2	no hyphae	*Candida parapsilosis*
RTT116	RTT	YRTT14†	16	> 8	0.125	1.04	1	no hyphae	*Candida parapsilosis*
RTT147	RTT	YRTT15	1	0.0156	0.125	0.68	2	Pseudohyphae	*Candida parapsilosis*
RTT147	RTT	YRTT16†‡	2	0.0156	0.125	0.93	2	no hyphae	*Candida parapsilosis*
RTT147	RTT	YRTT17	1	0.0156	0.125	1.46	2	Pseudohyphae	*Candida parapsilosis*
RTT147	RTT	YRTT18	1	0.0156	0.125	0.97	2	Pseudohyphae	*Candida parapsilosis*
RTT147	RTT	YRTT19	1	0.0156	0.125	0.82	2	Pseudohyphae	*Candida parapsilosis*
RTT163	RTT	YRTT20†	> 64	> 8	0.125	0.91	2	no hyphae	*Candida parapsilosis*
RTT199	RTT	YRTT21	> 64	> 8	0.125	0.94	3	Hyphae	*Trichosporon asteroides*
RTT199	RTT	YRTT22	> 64	> 8	0.125	0.74	2	no hyphae	*Candida pararugosa*

RTT199	RTT	YRTT23	> 64	> 8	0.125	0.83	0	no hyphae	*Candida glabrata*
RTT199	RTT	YRTT24†ǂ	0.125	0.0156	0.125	1.88	0	Pseudohyphae	*Candida parapsilosis*

†, isolates used for cytokine assays; ǂ, isolates analysed by flow cytometry; [§], MIC ranges: fluconazole, 0.125-64 μg/ml; itraconazole, 0.0156-8 μg/ml; 5-flucytosine, 0.125-64 μg/ml; [#], 0= non-invasive; 1= poor invasive; 2= invasive; 3= very invasive. *, measured by optical density at 570nm; NA, not applicable; nd, not detected.

Chapter 5

The importance of multiple sequence alignment in microbiome research: comparison of NAST algorithm implementations

This chapter will be shortly submitted for publication as an original research article:

Strati F, Albanese A, Donati C (2016). *The importance of multiple sequence alignment in microbiome research: comparison of NAST algorithm implementations.*

The bioinformatics pipelines widely used in microbiome research often, if not always, require the alignments of millions of sequences. However the MSA step represents a bottleneck in these pipelines since MSAs scale exponentially in memory and time in function of the number of sequences. In addition, "*de novo*" MSAs often lack the quality needed for phylogenetic analysis. The use of a reference MSA with fixed column positions makes the alignment process of candidate sequences readily scalable and allows the generation of very accurate alignments that incorporate the phylogenetically relevant information related to the secondary structure of the 16S rRNA gene, avoiding the excessive introduction of gaps in the alignment of non-conserved positions characteristic of *de novo* approaches [1]. The NAST algorithm is one of the most widely used reference-based MSA algorithm for metataxonomics analysis of 16S rRNA data [2]. Improved implementations of the NAST algorithm have been released but not comparatively tested. In this chapter we developed a computational testing procedure to compare the implementation of NAST included in the micca metagenomics data analysis pipeline with PyNAST [3] and the mothur's NAST [4], evaluating performances and alignment accuracy on synthetic and biological 16S rRNA sequence datasets. We showed that micca's NAST produced accurate MSAs and, by using a cross validation approach, we also observed that the alignments obtained with the micca's NAST retained highly conserved positions of the reference MSA better than the other implementations.

This work, in the framework of the research activities carried out within the "*Transdisciplinary program in Computational Biology*" gave me the possibility to improve my computational and programming skills learning the use of the Python programming language and addressing an important, but underrated issue in microbiome research: the importance to generate accurate MSAs for the correct interpretation of metagenomics data in microbial ecology. As author of this work, I participated to the design of the experiments, analysis of the data, figures generation and drafting of the manuscript.

References

1. Schloss PD: **A high-throughput DNA sequence aligner for microbial ecology studies**. *PloS one* 2009, **4**(12):e8230.
2. DeSantis T, Hugenholtz P, Keller K, Brodie E, Larsen N, Piceno Y, Phan R, Andersen GL: **NAST: a multiple sequence alignment server for comparative analysis of 16S rRNA genes**. *Nucleic acids research* 2006, **34**(suppl 2):W394-W399.
3. Caporaso JG, Bittinger K, Bushman FD, DeSantis TZ, Andersen GL, Knight R: **PyNAST: a flexible tool for aligning sequences to a template alignment**. *Bioinformatics* 2010, **26**(2):266-267.
4. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ: **Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities**. *Applied and environmental microbiology* 2009, **75**(23):7537-7541.

The importance of multiple sequence alignment in microbiome research: comparison of NAST algorithm implementations

Francesco Strati[1,2], Davide Albanese[1], Claudio Donati[1]

[1]Computational Biology Research Unit, Research and Innovation Centre, Fondazione Edmund Mach, Via E. Mach 1, 38010 San Michele all' Adige, Italy
[2]Centre for Integrative Biology, University of Trento, Via Sommarive 9, 38123 Trento, Italy

Abstract

In the present study we compared three different implementations of the NAST sequence aligner, *i.e.* PyNAST, the mothur's NAST and the recently introduced NAST algorithm in the micca pipeline, evaluating performances and alignment accuracy on synthetic and biological 16S rRNA sequence datasets. Although the accuracy and performances of MICCA, MOTHUR and PYNAST were quite similar, MICCA produced more robust MSAs compared to the other methods tested both in term of identity and retention of highly conserved positions in the alignment. Indeed MICCA, together with PYNAST, provided better results in the alignment of highly conserved positions compared to MOTHUR. Speed and memory benchmark analysis showed that MOTHUR is faster than MICCA and PYNAST although MICCA required minor memory usage and scaled better than MOTHUR in function of templates size. PYNAST, on the contrary, showed high speed at the expenses of larger memory requirements.

Introduction

The 16S rRNA gene sequence, of about 1.5 kb, contains highly conserved, variable and hypervariable regions that can be used as molecular markers to infer the phylogenetic structure of a microbial community and to distinguish bacteria in function of their evolutionary distance [1]. Since sequence similarity provides the only unambiguous definition of bacterial taxa, 16S rRNA sequences having at least 97% identity are grouped into operational taxonomic units (OTU) as a proxy for species [2, 3]. Multiple sequence alignment (MSA) is crucial step in order to calculate phylogenetic and pairwise genetic distances to be used for binning sequences into OTUs. Popular aligner tools such as ClustalW [4], MUSCLE [5] and T-Coffee [6] have been used to generate generic and multiple alignments, nevertheless these methods have two great limitations: i) they scale exponentially in time and memory in function of the length and number of sequences, ii) they do not incorporate information regarding

the secondary structure of the 16S rRNA gene which is important to retain conserved positional homologies in the MSA [7, 8].

Using curated reference alignments reduces the computational complexity of the MSA, and increases the alignment quality with respect to a *de novo* alignments avoiding the excessive introduction of gaps within the informative, non-conserved positions of the alignment. Currently RDP [9], SILVA [10] and Greengenes [11] are the 16S rRNA sequence databases most widely used in microbiome research. All of them provide curated MSAs of 16S rRNA sequences and their own aligners. Infernal, the RDP aligner [12], uses a model-based approach in which Hidden Markov Models are applied to 16S rRNA secondary structures to generate models to be applied to unaligned sequences. SINA [13], the aligner provided within the SILVA database, combines *k*-mer distance search with Partial Order Alignment [14] allowing high alignment accuracy and performance.

The NAST method used by Greengenes in its original implementation [15] uses BLAST [16] to obtain a pairwise alignment between the candidate sequence and the best match in the reference MSA, and subsequently introduces gaps in the candidate sequence to obtain a final alignment of the same length of the reference MSA. PyNAST [17] and the aligner within mothur [18] implement improved versions of the NAST algorithm. PyNAST relies on UCLUST [19] which uses a heuristic seed and extend aligner, while mothur uses an implementation of the *k*-mer strategy and Needleman-Wunsch algorithm. In addition to these NAST implementations, the micca pipeline (http://www.micca.org) [20] provides a new implementation of NAST based on VSEARCH [21]. VSEARCH is an open source tool, which uses an optimal global aligner based on a full dynamic programming Needleman-Wunsch. In the present study we compared the micca implementation of NAST with PyNAST and the mothur's NAST evaluating performances and alignment accuracy on synthetic and biological 16S rRNA sequence datasets.

Methods

Synthetic datasets

The synthetic datasets have been generated by random picking 100 near-full length 16S query sequences and 1000 reference sequences from the Greengenes and SILVA MSA templates clustered at different thresholds of similarity (for Greengenes, from 79% to 99%, for SILVA, from 90% to 99%). The gaps in the query sequences have been removed. For each reference set, 10 replicates have been created. The query replicate have been aligned against each corresponding template replicate by using micca's NAST, PyNAST and mothur's NAST.

For each replicate we assessed: the percentage of sequence aligned; the percentage of trimmed sequences (the NAST algorithm trims the candidate sequences to that which is bound by the

beginning and end points of the alignment span [15]); the average Hamming distance between the original query alignment versus the NAST alignment; the difference in length (average) between the original sequences and the (trimmed) NAST aligned sequences.

Biological Datasets

In order to evaluate the NAST performances on typical experimental settings nine datasets from the Qiita open-source microbial study management platform (https://qiita.ucsd.edu/), the red colobus dataset [22] (Table 1) and 11 samples of the V1-V3, V3-V5 and V6-V9 hypervariable regions of the bacterial 16S rRNA gene from the Human Microbiome Project (HMP) dataset [23] have been selected and aligned on the reference MSAs "Greengenes Core Set", "Greengenes 85% OTUs (v13_5)" and "PyNAST-curated Greengenes 85% OTUs". For the HMP datasets the sequences have been clustered before the NAST alignment by using a *de novo* greedy clustering method from the micca pipeline (`micca otu -i input.fasta --method denovo_greedy --id 0.97 -c`) while all the other datasets used were already clustered.

Table 1: Biological datasets used in this study

Dataset	Samples	16S region	Seq. Platform	Qiita ID*	Ref (PMID)	EBI
Brazilian Antarctic Station*	43	V4	Illumina	1033	NA	ERP016586
Cannabis Soil*	21	V4	Illumina	1001	NA	ERP016540
Columbia Mice*	31	V1-V3	454 pyroseq	107	21593810	NA
Great Lake Microbiome*	49	V4	Illumina	1041	NA	ERP016492
Green Hospital Air Sloan*	13	V1-V3	454 pyroseq	1345	22278670	NA
Jansson Twins IBD*	118	V3-V5	454 pyroseq	1070	20816835	NA
Koren Oral Gut Plaque*	73	V1-V3	454 pyroseq	349	20937873	NA
McGuire Nicaragua Soil*	63	V4	Illumina	1715	NA	NA
NICU contamination data*	30	V1-V3	454 pyroseq	386	NA	NA
Red Colobus	31	V1-V3	454 pyroseq	NA	26445280	PRJEB8977

*from https://qiita.ucsd.edu/

These biological datasets have been chosen because they represent (both in terms of sequencing platform used, variable regions of the 16S rRNA gene amplified and environmental origin of the samples) most of the variables affecting the results of a metataxonomics analysis.

For each replicate we assessed: the percentage of pairwise identities between the NAST aligned sequences; the percentage of sequences aligned; the percentage of trimmed sequences; the difference in length (average) between the original sequences and the (trimmed) NAST aligned sequences; the Shannon's entropy *per* position (only columns with coverage >= 75% in the NAST alignment were considered) for each datasets and the template MSA used (Greengenes "Core set").

Micca's NAST implementation

Micca's NAST is implemented in the `micca msa` command (option `--method nast`) and in principle it applies the same protocol of NAST in its original implementation [15]. However the micca's NAST implementation identifies the most similar sequence in the template alignment for each candidate sequence by using VSEARCH [21] (v1.9.5 in micca 1.5), with the options:

```
--usearch_global --maxaccepts 8 --maxrejects 32
```

NAST parameters

All the 16S rRNA sequence datasets have been aligned by using the NAST implementations of micca (v1.5), PyNAST (v1.2.1) and mothur (v1.37.3). The following parameters have been used:

- ```
  micca msa -m nast -i query.fasta -o msa.fasta \
  --nast-threads 1 --nast-template template.fasta \
  --nast-hits hits.txt --nast-nofilter \
  --nast-notaligned notaligned.fasta
  ```

- ```
  pynast -i query.fasta -a msa.fasta -t template.fasta \
  -l 1 –f notaligned.fasta -g hits.txt
  ```

- ```
  mothur "#align.seqs(candidate=query.fasta, \
  template=template.fasta, flip=t, processors=1)"
  ```

Speed and Memory benchmark

Ten replicate synthetic datasets have been generated by randomly picking 1000 near-full length 16S rRNA sequences (without gaps) from the Greengenes 99% OTUs and NAST aligned by using the Greengenes reference databases, from 79% to 99% OTUs, which are characterized by an increasing number of sequences. Speed and memory benchmarks have been performed by using the python package Memory Profiler (v0.41) in order to evaluate speed and memory requirements of the different NAST aligners when using an equal number of candidate sequences in function of the increasing size of the templates and a single thread. All tests were run on a Linux\Ubuntu 13.04, 64 bit machine equipped with Intel® Core™ i7-3770 CPU @ 3.40GHz × 8 and 16 GB RAM.
  ```

Results

Evaluation of NAST implementations comparing the accuracy of MSAs

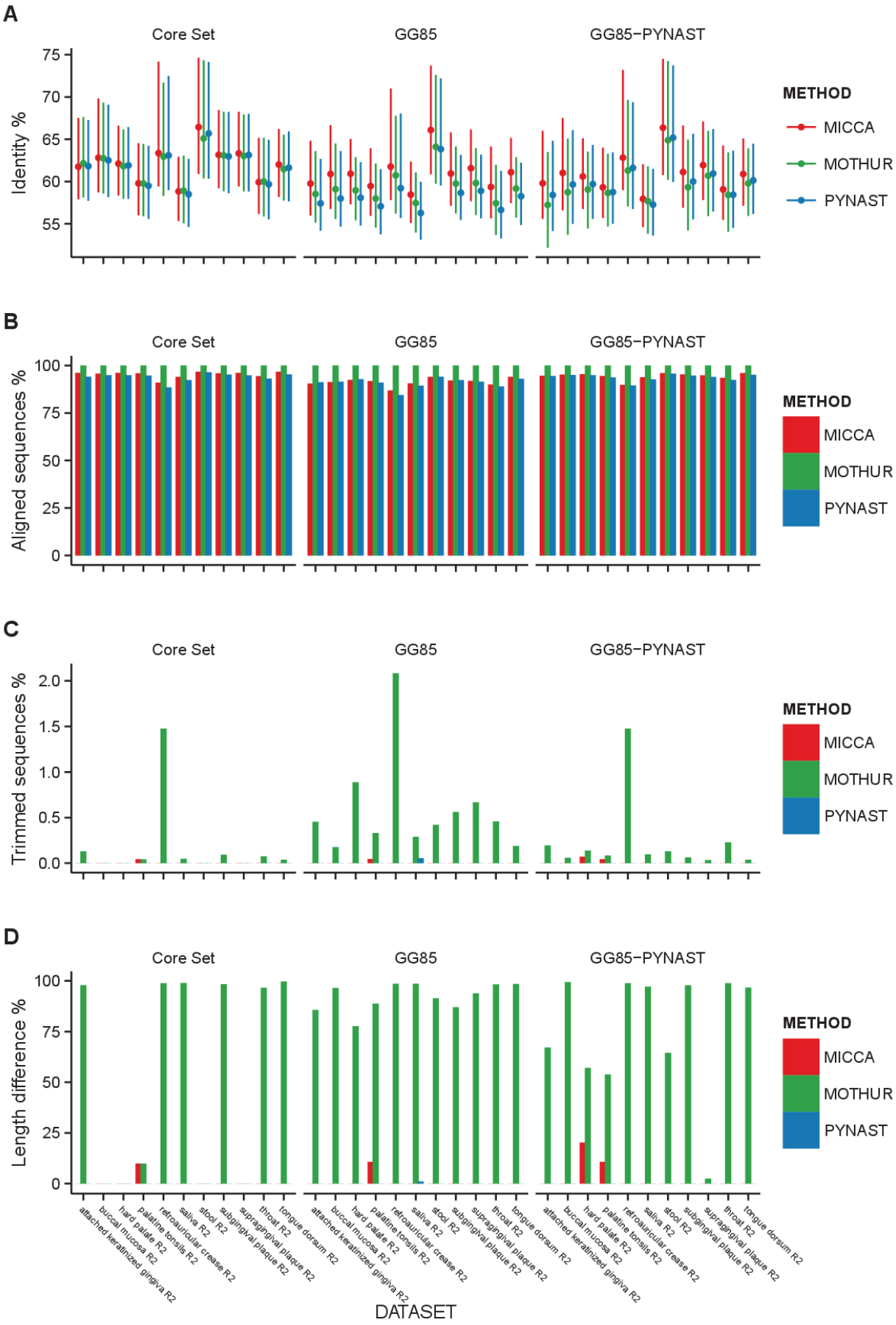
We first tested the alignment accuracy of the three used implementations of NAST, *i.e.* micca's NAST, PyNAST and mothur's NAST (hereinafter termed MICCA, PYNAST and MOTHUR) on the synthetic datasets comparing the NAST alignments with the original query alignments. The results reported in the supplementary figure 1 show that the average Hamming distances among the query alignments versus the NAST alignments did not vary substantially among MICCA, PYNAST and MOTHUR.

However, the evaluation of MICCA, PYNAST and MOTHUR alignments of the HMP datasets [23] and of the ten biological datasets (Table 1) revealed interesting clues. For these analysis we used as reference alignments the Greengenes "Core Set", because composed of *circa* 10.000 manually aligned non-chimeric sequences representative of a wide variety of bacterial and archaeal taxa, and the Greengenes 85% OTUs (v13_5) template because it is the default reference alignment used in the QIIME pipeline [24]. Nevertheless, it has been reported a bug in the PYNAST algorithm in which the sequences generated within the V1-V3 variable region of the 16S rRNA gene fail to align when using PYNAST with the default Greengenes 85% OTUs (v13_5) reference alignment. The bug has been fixed by the PYNAST developers generating a new template from the original Greengenes 85% OTUs (v13_5), here called "PyNAST-Greengenes 85% OTUs" that we included in our analysis for completeness.

We observed that PYNAST failed to align the vast majority of sequences from the HMP R1 dataset (V1-V3 16S region sequences) (Supplementary Fig. 2) with the default Greengenes 85% OTUs (v13_5). When aligning the HMP R2 datasets (V3-V5 16S region sequences) (Figure 1) MICCA showed slightly improved performances compared to MOTHUR and PYNAST while the latter aligned better the sequences from the HMP R3 datasets (V6-V9 16S region sequences) (Supplementary Fig. 3). Anyway, all the three NAST implementations had similar and better performances when using the "Core set" reference MSA compared to the less curated Greengenes 85% OTUs (v13_5) template, remarking that the use of well-curated templates can help in providing more accurate results.

Figure 1: Comparison of the alignment performances of MICCA, MOTHUR and PYNAST on the HMP R2 datasets. **A)** average Hamming distance between the query alignments versus the NAST alignments; **B)** sequences aligned (in %); **C)** trimmed sequences (in %). The NAST algorithm trims the candidate sequences to that which is bound by the beginning and end points of the alignment span; **D)** difference in length (average in %) between the candidate sequences and the (trimmed) NAST aligned sequences. (See figure 1 on next page).

Figure 1


A
Core Set
GG85
GG85−PYNAST
Identity %
75
70
65
60
55
METHOD
MICCA
MOTHUR
PYNAST
B
Aligned sequences %
100
75
50
25
0
C
Trimmed sequences %
2.0
1.5
1.0
0.5
0.0
D
Length difference %
attached keratinized gingiva R2
buccal mucosa R2
hard palate R2
palatine tonsils R2
retroauricular crease R2
saliva R2
stool R2
subgingival plaque R2
supragingival plaque R2
throat R2
tongue dorsum R2
DATASET

The analysis of the biological datasets (Figure 2) showed that MICCA and PYNAST alignments were slightly better than MOTHUR alignments in terms of pairwise identities (Figure 2A). For two datasets (*i.e.* the *Green Hospital Air Sloan* and the *Red Colobus*) PYNAST failed to align the vast majority of the sequences. Furthermore we observed that MOTHUR frequently trims the candidate sequences more than MICCA and PYNAST (Figure 1C and 2C) with a difference in length between the original sequences and the (trimmed) aligned sequences exceeding the 75% (Figure 1D and 2D).

This is due to the fact that MOTHUR always retains all the sequences of each dataset, even those low-quality sequences that both MICCA and PYNAST usually discard because not aligning over the reference MSAs.

Nevertheless, the use of pairwise identities as a measure of alignments accuracy is only as a coarse way to evaluate the goodness of a MSA, since the pairwise identity depends on the percentage of sequences retained in the alignment or how much the sequences were trimmed. Therefore we also evaluated the quality and the robustness of the NAST alignments by calculating how frequently a highly conserved position was retained in the alignment compared to the template. For this purpose we calculated the Shannon entropy of each column of the MSAs obtained with MICCA, MOTHUR and PYNAST and we compared them only for those columns with coverage ≥ 75%. As template we used the "Core Set" since it provided the best and homogeneous results with all the three NAST implementations. The alignments of the HMP datasets (Supplementary Fig. 4, 5 and 6) revealed no evident differences among MICCA, MOTHUR and PYNAST. However when using the biological datasets, which reflect better the empirical variability encountered in different experimental settings, MOTHUR was less able to align some highly conserved positions characterized by very low entropy. In the *Great Lake Microbiome* dataset, MOTHUR aligned approximately 10% less highly conserved positions compared to MICCA and PYNAST (Figure 3). These results were further supported by the observation that the *per* position entropy of the MOTHUR alignments was frequently higher than MICCA and PYNAST alignments when compared to the *per* position entropy of the "Core Set" reference MSA (Supplementary Fig. 7).

Figure 2: Comparison of the alignment performances of MICCA, MOTHUR and PYNAST on the real datasets. **A)** average Hamming distance between the query alignments versus the NAST alignments; **B)** sequences aligned (in %); **C)** trimmed sequences (in %). The NAST algorithm trims the candidate sequences to that which is bound by the beginning and end points of the alignment span; **D)** difference in length (average in %) between the candidate sequences and the (trimmed) NAST aligned sequences. (See figure 2 on next page).

Figure 2

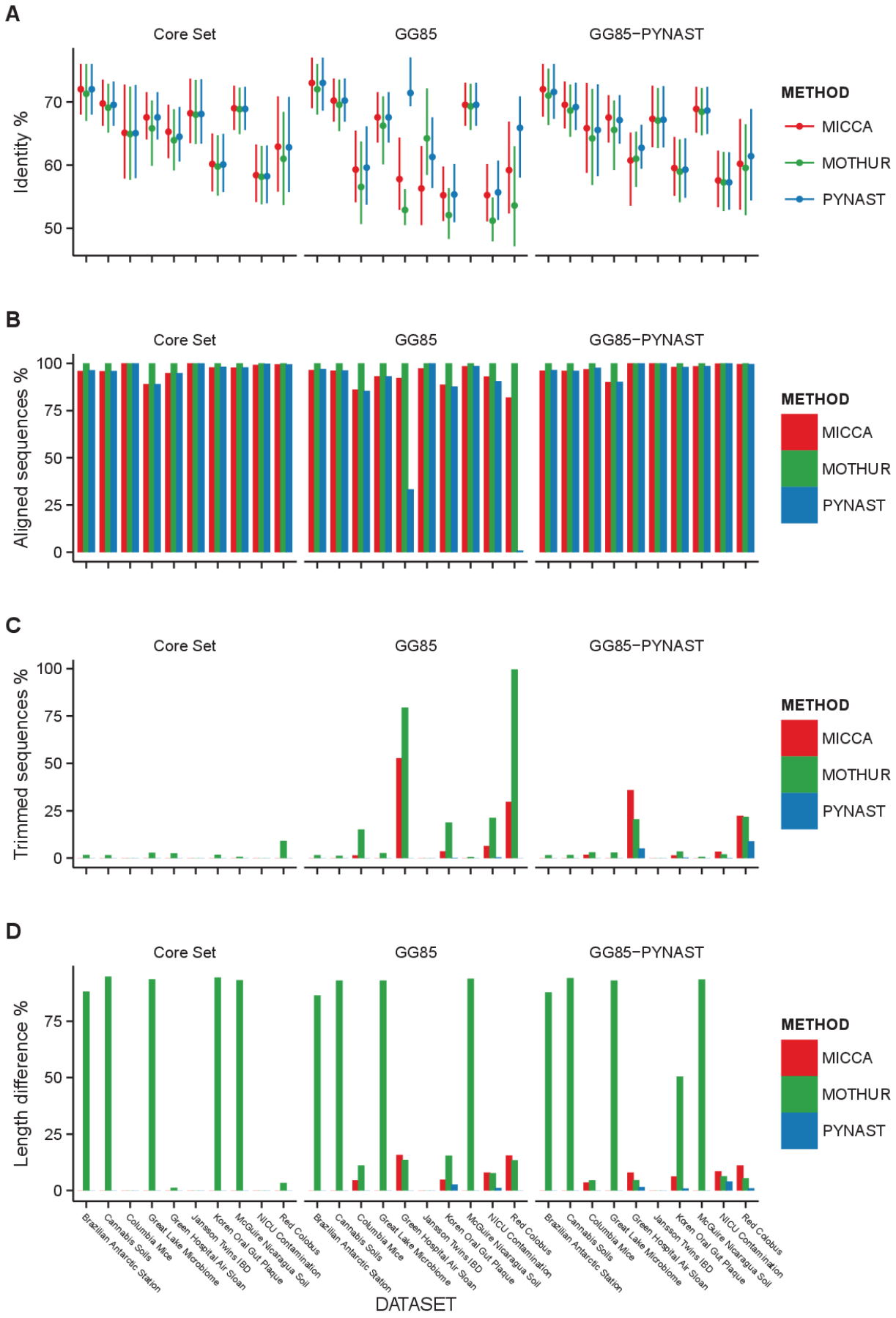

A
Core Set
GG85
GG85−PYNAST
Identity %
70
60
50
METHOD
MICCA
MOTHUR
PYNAST
B
Core Set
GG85
GG85−PYNAST
Aligned sequences %
100
75
50
25
0
METHOD
MICCA
MOTHUR
PYNAST
C
Core Set
GG85
GG85−PYNAST
Trimmed sequences %
100
75
50
25
0
METHOD
MICCA
MOTHUR
PYNAST
D
Core Set
GG85
GG85−PYNAST
Length difference %
75
50
25
0
Brazilian Antarctic Station
Cannabis Soils
Columbia Mice
Great Lake Microbiome
Green Hospital Air Sloan
Jansson Twins IBD
Koren Oral Gut Plaque
McGuire Nicaragua Soil
NICU Contamination
Red Colobus
DATASET
METHOD
MICCA
MOTHUR
PYNAST

Figure 3: Histogram-based estimation of the differential entropy of the MSAs for the real datasets. The entropy has been calculated only for those columns with coverage ≥ 75% in the NAST alignment. Only the columns with entropy ≤ 0.5 have been considered. (See figure 3 on next page).

Speed and Memory benchmark

We then benchmarked speed and memory requirements for MICCA, MOTHUR and PYNAST by aligning 1000 candidate sequences with different reference databases of increasing sizes. MOTHUR aligned the candidate sequences faster than MICCA and PYNAST although MICCA required minor memory usage and scaled better than MOTHUR in function of the increasing size of the reference alignments. PYNAST on the contrary showed high speed and memory requirements. (Figure 4).

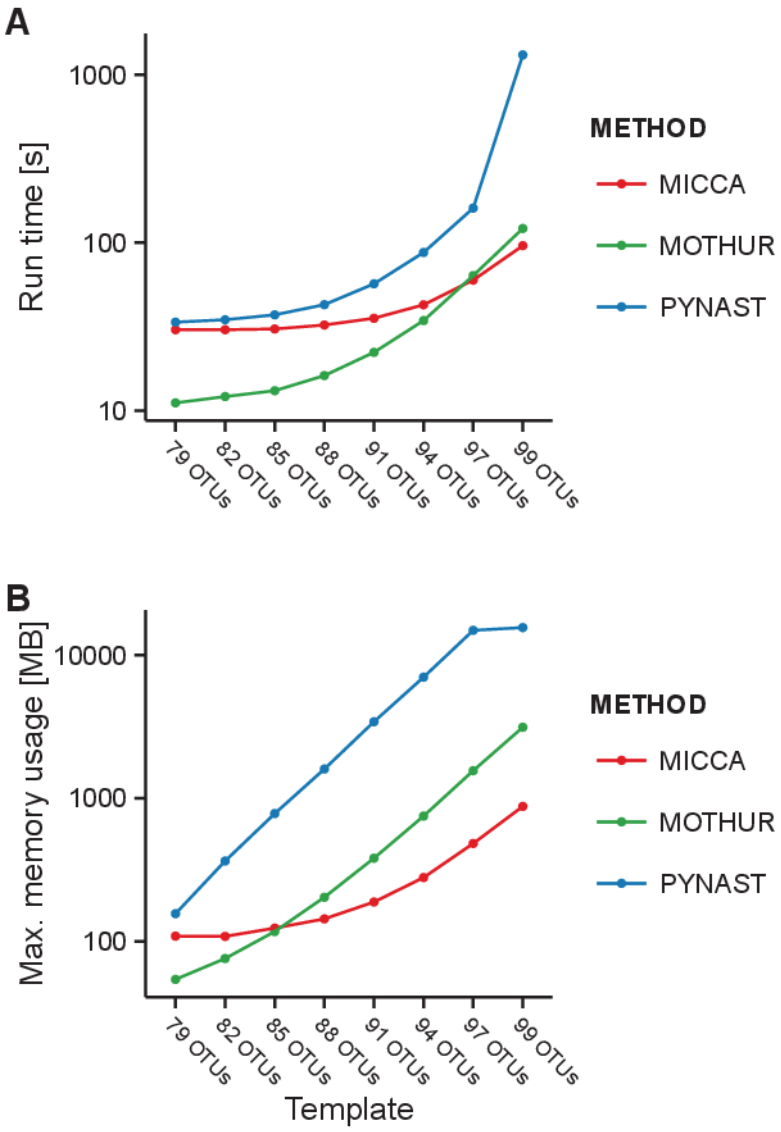
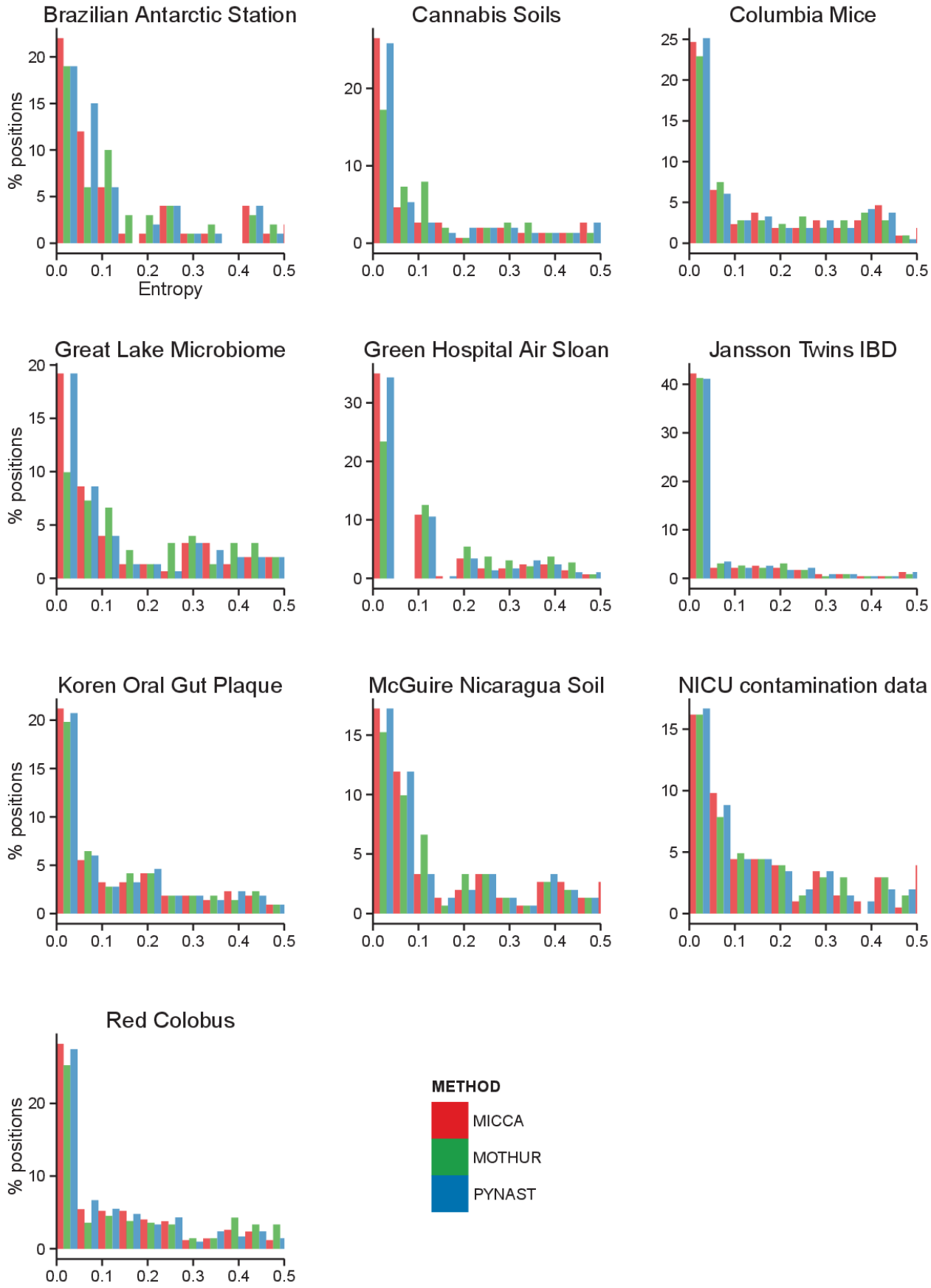


Figure 4: **A)** Run time of MICCA, MOTHUR and PYNAST; **B)** Max memory usage of MICCA, MOTHUR and PYNAST. The analysis have been performed by using a single thread.

Figure 3


Brazilian Antarctic Station
Cannabis Soils
Columbia Mice
Great Lake Microbiome
Green Hospital Air Sloan
Jansson Twins IBD
Koren Oral Gut Plaque
McGuire Nicaragua Soil
NICU contamination data
Red Colobus
% positions
Entropy
METHOD
MICCA
MOTHUR
PYNAST

Conclusions

The NAST algorithm aligns a candidate sequence and the best match in the template MSA and, by introducing gaps, finally guarantees an alignment of the same length of the template alignment. The advantage in using NAST is therefore its ability to align conserved positions avoiding the excessive introductions of gap extensions in the informative, non-conserved regions. By now, three implementations of NAST have been developed (*i.e.* MICCA, MOTHUR and PYNAST) but no one tested and compared the performances and the accuracy of the MSAs generated with these implementations. Here, we showed that micca (http://www.micca.org), a new bioinformatics pipeline for the processing of amplicon sequencing data improves the quality of NAST alignments by using a fast, multithread and memory efficient reimplementation of the NAST algorithm which relies completely on open source applications and therefore increasing its portability and accessibility.

Acknowledgements

We thank Unifarm S.p.A. for financial support to FS.

Author contribution

DA and CD conceived the study. FS, and DA analysed the data. FS wrote the manuscript.

Conflict of interest

Authors declare no conflict of interests.

References

1. Olsen GJ, Lane DJ, Giovannoni SJ, Pace NR, Stahl DA: **Microbial ecology and evolution: a ribosomal RNA approach**. *Annual review of microbiology* 1986, **40**:337-365.
2. Wayne L, Brenner D, Colwell R, Grimont P, Kandler O, Krichevsky M, Moore L, Moore W, Murray R, Stackebrandt E: **Report of the ad hoc committee on reconciliation of approaches to bacterial systematics**. *International Journal of Systematic and Evolutionary Microbiology* 1987, **37**(4):463-464.
3. Stackebrandt E, Goebel B: **Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology**. *International Journal of Systematic and Evolutionary Microbiology* 1994, **44**(4):846-849.
4. Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD: **Multiple sequence alignment with the Clustal series of programs**. *Nucleic acids research* 2003, **31**(13):3497-3500.
5. Edgar RC: **MUSCLE: multiple sequence alignment with high accuracy and high throughput**. *Nucleic acids research* 2004, **32**(5):1792-1797.
6. Notredame C, Higgins DG, Heringa J: **T-Coffee: A novel method for fast and accurate multiple sequence alignment**. *Journal of molecular biology* 2000, **302**(1):205-217.
7. Schloss PD: **A high-throughput DNA sequence aligner for microbial ecology studies**. *PloS one* 2009, **4**(12):e8230.

8. Gardner PP, Wilm A, Washietl S: **A benchmark of multiple sequence alignment programs upon structural RNAs**. *Nucleic acids research* 2005, **33**(8):2433-2439.
9. Wang Q, Garrity GM, Tiedje JM, Cole JR: **Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy**. *Applied and environmental microbiology* 2007, **73**(16):5261-5267.
10. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glockner FO: **The SILVA ribosomal RNA gene database project: improved data processing and web-based tools**. *Nucleic acids research* 2013, **41**(Database issue):D590-596.
11. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL: **Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB**. *Applied and environmental microbiology* 2006, **72**(7):5069-5072.
12. Nawrocki EP, Kolbe DL, Eddy SR: **Infernal 1.0: inference of RNA alignments**. *Bioinformatics* 2009, **25**(10):1335-1337.
13. Pruesse E, Peplies J, Glöckner FO: **SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes**. *Bioinformatics* 2012, **28**(14):1823-1829.
14. Lee C, Grasso C, Sharlow MF: **Multiple sequence alignment using partial order graphs**. *Bioinformatics* 2002, **18**(3):452-464.
15. DeSantis T, Hugenholtz P, Keller K, Brodie E, Larsen N, Piceno Y, Phan R, Andersen GL: **NAST: a multiple sequence alignment server for comparative analysis of 16S rRNA genes**. *Nucleic acids research* 2006, **34**(suppl 2):W394-W399.
16. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs**. *Nucleic acids research* 1997, **25**(17):3389-3402.
17. Caporaso JG, Bittinger K, Bushman FD, DeSantis TZ, Andersen GL, Knight R: **PyNAST: a flexible tool for aligning sequences to a template alignment**. *Bioinformatics* 2010, **26**(2):266-267.
18. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ: **Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities**. *Applied and environmental microbiology* 2009, **75**(23):7537-7541.
19. Edgar RC: **Search and clustering orders of magnitude faster than BLAST**. *Bioinformatics* 2010, **26**(19):2460-2461.
20. Albanese D, Fontana P, De Filippo C, Cavalieri D, Donati C: **MICCA: a complete and accurate software for taxonomic profiling of metagenomic data**. *Scientific reports* 2015, **5**:9743.
21. Rognes T, Flouri T, Nichols B, Quince C, Mahe F: **VSEARCH: a versatile open source tool for metagenomics**. *PeerJ* 2016, **4**:e2584.
22. Barelli C, Albanese D, Donati C, Pindo M, Dallago C, Rovero F, Cavalieri D, Tuohy KM, Hauffe HC, De Filippo C: **Habitat fragmentation is associated to gut microbiota diversity of an endangered primate: implications for conservation**. *Scientific reports* 2015, **5**:14862.
23. Human Microbiome Project C: **Structure, function and diversity of the healthy human microbiome**. *Nature* 2012, **486**(7402):207-214.
24. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI: **QIIME allows analysis of high-throughput community sequencing data**. *Nature methods* 2010, **7**(5):335-336.

Supplementary material chapter 5

Supplementary Figure 1: Comparison of the alignment performances of MICCA, MOTHUR and PYNAST on synthetic datasets. **A)** average Hamming distance between the query alignments versus the NAST alignments; **B)** sequences aligned (in %); **C)** trimmed sequences (in %). The NAST algorithm trims the candidate sequences to that which is bound by the beginning and end points of the alignment span; **D)** difference in length (average in %) between the candidate sequences and the (trimmed) NAST aligned sequences.

Supplementary Figure 2: Comparison of the alignment performances of MICCA, MOTHUR and PYNAST on the HMP R1 datasets. **A)** average Hamming distance between the query alignments versus the NAST alignments; **B)** sequences aligned (in %); **C)** trimmed sequences (in %). The NAST algorithm trims the candidate sequences to that which is bound by the beginning and end points of the alignment span; **D)** difference in length (average in %) between the candidate sequences and the (trimmed) NAST aligned sequences.

Supplementary Figure 3: Comparison of the alignment performances of MICCA, MOTHUR and PYNAST on the HMP R3 datasets. **A)** average Hamming distance between the query alignments versus the NAST alignments; **B)** sequences aligned (in %); **C)** trimmed sequences (in %). The NAST algorithm trims the candidate sequences to that which is bound by the beginning and end points of the alignment span; **D)** difference in length (average in %) between the candidate sequences and the (trimmed) NAST aligned sequences.

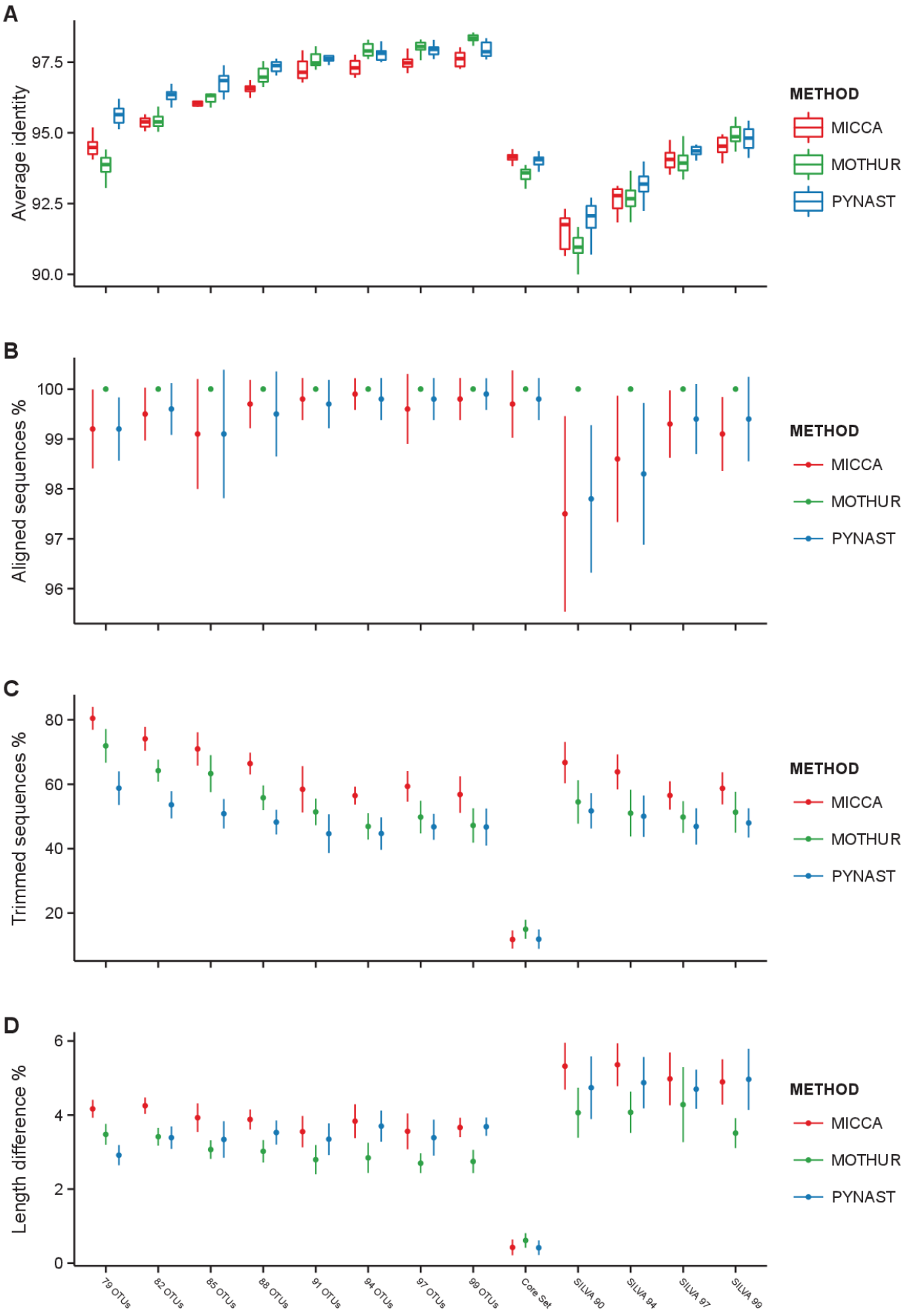
Supplementary Figure 4: Histogram -based estimation of the differential entropy of the MSAs for the HMP R1 datasets. The entropy has been calculated only for those columns with coverage $\geq 75\%$ in the NAST alignment. Only the columns with entropy ≤ 0.5 have been considered.

Supplementary Figure 5: Histogram -based estimation of the differential entropy of the MSAs for the HMP R2 datasets. The entropy has been calculated only for those columns with coverage $\geq 75\%$ in the NAST alignment. Only the columns with entropy ≤ 0.5 have been considered.

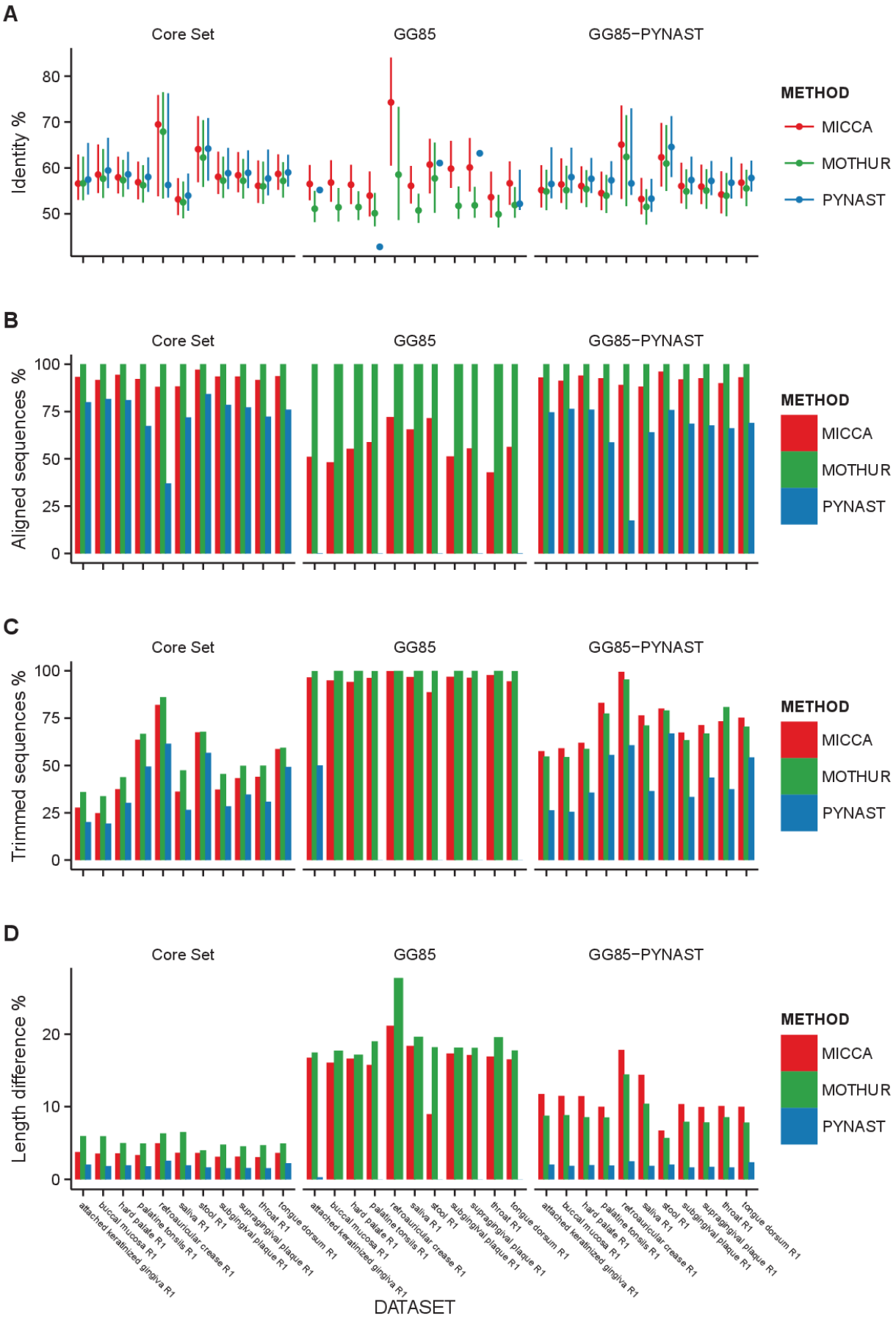
Supplementary Figure 6: Histogram -based estimation of the differential entropy of the MSAs for the HMP R3 datasets. The entropy has been calculated only for those columns with coverage ≥ 75% in the NAST alignment. Only the columns with entropy ≤ 0.5 have been considered.

Supplementary Figure 7: Entropy of the "Core Set" reference MSA vs entropy of the MICCA, MOTHUR and PYNAST MSAs by using the real datasets. Only the columns with entropy ≤ 0.3 have been considered.

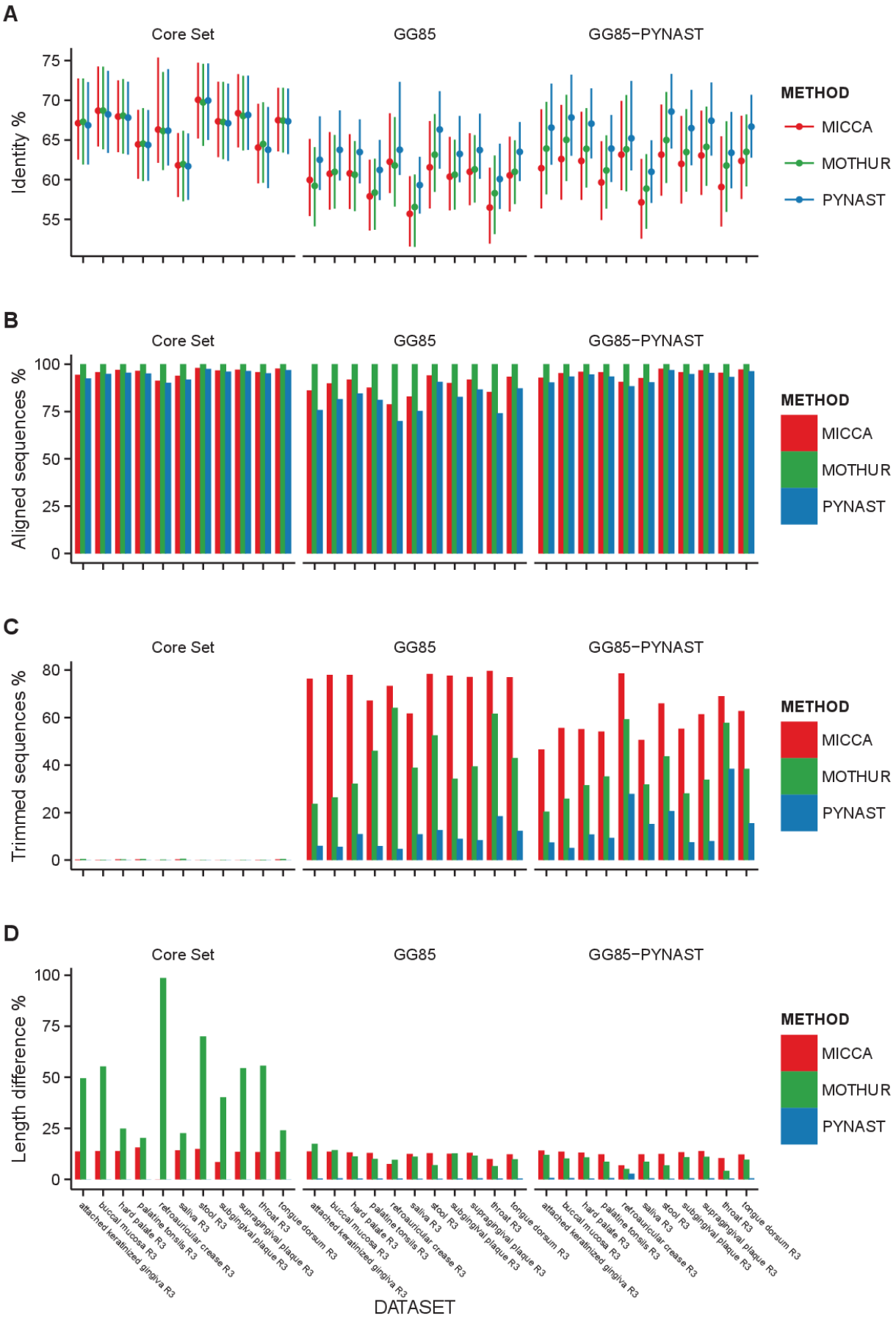
Supplementary Figure 1



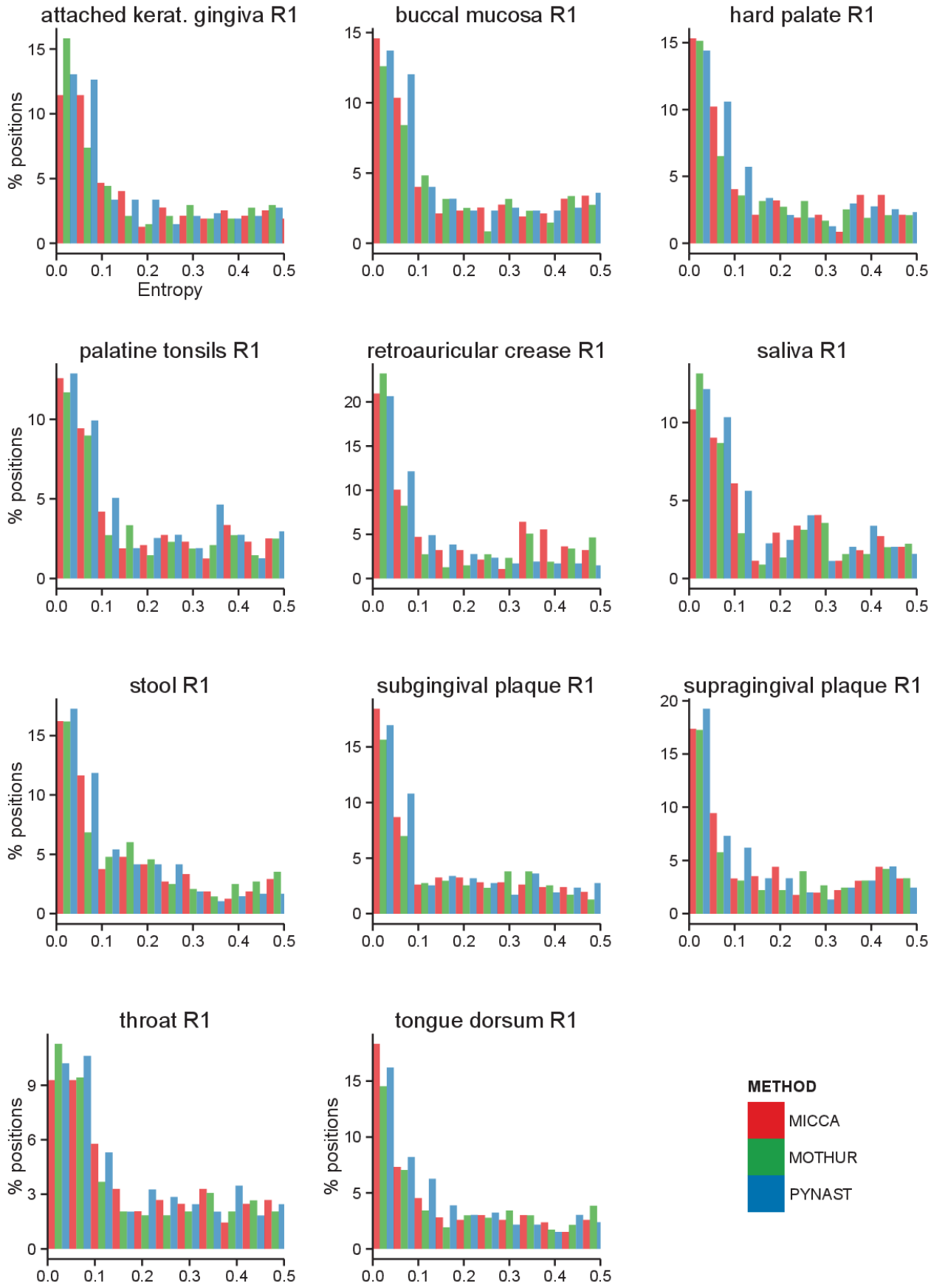
Supplementary Figure 2



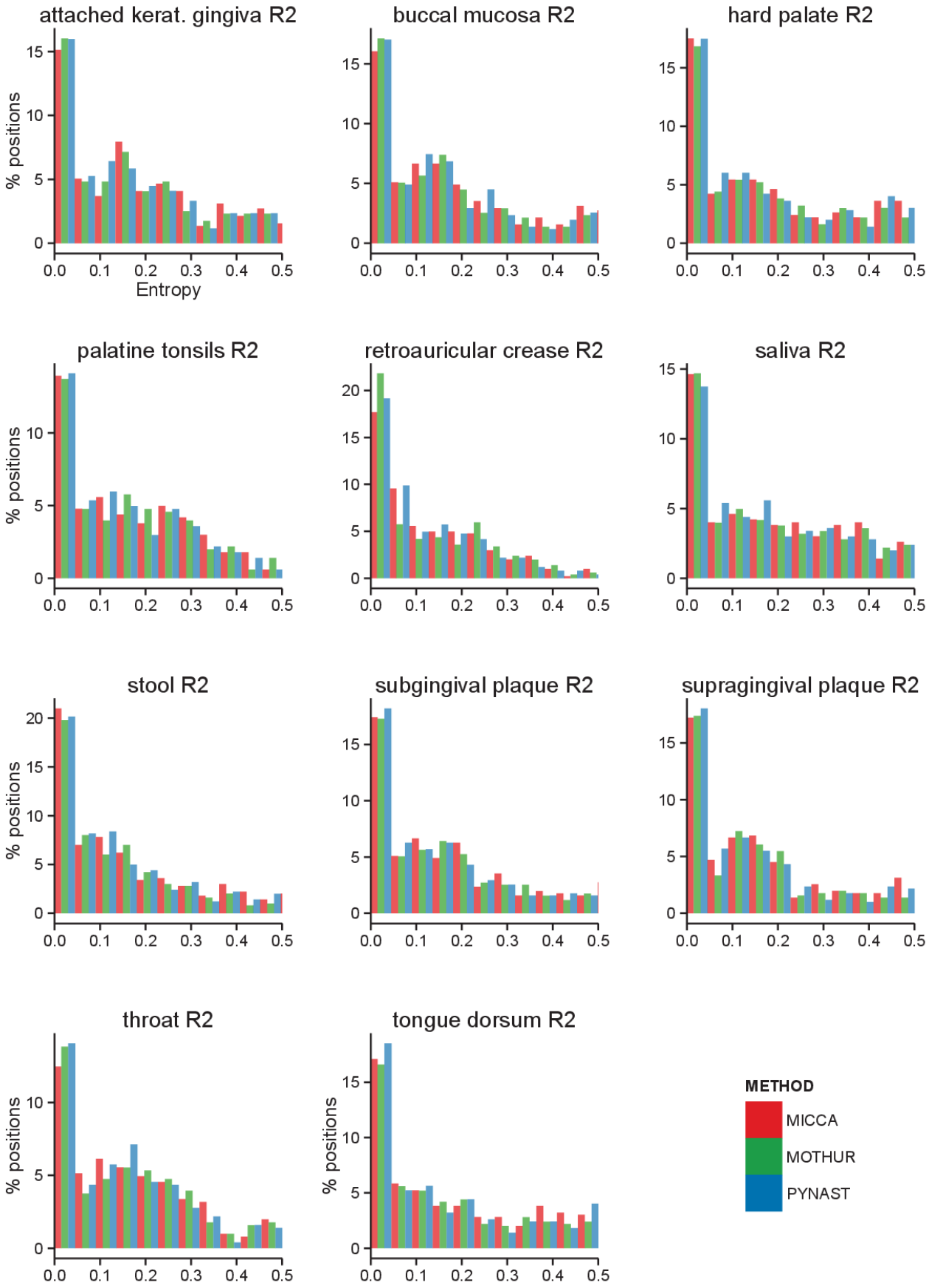
Supplementary Figure 3



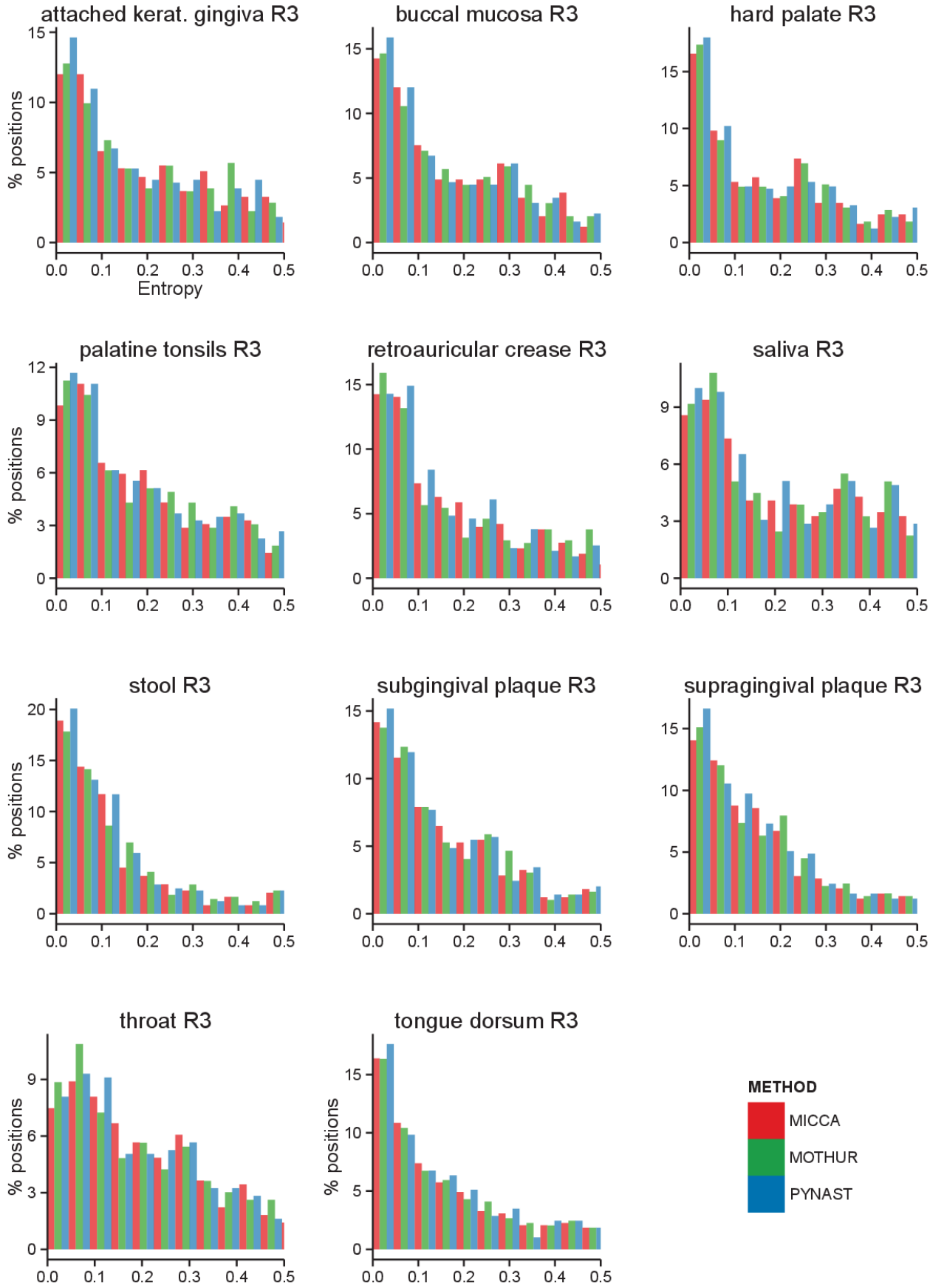
Supplementary Figure 4



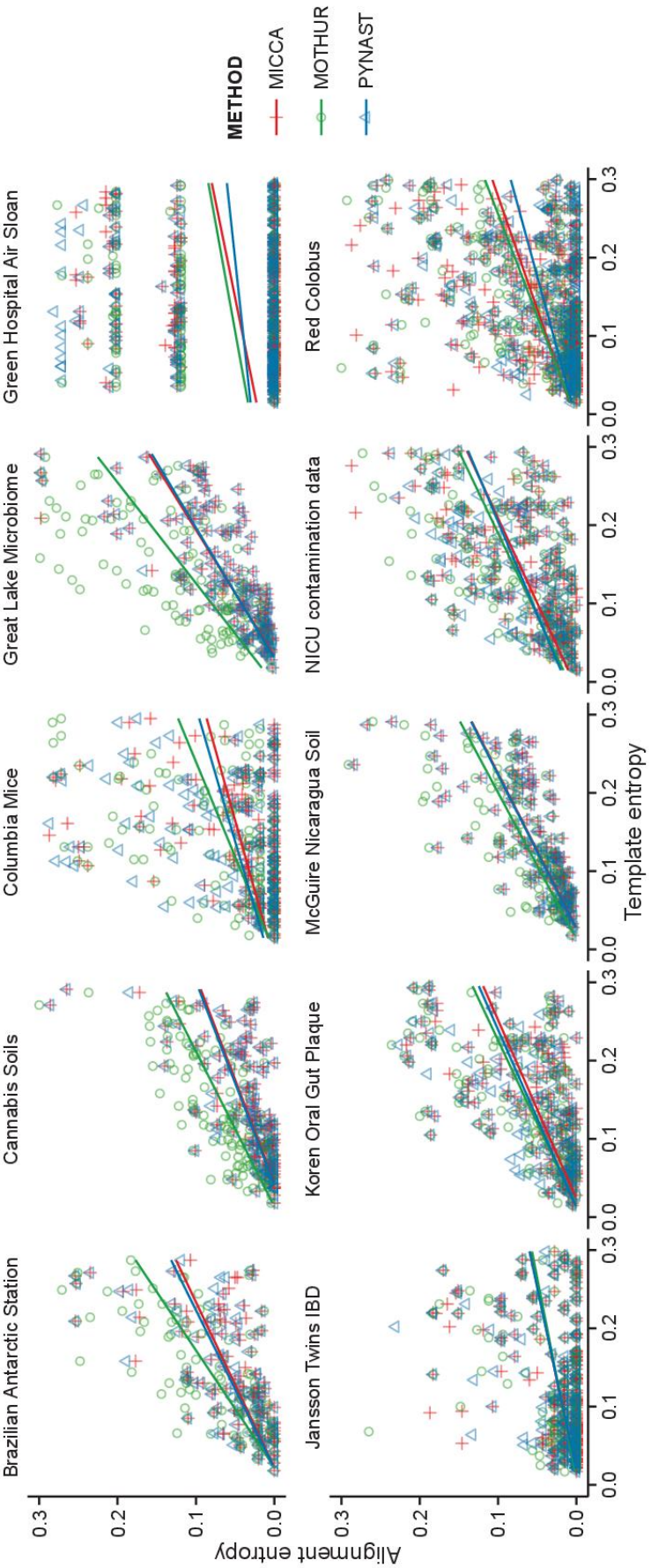
Supplementary Figure 5



Supplementary Figure 6



Supplementary Figure 7



Conclusions

The impact of the gut microbiota on metabolic processes, immune homeostasis and even neurological functions of the host is now well established. All these complex interactions are delicately balanced and disruption of this equilibrium can lead to pathologic states. Numerous studies over the years demonstrated the direct link between alterations of the gut microbiota and several pathologies, ranging from IBDs to ASDs. Nevertheless it is often not clear if the dysbiosis of the gut microbiota is a comorbidity or an aetiological factor in the pathophysiological mechanisms in which it is involved. The results obtained during this Ph.D. thesis suggest that the gastrointestinal symptoms, specifically constipation, commonly associated with neurological disorders such as ASDs and Rett syndrome are more likely a consequence of enteric nervous system dysfunctions rather than a direct effect of a dysbiotic gut microbiota. Indeed such neurological disorders have an important impact on the physiology of peripheral tissues (*e.g.* the gastrointestinal system) further triggering the observed alterations of the gut microbiota. A new picture is therefore emerging from this research work in which the intestinal microbial population of hosts affected by extra-intestinal, neurological disorders is shaped and adapted to the abnormal gastrointestinal physiology commonly associated with these pathologies, ultimately resulting in a positive feedback loop that amplify the gastrointestinal symptoms themselves. Accordingly, we provided the first evidence of an altered gut microbiota in Rett syndrome, hypothesising that this dysbiotic microbiota might reinforce the constipation status often observed in Rett syndrome by producing abnormal levels of SCFAs and therefore contributing to Rett syndrome's gastrointestinal physiopathology and intestinal inflammation. Moreover, we demonstrated that not only the intestinal bacterial communities, but also the gut mycobiota takes part to the complex set of interactions involving the host and the gut microbiota, requiring a further effort to better decipher host-microbial relationships in health and disease. In this sense, the observation that the gut mycobiota of Rett syndrome subjects is enriched in *Candida parapsilosis* isolates characterized by phenotypic and immunological properties favouring the low-grade intestinal inflammatory status observed in these subjects opens the possibility for a better understanding of the subclinical conditions derived from neurological disorders. Given the importance of the gut microbiota in host physiopathology other questions has to be addressed like whether alterations in microbiota-mediated immunomodulation leading to inflammatory disorders are associated with changes in microbiota-mediated neuromodulation and *vice-versa*. These new insights could stimulate the scientific and medical communities to develop novel therapeutic strategies, based on restoring the gut microbiota equilibrium, for the relief of gastrointestinal and behavioural abnormalities in Rett syndrome and other neurological disorders.

Acknowledgments

I am thankful to my supervisor Dr Claudio Donati (FEM) for his patient guidance and the support that I received. Thanks to his expertise and that of his collaborator Davide Albanese (FEM) I learnt joyfully the rigor of computational biology.

I wish to thank my co-advisor Dr Carlotta De Filippo (CNR-IBIMET) and Prof Duccio Cavalieri (UNIFI) for their continuous support and the possibility given to me to work independently during all this Ph.D. thesis. This project would never been existed without their foresight and expertise.

I wish to express my gratitude to Prof Antonio Calabrò (UNIFI) for his friendship and for the continuous exchange of scientific ideas.

Finally, I thank Prof Olivier Jousson (UNITN) for placing his trust in all the decisions that I have taken during this Ph.D. thesis.