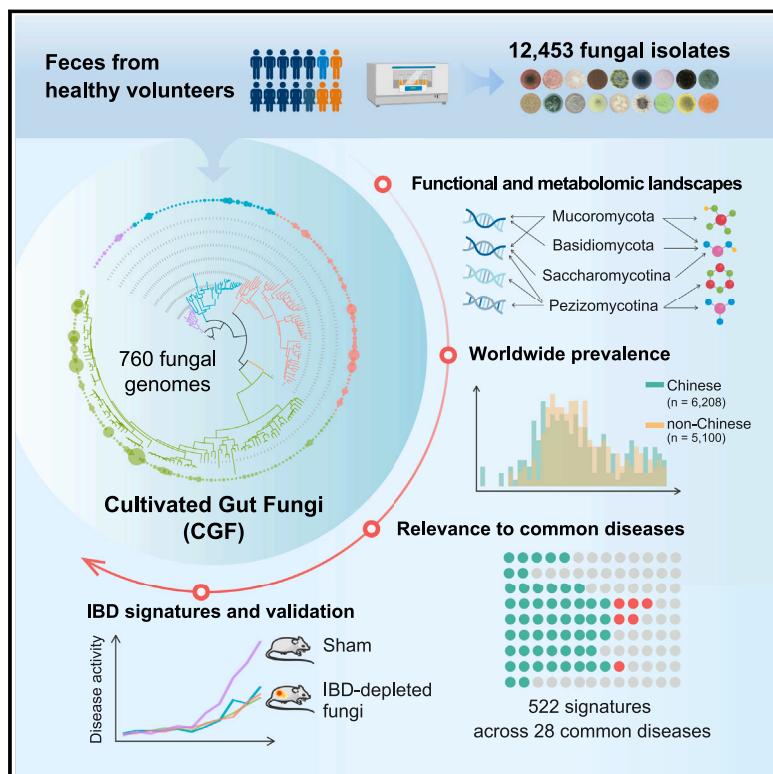


A genomic compendium of cultivated human gut fungi characterizes the gut mycobiome and its relevance to common diseases

Graphical abstract



Authors

Qiulong Yan, Shenghui Li,
Qingsong Yan, ..., De-an Guo,
Francis M. Martin, Xiaochi Ma

Correspondence

wach_edu@sina.com (C.W.),
daguo@simm.ac.cn (D.-a.G.),
francis.martin@inrae.fr (F.M.M.),
maxc1978@163.com (X.M.)

In brief

This study offers a catalog of cultivated gut fungi (CGFs) comprising 760 genomes derived from healthy individuals. This compendium highlights the taxonomic and functional diversity and metabolic potential of human gut fungi, aids in constructing a comprehensive phylogenetic representation of the gut mycobiome, and sheds light on common disease-related variations.

Highlights

- A cultivated gut fungi (CGF) catalog (760 genomes, 206 species) was established
- CGF catalog illustrates the functional and primary metabolic diversity of gut fungi
- CGF catalog characterizes the gut mycobiome in Chinese and non-Chinese populations
- Fungal signatures were identified in common diseases and validated in colitis mice



Article

A genomic compendium of cultivated human gut fungi characterizes the gut mycobiome and its relevance to common diseases

Qiulong Yan,^{1,2,4,12} Shenghui Li,^{5,7,12} Qingsong Yan,^{1,12} Xiaokui Huo,^{1,12} Chao Wang,^{1,2,3,*} Xifan Wang,^{7,8} Yan Sun,¹ Wenyu Zhao,² Zhenlong Yu,² Yue Zhang,⁵ Ruochun Guo,⁵ Qingbo Lv,⁵ Xin He,^{2,6} Changliang Yao,⁶ Zhiming Li,⁹ Fang Chen,⁴ Qianru Ji,⁵ Aiqin Zhang,⁵ Hao Jin,⁵ Guangyang Wang,⁴ Xiaoying Feng,¹ Lei Feng,¹ Fan Wu,¹ Jing Ning,² Sa Deng,² Yue An,¹ De-an Guo,^{6,*} Francis M. Martin,^{10,11,*} and Xiaochi Ma^{1,2,13,*}

¹Second Affiliated Hospital, Dalian Medical University, Dalian 116044, China

²Dalian Key Laboratory of Metabolic Target Characterization and Traditional Chinese Medicine Intervention, School of Pharmacy, Dalian Medical University, Dalian 116044, China

³First Affiliated Hospital, Dalian Medical University, Dalian 116044, China

⁴College of Basic Medical Sciences, Dalian Medical University, Dalian 116044, China

⁵Puensum Genetech Institute, Wuhan 430076, China

⁶Shanghai Research Center for Modernization of Traditional Chinese Medicine, National Engineering Laboratory for TCM Standardization Technology, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201210, China

⁷Key Laboratory of Precision Nutrition and Food Quality, Department of Nutrition and Health, China Agricultural University, Beijing 100091, China

⁸Department of Obstetrics and Gynecology, Columbia University, New York, NY 10027, USA

⁹BGI-Shenzhen, Shenzhen 518083, China

¹⁰Université de Lorraine, Institut national de recherche pour l'agriculture, l'alimentation et l'environnement, UMR Interactions Arbres/Microorganismes, Centre INRAE Grand Est-Nancy, Champenoux 54280, France

¹¹Beijing Advanced Innovation Center for Tree Breeding by Molecular Design, Beijing Forestry University, Beijing 100091, China

¹²These authors contributed equally

¹³Lead contact

*Correspondence: wach_edu@sina.com (C.W.), daguo@simm.ac.cn (D.-a.G.), francis.martin@inrae.fr (F.M.M.), maxc1978@163.com (X.M.)
<https://doi.org/10.1016/j.cell.2024.04.043>

SUMMARY

The gut fungal community represents an essential element of human health, yet its functional and metabolic potential remains insufficiently elucidated, largely due to the limited availability of reference genomes. To address this gap, we presented the cultivated gut fungi (CGF) catalog, encompassing 760 fungal genomes derived from the feces of healthy individuals. This catalog comprises 206 species spanning 48 families, including 69 species previously unidentified. We explored the functional and metabolic attributes of the CGF species and utilized this catalog to construct a phylogenetic representation of the gut mycobiome by analyzing over 11,000 fecal metagenomes from Chinese and non-Chinese populations. Moreover, we identified significant common disease-related variations in gut mycobiome composition and corroborated the associations between fungal signatures and inflammatory bowel disease (IBD) through animal experimentation. These resources and findings substantially enrich our understanding of the biological diversity and disease relevance of the human gut mycobiome.

INTRODUCTION

The human digestive system is inhabited by a diverse array of microorganisms, including bacteria, archaea, viruses, and fungi. Alongside the bacterial community, the fungal community, referred to as the gut mycobiome, forms an essential part of the healthy microbiome and exerts influence on various host processes,^{1–5} such as the metabolism of specialized dietary compounds. For the past decade, researchers have relied on DNA metabarcoding targeting the fungal 18S rRNA gene or ribosomal

DNA (rDNA) internal transcribed spacer (ITS) regions to study the gut mycobiome. This approach has greatly contributed to our current understanding of fungal phylogenetic composition in both healthy individuals and those with diseases.⁶ However, despite these advancements, our knowledge of genetic and functional variations within the human gut mycobiome remains limited. A major challenge stems from the lack of a comprehensive repertoire of gut fungal reference genomes, impeding the taxonomic classification of most gut fungal amplicons and restricting the survey of the entire fungal microbiota.^{7–9} Moreover,



the restricted set of reference genomes hampers the in-depth exploration of gene expression in gut fungal species using metatranscriptomics and metaproteomics.

The generation of reference genomes for human gut microbes has been achieved through cultivation efforts^{10–12} and direct assembly from massive fecal metagenomes.^{13–15} These endeavors have led to the creation of the Unified Human Gastrointestinal Genome (UHGG) collection¹⁶ and its expanded catalogs,^{17,18} encompassing over 230,000 gut microbial genomes from more than 5,400 species. However, it should be noted that these resources lack information pertaining to gut fungi. Although initiatives have been undertaken to cultivate gut fungi^{9,19,20} and reconstruct fungal genomes from fecal metagenomes,²¹ these covered only a minor fraction of the estimated 400 or more fungal species believed to inhabit healthy human guts.^{22,23} As of June 2023, approximately 4% of the fungal genomes available in the National Center for Biotechnology Information (NCBI) database were sourced from human habitats. These statistics underscore a significant gap in our understanding of this essential component of gut microbiota.

Recent studies have documented disease-induced alterations in gut microbiota across a range of gastrointestinal and systemic disorders, including inflammatory bowel disease (IBD),²⁴ colorectal cancer (CRC),²⁵ type 2 diabetes (T2D),²⁶ and liver diseases.^{27,28} These studies frequently indicated that individuals with a disease status typically harbor a dysbiotic gut mycobiome, characterized by reduced fungal diversity and an expansion of certain opportunistic pathogens, such as *Candida* and *Malassezia* spp.^{29–31} Despite these advances, our comprehension of the precise role of gut mycobiome in disease pathogenesis remains limited. The inadequacy of reference databases contributes to the less accurate quantification of gut fungal species in high-throughput sequencing datasets. Moreover, the highly variable experimental and analytical methods employed in different studies have hindered the comparison of fungal signatures across diverse disease states. Therefore, a comprehensive investigation of gut mycobiomes associated with various diseases necessitates a unified reference and analytic approach to elucidate the overarching patterns in disease-associated mycobiome shifts.

Here, we present a catalog of cultivated gut fungi (CGFs) that comprises 760 fungal genomes obtained from the feces of healthy volunteers. This catalog spans 48 families and 206 species, among which 69 are previously unidentified species without genomic information in existing databases. The CGF catalog significantly augments genomic resources for gut fungal species and protein families by over 4- and 2-fold, respectively. We characterized the core functions of CGF genomes and quantified the *in vitro* production of primary metabolites in different CGF species. Furthermore, we used the CGF catalog in conjunction with publicly available human-associated fungi (PHF) genome resources to conduct profiling of gut mycobiomes in over 6,000 fecal metagenomes from the Chinese population (covering 28 diseases or unhealthy states) and over 5,000 fecal metagenomes from non-Chinese populations. This allowed for a detailed exploration of the diversity and structural characteristics of the gut mycobiome in relation to various diseases. Disease-shared and disease-specific gut fungal signatures were

identified, and their association with IBD was further elucidated through animal experiments. This extensive CGF resource provides new insights into the biological significance of the human gut mycobiome, which promises to stimulate future experimental research in this field.

RESULTS

Cultivation and genome sequencing of gut fungi

To explore the genetic diversity of the gut fungal community, we initiated fungal cultivation from fresh fecal specimens of 135 healthy volunteers using multiple fungus-specific media (see **STAR Methods** section for details). Of the 12,453 fungal isolates obtained, 744 were selected for whole-genome sequencing based on their phylogenetic diversity (18S or ITS rDNA polymorphism) and mycelial morphology (**Figures S1A–S1D; Table S1**). Following *de novo* assembly of genomic sequencing reads, we identified 16 isolates with multiple genomes, which were further divided into 32 genomes. In total, 760 assembled genomes were obtained to form the CGF catalog (**Table S1**). The genome sizes and G+C contents of CGF genomes ranged from 7.6 to 65.1 Mbp (median 29.2 Mbp) and 24.8% to 63.8% (median 49.0%), respectively (**Figure S1E**). Completeness of the fungal genomes ranged from 51.1% to 100%, with a median completeness of 98.8% (**Table S1**).

The widely accepted criterion for a genome average nucleotide identity (ANI) of 95% for distinguishing species in prokaryotes³² was applied to CGFs, based on a survey of available fungal genomes in the NCBI database (containing 11,623 high-quality genomes; **Figure S1F; Table S2**). Similarly, the ANI threshold for distinguishing the genus boundaries was set at 75%. The sequenced CGF genomes were grouped into 206 distinct fungal species (**Table S1**), with 69 of them being newly sequenced because they did not match any reference genome within the NCBI database using the species-level ANI threshold. Among the newly sequenced species, 58 (comprising 94 genomes) were affiliated with known genera, whereas the remaining 11 (comprising 25 genomes) were classified at the family or higher taxonomic levels based on their ANI values and phylogenetic relationships with known fungi.

The taxonomic distribution of the 206 species is summarized in **Figure 1** and supported by the morphological and scanning electron microscopy photographs (**STAR Methods**). In CGFs, we identified three major fungal phyla: Ascomycota (32 families, 156 species, and 644 genomes), Basidiomycota (9 families, 36 species, and 85 genomes), and Mucoromycota (6 families, 14 species, and 31 genomes). Filamentous ascomycetes in the subphylum Pezizomycotina, such as *Aspergillus* (25 species and 196 genomes), *Penicillium* (21 species and 123 genomes), and *Talaromyces* (14 species and 48 genomes), dominated the CGF catalog. Other notable fungal taxa included typical yeasts (subphylum Saccharomycotina), such as *Candida* (6 species and 59 genomes) and *Pichia* (9 species and 48 genomes). Consistent with previous observations,^{33,34} members of filamentous fungi and yeasts exhibited distinct genomic features (**Figure S1G**). Notably, the 69 newly sequenced species were broadly distributed across taxonomic clades (**Figure 1**), suggesting that additional uncultured species were likely to be

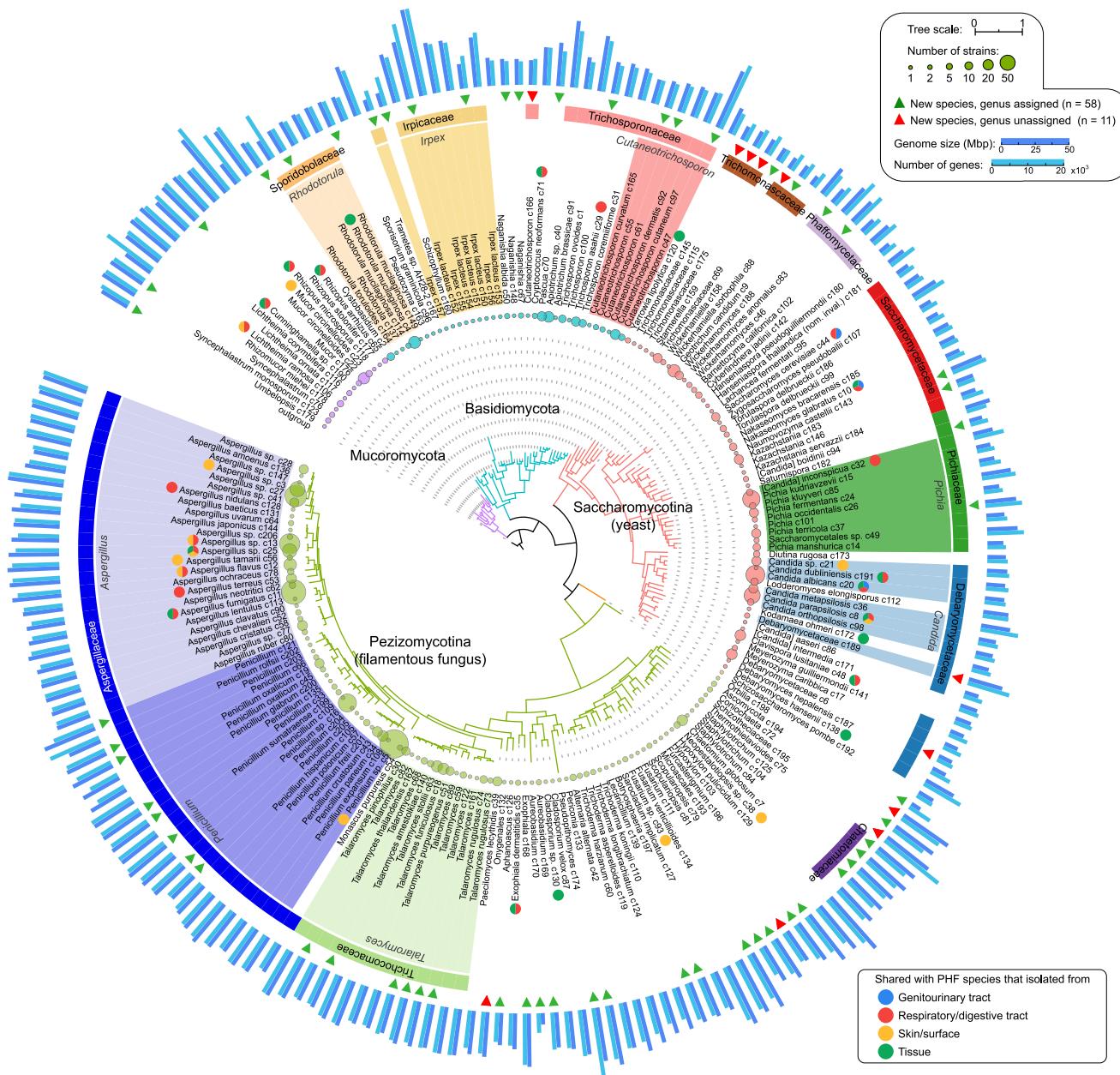


Figure 1. Genome-wide phylogeny of 206 fungal species cultivated from the human gut

Innermost circle: phylogenetic tree based on whole-proteome sequences using the feature frequency profile method (see STAR Methods). Colors represent four subphyla, and their names are labeled. Circle 2: the circle represents the number of strains for each species. Circles 3 and 4: taxonomic names of 206 species. The color represents the genus (circle 3) and family (circle 4) assignment of each species. Pies upon the taxonomic names indicate shared species between the cultivated gut fungi (CGF) and publicly available human-associated fungi (PHF) catalogs. PHF species isolated from different body sites are labeled by color. Circle 5: triangle represents newly sequenced species relative to available fungal genomes in the National Center of Biotechnology Information (NCBI) database. Colors indicate new species to the NCBI database at the genus level (green) or at the family or higher level (red). Outermost circle: genome size and number of genes in 206 species. Detailed information on the 206 species and original 760 genomes is available in Table S1. The maximum-likelihood tree is inferred with RAxML³⁵ based on 150 genes and visualized using the online iTOL tool.³⁶

See also Figure S1.

discovered in the human gut with further sampling and cultivation. Furthermore, metagenomic sequencing of fecal samples from volunteers revealed that 176 of the 206 CGF species were repeatedly detected in feces both at the time of cultivation

and over 3 years later, suggesting their residency in the human gut (Figure S1H; Table S1).

To further evaluate the uniqueness of the individual reference genomes in CGFs, we compared them with fungal strains

available from NCBI that originated from human sources. This catalog of PHFs comprised 502 genomes that were isolated from various tissues, such as the respiratory and digestive tract, genito-urinary tract, skin, and others (as listed in **Table S2**). When we examined the PHF genomes, we found that they were phylogenetically less diverse, encompassing 129 species under the same ANI threshold used for the CGF genomes. Of these, 42 species were previously isolated from the respiratory or digestive tracts (**Table S2**). Importantly, only 32 fungal species of the PHF catalog, including 20 from the respiratory or digestive tract, were shared with the CGF species (**Figure 1**). This finding indicates that our CGF catalog has significantly expanded the number of cultivated gut fungal species available, increasing it by over 4-fold (from 42 to a total of 228). The shared species between the two catalogs included members of typical yeast (11 species, including *Candida* spp., *Clavispora lusitaniae*, *Nakaseomyces glabratus*, and *Saccharomyces cerevisiae*), *Aspergillus* (8 species), and *Mucoromycota* (5 species). Notably, the majority of species within the *Penicillium* (20 out of 21 species) and *Pichia* (8 out of 9 species) genera, as well as all species in *Talaromyces* (11 species), *Irpea* (8 species), and *Cutaneotrichosporon* (6 species) genera, among other taxa, were exclusively identified in CGFs (**Figure 1**). While genomes for these fungi have not been previously released, many have been reported as human gut residents.^{7,37} However, some pathogenic fungi that may cause gastrointestinal disorders, such as *Malassezia* species,³⁰ were not present in the CGF catalog, probably because of their low prevalence in the gut microbiota of the individuals in this study. Furthermore, a high proportion (51.0%) of CGF species were found to be shared with available NCBI genomes isolated from animal and natural habitats (**Figure 1; Tables S1 and S2**), suggesting a potential fungal exchange between external environments and the human gut.

Functional configuration of the CGF genomes

The CGF genomes encoded 7,660,447 protein-coding genes (**Table S1**), which could be organized into 643,717 non-redundant protein clusters at a 50% average amino acid identity (AAI) threshold. Remarkably, approximately 77% of these proteins were absent in the respiratory/digestive tract fungi from the PHF catalog, indicating a substantial 247% increase (from 200,769 to a total of 696,865) in the known protein sequence diversity of human gut fungi (**Figure 2A**). Specifically, 31.6% (203,579/643,717) of the protein clusters in CGFs were not present in any of the available fungal genomes in the NCBI database (**Figure S2A**), highlighting the considerable novelty of the protein-coding content within the CGF catalog. The coding space, gene duplication rate, and proportion of taxonomically specific proteins varied greatly among fungi belonging to different clades (**Figure S2B**), suggesting that the CGF gene content is primarily shaped independently across various taxa and is likely influenced by ecological heterogeneity in the gut habitats.

On average, 83.3% and 52.1% of the CGF proteins exhibited sequence and domain homology in the eggNOG (evolutionary genealogy of genes: non-supervised orthologous groups) and KEGG (Kyoto encyclopedia of genes and genomes) databases, respectively (**Table S3**). Principal coordinate analysis (PCoA) of the eggNOG profiles (comprising 26,678 orthologous groups) revealed a distinct separation among the four fungal subphyla: Mu-

coromycota, Pezizomycotina, Saccharomycotina, and Basidiomycota (**Figure 2B**). Phylogenetic affiliation emerged as the primary factor in determining the functional configuration of CGFs across different taxonomic levels (class, order, family, and genus levels), accounting for 54.6%–81.6% of the functional variance (**Figure S2C**). For each of the 760 CGF genomes, core functional pathways were reconstructed based on the completeness ratio of KEGG modules (**Figure 2C**; see **Table S3** for details). The essential metabolic pathways, including glycolysis, the tricarboxylic acid cycle, fatty acid biosynthesis, β-oxidation, and nucleotide metabolism, were found to be largely complete across all species. While lipid and amino acid metabolism was prevalent in all subclades, Pezizomycotina species exhibited heightened functional diversity compared with the others (**Figure S2D**). Noteworthy variations included the presence of complete pathways of tyrosine degradation (M00044), methionine degradation (M00035), and glutathione biosynthesis (M00118) in most Pezizomycotina genomes but rare occurrences in other taxa, whereas the tryptophan degradation (M00038) pathway was only identified in Pezizomycotina species. The degradation of tyrosine and tryptophan gives rise to a series of important precursors of secondary metabolites.^{38,39} Additionally, triacylglycerol biosynthesis (M00089) and the methionine salvage pathway (M00034) were absent in Basidiomycota species, and the biosynthetic pathways for sphingosine (M00099), GABA (gamma-aminobutyric acid, M00135), and betaine (M00555) were absent in Saccharomycotina genomes, whereas the biosynthetic pathways for phenylalanine (M00024), tyrosine (M00025), and lysine (M00030) were absent in Mucoromycota genomes.

Owing to their larger gene repertoire (**Figure S2B**) and unique niche adaptation,^{40,41} Pezizomycotina species showed a broader array of affiliated metabolic pathways than other fungi (**Figure S2E**). For example, the degradation of pectin (M00081), D-galacturonate (M00630), and adenine/guanine ribonucleotide (M00958 and M00959) was common in Pezizomycotina species but rare in other fungi, while the biosynthetic pathways of secondary metabolites such as paspaline (M00661), fumiquinazoline (M00901), penicillin (M00672), and gibberellin A4/A1 (M00928) were exclusively encoded by some Pezizomycotina members. The nitrate assimilation pathway (M00531) was prevalent in almost all Pezizomycotina subclades and four other families (Ustilaginaceae, Irpicaceae, Filobasidiaceae, and Phaffomycetaceae). This aligns with the known capacity of filamentous fungi and certain Basidiomycota subclades to assimilate nitrate or nitrite,⁴² whereas certain yeasts, including Phaffomycetaceae, likely acquired this pathway through horizontal gene transfer from other taxa.^{43,44} Conversely, several pathways were notably absent in Pezizomycotina species within the CGFs. For example, tetrahydrofolate biosynthesis (M00126) was encoded in most Mucoromycota, Basidiomycota, and Saccharomycotina genomes but was rare in Pezizomycotina species (**Figures 2C and S2E**). Additionally, genes involved in lipoic acid biosynthesis (M00884) were uniquely encoded by Saccharomycotina yeasts, whereas those related to the biosynthesis of riboflavin (M00125) and tetrahydrobiopterin (M00842) were uniquely present in Mucoromycota genomes.

In the eggNOG and KEGG profiles, 76.2% and 33.8% of orthologous groups, respectively, were specifically encoded by one of the four subphyla (**Figure 2D**), in agreement with their

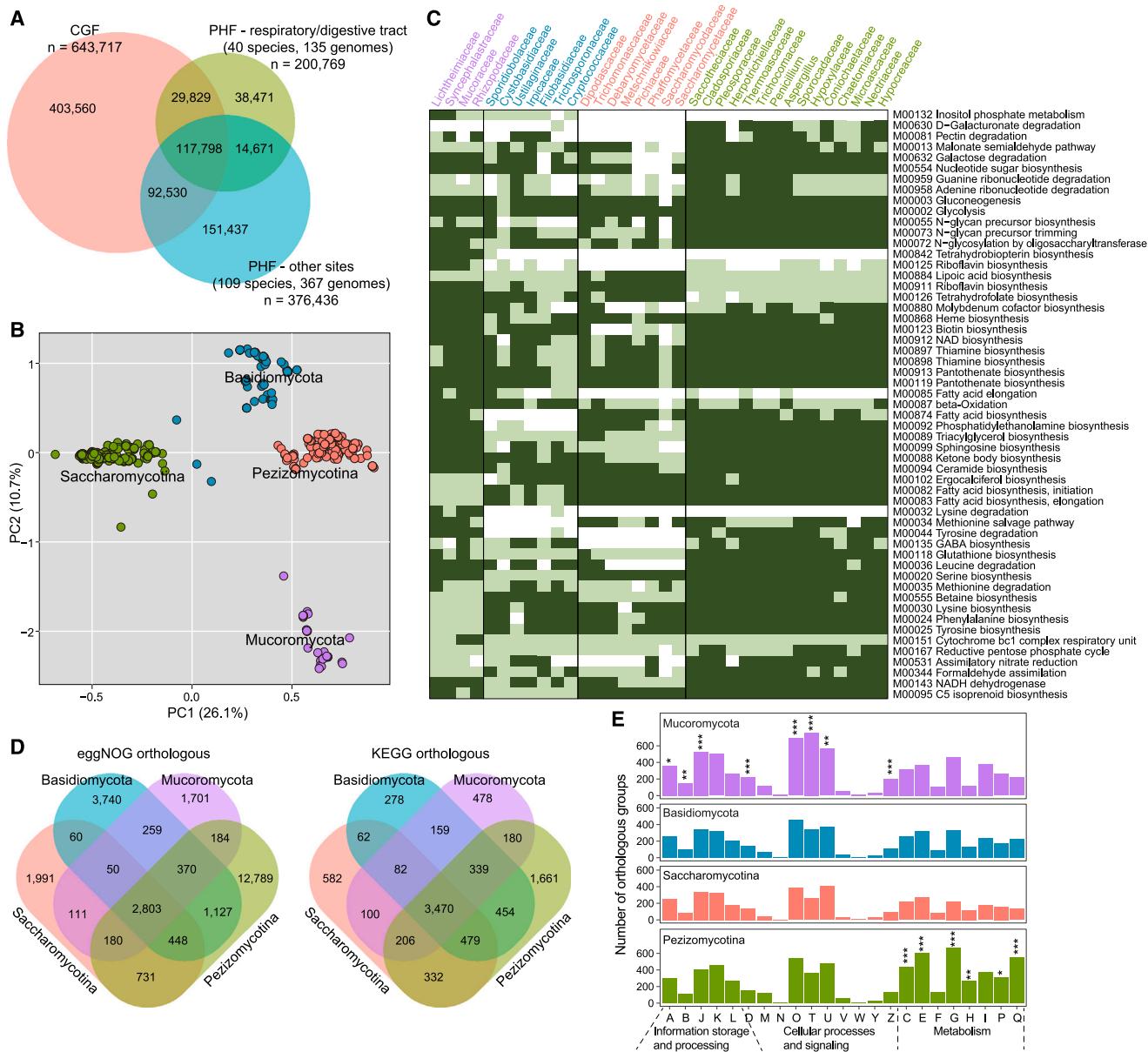


Figure 2. Overview of functional profiles of cultivated gut fungi

(A) Venn diagram of shared protein clusters between CGF and PHF catalogs.

(B) Multivariate analysis of functional profiles showing a clear separation of functions among different fungal subphyla. The analysis employed principal coordinate analysis (PCoA) on the composition of the eggNOG orthologous group of fungal genomes. The first two principal coordinates (PCs) and the ratio of the variance contribution are shown.

(C) Completeness ratio of KEGG metabolic modules for 32 dominant fungal clades and two genera of Aspergillaceae (*Aspergillus* and *Penicillium*). The color shows the median completeness ratio of the strains in each clade: dark green (presence of a complete module) and light green (largely present with only one enzyme missing). Only modules detected as complete in at least one clade are present, and the comprehensive results are available in Table S3.

(D) Venn diagram showing the comparisons of eggNOG orthologous groups (left) and KEGG orthologous groups (right) among the four fungal subphyla.

(E) Bar plots showing the distribution of eggNOG orthologous groups among species belonging to the four fungal subphyla. Functions are grouped according to the COG (clusters of orthologous groups) categories (<http://www.ncbi.nlm.nih.gov/COG/>). Fisher's exact test indicates the enrichment of categories among the four subphyla: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

See also Figure S2.

high level of specificity in protein clusters. Mucoromycota-specific functions were predominantly associated with genetic information processing, cellular processes, and signaling pathways,

whereas Pezizomycotina-specific functions were concentrated in secondary metabolism and transport/metabolism of amino acids and carbohydrates (Figure 2E).

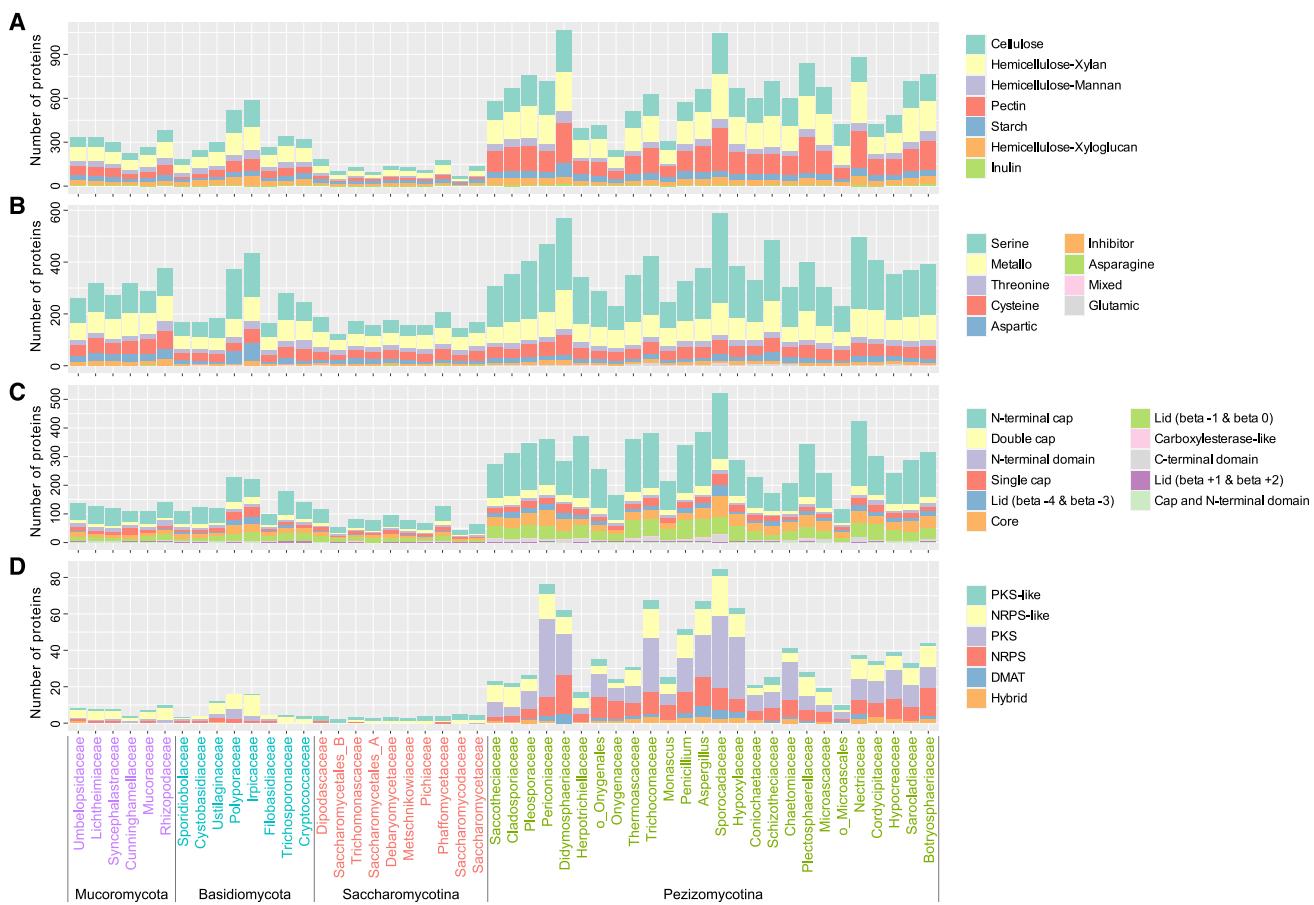


Figure 3. Degradation potential of fermentable substrates and biosynthesis of secondary metabolites by gut fungi

Bar plots showing the average number of carbohydrate-active enzymes (CAZymes) (A), proteases (B), lipases (C), and secondary metabolic gene clusters (SMGCs) (D) across 34 dominant fungal clades in CGFs. For CAZymes, the gene contents of seven target polysaccharides are shown. Backbone genes ($n = 26,355$) of the SMGCs are categorized by enzyme types: DMAT, dimethylallyl transferase; NRPS, non-ribosomal peptide synthetase; PKS, polyketide synthase; hybrid, the gene containing domains from NRPS and PKS backbones.

See also Figure S3.

Functional role of fungal genes in gut environment adaptation

Functions involved in the degradation of fermentable substrates (i.e., polysaccharides, proteins, and lipids) and biosynthesis of secondary metabolites are likely to play a key role in the ecological adaptation of fungi to the human gut environment.^{45,46} To explore the underlying genetic arsenal facilitating these adaptations, we first conducted an in-depth analysis of carbohydrate-active enzymes (CAZymes) in the 760 CGF genomes. A total of 532,809 CAZymes were predicted, representing 0.9%–9.3% (on average, 6.8%) of the total protein repertoire in CGF species (Figure 3A; Table S3). Pezizomycotina species encoded a significantly higher proportion of diverse CAZymes than the other fungal lineages (Wilcoxon rank-sum test, $p < 0.001$; Figure S3A). Specifically, these fungi expressed a plethora of plant cell wall-degrading enzymes (PCWDEs) that target cellulose, hemicellulose, starch, and pectin (Figures S3B and S3C). This observation strongly implies that Pezizomycotina species in the gut actively participate in the decomposition of plant polysaccharides, which is consistent with the findings from prior research on

Aspergillus spp.⁴⁷ Second, we detected 225,078 protease genes and 188,425 lipase genes in the CGF genomes, representing 1.3%–4.1% (on average, 2.9%) and 0.2%–4.3% (on average, 2.3%) of the total protein repertoire, respectively (Figures 3B and 3C). Enrichment analysis showed that Pezizomycotina species had the highest capacity for lipid catabolism compared with other subphyla, whereas Saccharomycotina species exhibited a remarkable proteolytic arsenal (Figure S3D). Finally, we identified a total of 22,762 secondary metabolic gene clusters (SMGCs) within the CGF genomes (Table S3). Based on a previously established method,⁴⁷ these SMGCs were clustered into 1,590 SMGC families, each anticipated to produce different compounds. Notably, Pezizomycotina genomes exhibited the highest proportion of SMGCs compared with other subphyla (Figure 3D), with 92.7% of the SMGC families being uniquely encoded (Figure S3E). The majority of SMGC families were specifically encoded by either a single species (45.3%) or genus (67.6%) within the CGF catalog (Figures S3F and S3G). This finding resonates with earlier surveys conducted in *Aspergillus* species,⁴⁷ suggesting that horizontal gene transfer may play a

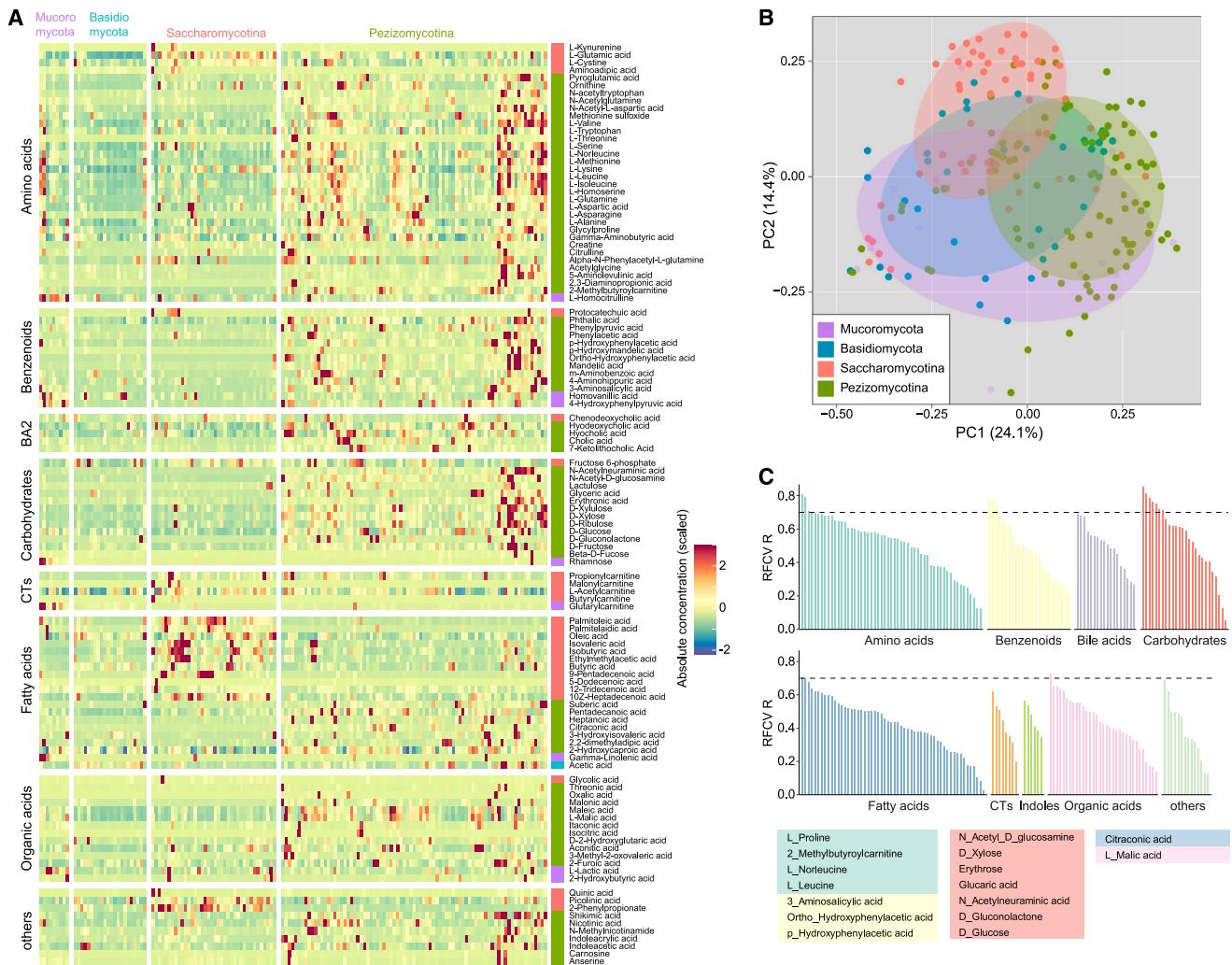


Figure 4. Overview of primary metabolite profiles of cultivated gut fungi

(A) Comparison of metabolites among four fungal subphyla. Heatmaps show the absolute concentrations of the significantly different metabolites. Enrichments are indicated (purple, enriched in Mucoromycota; blue, enriched in Basidiomycota; red, enriched in Saccharomycotina; green, enriched in Pezizomycotina), and metabolite names are provided on the right. BA, bile acid; CT, carnitine. The ordering of fungal species is shown in Table S4.

(B) Principal coordinate analysis (PCoA) of metabolomic profiles. The first two principal coordinates (PCs) and the ratio of the variance contribution are shown. Colored circles cover the fungal strains near the center of gravity for each subphylum.

(C) Distribution of the predicted effect sizes of metabolites. Spearman's correlation between measured values and 5-fold cross-validation random-forest-model predicted values (RFCV R) is used to estimate the effect size of fungal functional contents in each metabolite. Detailed information on the metabolites and effect sizes is provided in Table S4.

See also Figure S4.

significant role in the evolutionary dynamics of fungal secondary metabolism.

In vitro characterization of primary metabolites produced by CGF species

To characterize the metabolite content of the CGFs, we performed *in vitro* targeted metabolomic profiling of 199 of the 206 CGF species (a representative strain for each species was selected for analysis). The absolute concentrations of 241 primary metabolites, including 56 fatty acids and derivatives, 55 amino acids and derivatives, 33 organic acids, 26 carbohydrates, and 25 benzenoids, of lyophilized fungal powders were

measured. Of the 216 most prevalent metabolites, 166 showed significant differences in concentration among four subphyla (Kruskal-Wallis test, $q < 0.05$; Figure 4A; Table S4). Metabolites enriched in Pezizomycotina species included a diverse array of amino acids, benzenoids, carbohydrates, and organic acids, whereas Saccharomycotina-enriched metabolites comprised other amino acids and fatty acids. Basidiomycota-enriched metabolites are limited to certain organic acids.

PCoA showed significant divergence in the profiles of primary metabolic compounds between the fungal subphyla (Figure 4B). Nevertheless, phylogenetic affiliation only accounted for 8.2%–17.3% of the observed variances in metabolomic composition

at various taxonomic levels (28.0% in total; **Figure S4A**). Other genomic features, such as genome size and G+C content, contributed to an additional 5.2% of the metabolome variation (**Figure S4B**). This suggests that non-genomic environmental factors, including culture conditions and regulatory mechanisms, play a significant role in shaping the fungal metabolome. Furthermore, we applied random forest models to evaluate the influence of functional genome-coding capacity on metabolite concentrations. The predicted concentrations of each metabolite in 199 fungi were computed from the function-trained random forest model using a leave-one-out procedure (see **STAR Methods**). Here, 16 of the 241 analyzed metabolites, including several amino acids and carbohydrates, generated a correlation coefficient larger than 0.7 (correlation coefficient test $q < 0.0001$) between their predicted and targeted concentrations (**Figures 4C, S4C, and S4D**), suggesting that the concentrations of these metabolites are predominantly determined by the functional coding capacity of the genome. Conversely, the concentrations of some metabolites, exemplified by carnitine (CT), appeared to be independent of the encoded genomic functions.

Genome-reference-based delineation of the gut mycobiome from metagenomes

To develop a robust and accurate method for quantifying fungal species in fecal metagenomes, two broad strategies for bacterial composition profiling were initially evaluated: the clade-specific core gene approach (exemplified by MetaPhlAn4⁴⁸) and the direct mapping of reads to genomes using alignment algorithms, such as Burrow-Wheeler Aligner (BWA)⁴⁹ or Bowtie 2.⁵⁰ However, these strategies face challenges when applied to profile gut fungi because of insufficient genomic resources and difficulty in accurately discerning the origin of reads in low-abundance fungal environments. Combining the strengths of these approaches and drawing inspiration from a recently developed algorithm designed for high-precision quantification of ultra-low-abundance bacteria,⁵¹ we devised a similar algorithm that realizes fungal metagenomic profiling by considering the alignment of reads to both fungal species-specific and shared genes (a detailed step-by-step process of the algorithm is shown in **Figure S5A**). This algorithm was validated by testing two ultra-deep fecal metagenomic datasets from Chinese⁵² and Hadza⁵³ populations (**Table S5**) across various sequencing depths (1–200 million reads). As the number of reads increased, the observed fungal species and Shannon diversity index consistently increased, reaching saturation at 20–50 million reads, indicating a comprehensive coverage of fungal diversity within this condition (**Figure S5B**). Comparative analyses revealed high similarity in fungal composition in sequencing amounts greater than 10 million reads (Pearson correlation coefficient >0.9 in most samples; **Figures 5A and S5C**), signifying a robust and stable fungal community structure. Similarly, PCoA corroborated the resemblance of individual fungal compositions beyond 10 million reads (**Figure S5D**). Furthermore, examination of genome coverage for certain gut fungi revealed high uniformity across the entire genome zone, even at low sequencing amounts (**Figure S5E**). These findings underscore the reliability and accuracy of our algorithm, particularly when the sequencing amount is as low as 10–20 million reads.

Building on a fungal genomic reference database and leveraging quantitative tools, we aimed to explore the characteristics of the gut mycobiome and its potential links to common diseases. We collected 6,756 Chinese fecal metagenomes from 38 disease case-control studies, including three unpublished studies sequenced as part of this research (**Figures S6A and S6B**; **Table S5**). Additionally, we obtained 6,188 non-Chinese metagenomes from diverse studies spanning multiple countries, allowing us to investigate mycobiome features across different populations (**Table S5**). All samples underwent rigorous quality control using a standardized pipeline and were analyzed based on our fungal profiling algorithm. After excluding samples with a low sequencing depth and exceptionally low proportions of detectable fungal sequences (**Figure S6A**), we focused on the gut fungal profiles of 6,208 Chinese samples and 5,100 non-Chinese samples across 303 fungal species (comprising CGF and PHF species) for subsequent analysis.

In the Chinese population, approximately 20% (41/206) of the CGF species were identified as prevalent (defined by a prevalence rate $>10\%$; **Figure 5B**; **Table S5**). However, more than half of the CGF species (52.9%, 109/206) were rare and occurred in less than 1% of the entire population. The prevalence of PHF species was notably lower than that of CGF species; only 13.2% (17/129) of PHF species were prevalent, whereas 67.4% (87/129) were rare (**Figure 5B**). Notably, among the 17 prevalent PHF species, 13 were shared with CGF species (**Table S5**). In non-Chinese populations, we observed that 18.0% (37/206) of CGF species and 12.4% (16/129) of PHF species were prevalent, whereas 46.1% (95/206) of CGF species and 58.9% (76/129) of PHF species were rare (**Figure 5C**). Given that the CGF species were isolated from Chinese individuals, their prevalence rates were significantly higher in the Chinese population than in the non-Chinese population (paired Student's t test, $p = 0.001$). However, the prevalence of PHF species did not differ significantly between them ($p = 0.162$). Furthermore, despite the low prevalence of most fungi in the populations, we identified certain fungi that are commonly present in both Chinese and non-Chinese populations. A prominent example is *Saccharomyces cerevisiae* c44 (the term "c44" represents a unique species ID in the CGF and PHF catalogs, as detailed in **Tables S1 and S2**, with the same meaning below), which appeared in 63.0% of Chinese and 70.8% of non-Chinese samples. Several other fungi, including *Rhizopus stolonifer* c177, *Pseudopithomyces* c174, *Aspergillus* sp. c27, *Fusarium* sp. c93, *Lodderomyces elongisporus* c112, *Penicillium* sp. c16, *Pichia kudriavzevii* c15, *Barnettozyma californica* c102, and *Aspergillus* sp. c13, also exhibited prevalence rates exceeding 50% in both Chinese and non-Chinese populations (**Figure S6C**).

At the subphylum level, the composition patterns of the gut mycobiome were highly similar across all studies conducted on different populations and diseases (**Figure 5D**). The gut mycobiome was dominated by two Ascomycota subphyla, Saccharomycotina (average relative abundance, $57.1\% \pm 32.1\%$) and Pezizomycotina ($27.9\% \pm 28.0\%$), whereas Mucoromycota and Basidiomycota generally had lower abundances. At the species level, the fungal species displayed a broad range of average relative abundance (from 0% to 13.5%), followed by an approximately log-normal distribution across the entire population

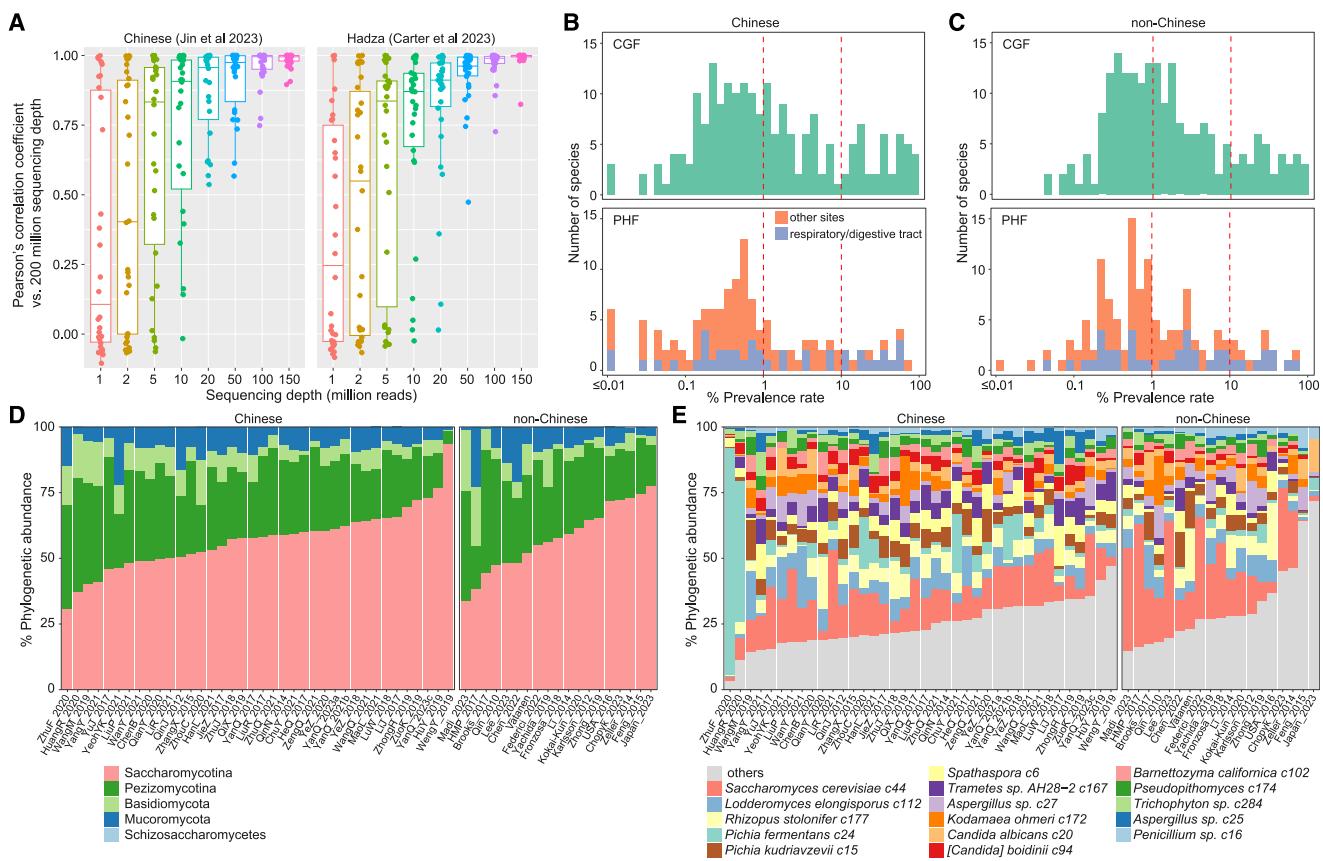


Figure 5. Delineation of the gut mycobiome via CGF and PHF catalogs

(A) Assessment of the updated fungal profiling algorithm in two ultra-deep fecal metagenomic datasets. Boxplot showing Pearson's correlation coefficient of gut fungal profiles across various sequencing depth gradients (from 1 to 150 million reads), compared with profiles at a sequencing depth of 200 million reads.

(B and C) Histograms showing the distribution of the prevalence rates of gut fungal species in the fecal metagenomes of the Chinese (B) and non-Chinese (C) populations. The upper and lower panels represent the prevalence rates of 206 CGF and 129 PHF species, respectively. PHF species isolated from the respiratory/digestive tract or other body sites are labeled with different colors.

(D and E) Phylogenetic compositions of the gut mycobiome at the subphylum (D) and species (E) levels. Fecal metagenomes are categorized into 38 studies in the Chinese population and 19 studies in non-Chinese populations. The bars indicate the average relative abundance of each subphylum/species and are colored according to their taxonomic assignments. Only species with an average relative abundance of >2% are presented.

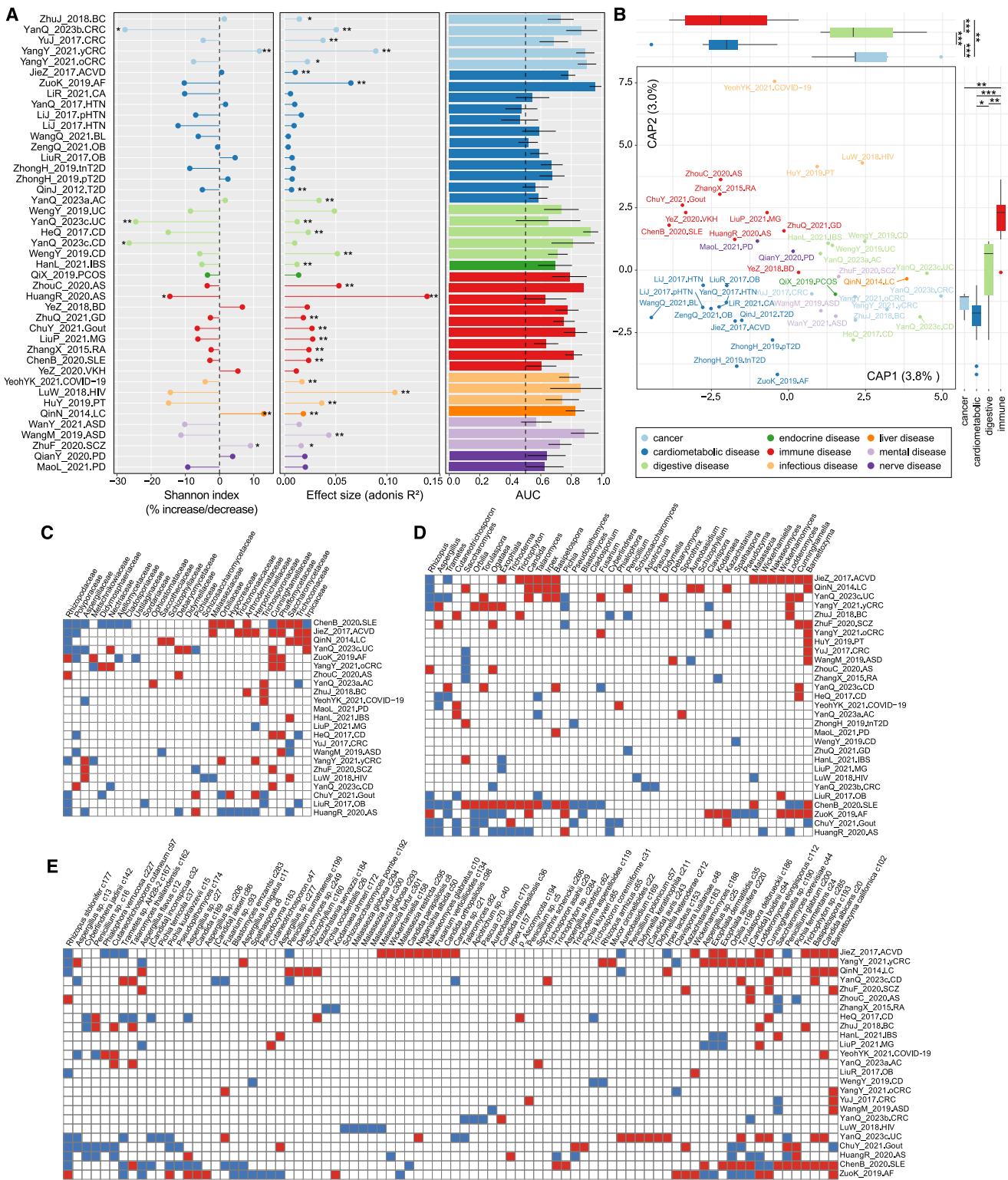
See also Figures S5 and S6.

(Figure S6D). Sixteen species, including *Saccharomyces cerevisiae* c44, *Lodderomyces elongisporus* c112, *Rhizopus stolonifer* c177, *Pichia fermentans* c24, and *Pichia kudriavzevii* c15, were considered as the core mycobiome of the human gut, as they comprised up to 80% of the abundance across all studies (Figure 5E). Unlike the gut bacteriome, which is stratified by enterotype,⁵⁴ we did not observe distinct community composition types in the mycobiome of either the Chinese or non-Chinese population (Figure S6E), suggesting that the gut mycobiome may be formed independently of the bacteriome. Interestingly, permutational multivariate analysis of variance (PERMANOVA) in the Chinese population revealed that the study cohort significantly contributed to shaping the gut mycobiomes across all investigated samples (*adonis* $R^2 = 9.0\%$, $p < 0.001$), possibly because of differences in geography or experimental methods (e.g., sample preparation and DNA extraction). However, even after adjusting for the impact of the study cohort and other confounders, disease status significantly influenced the gut myco-

biome (*adonis* $R^2 = 0.7\%$, $p < 0.001$). This finding suggests the presence of shared fungal signatures that are associated with a healthy state.

Diversity and structure of the gut mycobiome associate with common diseases

We first tested whether the diversity of the gut mycobiome differed between the disease and control groups, as observed for the gut bacteriome in multiple diseases.^{55,56} Among all investigated samples in the Chinese population, individuals with diseases exhibited significant reductions in the Shannon diversity index, compared with healthy controls (Figure S7A). This reduction trend was observed in 31 of 43 case-control comparisons and was particularly significant in patients with IBD (study YanQ_2023c; study IDs and abbreviations are detailed in Table S5), CRC (except young CRC from study YangY_2021), ankylosing spondylitis (AS) (study HuangR_2020), human immunodeficiency virus (HIV) infection, and pulmonary tuberculosis



(legend on next page)

(PT) (Figure 6A; Table S6). Conversely, patients with liver cirrhosis (LC), schizophrenia, and young CRC showed a noteworthy increase in fungal diversity, compared with their respective control groups.

Next, we applied PERMANOVA to the fungal species-level profiles to evaluate the effect size of each disease on the overall structure of the gut mycobiome. In 26 of the 43 case-control comparisons, disease status significantly altered the mycobiome structure (*adonis* $p < 0.05$), with effect sizes (*adonis* R^2) ranging from 0.7% to 13.9% (Figure 6A; Table S6). Patients with AS (study HuangR_2020), HIV infection, young CRC, atrial fibrillation (AF), and IBD (study WengY_2019) showed the most substantial variations in their gut mycobiomes. We found that digestive tract diseases and cancers were associated with more pronounced mycobiome alterations, followed by immune diseases, whereas most cardiometabolic diseases induced only minor alterations (Figure S7B). Furthermore, we built random forest classifiers for each disease to evaluate the effectiveness of fungal composition in distinguishing between cases and controls. The classifiers achieved high discriminatory abilities in 24 of the 43 case-control comparisons (area under the curve [AUC] > 0.70 ; Figure 6A; Table S6). Patients with AF (AUC = 0.96) showed the highest classifiability from controls, followed by those with IBD (AUC = 0.68–0.92 for five comparisons), CRC (AUC = 0.69–0.90 for four comparisons), autism spectrum disorder (ASD) (AUC = 0.89 for study WangM_2019), AS (AUC = 0.88 for study HuangR_2020), and HIV infection (AUC = 0.86).

Given that the degree of mycobiome alterations varies among different diseases, we next determined whether these alterations exhibited unique or overlapping patterns among the diseases. To accomplish this, we defined mycobiome alteration patterns for each disease as the fold-change profiles of the average relative abundance of 303 fungal species between the case and control groups. Multivariate analysis revealed that the diseases were clearly differentiated by their types at the first and second principal coordinates (PCs) (*adonis* $p = 0.008$; Figure 6B), suggesting that disease type strongly affected mycobiome alteration patterns in humans.

Shared and disease-specific gut fungal signatures

Next, we investigated fungal dysbiosis across a spectrum of diseases through disease-control comparisons using the Wilcoxon rank-sum test ($q < 0.05$). This analysis identified 522 distinct

fungal differential signatures, comprising 109 family-level, 171 genus-level, and 242 species-level signatures, spanning a wide array of diseases (Figures 6C–E; Table S7). Notably, certain taxa exhibited consistent trends, with four families, eight genera, and seven species demonstrating significant and uniform increases or decreases across multiple diseases. A subsequent random-effect meta-analysis confirmed substantial abundance variations in 18 of these 19 signatures across all studied diseases (Figure S7C; Table S7), suggesting their potential as universal disease-associated signatures. *Candida albicans* c20, a prominent member of these taxa, was significantly enriched in patients with atherosclerotic cardiovascular disease (ACVD), LC, systemic lupus erythematosus (SLE), ulcerative colitis (UC) (study YanQ_2023c), and COVID-19 (Wilcoxon rank-sum test $q < 0.05$) and demonstrated a consistent trend in various other diseases (Figure S7D). This finding aligns with those of previous studies indicating *Candida albicans* (*C. albicans*) overabundance in the stools of patients with various systemic disorders,^{57,58} suggesting that it may be a common pathogenic species in unhealthy individuals. Additionally, we identified disease-specific signatures comprising 9 families, 14 genera, and 40 species, which exhibited changes exclusive to specific disease comparisons (Table S7). Noteworthy signatures within this category include the enrichment of *Malassezia* in HIV-infected patients, a phenomenon contrary to observations in other studies,^{59,60} necessitating further investigation.

Despite the heterogeneous nature of the different immune and cardiometabolic diseases, a considerable level of consistency was observed (Figure 6B). A meta-analysis of disease types identified 32 fungal taxa as potentially shared signatures in immune diseases and 49 fungal taxa as shared signatures in cardiometabolic diseases (Table S7). This suggests that the existence of common fungal signatures transcends the traditionally distinct classifications of these disease categories.

Certain diseases may be influenced by the metabolic activity of the gut fungi. To assess this, we employed a predictive approach to elucidate the “primary metabolomic profile” (encompassing 241 primary metabolites) and “secondary metabolomic profile” (spanning 3,308 SMGC clusters from both CGF and PHF catalogs) across all fecal metagenomes. In 27 of the 43 disease-control comparisons, we observed significant alterations in their secondary metabolomic profiles (*adonis* $p < 0.05$), whereas only 5 comparisons showed significant changes in their primary metabolomic profiles (*adonis* $p < 0.05$;

Figure 6. Gut mycobiome characteristics and fungal signatures in common diseases

(A) Changes in gut fungal diversity (left), effect sizes of permutational multivariate analysis of variance (PERMANOVA) (middle), and within-study area under the curves (AUCs) for disease vs. control classification (right) across 43 case-control comparisons. Diseases are colored according to their respective types. Statistical significance is indicated by the Student's t test for the Shannon diversity index: * $p < 0.05$; ** $p < 0.01$. For effect size analysis, the *adonis* test with 1,000 permutations: * $p < 0.05$; ** $p < 0.01$.

(B) Distance-based redundancy analysis (dbRDA) on the Euclidean distance of gut mycobiome alteration patterns across 43 case-control comparisons. For each disease, the mycobiome alteration pattern was defined as the fold-change profile of the average relative abundance of 303 fungal species between the case and control groups. The first two primary constrained axes (CAPs, constrained analysis of principal coordinates) and their explained variances are presented. The upper and left boxplots show the disease scores in CAP1 and CAP2 (boxes show medians/quartiles; error bars extend to the most extreme values within 1.5 interquartile ranges). Wilcoxon rank-sum test: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

(C–E) Heatmap showing signatures of fungal families (C), genera (D), and species (E) in the 43 case-control comparisons. Colored squares indicate fungal taxa with significant differences (Wilcoxon rank-sum test, $q < 0.05$) in relative abundances between cases and controls in each case-control comparison: red, enriched in patients; blue, enriched in healthy controls. Additional details of these fungal signatures are provided in Table S7.

See also Figure S7.

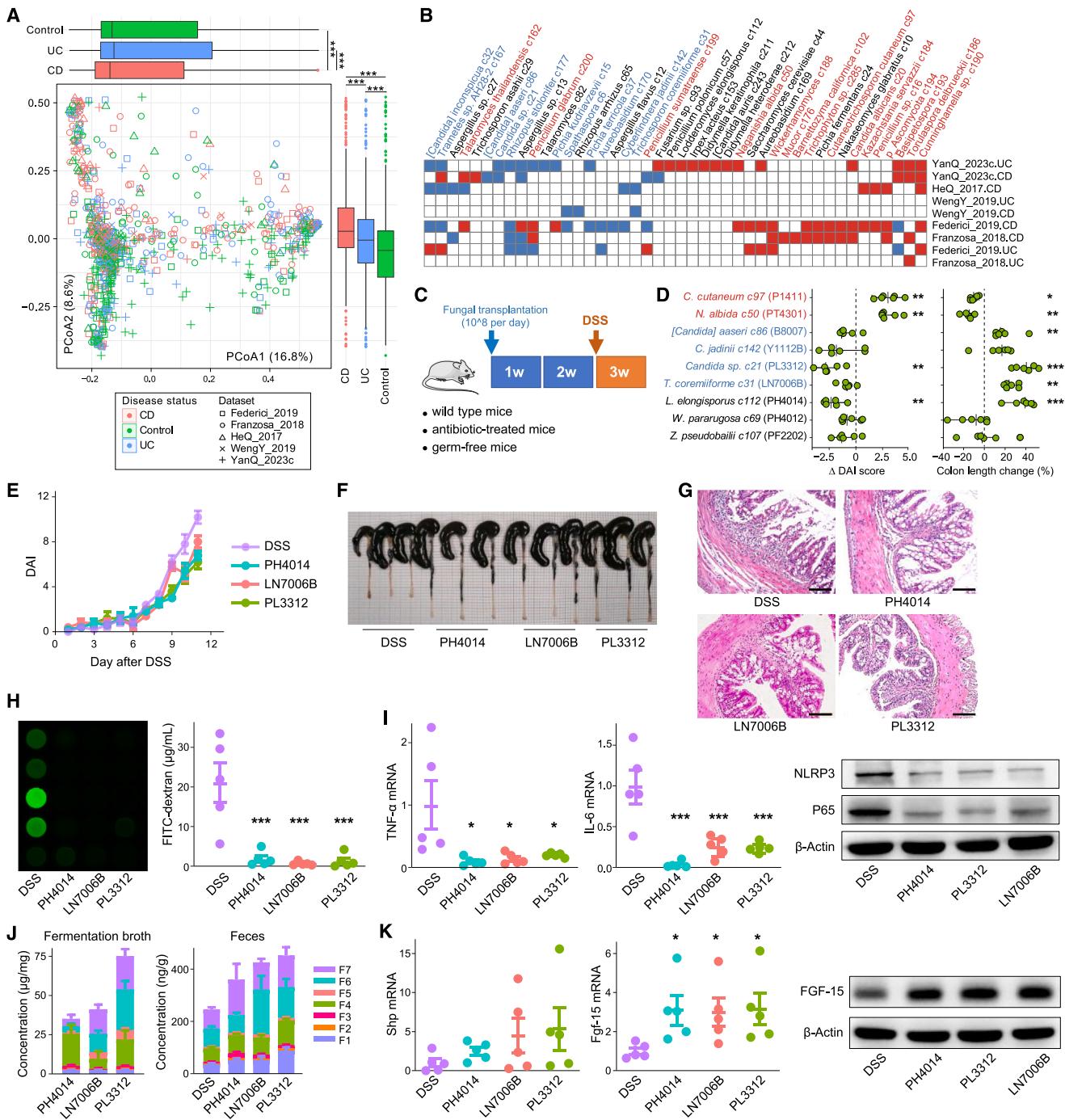


Figure 7. IBD-associated fungal signatures and the biological effects of fungal strains on DSS-induced colitis in mice

(A) Principal coordinate analysis (PCoA) of the gut mycobiome in Chinese (HeQ_2017, WengY_2019, and YanQ_2023c) and non-Chinese (Franzosa_2018 and Federici_2019) fecal metagenomic datasets. The first two principal coordinates (PCs) are shown, and their explained variances are presented. Upper and left boxplots show the disease scores in the first two PCs (boxes show medians/quartiles; error bars extend to the most extreme values within 1.5 interquartile ranges). Wilcoxon rank-sum test: *** $p < 0.001$.

(B) Heatmap showing IBD-associated gut fungal signatures. Colored squares indicate the fungal species that are significantly enriched in each case-control comparison (Wilcoxon rank-sum test: $q < 0.05$): red, enriched in IBD patients; blue, enriched in healthy controls. The colors of the fungal species names indicate enrichment in meta-analysis: red, enriched in diseases; blue, enriched in controls; and black, not significant.

(C) Schematic diagram of fungal colonization and the construction of a DSS-induced mouse model of colitis.

(legend continued on next page)

Figure S7E). This finding suggests that the effect of gut fungi on diseases may be more pronounced through secondary rather than primary metabolomes. Additionally, we observed a notable enrichment of the total SMGC gene abundance in the fecal metagenomes of patients with Parkinson's disease (PD) (study QianY_2020), PT, gout, SLE, and HIV infection, compared with that in healthy individuals (**Figure S7E**), which highlights the potential role of fungal secondary metabolism in contributing to the metabolic signatures associated with these diseases.

Exploration of IBD-associated gut fungi by global samples and animal experiments

To investigate the fungal signatures in individuals with IBD, we conducted an additional analysis of the gut mycobiome in fecal metagenomes from three Chinese datasets,^{61,62} including the YanQ_2023c dataset generated in this study (comprising 71 patients and 166 healthy controls, as detailed in **Table S5**), along with two non-Chinese datasets.^{63,64} Multivariate analysis revealed that the disease status of the subjects exhibited a weak but significant impact on the overall fungal composition across the different datasets (PERMANOVA effect size = 1.2%, $p < 0.001$; **Figure 7A**). Disease-control comparisons within each dataset allowed for the identification of a plethora of IBD-associated fungal signatures spanning 15 families, 24 genera, and 44 species (Wilcoxon rank-sum test, $q < 0.05$; **Figures 7B** and **S8A**). Notably, 62.7% (52/83) of these signatures were validated through a meta-analysis across five datasets, suggesting their potential as common fungal features of IBD (**Table S7**). These shared signatures demonstrated robust discriminatory power between patients and healthy individuals within each dataset, as evidenced by an average AUC of 0.80 (**Figure S8B**).

To verify the gut fungal signatures of IBD, nine cultivated strains from two IBD-enriched, four IBD-depleted, and three non-significant species were selected for oral transplantation into wild-type mice to evaluate their biological effects on dextran sulfate sodium (DSS)-induced colitis (**Figure 7C**). We then analyzed the observed colitis-related symptoms, such as disease activity index (DAI) and colon length, and found that the two IBD-enriched species significantly aggravated colitis in mice, whereas three out of four IBD-depleted species and one out of three non-significant species partly attenuated these symptoms (**Figures 7D** and **S8C**). The protective effects of three species, including the IBD-depleted *Trichosporon coremiiforme* c31 (strain LN7006B) and *Candida* sp. c21 (strain PL3312) and the non-significant *Lodderomyces elongisporus* c112 (strain

PH4014), against DSS-induced colitis were further confirmed in both germ-free mice and antibiotic-treated microbe-depleted mice (**Figures 7E–7I** and **S8D**). Consistently, the total metabolites prepared from the fermentation of these species also revealed their potential in the remission of DSS-induced colitis in mice (**Figure S8E**). To explore the bioactive molecules and action pathways of these fungi, we isolated secondary metabolites from their fermentation and identified seven acyclic sesquiterpenes that were significantly enriched in the feces of mice subjected to fungal transplantation (**Figures 7J** and **S8F**). As demonstrated by our previous study, fungal-derived acyclic sesquiterpenes attenuated experimental colitis by activating the farnesoid X receptor (FXR).⁶⁵ The current three fungal strains likely operate through a similar mechanism, as we observed significant upregulation of FXR target gene (i.e., *Fgf15* and *Shp*) expression in the colon tissues of fungal-transplanted germ-free and wild-type mice (**Figures 7K**, **S8C**, and **S8E**). Overall, the discovery and experimental validation of IBD-associated fungal signatures hold promise for advancing future research into their mechanistic and biological effects on diseases.

DISCUSSION

Despite comprising only a small portion (0.01%–1%) of the total microbiome, fungal populations in the human gut have a considerable influence on various physiological and pathological processes. To understand the significance of the gut mycobiome, it is necessary to identify, cultivate, and analyze the genomes of the fungal strains. Our extensive cultivation and sequencing efforts have led to the compilation of a large catalog of 760 annotated genomes from gut fungal species, which has enabled the identification of numerous genes, functions, and metabolites that contribute to the biological and ecological traits of this phylogenetically diverse group of microorganisms. Our CGF catalog has significantly expanded the genome repertoire of gut fungal species and provides a broad representation of the “known gut fungi” previously identified through amplicon- or metagenome-based studies.^{66,67}

In addition to commonly observed mucosa-enriched fungi such as *Candida* and *Saccharomyces* spp., the present CGF catalog includes a variety of luminal-associated fungi, including *Aspergillus* and *Cladosporium* spp. Although luminal-associated fungi are not highly active in the gut,⁶⁸ our genome-level functional analysis showed that some of them, primarily belonging to Pezizomycotina, may play pivotal metabolic roles. These fungi

(D) Evaluation of the disease activity index (DAI) and length change of colons in wild-type mice under fungal transplantation when compared with sham-gavaged mice. For each fungal species and the sham group ($n = 6$), mice were orally administered the fungus for a duration of 2 weeks, with gavage performed once daily at a dosage of 10^8 colony-forming units (CFUs)/mouse. Wilcoxon rank-sum test: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

(E) DAI is recorded and assessed daily in germ-free mice for fungal transplantation.

(F) Gross morphology of the colon in germ-free mice.

(G) Representative H&E staining images of the colon. Scale bars, 100 μm .

(H) Intestinal permeability is measured by the translocation of FITC-dextran from the intestinal lumen into the blood after oral administration for 4 h.

(I) Levels of inflammatory biomarkers in the colon.

(J) The metabolites (F1–F7) are identified in fungal fermentation broth (left) and the feces of germ-free mice (right).

(K) qPCR and western blot analyses of FXR target genes in the colon.

Data are presented as mean \pm SEM for (H)–(K) ($n = 5$). Ordinary one-way ANOVA is used to determine statistical significance, compared with the sham-gavaged DSS group: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

See also **Figure S8**.

exhibit a broad spectrum of specific metabolic functions such as nitrate assimilation and the degradation of certain amino acids (e.g., tyrosine, tryptophan, and methionine), plant polysaccharides, and lipases. In addition, they possess a predominant content of SMGCs that are involved in the production of secondary metabolites (i.e., mycotoxins), which may contribute to competition with other gut microbes and cause chronic inflammation or diseases.^{69–71}

Creating a fungal reference genome database presents a significant advantage for accurately measuring the composition of the fungal community in fecal metagenome datasets.⁷² By utilizing the CGF and PHF catalogs and an updated profiling approach, we conducted integrated analyses of gut mycobiome characteristics across over 11,000 fecal metagenomes from Chinese and non-Chinese populations. This allowed us to identify differences in mycobiome diversity and taxonomic and functional shifts between healthy individuals and patients across 28 disease or unhealthy statuses. Notably, only a small percentage (~20%) of fungal species were found to be prevalent, with the majority being limited to a few individuals. Conversely, several species in the CGFs were highly prevalent in the mycobiomes of both human populations, and compositional analysis revealed consistency between the two populations in the high-abundance species. These findings not only indicate the similarity of core fungi across different human populations but also demonstrate the broad applicability of combining CGFs with PHFs in profiling fecal metagenomes.

Our investigation revealed a reduced gut mycobiome diversity in patients with various common diseases. This decline in diversity has been consistently observed in prior studies,^{24,73–75} suggesting compromised resilience of the microbial ecosystem.⁷⁶ Significant mycobiome structural alterations have also been found across a spectrum of diseases, notably in patients with IBD, CRC, and infectious diseases. These trends align with existing literature reporting significant changes in the gut mycobiomes of IBD,^{24,77–79} CRC,^{25,80,81} and HIV infection.⁵⁹ Similarly, considerable variations have been observed in patients with immune diseases, such as AS, SLE, and gout, which have been minimally explored in terms of their gut mycobiome characteristics. In contrast to previous studies that have reported alterations in the mycobiome in obesity,⁸² T2D,²⁶ and cardiovascular diseases,⁸³ our results suggest that the extent of these changes is relatively limited. These insights enable the categorization of gut mycobiome-associated diseases according to the degree and nature of alterations, which is essential for understanding the disease etiology and developing effective treatments.

We identified a series of fungal taxa that exhibited distinct abundance patterns between patient and control mycobiomes in various disease datasets, thereby revealing disease-associated variations in the gut fungal community. More importantly, we identified fungal signatures that may be universal across a range of common diseases, characterized by the enrichment of opportunistic pathogens, such as *C. albicans*, and a decrease in beneficial microbes. Consistent with our study, the presence of *C. albicans* in the gut mycobiome has been observed in diseases including cancer,⁵⁸ schizophrenia,⁸⁴ COVID-19,⁸⁵ and alcohol-associated liver disease.⁸⁶ It is believed to play a role in disease pathogenesis by driving T helper 17 (Th17)-medi-

ated immune responses and disrupting the gut microbiome homeostasis.^{68,87} Additionally, we conducted a comprehensive examination of global fecal metagenomes and identified a distinct set of fungi associated with IBD. Animal experiments have revealed that certain IBD-enriched fungi can exacerbate colitis, while other IBD-depleted fungi can alleviate it, partially validating the accuracy of these signatures. Mechanistically, we found that the IBD-depleted species alleviate colitis possibly by secreting acyclic sesquiterpene metabolites as FXR agonists. This mechanism is consistent with our previous study on another fungus,⁶⁵ suggesting that the gut fungal-derived secondary metabolites may represent a widespread mechanism for influencing disease progression. Altogether, these findings not only validate the reliability and potential interpretability of our database and approaches but also pave the way for future studies across a broader spectrum of diseases, contributing to our understanding of variations in gut mycobiomes.

The current set of CGFs, along with their corresponding genomes and metabolome datasets, represents a largely untapped resource that could be utilized for further examination of the ecology and function of the human gut mycobiome. Our results, drawn from large fecal metagenomic datasets, revealed significant collections between the gut mycobiome and a number of prevalent human illnesses. Future investigations should focus on devising novel methods to harness the potential of the gut mycobiome, such as extracting nutrients from diverse food sources, collaborating with or competing against the bacterial microbiome, and safeguarding the host from potential diseases.

Limitations of the study

The deficiency of cultivated fungi from sources outside China is a notable limitation of the current CGF catalog. Our research has uncovered the extensive distribution of CGFs in non-Chinese populations, yet gut fungi unique to these human groups remain undiscovered, and the composition of the gut fungal community is undoubtedly influenced by a range of factors, including ethnicity, lifestyle, food, and geography. Additionally, the scarcity of fungal genomes from other regions restricts our ability to examine genomic polymorphisms such as mutations and regional variations in gut fungi. Additional experimental investigations are required to confirm and extend the connections observed between diseases and differential fungal signatures. This includes additional validation using independent cohorts, experimental verification, and mechanistic studies. Our examination of IBD-associated fungi serves as an illustration; however, for other diseases, particularly those with strong interactions with the gut mycobiome, such as immune and cardiometabolic conditions, further research is required.

STAR METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Materials availability
 - Data and code availability

- **EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS**
 - Volunteers and fecal sample collection
 - Cultivation of fungal isolates
- **METHOD DETAILS**
 - Taxonomic identification of isolated fungal strains by rDNA sequencing
 - Fecal DNA extraction and whole-metagenome shotgun sequencing
 - Fungal DNA preparation and whole-genome shotgun sequencing
 - qPCR for detection of fungal strains from original feces
 - Targeted quantification of fungal primary metabolites
 - NCBI reference and available human-associated fungi data
 - Genome assembly and analysis
 - Gene prediction and analysis
 - Gene functional annotation and analysis
 - Fungal profiling algorithm and validation
 - Collection and process of public datasets
 - Fecal metagenomic datasets of this study
 - The biological effects of fungi strain on DSS-induced colitis in mice
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
 - Multivariate analyses
 - Random forest analysis
 - Comparison analyses
 - Other statistical analyses

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.cell.2024.04.043>.

ACKNOWLEDGMENTS

This work was financially supported by the National Natural Science Foundation of China (no. 82225048, 81930112, and 82374297), Dalian Science and Technology Leading Talents Project (2019RD15 and 2020RJ09), Dalian Science and Technology Innovation Fund (2023JJ12SN033), Dalian Science and Technology Young Talents Project (2022RY18), and Dalian Science and Technology benefit people project (2023JJ13SN053). F.M.M. is supported by the Laboratory of Excellence ARBRE (grant no. ANR-11-LABX-0002-01), Huazhong Agricultural University, and Chinese Academy of Forestry Institute of Subtropical Forestry.

AUTHOR CONTRIBUTIONS

X.M., C.W., S.L., and D.-a.G. planned and designed the study. Qingsong Yan, X.M., X. He, Y.S., W.Z., Z.Y., Y.A., C.Y., and C.W. collected the fecal samples and isolated the fungal strains. Qiulong Yan, F.C., and G.W. sequenced the 18S rRNA gene and ITS for fungal identification. Q.J., A.Z., and H.J. performed DNA extraction and other related experiments. S.L., Y.Z., R.G., Q.L., Z.L., X.W., and H.J. coordinated and performed genome sequencing and data analysis. X. Huo, L.F., F.W., J.N., and S.D. performed the animal experiments. X.M., F.M.M., S.L., D.-a.G., X.F., Qiulong Yan, and C.W. jointly supervised the project and contributed to the discussion. X.M., S.L., Qiulong Yan, and C.W. wrote the manuscript. X.M., F.M.M., and D.-a.G. edited the manuscript. All authors approved the final version of the paper.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: March 24, 2023

Revised: February 17, 2024

Accepted: April 29, 2024

Published: May 21, 2024

REFERENCES

1. Huffnagle, G.B., and Noverr, M.C. (2013). The emerging world of the fungal microbiome. *Trends Microbiol.* 21, 334–341. <https://doi.org/10.1016/j.tim.2013.04.002>.
2. Huseyin, C.E., O'Toole, P.W., Cotter, P.D., and Scanlan, P.D. (2017). Forgotten fungi—the gut mycobiome in human health and disease. *FEMS Microbiol. Rev.* 41, 479–511. <https://doi.org/10.1093/femsre/fuw047>.
3. Chin, V.K., Yong, V.C., Chong, P.P., Amin Nordin, S., Basir, R., and Abdulla, M. (2020). Mycobiome in the gut: A multiperspective review. *Mediators Inflamm.* 2020, 9560684. <https://doi.org/10.1155/2020/9560684>.
4. Shuai, M., Fu, Y., Zhong, H.L., Gou, W., Jiang, Z., Liang, Y., Miao, Z., Xu, J.J., Huynh, T., Wahlqvist, M.L., et al. (2022). Mapping the human gut mycobiome in middle-aged and elderly adults: multiomics insights and implications for host metabolic health. *Gut* 71, 1812–1820. <https://doi.org/10.1136/gutjnl-2021-326298>.
5. Zhang, F., Aschenbrenner, D., Yoo, J.Y., and Zuo, T. (2022). The gut mycobiome in health, disease, and clinical applications in association with the gut bacterial microbiome assembly. *Lancet Microbe* 3, e969–e983. [https://doi.org/10.1016/S2666-5247\(22\)00203-8](https://doi.org/10.1016/S2666-5247(22)00203-8).
6. Wu, X., Xia, Y., He, F., Zhu, C., and Ren, W. (2021). Intestinal mycobiota in health and diseases: from a disrupted equilibrium to clinical opportunities. *Microbiome* 9, 60. <https://doi.org/10.1186/s40168-021-01024-x>.
7. Suhr, M.J., Banjara, N., and Hallen-Adams, H.E. (2016). Sequence-based methods for detecting and evaluating the human gut mycobiome. *Lett. Appl. Microbiol.* 62, 209–215. <https://doi.org/10.1111/lam.12539>.
8. Nash, A.K., Auchtung, T.A., Wong, M.C., Smith, D.P., Gesell, J.R., Ross, M.C., Stewart, C.J., Metcalf, G.A., Muzny, D.M., Gibbs, R.A., et al. (2017). The gut mycobiome of the Human Microbiome Project healthy cohort. *Microbiome* 5, 153. <https://doi.org/10.1186/s40168-017-0373-4>.
9. Huseyin, C.E., Rubio, R.C., O'Sullivan, O., Cotter, P.D., and Scanlan, P.D. (2017). The fungal frontier: A comparative analysis of methods used in the study of the human gut mycobiome. *Front. Microbiol.* 8, 1432. <https://doi.org/10.3389/fmicb.2017.01432>.
10. Human Microbiome Jumpstart Reference Strains Consortium, Nelson, K.E., Weinstock, G.M., Highlander, S.K., Worley, K.C., Creasy, H.H., Wortman, J.R., Rusch, D.B., Mireva, M., Sodergren, E., et al. (2010). A catalog of reference genomes from the human microbiome. *Science* 328, 994–999. <https://doi.org/10.1126/science.1183605>.
11. Zou, Y., Xue, W., Luo, G., Deng, Z., Qin, P., Guo, R., Sun, H., Xia, Y., Liang, S., Dai, Y., et al. (2019). 1,520 reference genomes from cultivated human gut bacteria enable functional microbiome analyses. *Nat. Biotechnol.* 37, 179–185. <https://doi.org/10.1038/s41587-018-0008-8>.
12. Forster, S.C., Kumar, N., Anonye, B.O., Almeida, A., Viciani, E., Stares, M.D., Dunn, M., Mkandawire, T.T., Zhu, A., Shao, Y., et al. (2019). A human gut bacterial genome and culture collection for improved metagenomic analyses. *Nat. Biotechnol.* 37, 186–192. <https://doi.org/10.1038/s41587-018-0009-7>.
13. Nielsen, H.B., Almeida, M., Juncker, A.S., Rasmussen, S., Li, J., Sunagawa, S., Plichta, D.R., Gautier, L., Pedersen, A.G., Le Chatelier, E., et al. (2014). Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. *Nat. Biotechnol.* 32, 822–828. <https://doi.org/10.1038/nbt.2939>.
14. Nayfach, S., Shi, Z.J., Seshadri, R., Pollard, K.S., and Kyrpides, N.C. (2019). New insights from uncultivated genomes of the global human gut microbiome. *Nature* 568, 505–510. <https://doi.org/10.1038/s41586-019-1058-x>.
15. Almeida, A., Mitchell, A.L., Boland, M., Forster, S.C., Gloor, G.B., Tarkowski, A., Lawley, T.D., and Finn, R.D. (2019). A new genomic blueprint of the human gut microbiota. *Nature* 568, 499–504. <https://doi.org/10.1038/s41586-019-0965-1>.

16. Almeida, A., Nayfach, S., Boland, M., Strozzi, F., Beracochea, M., Shi, Z.J., Pollard, K.S., Sakharova, E., Parks, D.H., Hugenholz, P., et al. (2021). A unified catalog of 204,938 reference genomes from the human gut microbiome. *Nat. Biotechnol.* 39, 105–114. <https://doi.org/10.1038/s41587-020-0603-3>.
17. Kim, C.Y., Lee, M., Yang, S., Kim, K., Yong, D., Kim, H.R., and Lee, I. (2021). Human reference gut microbiome catalog including newly assembled genomes from under-represented Asian metagenomes. *Genome Med.* 13, 134. <https://doi.org/10.1186/s13073-021-00950-7>.
18. Levitan, S., Shoer, S., Rothschild, D., Gorodetski, M., and Segal, E. (2022). An expanded reference map of the human gut microbiome reveals hundreds of previously unknown species. *Nat. Commun.* 13, 3863. <https://doi.org/10.1038/s41467-022-31502-1>.
19. Gouba, N., Raoult, D., and Drancourt, M. (2014). Eukaryote culturomics of the gut reveals new species. *PLOS One* 9, e106994. <https://doi.org/10.1371/journal.pone.0106994>.
20. Hamad, I., Ranque, S., Azhar, E.I., Yasir, M., Jiman-Fatani, A.A., Tissot-Dupont, H., Raoult, D., and Bittar, F. (2017). Culturomics and amplicon-based metagenomic approaches for the study of fungal population in human gut microbiota. *Sci. Rep.* 7, 16788. <https://doi.org/10.1038/s41598-017-17132-4>.
21. Olm, M.R., West, P.T., Brooks, B., Firek, B.A., Baker, R., Morowitz, M.J., and Banfield, J.F. (2019). Genome-resolved metagenomics of eukaryotic populations during early colonization of premature infants and in hospital rooms. *Microbiome* 7, 26. <https://doi.org/10.1186/s40168-019-0638-1>.
22. Suhr, M.J., and Hallen-Adams, H.E. (2015). The human gut mycobiome: pitfalls and potentials—a mycologist’s perspective. *Mycologia* 107, 1057–1073. <https://doi.org/10.3852/15-147>.
23. Auchting, T.A., Fofanova, T.Y., Stewart, C.J., Nash, A.K., Wong, M.C., Gesell, J.R., Auchting, J.M., Ajami, N.J., and Petrosino, J.F. (2018). Investigating colonization of the healthy adult gastrointestinal tract by fungi. *mSphere* 3, e00092-18. <https://doi.org/10.1128/mSphere.00092-18>.
24. Sokol, H., Leducq, V., Aschard, H., Pham, H.P., Jegou, S., Landman, C., Cohen, D., Liguori, G., Bourrier, A., Nion-Larmurier, I., et al. (2017). Fungal microbiota dysbiosis in IBD. *Gut* 66, 1039–1048. <https://doi.org/10.1136/gutjnl-2015-310746>.
25. Coker, O.O., Nakatsu, G., Dai, R.Z., Wu, W.K.K., Wong, S.H., Ng, S.C., Chan, F.K.L., Sung, J.J.Y., and Yu, J. (2019). Enteric fungal microbiota dysbiosis and ecological alterations in colorectal cancer. *Gut* 68, 654–662. <https://doi.org/10.1136/gutjnl-2018-317178>.
26. Jayasudha, R., Das, T., Kalyana Chakravarthy, S., Sai Prashanthi, G., Bhargava, A., Tyagi, M., Rani, P.K., Pappuru, R.R., and Shivaji, S. (2020). Gut mycobiomes are altered in people with type 2 diabetes mellitus and Diabetic Retinopathy. *PLoS One* 15, e0243077. <https://doi.org/10.1371/journal.pone.0243077>.
27. Demir, M., Lang, S., Hartmann, P., Duan, Y., Martin, A., Miyamoto, Y., Bondareva, M., Zhang, X., Wang, Y., Kasper, P., et al. (2022). The fecal mycobiome in non-alcoholic fatty liver disease. *J. Hepatol.* 76, 788–799. <https://doi.org/10.1016/j.jhep.2021.11.029>.
28. Jiang, L., Stärkel, P., Fan, J.G., Fouts, D.E., Bacher, P., and Schnabl, B. (2021). The gut mycobiome: a novel player in chronic liver diseases. *J. Gastroenterol.* 56, 1–11. <https://doi.org/10.1007/s00535-020-01740-5>.
29. Tiew, P.Y., Mac Aogain, M., Ali, N.A.B.M., Thng, K.X., Goh, K., Lau, K.J.X., and Chotirmall, S.H. (2020). The mycobiome in health and disease: emerging concepts, methodologies and challenges. *Mycopathologia* 185, 207–231. <https://doi.org/10.1007/s11046-019-00413-z>.
30. Spatz, M., and Richard, M.L. (2020). Overview of the potential role of *Malassezia* in gut health and disease. *Front. Cell. Infect. Microbiol.* 10, 201. <https://doi.org/10.3389/fcimb.2020.00201>.
31. Gutierrez, M.W., and Arrieta, M.C. (2021). The intestinal mycobiome as a determinant of host immune and metabolic health. *Curr. Opin. Microbiol.* 62, 8–13. <https://doi.org/10.1016/j.mib.2021.04.004>.
32. Jain, C., Rodriguez-R, L.M., Phillippe, A.M., Konstantinidis, K.T., and Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* 9, 5114. <https://doi.org/10.1038/s41467-018-07641-9>.
33. Kelkar, Y.D., and Ochman, H. (2012). Causes and consequences of genome expansion in fungi. *Genome Biol. Evol.* 4, 13–23. <https://doi.org/10.1093/gbe/evr124>.
34. Wisecaver, J.H., Slot, J.C., and Rokas, A. (2014). The evolution of fungal metabolic pathways. *PLoS Genet.* 10, e1004816. <https://doi.org/10.1371/journal.pgen.1004816>.
35. Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>.
36. Letunic, I., and Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 47, W256–W259. <https://doi.org/10.1093/nar/gkz239>.
37. Raimondi, S., Amaretti, A., Gozzoli, C., Simone, M., Righini, L., Candeliere, F., Brun, P., Ardizzone, A., Colombari, B., Paulone, S., et al. (2019). Longitudinal survey of fungi in the human gut: ITS Profiling, phenotyping, and colonization. *Front. Microbiol.* 10, 1575. <https://doi.org/10.3389/fmicb.2019.01575>.
38. Khalidi, N., Seifuddin, F.T., Turner, G., Haft, D., Nierman, W.C., Wolfe, K.H., and Fedorova, N.D. (2010). Smurf: genomic mapping of fungal secondary metabolite clusters. *Fungal Genet. Biol.* 47, 736–741. <https://doi.org/10.1016/j.fgb.2010.06.003>.
39. Fan, A., Chen, H., Wu, R., Xu, H., and Li, S.M. (2014). A new member of the DMATS superfamily from *Aspergillus niger* catalyzes prenylations of both tyrosine and tryptophan derivatives. *Appl. Microbiol. Biotechnol.* 98, 10119–10129. <https://doi.org/10.1007/s00253-014-5872-7>.
40. Soanes, D.M., Alam, I., Cornell, M., Wong, H.M., Hedeler, C., Paton, N.W., Rattray, M., Hubbard, S.J., Oliver, S.G., and Talbot, N.J. (2008). Comparative genome analysis of filamentous fungi reveals gene family expansions associated with fungal pathogenesis. *PLoS One* 3, e2300. <https://doi.org/10.1371/journal.pone.0002300>.
41. de Vries, R.P., Riley, R., Wiebenga, A., Aguilar-Osorio, G., Amillis, S., Uchima, C.A., Anderluh, G., Asadollahi, M., Askin, M., Barry, K., et al. (2017). Comparative genomics reveals high biological diversity and specific adaptations in the industrially and medically important fungal genus *Aspergillus*. *Genome Biol.* 18, 28. <https://doi.org/10.1186/s13059-017-1151-0>.
42. Gorfer, M., Blumhoff, M., Klaubauf, S., Urban, A., Inselsbacher, E., Badiani, D., Mitter, B., Sessitsch, A., Wanek, W., and Strauss, J. (2011). Community profiling and gene expression of fungal assimilatory nitrate reductases in agricultural soil. *ISME J.* 5, 1771–1783. <https://doi.org/10.1038/ismej.2011.53>.
43. Shen, X.X., Opulente, D.A., Kominek, J., Zhou, X., Steenwyk, J.L., Buh, K.V., Haase, M.A.B., Wisecaver, J.H., Wang, M., Doering, D.T., et al. (2018). Tempo and mode of genome evolution in the budding yeast subphylum. *Cell* 175, 1533–1545.e20. <https://doi.org/10.1016/j.cell.2018.10.023>.
44. Ocaña-Pallarès, E., Najle, S.R., Scazzocchio, C., and Ruiz-Trillo, I. (2019). Reticulate evolution in eukaryotes: origin and evolution of the nitrate assimilation pathway. *PLoS Genet.* 15, e1007986. <https://doi.org/10.1371/journal.pgen.1007986>.
45. Solomon, K.V., Haitjema, C.H., Henske, J.K., Gilmore, S.P., Borges-Rivera, D., Lipzen, A., Brewer, H.M., Purvine, S.O., Wright, A.T., Theodorou, M.K., et al. (2016). Early-branching gut fungi possess a large, comprehensive array of biomass-degrading enzymes. *Science* 351, 1192–1195. <https://doi.org/10.1126/science.aad1431>.
46. Moglinicka, I., and Ufnal, M. (2019). Gut mycobiota and fungal metabolites in human homeostasis. *Curr. Drug Targets* 20, 232–240. <https://doi.org/10.2174/1389450119666180724125020>.

47. Vesth, T.C., Nybo, J.L., Theobald, S., Frisvad, J.C., Larsen, T.O., Nielsen, K.F., Hoof, J.B., Brandl, J., Salamov, A., Riley, R., et al. (2018). Investigation of inter- and intraspecies variation through genome sequencing of *Aspergillus* section Nigri. *Nat. Genet.* 50, 1688–1695. <https://doi.org/10.1038/s41588-018-0246-1>.
48. Blanco-Míguez, A., Beghini, F., Cumbo, F., McIver, L.J., Thompson, K.N., Zolfo, M., Manghi, P., Dubois, L., Huang, K.D., Thomas, A.M., et al. (2023). Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4. *Nat. Biotechnol.* 41, 1633–1644. <https://doi.org/10.1038/s41587-023-01688-w>.
49. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>.
50. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. <https://doi.org/10.1038/nmeth.1923>.
51. Cheng, A.G., Ho, P.Y., Aranda-Díaz, A., Jain, S., Yu, F.B., Meng, X., Wang, M., Iakivik, M., Nagashima, K., Zhao, A., et al. (2022). Design, construction, and in vivo augmentation of a complex gut microbiome. *Cell* 185, 3617–3636.e19. <https://doi.org/10.1016/j.cell.2022.08.003>.
52. Jin, H., Quan, K., He, Q., Kwok, L.Y., Ma, T., Li, Y., Zhao, F., You, L., Zhang, H., and Sun, Z. (2023). A high-quality genome compendium of the human gut microbiome of Inner Mongolians. *Nat. Microbiol.* 8, 150–161. <https://doi.org/10.1038/s41564-022-01270-1>.
53. Carter, M.M., Olm, M.R., Merrill, B.D., Dahan, D., Tripathi, S., Spencer, S.P., Yu, F.B., Jain, S., Neff, N., Jha, A.R., et al. (2023). Ultra-deep sequencing of Hadza hunter-gatherers recovers vanishing gut microbes. *Cell* 186, 3111–3124.e13. <https://doi.org/10.1016/j.cell.2023.05.046>.
54. Costea, P.I., Hildebrand, F., Arumugam, M., Bäckhed, F., Blaser, M.J., Bushman, F.D., de Vos, W.M., Ehrlich, S.D., Fraser, C.M., Hattori, M., et al. (2018). Enterotypes in the landscape of gut microbial community composition. *Nat. Microbiol.* 3, 8–16. <https://doi.org/10.1038/s41564-017-0072-8>.
55. Pickard, J.M., Zeng, M.Y., Caruso, R., and Núñez, G. (2017). Gut microbiota: role in pathogen colonization, immune responses, and inflammatory disease. *Immunol. Rev.* 279, 70–89. <https://doi.org/10.1111/imr.12567>.
56. Gupta, V.K., Kim, M., Bakshi, U., Cunningham, K.Y., Davis, J.M., 3rd, Lazaridis, K.N., Nelson, H., Chia, N., and Sung, J. (2020). A predictive index for health status using species-level gut microbiome profiling. *Nat. Commun.* 11, 4635. <https://doi.org/10.1038/s41467-020-18476-8>.
57. Kumamoto, C.A., Gresnigt, M.S., and Hube, B. (2020). The gut, the bad and the harmless: *Candida albicans* as a commensal and opportunistic pathogen in the intestine. *Curr. Opin. Microbiol.* 56, 7–15. <https://doi.org/10.1016/j.mib.2020.05.006>.
58. Ho, J., Camilli, G., Griffiths, J.S., Richardson, J.P., Kichik, N., and Naglik, J.R. (2021). *Candida albicans* and candidalysin in inflammatory disorders and cancer. *Immunology* 162, 11–16. <https://doi.org/10.1111/imm.13255>.
59. Hamad, I., Abou Abdallah, R., Ravaux, I., Mokhtari, S., Tissot-Dupont, H., Michelle, C., Stein, A., Lagier, J.C., Raoult, D., and Bittar, F. (2018). Metabarcoding analysis of eukaryotic microbiota in the gut of HIV-infected patients. *PLoS One* 13, e0191913. <https://doi.org/10.1371/journal.pone.0191913>.
60. Gosálbez, M.J., Jiménez-Hernández, N., Moreno, E., Artacho, A., Pons, X., Ruiz-Pérez, S., Navia, B., Estrada, V., Manzano, M., Talavera-Rodríguez, A., et al. (2022). Interactions among the mycobiome, bacteriome, inflammation, and diet in people living with HIV. *Gut Microbes* 14, 2089002. <https://doi.org/10.1080/19490976.2022.2089002>.
61. He, Q., Gao, Y., Jie, Z., Yu, X., Laursen, J.M., Xiao, L., Li, Y., Li, L., Zhang, F., Feng, Q., et al. (2017). Two distinct metacommunities characterize the gut microbiota in Crohn's disease patients. *GigaScience* 6, 1–12. <https://doi.org/10.1093/gigascience/gix050>.
62. Weng, Y.J., Gan, H.Y., Li, X., Huang, Y., Li, Z.C., Deng, H.M., Chen, S.Z., Zhou, Y., Wang, L.S., Han, Y.P., et al. (2019). Correlation of diet, microbiota and metabolite networks in inflammatory bowel disease. *J. Dig. Dis.* 20, 447–459. <https://doi.org/10.1111/1751-2980.12795>.
63. Franzosa, E.A., Sirota-Madi, A., Avila-Pacheco, J., Fornelos, N., Haiser, H.J., Reinker, S., Vatanen, T., Hall, A.B., Mallick, H., McIver, L.J., et al. (2019). Gut microbiome structure and metabolic activity in inflammatory bowel disease. *Nat. Microbiol.* 4, 293–305. <https://doi.org/10.1038/s41564-018-0306-4>.
64. Federici, S., Kredo-Russo, S., Valdés-Mas, R., Kviatcovsky, D., Weinstock, E., Matiuhin, Y., Silberberg, Y., Atarashi, K., Furuichi, M., Oka, A., et al. (2022). Targeted suppression of human IBD-associated gut microbiota commensals by phage consortia for treatment of intestinal inflammation ibd. *Cell* 185, 2879–2898.e24. <https://doi.org/10.1016/j.cell.2022.07.003>.
65. Huo, X., Li, D., Wu, F., Li, S., Qiao, Y., Wang, C., Wang, Y., Zhou, C., Sun, L., Luan, Z., et al. (2022). Cultivated human intestinal fungus *Candida metapsilosis* M2006B attenuates colitis by secreting acyclic sesquiterpenoids as FXR agonists. *Gut* 71, 2205–2217. <https://doi.org/10.1136/gutjnl-2021-325413>.
66. Sun, Y., Zuo, T., Cheung, C.P., Gu, W., Wan, Y., Zhang, F., Chen, N., Zhan, H., Yeoh, Y.K., Niu, J., et al. (2021). Population-level configurations of gut mycobiome across 6 ethnicities in Urban and rural China. *Gastroenterology* 160, 272–286.e11. <https://doi.org/10.1053/j.gastro.2020.09.014>.
67. Hoffmann, C., Dollive, S., Grunberg, S., Chen, J., Li, H., Wu, G.D., Lewis, J.D., and Bushman, F.D. (2013). Archaea and fungi of the human gut microbiome: correlations with diet and bacterial residents. *PLoS One* 8, e66019. <https://doi.org/10.1371/journal.pone.0066019>.
68. Leonardi, I., Gao, I.H., Lin, W.Y., Allen, M., Li, X.V., Fiers, W.D., De Celie, M.B., Putzel, G.G., Yantiss, R.K., Johnnilla, M., et al. (2022). Mucosal fungi promote gut barrier function and social behavior via Type 17 immunity. *Cell* 185, 831–846.e14. <https://doi.org/10.1016/j.cell.2022.01.017>.
69. Kraska, R., Schubert-Ullrich, P., Molinelli, A., Sulyok, M., MacDonald, S., and Crews, C. (2008). Mycotoxin analysis: an update. *Food Addit. Contam. Part A Chem. Anal. Control Expo. Risk Assess.* 25, 152–163. <https://doi.org/10.1080/02652030701765723>.
70. Maresca, M., and Fantini, J. (2010). Some food-associated mycotoxins as potential risk factors in humans predisposed to chronic intestinal inflammatory diseases. *Toxicon* 56, 282–294. <https://doi.org/10.1016/j.toxicon.2010.04.016>.
71. Liew, W.P.P., and Mohd-Redzwan, S. (2018). Mycotoxin: its impact on gut health and microbiota. *Front. Cell. Infect. Microbiol.* 8, 60. <https://doi.org/10.3389/fcimb.2018.00060>.
72. Lind, A.L., and Pollard, K.S. (2021). Accurate and sensitive detection of microbial eukaryotes from whole metagenome shotgun sequencing. *Microbiome* 9, 58. <https://doi.org/10.1186/s40168-021-01015-y>.
73. Chehoud, C., Albenberg, L.G., Judge, C., Hoffmann, C., Grunberg, S., Bittinger, K., Baldassano, R.N., Lewis, J.D., Bushman, F.D., and Wu, G.D. (2015). Fungal signature in the gut microbiota of pediatric patients with inflammatory bowel disease. *Inflamm. Bowel Dis.* 21, 1948–1956. <https://doi.org/10.1097/MIB.0000000000000454>.
74. Li, M., Dai, B., Tang, Y., Lei, L., Li, N., Liu, C., Ge, T., Zhang, L., Xu, Y., Hu, Y., et al. (2019). Altered bacterial-fungal interkingdom networks in the guts of ankylosing spondylitis patients. *mSystems* 4, e00176-18. <https://doi.org/10.1128/mSystems.00176-18>.
75. Cirstea, M.S., Sundvick, K., Golz, E., Yu, A.C., Boutin, R.C.T., Kliger, D., Finlay, B.B., and Appel-Cresswell, S. (2021). The gut mycobiome in Parkinson's disease. *J. Parkinsons Dis.* 11, 153–158. <https://doi.org/10.3233/JPD-202237>.
76. Larsen, O.F.A., and Claassen, E. (2018). The mechanistic link between health and gut microbiota diversity. *Sci. Rep.* 8, 2183. <https://doi.org/10.1038/s41598-018-20141-6>.

77. Li, Q., Wang, C., Tang, C., He, Q., Li, N., and Li, J. (2014). Dysbiosis of gut fungal microbiota is associated with mucosal inflammation in Crohn's disease. *J. Clin. Gastroenterol.* 48, 513–523. <https://doi.org/10.1097/MCG.0000000000000035>.
78. Hoarau, G., Mukherjee, P.K., Gower-Rousseau, C., Hager, C., Chandra, J., Retuerto, M.A., Neut, C., Vermeire, S., Clemente, J., and Colombel, J.-F. (2016). Bacteriome and mycobiome interactions underscore microbial dysbiosis in familial Crohn's disease. *mBio* 7, e01250-16. <https://doi.org/10.1128/mBio.01250-16>.
79. Lam, S., Zuo, T., Ho, M., Chan, F.K.L., Chan, P.K.S., and Ng, S.C. (2019). Review article: fungal alterations in inflammatory bowel diseases. *Aliment. Pharmacol. Ther.* 50, 1159–1171. <https://doi.org/10.1111/apt.15523>.
80. Luan, C., Xie, L., Yang, X., Miao, H., Lv, N., Zhang, R., Xiao, X., Hu, Y., Liu, Y., Wu, N., et al. (2015). Dysbiosis of fungal microbiota in the intestinal mucosa of patients with colorectal adenomas. *Sci. Rep.* 5, 7980. <https://doi.org/10.1038/srep07980>.
81. Gao, R., Kong, C., Li, H., Huang, L., Qu, X., Qin, N., and Qin, H. (2017). Dysbiosis signature of mycobiota in colon polyp and colorectal cancer. *Eur. J. Clin. Microbiol. Infect. Dis.* 36, 2457–2468. <https://doi.org/10.1007/s10096-017-3085-6>.
82. Mar Rodríguez, M., Pérez, D., Javier Chaves, F., Esteve, E., Marin-Garcia, P., Xifra, G., Vendrell, J., Jové, M., Pamplona, R., Ricart, W., et al. (2015). Obesity changes the human gut mycobiome. *Sci. Rep.* 5, 14600. <https://doi.org/10.1038/srep14600>.
83. Chacón, M.R., Lozano-Bartolomé, J., Portero-Otín, M., Rodríguez, M.M., Xifra, G., Puig, J., Blasco, G., Ricart, W., Chaves, F.J., and Fernández-Real, J.M. (2018). The gut mycobiome composition is linked to carotid atherosclerosis. *Benefit. Microbes* 9, 185–198. <https://doi.org/10.3920/BM2017.0029>.
84. Severance, E.G., Gressitt, K.L., Stallings, C.R., Katsafanas, E., Schweinfurth, L.A., Savage, C.L., Adamos, M.B., Sweeney, K.M., Origoni, A.E., Khushalani, S., et al. (2016). *Candida albicans* exposures, sex specificity and cognitive deficits in schizophrenia and bipolar disorder. *NPJ Schizophr.* 2, 16018. <https://doi.org/10.1038/npszch.2016.18>.
85. Zuo, T., Zhan, H., Zhang, F., Liu, Q., Tso, E.Y.K., Lui, G.C.Y., Chen, N., Li, A., Lu, W., Chan, F.K.L., et al. (2020). Alterations in fecal fungal microbiome of patients with COVID-19 during time of hospitalization until discharge. *Gastroenterology* 159, 1302–1310.e5. <https://doi.org/10.1053/j.gastro.2020.06.048>.
86. Chu, H., Duan, Y., Lang, S., Jiang, L., Wang, Y., Llorente, C., Liu, J., Mogavero, S., Bosques-Padilla, F., Abraldes, J.G., et al. (2020). The *Candida albicans* exotoxin candidalysin promotes alcohol-associated liver disease. *J. Hepatol.* 72, 391–400. <https://doi.org/10.1016/j.jhep.2019.09.029>.
87. Hernández-Santos, N., and Gaffen, S.L. (2012). Th17 cells in immunity to *Candida albicans*. *Cell Host Microbe* 11, 425–435. <https://doi.org/10.1016/j.chom.2012.04.008>.
88. Raja, H.A., Miller, A.N., Pearce, C.J., and Oberlies, N.H. (2017). Fungal identification using molecular tools: A primer for the natural products research community. *J. Nat. Prod.* 80, 756–770. <https://doi.org/10.1021/acs.jnatprod.6b01085>.
89. Hamal, P., Vavrova, A., Mrazek, J., and Svobodova, L. (2022). Identification of filamentous fungi including dermatophytes using MALDI-TOF mass spectrometry. *Folia Microbiol.* 67, 55–61. <https://doi.org/10.1007/s12223-021-00917-6>.
90. Khodadadi, H., Karimi, L., Jalalizand, N., Adin, H., and Mirhendi, H. (2017). Utilization of size polymorphism in ITS1 and ITS2 regions for identification of pathogenic yeast species. *J. Med. Microbiol.* 66, 126–133. <https://doi.org/10.1099/jmm.0.000426>.
91. Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F.O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. <https://doi.org/10.1093/nar/gks1219>.
92. Abarenkov, K., Henrik Nilsson, R., Larsson, K.H., Alexander, I.J., Eberhardt, U., Erland, S., Høiland, K., Kjøller, R., Larsson, E., Pennanen, T., et al. (2010). The UNITE database for molecular identification of fungi—recent updates and future perspectives. *New Phytol.* 186, 281–285. <https://doi.org/10.1111/j.1469-8137.2009.03160.x>.
93. Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forsslund, S.K., Cook, H., Mende, D.R., Letunic, I., Rattei, T., Jensen, L.J., et al. (2019). eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* 47, D309–D314. <https://doi.org/10.1093/nar/gky1085>.
94. Mitchell, A.L., Attwood, T.K., Babbitt, P.C., Blum, M., Bork, P., Bridge, A., Brown, S.D., Chang, H.Y., El-Gebali, S., Fraser, M.I., et al. (2019). InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res.* 47, D351–D360. <https://doi.org/10.1093/nar/gky1100>.
95. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 45, D353–D361. <https://doi.org/10.1093/nar/gkw1092>.
96. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., and Henriksen, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, D490–D495. <https://doi.org/10.1093/nar/gkt1178>.
97. Rawlings, N.D., Barrett, A.J., and Finn, R. (2016). Twenty years of the *Merops* database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* 44, D343–D350. <https://doi.org/10.1093/nar/gkv1118>.
98. Fischer, M., and Pleiss, J. (2003). The Lipase Engineering Database: a navigation and analysis tool for protein families. *Nucleic Acids Res.* 31, 319–321. <https://doi.org/10.1093/nar/gkg015>.
99. Chen, B.D., Jia, X.M., Xu, J.Y., Zhao, L.D., Ji, J.Y., Wu, B.X., Ma, Y., Li, H., Zuo, X.X., Pan, W.Y., et al. (2021). An autoimmunogenic and proinflammatory profile defined by the gut microbiota of patients with untreated systemic lupus erythematosus. *Arthritis Rheumatol.* 73, 232–243. <https://doi.org/10.1002/art.41511>.
100. Chu, Y., Sun, S., Huang, Y., Gao, Q., Xie, X., Wang, P., Li, J., Liang, L., He, X., Jiang, Y., et al. (2021). Metagenomic analysis revealed the potential role of gut microbiome in gout. *npj Biofilms Microbiomes* 7, 66. <https://doi.org/10.1038/s41522-021-00235-2>.
101. Gao, R., Zhu, Y., Kong, C., Xia, K., Li, H., Zhu, Y., Zhang, X., Liu, Y., Zhong, H., Yang, R., et al. (2021). Alterations, interactions, and diagnostic potential of gut bacteria and viruses in colorectal cancer. *Front. Cell. Infect. Microbiol.* 11, 657867. <https://doi.org/10.3389/fcimb.2021.657867>.
102. Huang, R., Li, F., Zhou, Y., Zeng, Z., He, X., Fang, L., Pan, F., Chen, Y., Lin, J., Li, J., et al. (2020). Metagenome-wide association study of the alterations in the intestinal microbiome composition of ankylosing spondylitis patients and the effect of traditional and herbal treatment. *J. Med. Microbiol.* 69, 797–805. <https://doi.org/10.1099/jmm.0.001107>.
103. Hu, Y., Feng, Y., Wu, J., Liu, F., Zhang, Z., Hao, Y., Liang, S., Li, B., Li, J., Lv, N., et al. (2019). The gut microbiome signatures discriminate healthy from pulmonary tuberculosis patients. *Front. Cell. Infect. Microbiol.* 9, 90. <https://doi.org/10.3389/fcimb.2019.00090>.
104. Jie, Z., Xia, H., Zhong, S.-L., Feng, Q., Li, S., Liang, S., Zhong, H., Liu, Z., Gao, Y., Zhao, H., et al. (2017). The gut microbiome in atherosclerotic cardiovascular disease. *Nat. Commun.* 8, 845. <https://doi.org/10.1038/s41467-017-00900-1>.
105. Li, J., Zhao, F., Wang, Y., Chen, J., Tao, J., Tian, G., Wu, S., Liu, W., Cui, Q., Geng, B., et al. (2017). Gut microbiota dysbiosis contributes to the development of hypertension. *Microbiome* 5, 14. <https://doi.org/10.1186/s40168-016-0222-x>.
106. Liu, R., Hong, J., Xu, X., Feng, Q., Zhang, D., Gu, Y., Shi, J., Zhao, S., Liu, W., Wang, X., et al. (2017). Gut microbiome and serum metabolome

- alterations in obesity and after weight-loss intervention. *Nat. Med.* 23, 859–868. <https://doi.org/10.1038/nm.4358>.
107. Lu, W., Feng, Y., Jing, F., Han, Y., Lyu, N., Liu, F., Li, J., Song, X., Xie, J., Qiu, Z., et al. (2018). Association between gut microbiota and CD4 recovery in HIV-1 infected patients. *Front. Microbiol.* 9, 1451. <https://doi.org/10.3389/fmicb.2018.01451>.
108. Qian, Y., Yang, X., Xu, S., Huang, P., Li, B., Du, J., He, Y., Su, B., Xu, L.-M., Wang, L., et al. (2020). Gut metagenomics-derived genes as potential biomarkers of Parkinson's disease. *Brain* 143, 2474–2489. <https://doi.org/10.1093/brain/awaa201>.
109. Qin, J., Li, Y., Cai, Z., Li, S., Zhu, J., Zhang, F., Liang, S., Zhang, W., Guan, Y., Shen, D., et al. (2012). A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* 490, 55–60. <https://doi.org/10.1038/nature11450>.
110. Qin, N., Yang, F., Li, A., Prifti, E., Chen, Y., Shao, L., Guo, J., Le Chatelier, E., Yao, J., Wu, L., et al. (2014). Alterations of the human gut microbiome in liver cirrhosis. *Nature* 513, 59–64. <https://doi.org/10.1038/nature13568>.
111. Qi, X., Yun, C., Sun, L., Xia, J., Wu, Q., Wang, Y., Wang, L., Zhang, Y., Liang, X., Wang, L., et al. (2019). Gut microbiota-bile acid-interleukin-22 axis orchestrates polycystic ovary syndrome. *Nat. Med.* 25, 1225–1233. <https://doi.org/10.1038/s41591-019-0509-0>.
112. Wang, Q., Sun, Q., Li, X., Wang, Z., Zheng, H., Ju, Y., Guo, R., Peng, S., and Jia, H. (2021). Linking gut microbiome to bone mineral density: a shotgun metagenomic dataset from 361 elderly women. *Gigabyte* 2021, gigabyte12. <https://doi.org/10.46471/gigabyte.12>.
113. Wen, C., Zheng, Z., Shao, T., Liu, L., Xie, Z., Le Chatelier, E., He, Z., Zhong, W., Fan, Y., Zhang, L., et al. (2017). Quantitative metagenomics reveals unique gut microbiome biomarkers in ankylosing spondylitis. *Genome Biol.* 18, 142. <https://doi.org/10.1186/s13059-017-1271-6>.
114. Yan, Q., Gu, Y., Li, X., Yang, W., Jia, L., Chen, C., Han, X., Huang, Y., Zhao, L., Li, P., et al. (2017). Alterations of the gut microbiome in hypertension. *Front. Cell. Infect. Microbiol.* 7, 381. <https://doi.org/10.3389/fcimb.2017.00381>.
115. Yeoh, Y.K., Zuo, T., Lui, G.C.-Y., Zhang, F., Liu, Q., Li, A.Y., Chung, A.C., Cheung, C.P., Tso, E.Y., Fung, K.S., et al. (2021). Gut microbiota composition reflects disease severity and dysfunctional immune responses in patients with COVID-19. *Gut* 70, 698–706. <https://doi.org/10.1136/gutjnl-2020-323020>.
116. Ye, Z., Zhang, N., Wu, C., Zhang, X., Wang, Q., Huang, X., Du, L., Cao, Q., Tang, J., Zhou, C., et al. (2018). A metagenomic study of the gut microbiome in Behcet's disease. *Microbiome* 6, 135. <https://doi.org/10.1186/s40168-018-0520-6>.
117. Ye, Z., Wu, C., Zhang, N., Du, L., Cao, Q., Huang, X., Tang, J., Wang, Q., Li, F., Zhou, C., et al. (2020). Altered gut microbiome composition in patients with Vogt-Koyanagi-Harada disease. *Gut Microbes* 11, 539–555. <https://doi.org/10.1080/19490976.2019.1700754>.
118. Yu, J., Feng, Q., Wong, S.H., Zhang, D., Liang, Q.Y., Qin, Y., Tang, L., Zhao, H., Stenvang, J., Li, Y., et al. (2017). Metagenomic analysis of faecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer. *Gut* 66, 70–78. <https://doi.org/10.1136/gutjnl-2015-309800>.
119. Zeng, Q., Yang, Z., Wang, F., Li, D., Liu, Y., Wang, D., Zhao, X., Li, Y., Wang, Y., Feng, X., et al. (2021). Association between metabolic status and gut microbiome in obese populations. *Microb. Genom.* 7, 000639. <https://doi.org/10.1099/mgen.0.000639>.
120. Zhang, X., Zhang, D., Jia, H., Feng, Q., Wang, D., Liang, D., Wu, X., Li, J., Tang, L., Li, Y., et al. (2015). The oral and gut microbiomes are perturbed in rheumatoid arthritis and partly normalized after treatment. *Nat. Med.* 21, 895–905. <https://doi.org/10.1038/nm.3914>.
121. Zhong, H., Ren, H., Lu, Y., Fang, C., Hou, G., Yang, Z., Chen, B., Yang, F., Zhao, Y., Shi, Z., et al. (2019). Distinct gut metagenomics and metaproteomics signatures in prediabetics and treatment-naïve type 2 diabetics. *EBioMedicine* 47, 373–383. <https://doi.org/10.1016/j.ebiom.2019.08.048>.
122. Zhou, C., Zhao, H., Xiao, X.-Y., Chen, B.D., Guo, R.-J., Wang, Q., Chen, H., Zhao, L.-D., Zhang, C.-C., Jiao, Y.-H., et al. (2020). Metagenomic profiling of the pro-inflammatory gut microbiota in ankylosing spondylitis. *J. Autoimmun.* 107, 102360. <https://doi.org/10.1016/j.jaut.2019.102360>.
123. Zhu, F., Ju, Y., Wang, W., Wang, Q., Guo, R., Ma, Q., Sun, Q., Fan, Y., Xie, Y., Yang, Z., et al. (2020). Metagenome-wide association of gut microbiome features for schizophrenia. *Nat. Commun.* 11, 1612. <https://doi.org/10.1038/s41467-020-15457-9>.
124. Zhu, J., Liao, M., Yao, Z., Liang, W., Li, Q., Liu, J., Yang, H., Ji, Y., Wei, W., Tan, A., et al. (2018). Breast cancer in postmenopausal women is associated with an altered gut metagenome. *Microbiome* 6, 136. <https://doi.org/10.1186/s40168-018-0515-3>.
125. Zhu, Q., Hou, Q., Huang, S., Ou, Q., Huo, D., Vázquez-Baeza, Y., Cen, C., Cantu, V., Estaki, M., Chang, H., et al. (2021). Compositional and genetic alterations in Graves' disease gut microbiome reveal specific diagnostic biomarkers. *ISME J.* 15, 3399–3411. <https://doi.org/10.1038/s41396-021-01016-7>.
126. Zuo, K., Li, J., Li, K., Hu, C., Gao, Y., Chen, M., Hu, R., Liu, Y., Chi, H., Wang, H., et al. (2019). Disordered gut microbiota and alterations in metabolic patterns are associated with atrial fibrillation. *GigaScience* 8, giz058. <https://doi.org/10.1093/gigascience/giz058>.
127. Brooks, B., Olm, M.R., Firek, B.A., Baker, R., Thomas, B.C., Morowitz, M.J., and Banfield, J.F. (2017). Strain-resolved analysis of hospital rooms and infants reveals overlap between the human and room microbiome. *Nat. Commun.* 8, 1814. <https://doi.org/10.1038/s41467-017-02018-w>.
128. Chen, L., Zheng, T., Yang, Y., Chaudhary, P.P., Teh, J.P.Y., Cheon, B.K., Moses, D., Schuster, S.C., Schlundt, J., Li, J., and Conway, P.L. (2022). Integrative multiomics analysis reveals host-microbe-metabolite interplays associated with the aging process in Singaporeans. *Gut Microbes* 14, 2070392. <https://doi.org/10.1080/19490976.2022.2070392>.
129. Chopyk, J., Cobán Güemes, A.G., Ramirez-Sánchez, C., Attai, H., Ly, M., Jones, M.B., Liu, R., Liu, C., Yang, K., Tu, X.M., et al. (2023). Common antibiotics, azithromycin and amoxicillin, affect gut metagenomics within a household. *BMC Microbiol.* 23, 206. <https://doi.org/10.1186/s12866-023-02949-z>.
130. Feng, Q., Liang, S., Jia, H., Stadlmayr, A., Tang, L., Lan, Z., Zhang, D., Xia, H., Xu, X., Jie, Z., et al. (2015). Gut microbiome development along the colorectal adenoma-carcinoma sequence. *Nat. Commun.* 6, 6528. <https://doi.org/10.1038/ncomms7528>.
131. Lloyd-Price, J., Mahurkar, A., Rahnavard, G., Crabtree, J., Orvis, J., Hall, A.B., Brady, A., Creasy, H.H., McCracken, C., Giglio, M.G., et al. (2017). Strains, functions and dynamics in the expanded Human Microbiome Project. *Nature* 550, 61–66. <https://doi.org/10.1038/nature23889>.
132. Karlsson, F.H., Tremaroli, V., Nookaew, I., Bergström, G., Behre, C.J., Fagerberg, B., Nielsen, J., and Bäckhed, F. (2013). Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature* 498, 99–103. <https://doi.org/10.1038/nature12198>.
133. Kokai-Kun, J.F., Le, C., Trout, K., Cope, J.L., Ajami, N.J., Degar, A.J., and Connelly, S. (2020). Ribaxamase, an orally administered beta-lactamase, diminishes changes to acquired antimicrobial resistance of the gut resistome in patients treated with ceftriaxone. *Infect. Drug Resist.* 13, 2521–2535. <https://doi.org/10.2147/IDR.S260258>.
134. Lee, J.W.J., Plichta, D.R., Asher, S., Delsignore, M., Jeong, T., McGoldrick, J., Staller, K., Khalili, H., Xavier, R.J., and Chung, D.C. (2023). Association of distinct microbial signatures with premalignant colorectal adenomas. *Cell Host Microbe* 31, 827–838.e3. <https://doi.org/10.1016/j.chom.2023.04.007>.
135. Li, J., Jia, H., Cai, X., Zhong, H., Feng, Q., Sunagawa, S., Arumugam, M., Kultima, J.R., Prifti, E., Nielsen, T., et al. (2014). An integrated catalog of reference genes in the human gut microbiome. *Nat. Biotechnol.* 32, 834–841. <https://doi.org/10.1038/nbt.2942>.

136. Madi, N., Cato, E.T., Sayeed, M.A., Islam, K., Khabir, M.I.U., Bhuiyan, M.T., Begum, Y., Rashid, M., Creasy-Marrazzo, A., and Brinkley, L.J.b. (2023). Phage predation and antibiotic exposure are inversely associated with disease severity and shape pathogen genetic diversity in cholera patients. Preprint at bioRxiv. <https://doi.org/10.1101/2023.06.14.544933>.
137. Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., et al. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464, 59–65. <https://doi.org/10.1038/nature08821>.
138. Vatanen, T., Kostic, A.D., d'Hennezel, E., Siljander, H., Franzosa, E.A., Yassour, M., Kolde, R., Vlamakis, H., Arthur, T.D., Hämäläinen, A.M., et al. (2016). Variation in microbiome LPS immunogenicity contributes to autoimmunity in humans. *Cell* 165, 1551. <https://doi.org/10.1016/j.cell.2016.05.056>.
139. Yachida, S., Mizutani, S., Shiroma, H., Shiba, S., Nakajima, T., Sakamoto, T., Watanabe, H., Masuda, K., Nishimoto, Y., Kubo, M., et al. (2019). Metagenomic and metabolomic analyses reveal distinct stage-specific phenotypes of the gut microbiota in colorectal cancer. *Nat. Med.* 25, 968–976. <https://doi.org/10.1038/s41591-019-0458-7>.
140. Zeller, G., Tap, J., Voigt, A.Y., Sunagawa, S., Kultima, J.R., Costea, P.I., Amiot, A., Böhm, J., Brunetti, F., Habermann, N., et al. (2014). Potential of fecal microbiota for early-stage detection of colorectal cancer. *Mol. Syst. Biol.* 10, 766. <https://doi.org/10.1525/msb.20145645>.
141. Zhong, H., Penders, J., Shi, Z., Ren, H., Cai, K., Fang, C., Ding, Q., Thijs, C., Blaak, E.E., Stehouwer, C.D.A., et al. (2019). Impact of early events and lifestyle on the gut microbiota and metabolic phenotypes in young school-age children. *Microbiome* 7, 2. <https://doi.org/10.1186/s40168-018-0608-z>.
142. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009). Blast+: architecture and applications. *BMC Bioinformatics* 10, 421. <https://doi.org/10.1186/1471-2105-10-421>.
143. Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prijibelski, A.D., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. <https://doi.org/10.1089/cmb.2012.0021>.
144. Boetzer, M., Henkel, C.V., Jansen, H.J., Butler, D., and Pirovano, W. (2011). Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27, 578–579. <https://doi.org/10.1093/bioinformatics/btq683>.
145. Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>.
146. Seppey, M., Manni, M., and Zdobnov, E.M. (2019). BUSCO: assessing genome assembly and annotation completeness. *Methods Mol. Biol.* 1962, 227–245. https://doi.org/10.1007/978-1-4939-9173-0_14.
147. Lagesen, K., Hallin, P., Rødland, E.A., Staerfeldt, H.H., Rognes, T., and Ussery, D.W. (2007). RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35, 3100–3108. <https://doi.org/10.1093/nar/gkm160>.
148. Keller, O., Kollmar, M., Stanke, M., and Waack, S. (2011). A novel hybrid gene prediction method employing protein multiple sequence alignments. *Bioinformatics* 27, 757–763. <https://doi.org/10.1093/bioinformatics/btr010>.
149. Brúna, T., Lomsadze, A., and Borodovsky, M. (2020). GeneMark-EP+: eukaryotic gene prediction with self-training in the space of genes and proteins. *NAR Genom. Bioinform.* 2, lqaa026. <https://doi.org/10.1093/nargab/lqaa026>.
150. Lomsadze, A., Ter-Hovhannisyan, V., Chernoff, Y.O., and Borodovsky, M. (2005). Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Res.* 33, 6494–6506. <https://doi.org/10.1093/nar/gki937>.
151. Holt, C., and Yandell, M. (2011). MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* 12, 491. <https://doi.org/10.1186/1471-2105-12-491>.
152. Steinegger, M., and Söding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.* 35, 1026–1028. <https://doi.org/10.1038/nbt.3988>.
153. Huerta-Cepas, J., Forslund, K., Coelho, L.P., Szklarczyk, D., Jensen, L.J., von Mering, C., and Bork, P. (2017). Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Mol. Biol. Evol.* 34, 2115–2122. <https://doi.org/10.1093/molbev/msx148>.
154. Buchfink, B., Xie, C., and Huson, D.H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60. <https://doi.org/10.1038/nmeth.3176>.
155. Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. <https://doi.org/10.1093/molbev/mst010>.
156. Price, M.N., Dehal, P.S., and Arkin, A.P. (2010). FastTree 2 – Approximately maximum-likelihood trees for large alignments. *PLoS One* 5, e9490. <https://doi.org/10.1371/journal.pone.0009490>.
157. Dray, S., and Dufour, A.-B. (2007). The ade4 package: implementing the duality diagram for ecologists. *J. Stat. Software* 22, 1–20. <https://doi.org/10.18637/jss.v022.i04>.
158. Villanueva, R.A.M., and Chen, Z.J. (2019). *ggplot2: Elegant Graphics for Data Analysis* (Taylor & Francis).
159. Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *J. Veg. Sci.* 14, 927–930. <https://doi.org/10.1111/j.1654-1103.2003.tb02228.x>.
160. Willis, K.A., Purvis, J.H., Myers, E.D., Aziz, M.M., Karabayir, I., Gomes, C.K., Peters, B.M., Akbilgic, O., Talati, A.J., and Pierre, J.F. (2019). Fungi form interkingdom microbial communities in the primordial human gut that develop with gestational age. *FASEB J.* 33, 12825–12837. <https://doi.org/10.1096/fj.201901436RR>.
161. Timm, C.M., Loomis, K., Stone, W., Mehoke, T., Brensingler, B., Pellicore, M., Staniczenko, P.P.A., Charles, C., Nayak, S., and Karig, D.K. (2020). Isolation and characterization of diverse microbial representatives from the human skin microbiome. *Microbiome* 8, 58. <https://doi.org/10.1186/s40168-020-00831-y>.
162. Leonardi, I., Li, X., Semon, A., Li, D., Doron, I., Putzel, G., Bar, A., Prieto, D., Rescigno, M., McGovern, D.P.B., et al. (2018). CX3CR1(+) mononuclear phagocytes control immunity to intestinal fungi. *Science* 359, 232–236. <https://doi.org/10.1126/science.aao1503>.
163. Xie, G., Wang, L., Chen, T., Zhou, K., Zhang, Z., Li, J., Sun, B., Guo, Y., Wang, X., Wang, Y., et al. (2021). A metabolite array technology for precision medicine. *Anal. Chem.* 93, 5709–5717. <https://doi.org/10.1021/acs.analchem.0c04686>.
164. Wang, Z., Huang, P., You, R., Sun, F., and Zhu, S. (2023). MetaBinner: a high-performance and stand-alone ensemble binning method to recover individual genomes from complex microbial communities. *Genome Biol.* 24, 1. <https://doi.org/10.1186/s13059-022-02832-6>.
165. Gregory, A.C., Zayed, A.A., Conceição-Neto, N., Temperton, B., Bolduc, B., Alberti, A., Ardyna, M., Arkhipova, K., Carmichael, M., Cruaud, C., et al. (2019). Marine DNA viral macro- and microdiversity from pole to pole. *Cell* 177, 1109–1123.e14. <https://doi.org/10.1016/j.cell.2019.03.040>.
166. Tettelin, H., Riley, D., Cattuto, C., and Medini, D. (2008). Comparative genomics: the bacterial pan-genome. *Curr. Opin. Microbiol.* 11, 472–477. <https://doi.org/10.1016/j.mib.2008.09.006>.
167. Rawlings, N.D., Barrett, A.J., Thomas, P.D., Huang, X., Bateman, A., and Finn, R.D. (2018). The Merops database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res.* 46, D624–D632. <https://doi.org/10.1093/nar/gkx1134>.

168. Kjærboelling, I., Vesth, T., Frisvad, J.C., Nybo, J.L., Theobald, S., Kildgaard, S., Petersen, T.I., Kuo, A., Sato, A., Lyhne, E.K., et al. (2020). A comparative genomics study of 23 *Aspergillus* species from section Flavi. *Nat. Commun.* 11, 1106. <https://doi.org/10.1038/s41467-019-14051-y>.
169. Madeira, F., Park, Y.M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A.R.N., Potter, S.C., Finn, R.D., and Lopez, R. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* 47, W636–W641. <https://doi.org/10.1093/nar/gkz268>.
170. Maaser, C., Sturm, A., Vavricka, S.R., Kucharzik, T., Fiorino, G., Annese, V., Calabrese, E., Baumgart, D.C., Bettenworth, D., Borralho Nunes, P., et al. (2019). ECCO-ESGAR Guideline for Diagnostic Assessment in IBD Part 1: Initial diagnosis, monitoring of known IBD, detection of complications ibd. *J. Crohns Colitis* 13, 144–164. <https://doi.org/10.1093/ecco-jcc/jjy113>.
171. Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B Methodol.* 57, 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>.
172. Wirbel, J., Pyl, P.T., Kartal, E., Zych, K., Kashani, A., Milanese, A., Fleck, J.S., Voigt, A.Y., Palleja, A., Ponnudurai, R., et al. (2019). Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer. *Nat. Med.* 25, 679–689. <https://doi.org/10.1038/s41591-019-0406-6>.
173. Kuznetsova, A., Brockhoff, P.B., and Christensen, R.H.B. (2017). lmerTest package: tests in linear mixed effects models. *J. Stat. Software* 82, 1–26. <https://doi.org/10.18637/jss.v082.i13>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological Samples		
Human feces	This paper	N/A
Oligonucleotides		
NS1 (GTAGTCATATGCTTGTCTC)	Raja et al. ⁸⁸	18S primer
NS4 (CTTCCGTCAATTCCCTTAAG)	Raja et al. ⁸⁸	18S primer
ITS1F (TCCGTAGGTGAAACCTGCGG)	Hamal et al. ⁸⁹	ITS primer
ITS4 (TCCTCCGCTTATTGATATGC)	Hamal et al. ⁸⁹	ITS primer
ITS1 (TCCGTAGGTGAAACCTGCGG)	Khodadadi et al. ⁹⁰	ITS1 primer
ITS2 (GCTGCGTTCTCATCGATGC)	Khodadadi et al. ⁹⁰	ITS1 primer
NL1 (GCATATCAATAAGCGGAGGAAAAG)	Hamal et al. ⁸⁹	28S primer
NL4 (GGTCCGTGTTCAAGACGG)	Hamal et al. ⁸⁹	28S primer
Deposited Data		
NCBI fungal genomes	NCBI	https://ftp.ncbi.nlm.nih.gov/genomes/genbank/
SILVA database	Quast et al. ⁹¹	https://www.arb-silva.de/download/archive/
UNITE database	Abarenkov et al. ⁹²	https://unite.ut.ee/repository.php#uchime
UHGG database	Almeida et al. ¹⁶	http://ftp.ebi.ac.uk/pub/databases/metagenomics/mgnify_genomes
eggNOG database (v 5.0)	Huerta-Cepas et al. ⁹³	http://eggnog5.embl.de/#/app/downloads
InterPro & InterProScan	Mitchell et al. ⁹⁴	http://www.ebi.ac.uk/interpro/download/
KEGG database	Kanehisa et al. ⁹⁵	https://www.genome.jp/kegg/kegg1.html
CAZy database	Lombard et al. ⁹⁶	https://bcb.unl.edu/dbCAN2/download/
MEROPS database	Rawlings et al. ⁹⁷	https://www.ebi.ac.uk/merops
Lipase Engineering database (v4.0)	Fischer and Pleiss ⁹⁸	https://led.biocatnet.de/sequence-browser
ChenB_2020 sequencing reads	Chen et al. ⁹⁹	NCBI: PRJNA532888
ChuY_2021 sequencing reads	Chu et al. ¹⁰⁰	CNGB: CNP0000284
GaoR_2021 sequencing reads	Gao et al. ¹⁰¹	NCBI: PRJNA514108
HeQ_2017 sequencing reads	He et al. ⁶¹	NCBI: PRJEB15371
HuangR_2020 sequencing reads	Huang et al. ¹⁰²	NCBI: PRJEB28545
HuY_2019 sequencing reads	Hu et al. ¹⁰³	NCBI: PRJNA401385
JieZ_2017 sequencing reads	Jie et al. ¹⁰⁴	NCBI: PRJEB21528
LiJ_2017 sequencing reads	Li et al. ¹⁰⁵	NCBI: PRJEB13870
LiuR_2017 sequencing reads	Liu et al. ¹⁰⁶	NCBI: PRJEB12123
LuW_2018 sequencing reads	Lu et al. ¹⁰⁷	NCBI: PRJNA391226
QianY_2020 sequencing reads	Qian et al. ¹⁰⁸	NCBI: PRJNA433459
QinJ_2012 sequencing reads	Qin et al. ¹⁰⁹	NCBI: PRJNA422434
QinN_2014 sequencing reads	Qin et al. ¹¹⁰	NCBI: PRJEB6337
QiX_2019 sequencing reads	Qi et al. ¹¹¹	NCBI: PRJNA530971
WangQ_2021 sequencing reads	Wang et al. ¹¹²	NCBI: PRJNA530339
WenC_2017 sequencing reads	Wen et al. ¹¹³	NCBI: SRP100575
WengY_2019 sequencing reads	Weng et al. ⁶²	NCBI: PRJNA429990
YanQ_2017 sequencing reads	Yan et al. ¹¹⁴	EBI: ERP023883
YeohYK_2021 sequencing reads	Yeoh et al. ¹¹⁵	NCBI: PRJNA650244
YeZ_2018 sequencing reads	Ye et al. ¹¹⁶	NCBI: PRJNA431482

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
YeZ_2020 sequencing reads	Ye et al. ¹¹⁷	NCBI: PRJNA356225
YuJ_2017 sequencing reads	Yu et al. ¹¹⁸	NCBI: PRJEB10878
ZengQ_2021 sequencing reads	Zeng et al. ¹¹⁹	NCBI: PRJNA539850
ZhangX_2015 sequencing reads	Zhang et al. ¹²⁰	NCBI: PRJEB6997
ZhongH_2019 sequencing reads	Zhong et al. ¹²¹	CNGB: CNP0000175
ZhouC_2020 sequencing reads	Zhou et al. ¹²²	NCBI: PRJEB29373
ZhuF_2020 sequencing reads	Zhu et al. ¹²³	NCBI: PRJEB29127
ZhuJ_2018 sequencing reads	Zhu et al. ¹²⁴	NCBI: PRJNA453965
ZhuQ_2021 sequencing reads	Zhu et al. ¹²⁵	NCBI: PRJNA602729
ZuoK_2019 sequencing reads	Zuo et al. ¹²⁶	NCBI: PRJEB28384
Brooks_2017 sequencing reads	Brooks et al. ¹²⁷	NCBI: PRJNA376566
Chen_2022 sequencing reads	Chen et al. ¹²⁸	NCBI: PRJNA762543
Chopyk_2023 sequencing reads	Chopyk et al. ¹²⁹	NCBI: PRJNA715245
Federici_2022 sequencing reads	Federici et al. ⁶⁴	NCBI: PRJEB50555
Feng_2015 sequencing reads	Feng et al. ¹³⁰	NCBI: PRJEB7774
Franzosa_2018 sequencing reads	Franzosa et al. ⁶³	NCBI: PRJNA400072
HMP_2017 sequencing reads	Lloyd-Price et al. ¹³¹	NCBI: PRJNA48479
Japan_2023 sequencing reads	Unpublished	NCBI: PRJDB13214
Karlsson_2012 sequencing reads	Karlsson et al. ¹³²	NCBI: PRJEB1786
Kokai-Kun_2020 sequencing reads	Kokai-Kun et al. ¹³³	NCBI: PRJNA589866
Lee_2023 sequencing reads	Lee et al. ¹³⁴	NCBI: PRJNA784939
Li_2014 sequencing reads	Li et al. ¹³⁵	NCBI: PRJEB5224
Madi_2023 sequencing reads	Madi et al. ¹³⁶	NCBI: PRJNA976726
Qin_2010 sequencing reads	Qin et al. ¹³⁷	NCBI: PRJEB2054
USA_2016 sequencing reads	Unpublished	NCBI: PRJNA352220
Vatanen_2016 sequencing reads	Vatanen et al. ¹³⁸	NCBI: PRJNA290380
Yachida_2019 sequencing reads	Yachida et al. ¹³⁹	NCBI: PRJDB4176
Zeller_2014 sequencing reads	Zeller et al. ¹⁴⁰	NCBI: PRJEB6070
Zhong_2019 sequencing reads	Zhong et al. ¹⁴¹	NCBI: PRJEB26795
YanQ_2023a sequencing reads	This paper	NCBI: PRJEB36300
YanQ_2023b sequencing reads	This paper	NCBI: PRJEB71419
YanQ_2023c sequencing reads	This paper	NCBI: PRJEB67456
Fungal genomes	This paper	NCBI: PRJNA833221
Fecal metagenomes	This paper	NCBI: PRJNA835516
Morphological and SEM images	This paper	https://doi.org/10.5281/zenodo.10802646
Primary metabolome data	This paper	https://doi.org/10.5281/zenodo.10417436
Software and Algorithms		
BLAST 2.5.0+	Camacho et al. ¹⁴²	http://ftp.ncbi.nlm.nih.gov/blast/executables/blast/
SPAdes	Bankevich et al. ¹⁴³	http://cab.spbu.ru/files/release3.11.1/manual.html
SSPACE	Boetzer et al. ¹⁴⁴	https://github.com/nsoranzo/sspace_basic/releases/tag/v2.1.1
fastp 0.23.0	Chen et al. ¹⁴⁵	https://github.com/OpenGene/fastp
BUSCO	Seppey et al. ¹⁴⁶	https://gitlab.com/ezlab/busco/issues
FASTANI	Jain et al. ³²	https://github.com/ParBLiSS/FastANI
RNAmer	Lagesen et al. ¹⁴⁷	https://services.healthtech.dtu.dk/service.php?RNAmmer-1.2
bowtie 2.4.4	Langmead and Salzberg ⁵⁰	http://bowtie-bio.sourceforge.net/index.shtml

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
AUGUSTUS	Keller et al. ¹⁴⁸	http://augustus.gobics.de/binaries/
GeneMark-EP+	Bruna et al. ¹⁴⁹	https://github.com/gatech-genemark/GeneMark-EP-plus
GeneMark-ES	Lomsadze et al. ¹⁵⁰	http://topaz.gatech.edu/GeneMark/license_download.cgi
MAKER2	Holt and Yandell ¹⁵¹	http://www.yandell-lab.org/software/maker.html
MMseqs2	Steinegger and Söding ¹⁵²	https://github.com/soedinglab/mmseqs2
eggNOG-mapper	Huerta-Cepas et al. ¹⁵³	http://eggnog-mapper.embl.de
DIAMOND	Buchfink et al. ¹⁵⁴	https://github.com/bbuchfink/diamond
mafft (v7.475)	Katoh and Standley ¹⁵⁵	https://mafft.cbrc.jp/alignment/software/
FastTree (2.1.10)	Price et al. ¹⁵⁶	http://www.microbesonline.org/fasttree/
RAxML	Stamatakis ³⁵	https://github.com/stamatak/standard-RAxML
iTOL	Letunic and Bork ³⁶	https://itol.embl.deitol.cgi
ade4 (R package)	Dray and Dufour ¹⁵⁷	http://pbil.univ-lyon1.fr/ADE-4/
randomForest (R package)	R project	https://cran.r-project.org/web/packages/randomForest/index.html
ggpubr (R package)	R project	https://cran.r-project.org/web/packages/ggpubr/index.html
ggplot2 (R package)	Villanueva and Chen ¹⁵⁸	https://cran.r-project.org/web/packages/ggplot2/index.html
vegan (R package)	Dixon ¹⁵⁹	https://cran.r-project.org/web/packages/vegan/index.html
fpc (R package)	R project	https://cran.r-project.org/web/packages/fpc/index.html
coin (R package)	R project	https://cran.r-project.org/web/packages/coin/index.html
Analysis scripts and input data	This paper	https://github.com/yexianingyue/Cultivated-Gut-Fungi

RESOURCE AVAILABILITY**Lead contact**

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Xiaochi Ma (maxc1978@163.com).

Materials availability

The fungal strains cultivated and sequenced in this study have been deposited at the Guangdong Microbial Culture Collection Center (GDMCC; <https://gdmcc.net/#/index>) with the corresponding accession IDs provided in Table S1. These strains are openly available for non-commercial use by the public. Additionally, all cultivated fungal strains are maintained at Dalian Medical University and can be obtained from the corresponding authors upon request.

Data and code availability

- The authors declare that the data supporting the findings of this study are available within the paper and its [supplemental information](#) files. The assembled genomes of the cultivated fungal strains have been deposited in the National Center for Biotechnology Information (NCBI) database with BioProject accession ID PRJNA833221 (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA833221>), and are publicly available as of the date of publication. Fecal metagenomic datasets generated in this study have been submitted to the NCBI database with project ID PRJNA835516 (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA835516>; human feces from healthy individuals) and the European Bioinformatics Institute (EBI) database with project IDs PRJEB36300 (<https://www.ebi.ac.uk/ena/browser/view/PRJEB36300>; YanQ_2023a), PRJEB71419 (<https://www.ebi.ac.uk/ena/browser/view/PRJEB71419>; YanQ_2023b) and PRJEB67456 (<https://www.ebi.ac.uk/ena/browser/view/PRJEB67456>; YanQ_2023c). All these datasets are publicly available as of the date of publication.

- Metadata for the cultivated fungal strains can be accessed through corresponding NCBI BioSample items, and detailed information are also provided in **Table S1** (e.g., culture medium, phenotype, and genomic characteristics, etc.). Morphological and scanning electron microscopy (SEM) images of fungal species have been deposited in the Zenodo database under accession ID 10802646 (<https://doi.org/10.5281/zenodo.10802646>), and are publicly available as of the date of publication. Primary metabolome data of fungal species were deposited in the Zenodo database under accession ID 10417436 (<https://doi.org/10.5281/zenodo.10417436>), and are publicly available as of the date of publication.
- The data processing and analysis codes, genomic and functional annotations, gut mycobiome profiles, and statistical scripts used in this study are publicly available at <https://github.com/yexianingyue/Cultivated-Gut-Fungi>.
- Any additional information required to reanalyze the data reported in this paper is available from the **lead contact** upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Volunteers and fecal sample collection

The study was conducted in accordance with the Declaration of Helsinki and was approved by the local ethics committee of the Second Affiliated Hospital of Dalian Medical University. Informed consent was obtained from all volunteers, and the experiments were performed in accordance with relevant guidelines and regulations. A total of 135 healthy volunteers (74 female, 61 male), aged from 18 to 42 years, with no gastrointestinal disorders or chronic health conditions were recruited; their detailed information is provided in **Table S1**. All volunteers were of Chinese origin. No antibiotic, antifungal, or antiparasitic therapy was administered to any enrolled subject at the time of sample collection or in the previous two months. At the appointed lab, fresh feces were collected from donors and stored in a sterile container, and the same process was followed with a sterile diluent buffer in the container as a negative control for the cultivation of gut fungi. Part of the fresh feces was sub-packed into sterile tubes (1.8 mL×5) inside the biosafety cabinet and immediately stored at -80 °C until further analysis. The remaining fresh samples were placed in a 50 mL centrifuge tube containing 20 mL of sterilized water and homogenized.

During the whole experimental process, we performed stringent quality control to avoid any spurious contaminating passenger fungi in our fecal fungal isolates by: (1) all experiments were performed in a clean laboratory with an independent filter ventilation system, (2) all sample collection consumables were disposable and sterile, (3) fresh feces were collected promptly from donors at the closest toilet to the laboratory and temporarily stored in a sterile disposable container for transfer to the laboratory, (4) after arriving in the biosafety cabinet, samples were processed using a sterile disposable tongue depressor in such a way that the outer surface of the fecal sample was peeled out and the internal feces were collected into a new sterile tube for further experiments, and (5) to ensure that the cultivated fungi were exclusively derived from gut microbiota samples, sterile dilution buffer was used in parallel as a negative control by streaking it in culture plates under the same growth conditions.

Cultivation of fungal isolates

A total of 1 mL of suspension was inoculated immediately onto each agar plate containing fungal culture medium, which was incubated until colonies were observed under aerobic and anaerobic conditions (32 °C).¹⁶⁰ The period for the collection of fresh feces from a donor, frozen storage, and incubation in culture medium was no more than two hours. Detailed information of the culture media (n = 26) for the isolation of gut fungi is listed in **Table S1**, all of which contained penicillin 100 U/mL and streptomycin 100 µg/mL. After incubation for 3–14 days, various fungal colonies were observed on the agar plates. Moreover, no fungal colonies were observed on agar plates when sterile diluent buffer was used as the negative control. Phenotypically distinct colonies were picked from each incubated agar plate on a fresh medium for further purification. A single colony was picked and re-streaked on the corresponding medium to purify the fungal strains. The purified fungal strains for identification and storage were inoculated in 5 mL of martin broth modified (MtB) medium under the same culture conditions (32 °C, 130 rpm shaking). A clean fungal precipitate was collected for taxonomic classification. Scanning electron microscopy (SEM) images of the isolated fungal strains were obtained from the Wuhan Servicebio Technology Co. Ltd., China. A flowchart of the cultivation process is shown in **Figure S1A**. All isolated fungal strains were stored at -80 °C in 25% glycerol.¹⁶¹

METHOD DETAILS

Taxonomic identification of isolated fungal strains by rDNA sequencing

The centrifuged fungal pellets from 5 mL of liquid medium were fragmented with 200 µL of sterile water and zirconium beads in a bead beater and boiled for 10 minutes. The supernatant was used as the PCR template. Four pairs of universal primers, including 18S primers NS1(GTAGTCATATGCTTGTCTC)/NS4(CTTCCGTCAATTCTTTAAG),⁸⁸ ITS primers ITS1FT(TCCGTAGGTGAACCTGCGG)/ITS4(TCCTCCGCTTATTGATATGC)⁸⁹ (targeting the ITS1-5.8S-ITS2 region), ITS1 primers ITS1(TCCGTAGGTGAACCTGCGG)/ITS2(GCTGCGTTCTCATCGATGC),⁹⁰ and 28S primers NL1(GCATATCAATAAGCGGAGGAAAG)/NL4(GGTCCGTGTTCAAGA CGG),⁸⁹ were used sequentially for Sanger sequencing, which means that the next pair of primers would be used if the previous pair of primers failed. Each reaction was performed in a final volume of 50 µL with 25 µL of 2× Taq PCR Master Mix (Takara, Japan), 1.5 µL of forward primer (10 µM), 1.5 µL of reverse primer (10 µM), 20 µL of sterile water and 2 µL of template DNA. In addition, to avoid

contamination of the PCR system when identifying cultivated fungi, we ensured that no 18S rDNA or ITS target fragment DNA product could be amplified from a parallel template of an equal volume of nuclease-free water for quality control. Amplifications were performed as follows: 94 °C for 4 min, 30 cycles of 94 °C for 30 s, 52 °C for 30 s, and 72 °C for 2 min; and a final extension at 72 °C for 10 min. Automated Sanger sequencing was performed on an ABI 3730XL platform (Applied Biosystems, USA) using the same primers as described above. These PCR product sequences were then aligned to the NCBI NT databases using the BLASTN program.

Fecal DNA extraction and whole-metagenome shotgun sequencing

Fecal DNA extraction was performed using DNA extraction kits (TIANamp Stool DNA Kit, TIANGEN, China) following the manufacturer's instructions. Briefly, approximately 170 mg of stool or the centrifugal pellet from 5 mL of fungal liquid was added to a Bead Beater Tube with an equal weight of zirconium beads and 500 µL of lysis solution, and then processed at 6.5 m/s speed for 9 min. The lysates were centrifuged at $\geq 13,400 \times g$ for 3 min. The supernatant was transferred to a spin column in a collection tube and centrifuged at $13,000 \times g$ for 1 min. Approximately 500 µL of binding buffer and 700 µL of wash buffer were added to the filtrate in the collection tube followed by concentration in turn. Finally, a total of 50 µL of eluted DNA with double-distilled water was prepared for each sample. DNA extracts were stored at -80 °C until use for whole-metagenome shotgun sequencing or quantitative PCR (qPCR). A total amount of 1 µg of DNA was used as the input material for sequencing library preparations. Sequencing libraries were prepared using the NEBNext® Ultra™ DNA Library Prep Kit for Illumina (New England BioLabs, USA), following the manufacturer's recommendations, and index codes were added to attribute sequences to each sample. The Illumina NovaSeq platform was used for 2 × 150 bp paired-end sequencing.

Fungal DNA preparation and whole-genome shotgun sequencing

In total, 744 isolated fungal strains were selected for whole-genome sequencing after dereplication at 97% nucleotide similarity to their 18S or ITS sequences. Manual selection was performed based on the physical appearance, such as color, texture, colony topography, and diffusible pigments. The centrifuged fungal pellets were fragmented with zirconium beads in bead beaters and used for DNA extraction using commercial DNA extraction kits (Qiagen QIAamp DNA Mini Kit), following the manufacturer's instructions. Extracted genomic DNA of the isolated fungi was stored at -80 °C until use as a template for subsequent qPCR amplification or shotgun sequencing. The whole-genome shotgun sequencing method was the same as that used for whole-metagenome shotgun sequencing. High-quality reads were filtered from the raw Illumina data by trimming the low-quality ($Q < 30$) bases on the end of the reads and filtering 'N'-containing, adapter contaminated, or short length (<90 bp) reads using fastp.¹⁴⁵

qPCR for detection of fungal strains from original feces

Due to the characteristics of the cluster growth of some molds, it is difficult to accurately compute their colony-forming units (CFU) and quantify the number of corresponding fungal species in feces. To further verify that the isolated fungi were exclusively derived from feces, we performed a relative quantification analysis by quantitative PCR (ABI StepOne, US). The number of fungal cells was determined using a blood counting chamber at $\times 400$ magnification prior to fungal genome extraction. Eleven randomly selected fungal strains were examined in terms of their presence or absence in feces. Based on the genomic data of each strain, we screened unique amplified sequence regions and designed strain-specific primers to determine their specificity by blasting the NCBI NT database. The 18S rRNA gene was used as a reference gene with the primers listed below: forward, ATTGGAGGGCAAGTCTGGTG; reverse, CCGATCCCTAGTCGGCATAG. The isolated fungal DNA extracted from 5 mL of liquid enrichment culture medium was used as a reference sample, and the total DNA extracted from 150 mg of corresponding host feces was used as a test sample. Chemical reagents for the amplification of 18S rRNA and genomic DNA were used according to the manufacturer's instructions (Innogene, China). Each reaction was performed in triplicates. A 25 µL reaction system contained 1 µL of template DNA or sterile double distilled water (negative control), 12.5 µL of SYBR Premix (Innogene, China), 0.5 µL of forward primer, 0.5 µL of reverse primer, and 10.5 µL of water. Amplifications were performed under the following reaction conditions: one cycle of pre-denaturation at 95 °C for 10 min, followed by 50 cycles of denaturation at 95 °C for 30 s, and annealing at 50 °C for 1 min. Primer specificity was confirmed against the non-target test fungal strains. Each sample was repeated 3 times. The threshold cycle for each sample was determined for a specific fragment and normalized to the CT values of all the fungal 18S rRNA genes. The relative quantity of specific fungi in the test feces compared to the reference sample was calculated using the $2^{-\Delta\Delta CT}$ method.^{24,162}

Targeted quantification of fungal primary metabolites

The 199 fungal species (a representative strain for each species was selected for analysis) were inoculated into MtB medium (400 mL) in triplicate and incubated at 32 °C with shaking for 5 days. The MtB medium was then discarded, and the fungal pellets were washed with methanol-water at -40 °C (v:v=60:40) by centrifugation at 3,000 rpm for 3 min. The fungal pellets were lyophilized using a freeze dryer and subjected to targeted metabolome measurements using metabolite array technology¹⁶³ (Metabo-Profile, Shanghai, China). Briefly, 310 standards (mostly primary metabolites) were applied to establish the quantification method using ultra-performance liquid chromatography coupled to tandem mass spectrometry (UPLC-MS/MS) (ACQUITY UPLC-Xevo TQ-S, Waters Corp., Milford, USA). All the standards were accurately weighed and dissolved to obtain individual stock solutions (5 mg/mL). Appropriate amounts of each stock solution were mixed to create the stock calibration solutions. Dried fungal pellets (5 mg) were

homogenized in water using zirconium oxide beads for 3 min and extracted using cold methanol with an internal standard mix. After centrifugation ($18,000 \times g$, 20 min), the supernatant containing the metabolites was transferred to a 96-well plate, which was used for derivatization using a Biomek 4000 workstation (Beckman Coulter, Inc., California, USA) in the presence of 20 μL of freshly prepared derivative reagents ($30^{\circ}C$, 60 min). The sample was diluted with ice-cold 50% methanol (350 μL), and stored at $-20^{\circ}C$ for 20 min, followed by centrifugation at $4,000 \times g$ at $4^{\circ}C$. Finally, the supernatant was transferred to a new 96-well plate, together with the internal standards in each well, which was sealed for LC-MS analysis. For metabolite determination, a comprehensive set of rigorous quality control/assurance procedures was employed to ensure consistently high-quality analytical results by controlling every single step from sample receipt at the laboratory to final deliverables.

NCBI reference and available human-associated fungi data

The available NCBI fungal genomes were downloaded in June 2023. The raw fungal genomes included 12,587 genomes, 964 of which were removed because of 1) extremely low assembly quality (N50 length <2,000 bp or number of scaffolds >10,000), or 2) a mixture of multiple genomes, with the remaining 11,623 genomes retained as a reference for further analyses (Table S2). To construct the publicly available human-associated fungi catalogue, we searched the NCBI BioSample metadata of all fungal strains and identified totalling 502 strains isolated or sourced from human body sites or clinical samples, including 156 strains from the respiratory or digestive tract, 152 from the skin or body surface, 160 from tissue specimens, and 32 from the genitourinary tract (e.g., urine and vagina) (Table S2), named as the publicly accessible catalogue of human-associated fungi (PHF). These strains clustered into 129 PHF species following an ANI threshold of 95% (see below).

Genome assembly and analysis

Genome assembly and quality assessment

Shotgun sequencing reads for each fungal isolate were *de novo* assembled by SPAdes¹⁴³ using the k-mer parameter “21, 33, 55, 77” and mismatch correlation mode (–careful). The raw assembled sequences were processed by contig extension and scaffolding using SSPPACE¹⁴⁴ (iteratively run until the best assembly result was generated), and the shortest scaffolds were filtered with a minimum length threshold of 500 bp. After assembly, the scaffolds for each sample were checked to determine if they included multiple fungal genomes based on 1) the taxonomic assignment of each scaffold when aligned to the NCBI NT database and 2) the distribution of G+C percentage vs. sequencing depth of all scaffolds. Scaffolds in multiple genomes were split into single genomes based on the MetaBinner algorithm¹⁶⁴ and confirmed by manual inspection. The reads that could be mapped to each split single genome were collected and reassembled to result in a higher-quality genome. The quality of fungal genome assemblies was assessed by quantifying their completeness using Benchmarking Universal Single-Copy Orthologs (BUSCO).¹⁴⁶ BUSCO uses a set of 758 universal fungal single-copy orthologs (version 2020-09-10) to infer the completeness of a query fungus. However, this method may generate low completeness for some novel fungi because their genes may have low homology to the reference orthologs.

Average nucleotide identity for species and genus definition

The ANI of fungal genomes was calculated using FASTANI³² with default parameters. To explore the species/genus boundary of fungi, we generated pairwise ANIs for all fungal genomes from NCBI and evaluated the accuracy of species/genus definition at the ANI threshold of 70%–100%. The results are shown in Figure S1F. When setting the ANI threshold to 96%, we observed that 98.76% of intra-species pairs met or exceeded this value, with only 0.14% of inter-species pairs achieving this threshold (equivalent to false positives). Similarly, at an ANI threshold of 95%, 99.09% of intra-species pairs reached this criterion, whereas only 0.16% of inter-species pairs did. Considering that an ANI of 95% is widely employed as a standard in prokaryotic and viral genome studies,^{32,165} we adopted the 95% threshold as the definition for fungal species in the revised manuscript. At the genus level, an ANI threshold of 75% and 82.45% of intra-genus pairs reached this criterion, whereas only 0.13% of inter-genus pairs did, suggesting high accuracy for genus definition.

Annotation of rDNA sequences and assignment to reference database

The rDNA sequences (18S, 28S, 5.8S, ITS1, and ITS2) of fungal genomes were predicted using RNAmmer,¹⁴⁷ a program based on hidden Markov models for annotation of rRNA genes and ITS zones. For each fungus, the rDNA sequences were aligned against the reference rDNA sequence database integrated from the NCBI RefSeq, SILVA,⁹¹ and UNITE⁹² databases. Sequences with similarity >97% at the nucleotide level were considered to belong to the same species.

Phylogenetic analyses

The maximum-likelihood phylogenetic tree of 206 fungal species shown in Figure 1 was inferred using RAxML software³⁵ based on 150 genes that appeared in all species, and the tree was visualized using the online iTOL tool.³⁶

Gene prediction and analysis

Gene prediction and protein clustering

Protein-coding genes were predicted from the fungal genomes using GeneMark-ES.¹⁵⁰ Genes with lengths less than 75 bp were removed for further analysis. Orthologous gene clustering of the fungal genomes was performed to generate protein clusters, using the MMseqs2 algorithm¹⁵² with options “–min-seq-id 0.5 -c 0.9” (similarity 50% and minimum coverage threshold of 90% of the length of the shortest sequence) at the protein level. For each fungal genome, the number of duplicate genes was defined as the

number of genes minus the number of gene clusters, and the proportion of duplicate genes (also referred to as the gene duplication rate) was calculated by dividing the number of duplicate genes by the number of genes.

Analysis of gene diversity

Core and pangenome analyses of the fungal species were performed using in-house scripts. The genes shared by all strains within a species were defined as the core genes, whereas genes partially shared in some strains (accessory genes) and unique to a single strain (strain-specific genes) in a species were defined as dispensable genes. When the gene accumulation curve fits Heap's law ($n = \kappa N^\gamma$) with parameter $\gamma < 0$, or the occurrence of new genes fits the power law ($n = N^{-a}$) with exponent $a > 1$, a species was defined as having a "closed" pangenome.¹⁶⁶

Gene functional annotation and analysis

To predict the function of the protein repertoire of 760 fungal genomes, we assigned protein sequences to the eggNOG (evolutionary genealogy of genes: Non-supervised Orthologous Groups, v5.0)⁹³ and InterPro (version 77.0)⁹⁴ databases using eggNOG-mapper¹⁵³ and InterProScan, respectively. KEGG (Kyoto Encyclopedia of Genes and Genomes)⁹⁵ annotation was performed based on searching against the KEGG database (downloaded on June 2023) using DIAMOND¹⁵⁴ (score threshold 60, over 50% coverage). Each protein was assigned an eggNOG, InterPro, or KEGG ortholog based on the best-hit gene in the corresponding database. CAZymes, proteases, and lipases of 760 fungal genomes were annotated using the carbohydrate-active enzymes (CAZy, downloaded on June 2023),⁹⁶ MEROPS,¹⁶⁷ and Lipase Engineering Database (LED, v4.0)⁹⁸ databases, respectively. PCWDEs were determined using the CAZyme-based ranking of the fungi (CBRF) web page (<http://13.58.192.177/RankEnzymes/PlantCellWallComponents>). The identification of secondary metabolism gene clusters (SMGCs) was carried out using the SMURF algorithm³⁸ and a modified approach derived from previous studies on *Aspergillus* fungal genomes.^{47,168} Briefly, we first aligned all fungal genes against the "secondary metabolite-specific PFAM domains" sourced from the SMURF article (comprising 27 domains) using the PfamScan tool.¹⁶⁹ Genes that matched the criteria were designated as "backbone" genes, while other genes that aligned with the "secondary metabolite-specific domains" were considered as "decorating" genes. We then searched for the decorating genes within the 10 genes upstream and downstream of each backbone gene to determine the boundaries of the SMGC. A valid SMGC was required to satisfy both conditions: 1) a maximum intergenic length of <3 kb, and 2) no more than 6 genes without secondary metabolite-specific domains. Using this approach, we identified a total of 22,762 SMGCs (encompassing 26,355 backbone genes) across 760 fungal genomes in the CGF catalogue.

Fungal profiling algorithm and validation

To develop a robust and accurate method for quantifying fungal species in fecal metagenomes, we assessed two strategies commonly applied to profile bacterial compositions: the clade-specific core gene strategy (exemplified by MetaPhlAn 4⁴⁸) and the direct mapping of reads to genomes (using alignment algorithms such as BWA⁴⁹ or Bowtie 2⁵⁰). For the first strategy, the determination of species-specific core genes for gut fungi is hampered due to the lack of sufficient genomic resources. The second strategy, relying on direct read mapping, struggles to accurately differentiate whether a read originates from a target fungus or other closely related species. Given the typically low abundance of fungi in feces, such confusion leads to significant inaccuracies. To address these challenges, drawing upon the strengths of these existing strategies and leveraging a recently developed algorithm designed for high-precision quantification of ultra-low abundance bacteria in metagenomes,⁵³ we have devised an updated algorithm for profiling fungi in metagenomes. In summary, firstly we integrated a gut fungal gene set (totalling ~11.7 million genes) from the CGF and PHF collections, performed a pairwise nucleotide-level alignment for all genes, and identified ~9.6 million genes as species-specific and ~2.1 million genes as shared among multiple species (Figure S5A). Then, to avoid potential contamination from human sequences or other gut microbes (i.e., bacteria, archaea, and viruses), the reads from fecal metagenomes were aligned against 1) all human genome sequences extracted from the NCBI-NT database, 2) all bacterial, archaeal, or viral sequences extracted from the NCBI-NT database, and 3) 4,644 prokaryotic genomes from the Unified Human Gastrointestinal Genome (UHGG) collections,¹⁶ and the contamination reads were removed accordingly. Reads that mapped to fungal rDNA sequences, tRNA sequences, or potential horizontal gene transfer regions were dismissed. Filtered reads were aligned with the gut fungal gene set. Reads uniquely matching species-specific genes were unambiguously assigned to that species, whereas reads matching multiple species-specific genes or shared genes were temporarily held in escrow (Figure S5A). After alignment, unambiguous reads were used to calculate the initial read count for each species, and the reads in escrow were fractionally assigned in proportion to the count of each species, generating the final read count. Species' relative abundance was determined by normalizing the final read counts according to the average genome size of each species. To evaluate the performance of our fungal profiling algorithm, we downloaded raw datasets from ultra-deep fecal metagenomes from the Chinese⁵² (NCBI accession ID: PRJNA763692) and Hazra⁵³ (NCBI accession ID: PRJEB49206) populations.

Collection and process of public datasets

We searched the published gut microbiome studies in PubMed and Google Scholar databases based on exhaustive keywords such as "gut metagenome", "gut microbiota/microbiome", and "shotgun sequencing" (as of January 2022). The materials of each study were manually reviewed, and 46 studies were included under the following criteria: 1) case-control study of disease, 2) samples from Chinese individuals, and 3) availability of fecal metagenomic data. Eight of these studies were excluded because they included less

than 50 samples or the case and control samples were not from the same batch (Figure S6A). Raw whole-metagenome shotgun sequencing datasets from the remaining studies were downloaded from the NCBI Sequence Read Archive (SRA), European Nucleotide Archive (ENA), and China National GeneBank (CNCB) databases. Metadata of the studies were obtained from the original articles and materials, NCBI/EBI/CNCB sample information, or by contacting the corresponding authors (Table S5). Within each metadata-available study, samples were excluded based on the following criteria: 1) nonstandard definitions of disease or unhealthy statuses (e.g., hypertension/diabetes were defined according to the latest standards), 2) non-baseline samples for longitudinal sampling, and 3) body mass index (BMI) $<17 \text{ kg/m}^2$ or $>30 \text{ kg/m}^2$ (for samples with available phenotypic data, except for the obesity and diabetes studies). Additionally, to compare the gut mycobiome characteristics between the Chinese and non-Chinese populations, we downloaded raw fecal metagenomic datasets from 19 studies of non-Chinese origin (Table S5) and processed the data in the same way as the Chinese population.

Raw metagenomic reads were processed for quality control using fastp.¹⁴⁵ Low quality (>45 bases with quality score <20 , or >5 'N' bases), low complexity, and adapter-containing reads were removed. The remaining reads were trimmed at the tails for low quality ($<Q20$) or 'N' bases, and the trimmed reads with length <45 bases were also removed. Human contaminating sequences were eliminated via mapping against the reference human genome (GRCh38) using Bowtie 2. Samples with fewer than 20 million high-quality non-human reads or an extremely low proportion of fungi (<100 reads) were removed.

Fecal metagenomic datasets of this study

Three independent studies, including acute cholecystitis (AC) (YanQ_2023a), colorectal cancer (CRC) (YanQ_2023b), and inflammatory bowel disease (IBD) (YanQ_2023c) were included in this study. The ethics committee of Dalian Medical University approved this study, and each participant signed an informed consent agreement. Participants (YanQ_2023a, 47 AC patients vs. 58 healthy controls; YanQ_2023b, 27 CRC patients vs. 28 healthy controls; and YanQ_2023c, 71 IBD patients vs. 166 healthy controls) were recruited from the Dalian University Affiliated Xinhua Hospital, Second Affiliated Hospital of Dalian Medical University, and Dalian Medical University. Patients with IBD patients (including 15 with Crohn's disease and 56 with ulcerative colitis) were enrolled with inclusion criteria requiring a clear diagnosis by a licensed physician in strict accordance with the ECCO-ESGAR Guideline for Diagnostic Assessment in Inflammatory Bowel Disease.¹⁷⁰ The exclusion criteria for healthy volunteers were diabetes, severe hypertension, metabolic syndrome, IBD, cancers, abnormal liver or kidney function, and dyslipidemia. Individuals were excluded if they had taken antibiotics or probiotic products within 4 weeks.

Fecal specimens were collected from participants, temporarily stored on dry ice, transported to the laboratory within 24 h, and stored at -80°C for further analysis. DNA was extracted from fecal samples using the TIANamp Stool DNA Kit (TIANGEN, China). DNA quality was assessed using the Qubit 2.0. The extracted DNA samples were stored at -80°C until use. The sequencing library was generated using the NEB Next® Ultra™ DNA Library Prep Kit (NEB, USA), following the manufacturer's recommendations, and index codes were added to each sample. Library quality was confirmed using an Agilent 2100. Clustering of the index-coded samples was performed on a cBot Cluster Generation System using an Illumina PE Cluster Kit (Illumina, USA) according to the manufacturer's instructions. After cluster generation, the DNA libraries were sequenced on the Illumina NovaSeq platform and 150 bp paired-end reads were generated. Quality control and human contaminant were processed using the aforementioned pipeline for publicly available samples.

The biological effects of fungi strain on DSS-induced colitis in mice

Male C57BL/6J mice (6–8 weeks old) were purchased from the Experimental Animal Center of Dalian Medical University and were raised under SPF conditions. Germ-free mice, raised in gnotobiotic isolators and fed sterilized food and water, were provided by Cya-gen Biosciences (Suzhou, China). Their feces, food, and padding were periodically monitored to evaluate bacterial contamination. All mice were housed under a regular 12-h light/dark cycle and allowed free access to food and water during the experiments. All animal experiments were performed in accordance with the Guidelines for the Care and Use of Laboratory Animals of the National Institute of Health. The study protocol was approved by the Ethics Committee of Dalian Medical University (Approval No.: AEE19046).

A mouse model of colitis was established by the administration of dextran sulfate sodium (DSS) in drinking water for 5–7 days, followed by 2 days of regular drinking water before sacrifice. Disease activity index (DAI) was recorded and assessed daily after DSS administration. DAI consists of body weight loss, diarrhoea, and rectal bleeding. All were assigned scores on a 0–4 scale. Body weight: 0, body weight loss ≤ 0 ; 1, body weight loss ≤ 5 ; 2, body weight loss ≤ 10 ; 3, body weight loss ≤ 15 ; 4, body weight loss > 15 . Diarrhoea: 0, normal stool; 1, Soft but still formed stool; 2, soft stool; 3, very soft and wet stool; 4, Watery stool. Rectal bleeding: 0, No bleeding; 1, positive hemoccult; 2, visible blood traces in stool; 3, visible blood traces that adhered to the anus; 4, gross bleeding. Severity of colitis was evaluated by colonoscopy using high-resolution mini-endoscopy. Intestinal permeability was assessed by measuring the translocation of FITC-labelled dextran from the intestinal lumen into the blood after oral administration for 4 h. The blood, ileum, and colon were sampled for further examination.

To evaluate the effect of differential fungi between the healthy population and IBD patients on colitis, mice were orally administered fungi for two weeks at a dose of 10^8 CFU/mouse per day (1 time per day) before DSS. The fungi tested included two IBD-enriched species: *Cutaneotrichosporon cutaneum* c97 (strain P1411) and *Naganishia albida* c50 (strain PT4301); 4 IBD-reduced species: [*Candida*] *aaseri* c86 (strain B8007), *Cyberlindnera jadinii* c142 (strain Y1112B), *Candida* sp. c21 (strain PL3312), and *Trichosporon coremiiforme* c31 (strain LN7006B); and 3 non-significant species *Zygosaccharomyces pseudobailii* c107 (strain PF2202),

Wickerhamiella pararugosa c69 (strain PH4012), and *Lodderomyces elongisporus* c112 (strain PH4014). The metabolites were also evaluated for their potential protective effects against DSS-induced colitis in mice. Metabolites were prepared by extracting the fungal fermentation broth and orally administered at 100 mg/kg.

To exclude the effects of the gut microbiota, antibiotic-treated and germ-free mice were used for DSS-induced colitis. The gut microbiota was depleted by adding an antibiotic cocktail to the drinking water for 7 days (neomycin 1 mg/mL, streptomycin 1 mg/mL, and bacitracin 1 mg/mL). Then, antibiotic-treated and germ-free mice were subjected to fungal colonization and DSS challenge.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analyses were conducted using the R 4.0.1 platform (<https://www.r-project.org/>), and the statistical scripts used in this study are available on the GitHub website (<https://github.com/yexianingyue/Cultivated-Gut-Fungi>).

Multivariate analyses

Principal coordinate analysis (PCoA) was performed on the Bray-Curtis dissimilarity using the R *vegan* package¹⁵⁹ and visualized using the R *ade4* package. Distance-based redundancy analysis (dbRDA) was performed using the R *vegan* package, based on Bray-Curtis dissimilarity. Permutational multivariate analysis of variance (PERMANOVA) was realized with the R *vegan* package, for which the effect size (R^2) of disease status on mycobiome variation was calculated using the *adonis* function, and the *P*-value was generated based on 1,000 permutations.

Random forest analysis

Random forest analysis was applied using the R *randomForest* package. A random forest model was trained for each metabolite based on their concentrations in 150 fungi, and the predicted values were generated using a leave-one-out cross-validation (RFCV) procedure. The performance of the models was evaluated using Spearman's correlation coefficient between the predicted and actual values. The *P*-value was generated using the Wilcoxon rank-sum test unless specifically mentioned, and the *q*-value was calculated to evaluate the false discovery rate (FDR) for correction of multiple comparisons. For case-control comparisons, in which we aimed to assess the efficacy of fungal compositions in distinguishing cases and controls using the random forest model, we adopted a 5-fold cross-validation approach. In the case of the 5 global IBD datasets, where we utilized IBD-associated fungal signatures to differentiate between patients and healthy individuals within each dataset, we employed a leave-one-dataset-out strategy to test the performance of the random forest models.

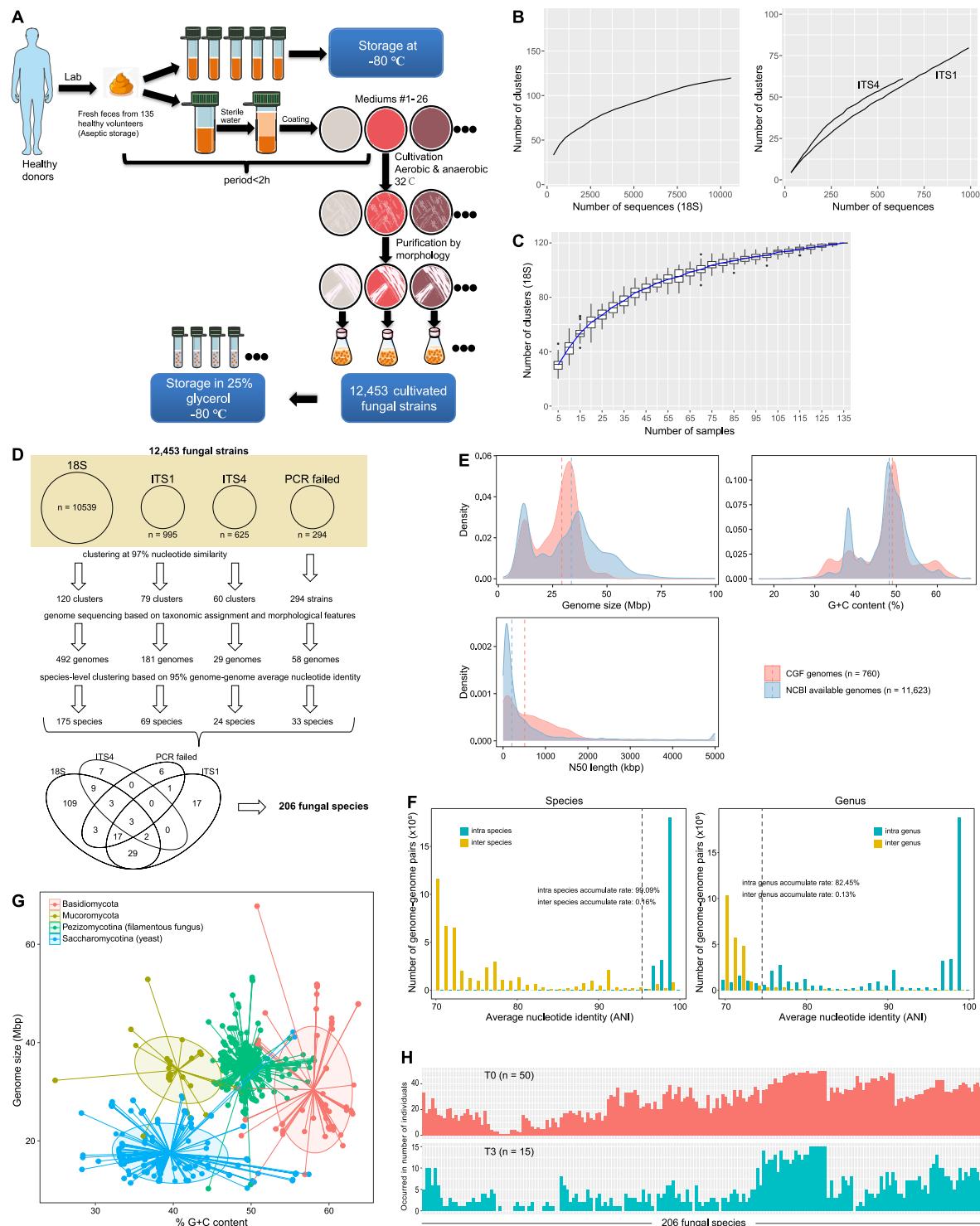
Comparison analyses

For case-control comparison analysis within each dataset, *P*-values were calculated using the Wilcoxon rank-sum test. The *q* value was used to evaluate FDR for correction of multiple comparisons and was calculated based on the R *p.adjust* function based on the Benjamini-Hochberg algorithm.¹⁷¹ Random effects meta-analysis was implemented based on the previously developed algorithm in gut microbiome datasets.¹⁷² For meta-analysis, the species relative abundances were converted to \log_{10} -transformed proportions. The *escalc* function from the R *metafor* package was employed the Hedges' *g* standardized mean difference statistic to calculate the pooled effect size using a random effects model. Between-study heterogeneity was assessed using Cochrane's Q test and the I^2 index, and the meta-analysis *P*-values were obtained from the random effects models. Linear mixed effects model analysis was conducted using the *lmer* function of the *LmerTest* package¹⁷³ with study as a random effect, and the significance of the models was calculated using the *anova* function. For the meta-analysis, we also employed blocked Wilcoxon rank-sum tests (equivalent to the Kruskal-Wallis test for binary data) using the *wilcox_test* function of the *coin* package. Following adjustment for study confounders, we obtained *p*-values for disease and healthy control comparisons. It is important to note that the fungal taxa must satisfy both meta-analysis *P*<0.05 and blocked Wilcoxon rank-sum test *P*<0.05 criteria to be considered as shared signatures across diseases.

Other statistical analyses

Clustering analysis of the gut mycobiomes was performed based on the species-level relative abundance profiles using the Jensen-Shannon divergence and partitioning around medoids (PAM) clustering algorithm, following the tutorial of gut microbiome enterotype analysis.⁵⁴ The average silhouette width (ASW) was used to determine the optimal number of clusters, and an ASW less than 0.3 suggested poor support for the existence of discrete clusters. Receiver operating characteristic (ROC) analysis was implemented using the R *pROC* package, and the area under the ROC curve (AUC) was calculated accordingly. Spearman's correlation coefficients and their significance were assessed using the *cor.test* function.

Supplemental figures



(legend on next page)

Figure S1. Overview of gut fungal cultivation and genome information, related to Figure 1

(A) Overview of the gut fungal cultivation processes. A total of 12,453 fungal strains were isolated and stored in this study. A full description of the experimental methods is provided in [STAR Methods](#).

(B) Clustering analysis of 18S ($n = 10,539$ strains), ITS1 ($n = 995$ strains), and ITS4 ($n = 625$ strains) ribosomal DNA (rDNA) sequences obtained from the isolated fungi. For all isolated fungi, ITS1 and ITS4 rDNA were sequenced in cases in which 18S rDNA sequencing failed. Sequences were clustered at the 97% similarity threshold at the nucleotide level. The accumulation curves depict the number of clusters as a function of the number of rDNA sequences.

(C) Rarefaction analysis of sampled fungal species (the fungal species-sample size saturation curve). For a given number of individual samples, we performed random sampling 30 times in the cohort with replacement and estimated the total number of 18S clusters (representing fungal species) that could be identified from these samples. The boxplot shows the median and interquartile ranges of the number of clusters at each sample size, and the blue line connects the medians.

(D) Overview of the selection processes for fungal strains chosen for whole-genome shotgun sequencing and the ultimate source of the fungal species.

(E) Genome information of the CGF catalog compared with the NCBI available genomes. Histogram plots show the distribution of genome size (upper), % G+C content (right), and scaffold N50 length (bottom) for both the CGF genomes and the NCBI available fungal genomes. Dotted lines represent the medians for each respective catalog.

(F) Calibration of fungal species and genus boundaries based on the NCBI available genomes. Pairwise average nucleotide identity (ANI) calculations were performed on 11,623 fungal genomes available in the NCBI database. The resulting genome-genome ANI pairs were divided into “intra-species/genus” and “inter-species/genus” groups based on the known species/genus assignments of the fungi. At the species level, with an ANI threshold set at 95%, 99.09% of intra-species pairs met or exceeded this threshold (equivalent to true positives), with only 0.16% of inter-species pairs achieving it (indicating false positives). Similarly, at the genus level, with an ANI threshold set at 75%, 82.45% of intra-genus pairs reached this criterion, whereas only 0.13% of inter-genus pairs did. Thus, the chosen standards of 95% and 75% were highly accurate for species and genus definition, respectively.

(G) Distribution of genome size and % G+C content of the CGF genomes. The scatterplot illustrates the distinctive genomic features of Saccharomycotina and Pezizomycotina members, showcasing the distribution of genome size and % G+C content among CGF genomes.

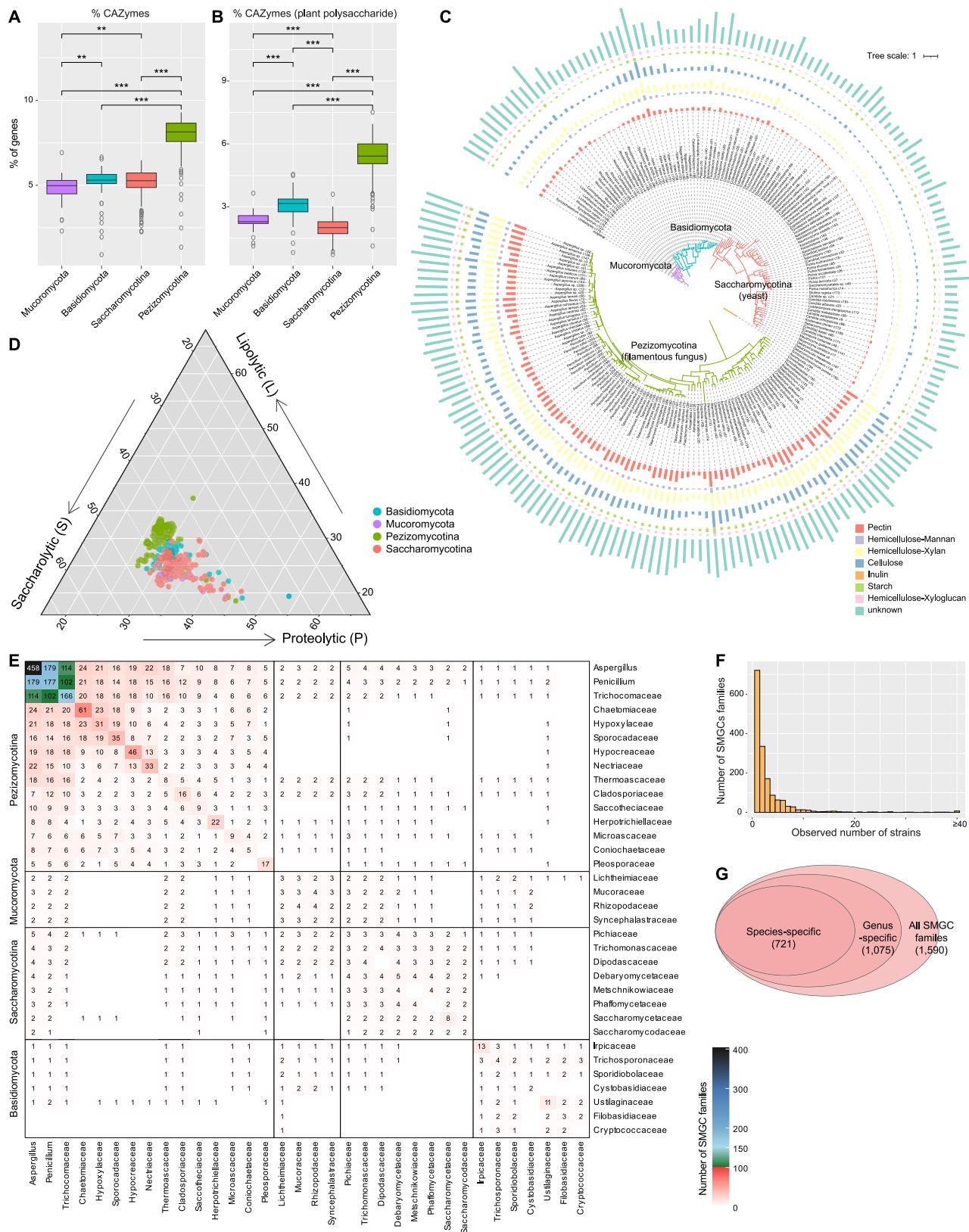
(H) Presence of 206 CGF species in human fecal samples. The x axis represents each species in the CGF catalog, while the y axis indicates the number of individuals' fecal metagenomes where the species is present. The red bars correspond to 50 fecal samples collected at the time of fungal cultivation (T0), and the green bars represent 15 fecal samples collected 3 years later (T3).



(legend on next page)

Figure S2. Supplement of the functional analyses of cultivated gut fungi, related to Figure 2

- (A) Venn diagram of shared protein clusters between CGF, PHF, and the NCBI available fungal genomes.
- (B) Coding density, gene duplication rate, and proportion of clade-specific protein clusters in the CGF genomes. The left panel illustrates the coding density (length of coding area vs. length of the genome), the middle panel represents the gene duplication rate (number of protein clusters vs. number of all genes), and the right panel displays the proportion of clade-specific proteins for the dominant 32 families and 2 Aspergillaceae genera (*Aspergillus* and *Penicillium*) in the CGF catalog. Text colors indicate the respective subphyla for fungal clades: purple, Mucoromycota; blue, Basidiomycota; red, Saccharomycotina; green, Pezizomycotina.
- (C) Influence of phylogenetic affiliation for the functional configuration of CGFs. Upper: permutational multivariate analysis of variance (PERMANOVA) reveals the effect size (R^2) and p value of fungal phylogeny on their functional profiles, calculated based on 10,000 permutations. Bottom: principal coordinate analysis (PCoA) of functional profiles within four subphyla at the genus level. Each subphylum was subjected to PCoAs utilizing genus-level taxonomic information for strain classification. The first two principal coordinates and their respective variance contribution ratios are shown.
- (D and E) The completeness ratio of amino acid and lipid metabolism (D) and several important metabolism pathways (E) in gut fungi. Barplot showing the prevalence rate of metabolic modules for four fungal subphyla. Colors indicate the completeness ratio of each metabolic module: dark green indicates the presence of a complete module, while light green represents modules that are largely present (with only one enzyme not found).



(legend on next page)

Figure S3. Supplement of the functional analyses of cultivated gut fungi, related to Figure 3

(A and B) Boxplots showing the comparison of all CAZymes (A) and plant cell wall-degrading enzymes (B) among four fungal subphyla. Wilcoxon rank-sum test with false discovery rate (FDR) correction: ** $q < 0.01$; *** $q < 0.001$.

(C) For each species, the average numbers of enzymes and the gene contents of target polysaccharides in genomes are shown.

(D) Enrichment analysis of saccharolytic, proteolytic, and lipolytic abilities of gut fungi. Triangle plot showing the enrichment of 760 fungal genomes in saccharolytic, proteolytic, and lipolytic abilities.

(E) Distribution of secondary metabolic gene cluster (SMGC) families in gut fungi. The number of shared SMGC families among 34 dominant fungal clades. The numbers in the diagonal line indicate the numbers of clade-specific SMGC families.

(F) Barplot showing the number of SMGC families that occur in a certain number of strains.

(G) The number of species-specific and genus-specific SMGC families.

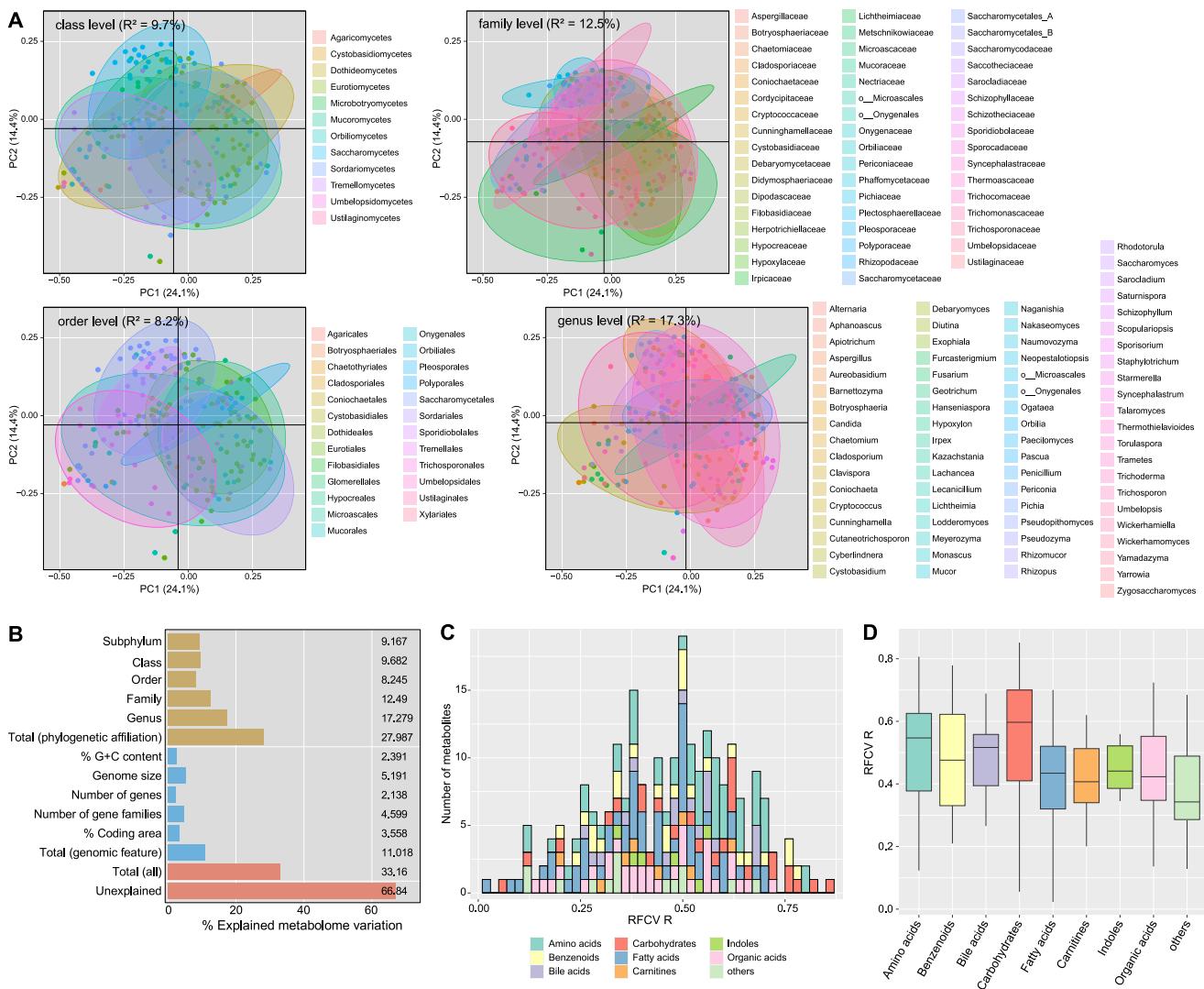
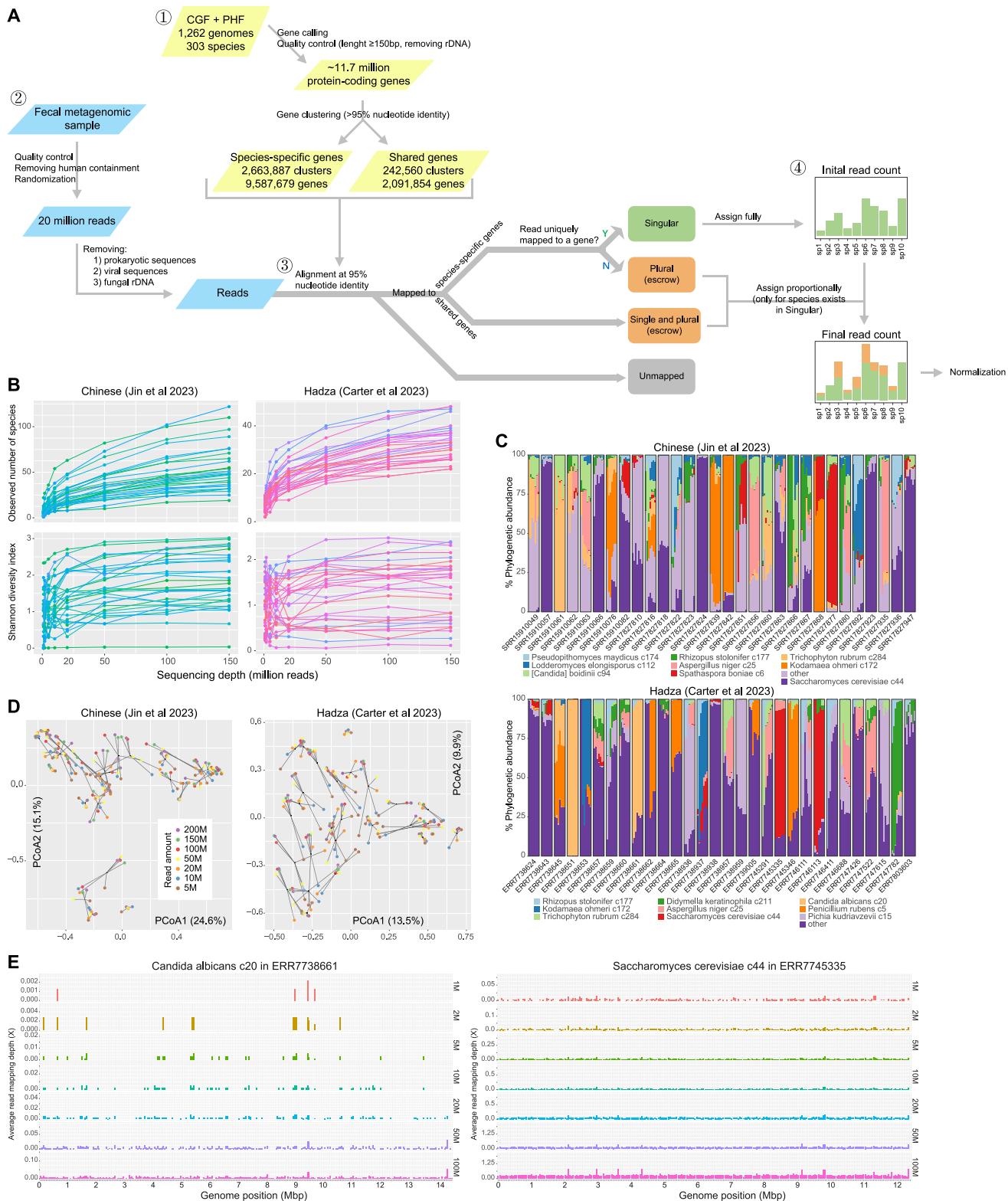


Figure S4. Influence of phylogenetic affiliation and genomic feature for the primary metabolome of CGFs, related to Figure 4

(A) Principal coordinate analysis (PCoA) of the metabolomic profile at the class, order, family, and genus levels. The first two principal coordinates and their respective variance contribution ratios are shown.

(B) Permutational multivariate analysis of variance (PERMANOVA) reveals the effect size (% R^2) of fungal phylogeny and genomic features on primary metabolome profiles.

(C and D) Distribution of predictable effect size of metabolites grouped by type. The effect size of fungal functional contents in each metabolite was estimated using Spearman's correlation between the measured value and the 5-fold cross-validation random-forest-model predicted value (RFCV R).



(legend on next page)

Figure S5. The workflow and assessment of the algorithm for profiling of gut fungi in fecal metagenomes, related to Figure 5

(A) Workflow of the updated algorithm for profiling of gut fungi in fecal metagenomes. Step 1: process of the fungal genome reference. The gut fungal gene set (totaling ~11.7 million genes) was integrated from the CGF and PHF catalogs. A pairwise nucleotide-level alignment was performed on these genes, identifying ~9.6 million genes as species specific and ~2.1 million genes as shared among multiple species. Step 2: fecal metagenomic sample processing. For each fecal metagenome, 20 million reads were randomly selected, and reads potentially originating from prokaryotes, viruses, or fungal rDNA sequences were removed. Step 3: read mapping processing. Filtered reads were mapped into the fungal gene set at 95% identity at the nucleotide acid level. Reads uniquely matching species-specific genes were unambiguously assigned to that species, whereas reads matching multiple species-specific genes or shared genes were temporarily held in escrow. Step 4: processes of read assignment. After alignment, unambiguous reads were used to calculate an initial read count for each species, and reads in escrow were fractionally assigned in proportion to the count of each species, generating the final read count. Species' relative abundances were determined by normalizing the final read counts according to the average genome size of each species.

(B) Evaluation of fungal species diversity at different sequencing depths in two ultra-deep fecal metagenomic datasets. Each line represents a fecal metagenomic sample. The x axis indicates different sequencing amounts, while the y axis represents fungal richness (represented by the observed number of species) (upper) and Shannon diversity index (bottom) at that sequencing depth.

(C) Barplot showing the fungal species composition at different sequencing depths. For each sample, the species-level fungal composition is presented at sequencing depths of 5, 10, 20, 50, 100, 150, and 200 million (M) reads.

(D) Principal coordinate analysis (PCoA) of fungal compositional profiles at different sequencing depths. The first two principal coordinates (PCs) are shown, and their explained variances are presented. Sample points belonging to the same metagenomic sample are connected by lines.

(E) Barplot showing the genome coverage of two gut fungal species in the corresponding fecal metagenomes.

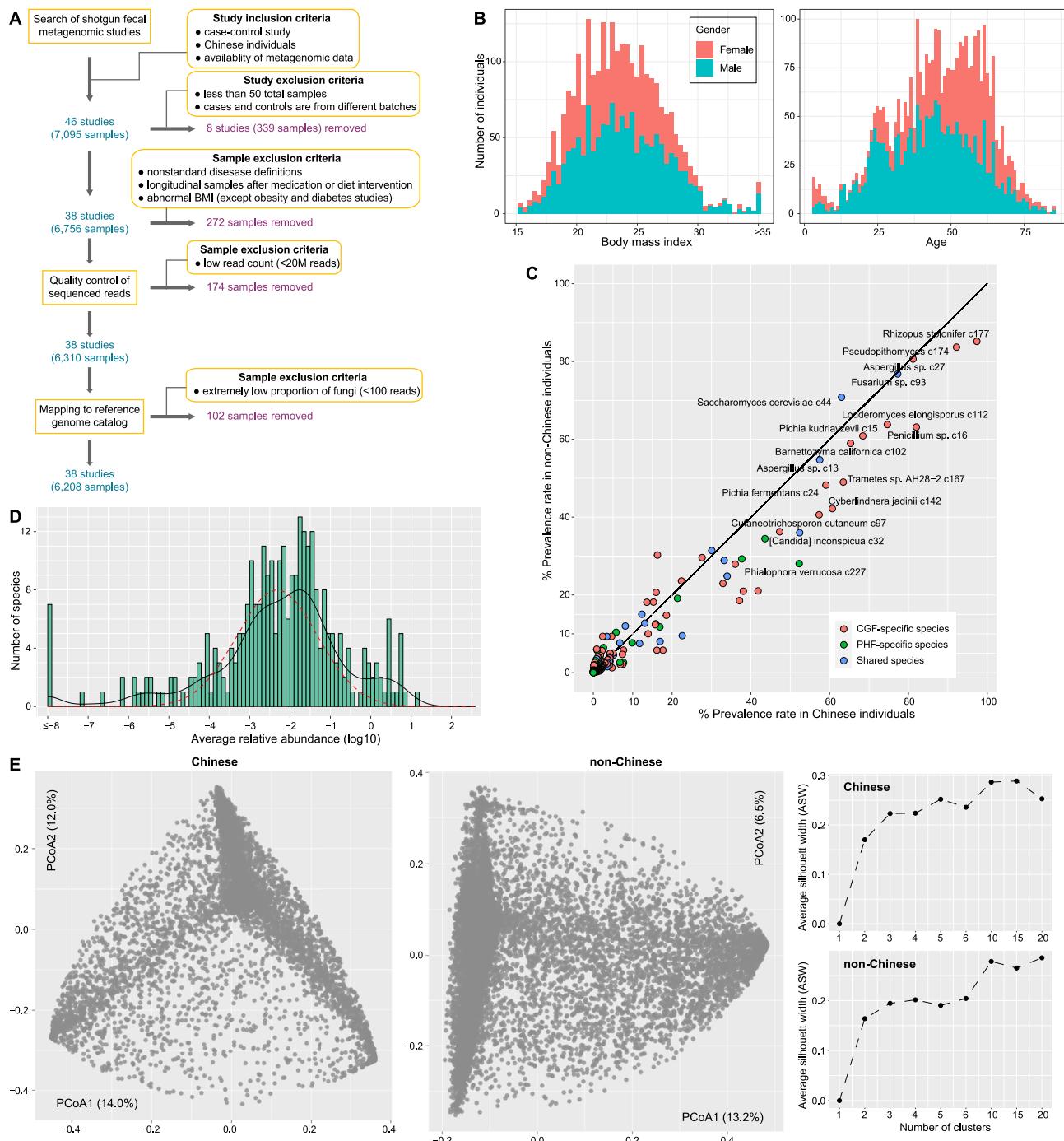


Figure S6. Prevalence and clustering analyses of gut fungal species in Chinese and non-Chinese populations, related to Figure 5

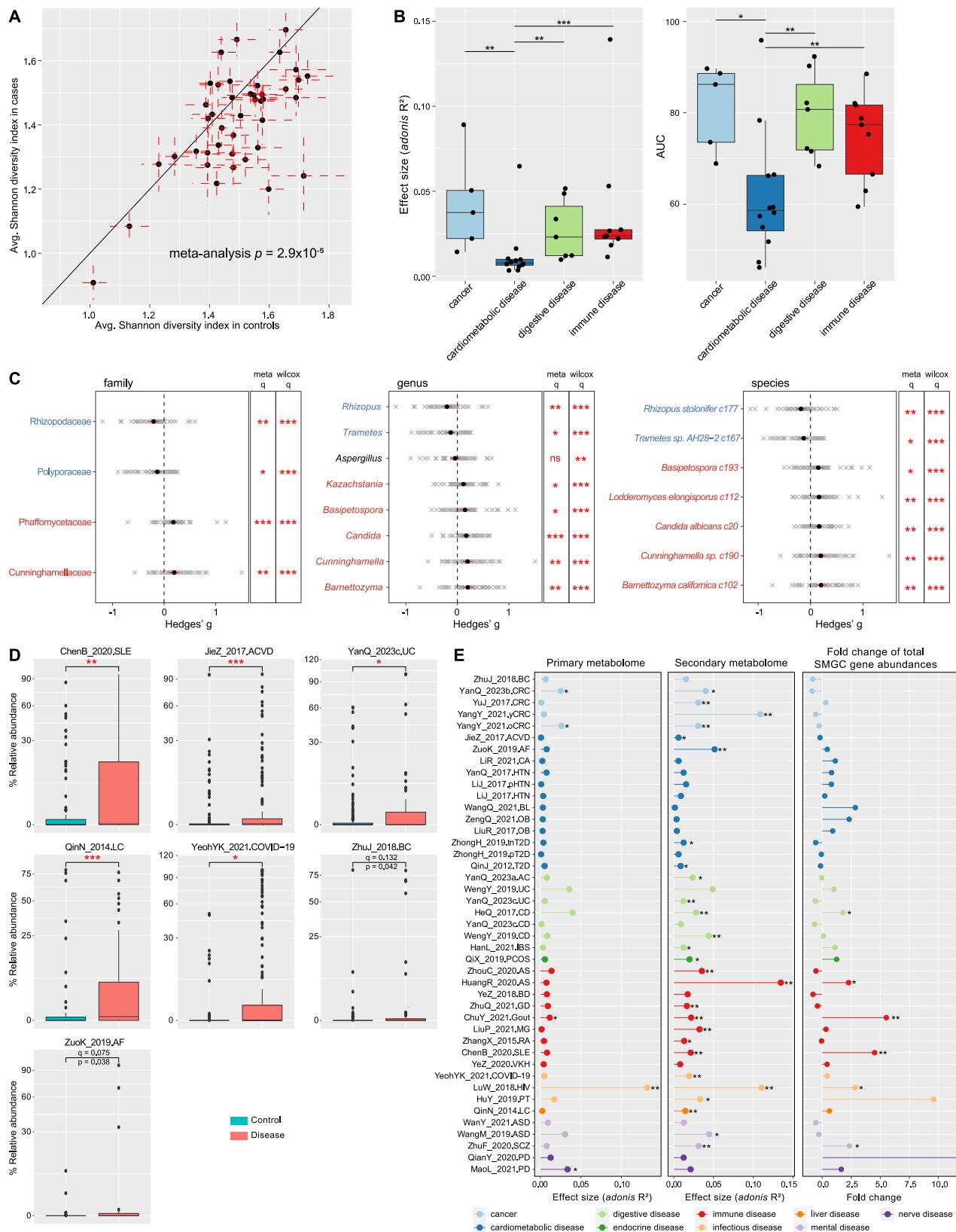
(A) Flow diagram showing the criteria for inclusion and exclusion of studies and samples from the Chinese population.

(B) Distributions of host body mass index (BMI) and age of fecal samples of the Chinese population. Individual properties (i.e., sex, BMI, and age) were collected for approximately half of the samples from the original studies. These individuals included 45% females, with an average age of 46 ± 15 years and an average BMI of 23.6 ± 3.7 kg/m².

(C) Prevalence rates of gut fungal species in Chinese and non-Chinese populations. Scatterplot depicting the prevalence of 303 fungal species in the two populations. Taxonomic names of highly prevalent species are indicated.

(legend continued on next page)

(D) Distribution of the average relative abundances of 303 CGF species in fecal metagenomes. The dotted line indicates the fit curve of the log-normal distribution.
(E) Clustering analysis of the gut mycobiome in Chinese and non-Chinese populations. Principal coordinate analysis (PCoA) of the Bray-Curtis distances between gut mycobiome profiles in Chinese and non-Chinese populations. The first two principal coordinates (PCs) and their explained variances are shown. The right panels show the average silhouette width (ASW) of the optimal number of clusters for clustering analysis of the gut mycobiome. The result shows poor support for the existence of discrete mycobiome clusters (ASW < 0.3 in all cases) in both Chinese and non-Chinese populations.



(legend on next page)

Figure S7. Supplemental analyses of the gut mycobiome and metabolome associated with common diseases, related to Figure 6

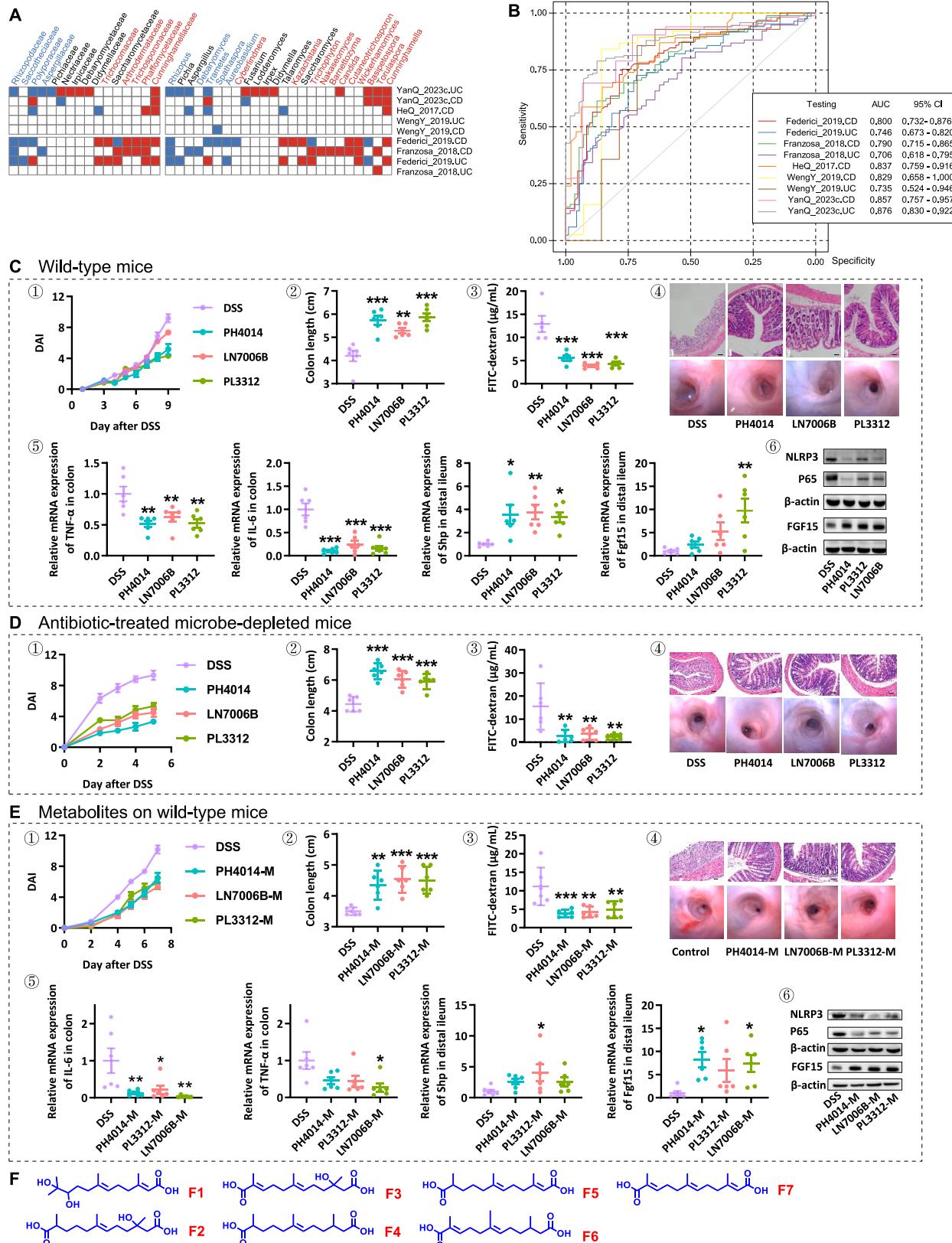
(A) Random effects meta-analysis of Shannon diversity index across 43 case-control comparisons. Scatterplot shows the average Shannon diversity index in cases and controls across 43 comparisons. Dotted red lines indicate the standard error of the mean (SEM).

(B) Comparison of the degree of gut mycobiome alteration among four major disease types. Boxplots show the comparison of the effect sizes of PERMANOVA analysis (left) and within-study AUCs for disease vs. control classification (right) across four types of diseases. Boxes show medians/quartiles and error bars extend to the most extreme values within 1.5 interquartile ranges. Wilcoxon rank-sum test with false discovery rate (FDR) correction: * $q < 0.05$; ** $q < 0.01$; *** $q < 0.001$.

(C) Meta-analysis of the gut fungal signatures among common diseases. Forest plot displays the Hedges' g standardized mean differences of gut fungal signatures among all investigated diseases. Solid red lines indicate the 95% confidence intervals. The colors of the fungal taxonomic names indicate the enrichment in meta-analysis: red, enriched in diseases; blue, enriched in controls; black, not significant. FDR-corrected p values (q) are calculated based on random effects meta-analysis (meta q) and blocked Wilcoxon rank-sum test (wilcox q): * $q < 0.05$; ** $q < 0.01$; *** $q < 0.001$; ns, not significant.

(D) Boxplot showing the relative abundances of *Candida albicans* c20 in samples of seven significant diseases. Boxes show medians/quartiles; error bars extend to the most extreme values within 1.5 interquartile ranges. p values are calculated based on the Wilcoxon rank-sum test: * $q < 0.05$; ** $q < 0.01$; *** $q < 0.001$.

(E) Analysis of the fungus-predicted primary and secondary metabolomes across common diseases. Effect sizes of permutational multivariate analysis of variance (PERMANOVA) for the primary and secondary metabolomes in disease vs. control classifications among 43 case-control comparisons. Diseases are color-coded by their respective types. The analysis is conducted using the *adonis* test with 1,000 permutations. Fold changes of total SMGC gene abundances between cases and controls across 43 comparisons. Statistical significance is denoted as follows: * $q < 0.05$; ** $q < 0.01$.



(legend on next page)

Figure S8. Analysis of IBD-associated gut fungal signatures and experimental validation of three fungal species, related to Figure 7

(A) Heatmap showing the IBD-associated gut fungal signatures in Chinese (HeQ_2017, WengY_2019, and YanQ_2023c) and non-Chinese (Franzosa_2018 and Federici_2019) fecal metagenomic datasets. Colored squares indicate the fungal families and genera that are significantly enriched in each case-control comparison (Wilcoxon rank-sum test with false discovery rate correction $q < 0.05$): red, enriched in IBD patients; blue, enriched in healthy controls. The colors of the fungal taxonomic names indicate the enrichment in meta-analysis: red, enriched in diseases; blue, enriched in controls; black, not significant.

(B) Receiver operating characteristic (ROC) analysis of the classification of IBD status using random forest models. The models are trained and tested based on a leave-one-dataset-out strategy, in which models were built on four datasets combined and validated on the left-out dataset, for each dataset in turn. The performances of models are assessed by the area under the curve (AUC).

(C and D) Protective effect of gut fungi *Trichosporon coremiiforme* c31 (LN7006B), *Candida* sp. C21 (PL3312), and *Lodderomyces elongisporus* c112 (PH4014) against DSS-induced colitis in wild-type (C) and antibiotic-treated microbe-depleted mice (D).

(E) Protective effect of fungal metabolites of three species against DSS-induced colitis in wild-type mice. For (C)–(E), the panels indicate as follows: (1) the changes in disease activity index (DAI) are recorded and assessed daily. (2) Length of the colons. (3) Intestinal permeability measured by translocation of FITC-dextran from the intestinal lumen into the blood after oral administration for 4 h. (4) Representative images of H&E staining and colonoscopy of the colon. Scale bars, 100 μ m. (5) Levels of inflammatory cytokines (tumor necrosis factor alpha [TNF- α] and interleukin-6 [IL-6]) in the colon and FXR target genes (*Shp* and *Fgf15*) in the distal ileum by qPCR analysis. (6) Analysis of inflammatory biomarkers (nucleotide-binding oligomerization domain-like receptor family, pyrin domain-containing 3, NLRP3 and P65) and FXR target gene (*Fgf15*) in the colons by western blot. Data are presented as mean \pm SEM ($n = 6$), and ordinary one-way ANOVA is used to determine statistical significance, compared with the sham-gavaged DSS group: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

(F) The chemical structures of fungal metabolites (F1–F7) identified in the fermentations of *Trichosporon coremiiforme* c31 (LN7006B), *Candida* sp. c21 (PL3312), and *Lodderomyces elongisporus* c112 (PH4014).