Integrative functional analysis uncovers metabolic differences between *Candida* species

Neelu Begum[1], Sunjae Lee[1], Aize Pellon[1], Shervin Sadeghi Nasab[1], Jens Nieslen[2,3], Mathias Uhlen[4], David Moyes[1*], Saeed Shoaie[1,4*]

[1] Centre for Host-Microbiome Interactions, Faculty of Dentistry, Oral & Craniofacial Sciences, King's College London, SE1 9RT, UK.

[2] Department of Biology and Biological Engineering, Kemivägen 10, Chalmers University of Technology, SE-412 96, Gothenburg, Sweden.

[3] BioInnovation Institute, Ole Maaløes Vej 3, DK2200 Copenhagen N, Denmark

[4] Science for Life Laboratory, KTH – Royal Institute of Technology, Stockholm, SE-171 21, Sweden.

*corresponded to: david.moyes@kcl.ac.uk; saeed.shoaie@kcl.ac.uk

Abstract

Candida species are a dominant constituent of the human mycobiome and a better understanding of their metabolism from a fungal perspective can provide key insights into their ability to cause pathogenesis. Here, we have developed the BioFung database – a fungal specific tool for functional annotation using the KEGG database that provides an efficient method for annotation of protein-encoding gene. Analysis of carbohydrate-active enzyme (CAZymes) and BioFung, uncovered core and accessory features across *Candida* species demonstrating plasticity, adaptation to the environment and acquired features. Integerative functional analysis revealed that all *Candida species* can employ amino acid metabolism. However, metabolomics revealed that only a specific cluster of species (AGAu species - *C. albicans, C. glabrata and C. auris*) utilised amino acid metabolism. We identified critical metabolic pathways in the AGAu clusters with biomarkers and anti-fungal target potential in the CAZyme profile, polyamine, choline and fatty acid biosynthesis pathways. This study, combining genomic analysis, metabolomics and gene expression validation, highlights the metabolic diversity within AGAu species that underlies their remarkable ability to dominate the mycobiome and cause disease.

Introduction

Fungal infections affect around 7.5 million people around the world every year. Within human fungal communities (mycobiome), with the notable exception of the skin, *Candida* species are the most common group[1–3]. These species are generally pathobionts, being the most common human fungal pathogens, despite also being commensal organisms[4]. *Candida* infections are becoming increasingly concerning, and the World Health Organisation (WHO) has recently emphasised international surveillance for diagnosis and management of fungal infection, particularly *C. albicans* infection[5–7]. Recently, a novel *Candida species*, *Candida auris* has been identified with significant mortality and morbidity, as well as a high degree of antifungal resistance[8,9]. There are over 200 *Candida* species currently identified, but only a handful of these are present in the human microbiota with the ability to cause infection and pathology. Most notable examples of these include *C. albicans, C. glabrata, C. dubliniensis, C. tropicalis* and *C. auris*[9–15]. *Candida* species, most notably *C. albicans* and *C. glabrata,* can give rise to a variety of superficial infections, including oral thrush and vulvovaginal candidiasis, but are also capable of causing a systemic infection with significant mortality[16–20]. As well as direct infections, fungi such as *Candida* species have also been associated with oncogenesis through complement activation, demonstrating potential effects of the interaction of fungal species with human host [21].

An essential virulence determinant of fungi is their metabolic plasticity[22]. Fungal are significant in their ability to utilise numerous different anabolic and catabolic sources in their metabolic processes, attributable to switching between carbon and nitrogen sources[23]. Nutritional availability, environmental factors, competition and pathogenic factors all influence this plasticity[24,25]. Investigations of *Candida* species-specific transcriptional regulators of glycolytic genes (e.g. Tye2 and Gal4) and enzymes of the glycolytic pathway (hexose catabolism), indicate these factors play an essential role in central carbon metabolism commonly applied during infection events[22,24,26]. Glycolytic metabolism can activate virulence factors that initiate hypha formation, activate fermentative pathways, repress gluconeogenesis, and the TCA cycle[27–30]. Alternatively, *C. albicans* can switch to gluconeogenesis and the glyoxylate cycle to confer full pathogenesis during systemic candidiasis[31–34]. Carbohydrate metabolism is coupled with changes of cell wall architecture, host immune response modulation, as well as adherence, biofilm formation, stress response and drug resistance[24,35–37]. If carbohydrate sources are limited, *Candida* species can use amino acids and lipids as supplementation for metabolic adaptation[22,38–40]. Amino acids

produced by *C. albicans* have been shown to drive tissue damage by initiating stress responses and adjusting the surrounding environmental pH, helping induce of host invasion processes [35,41–47]. Very little is known about the regulation, process and utilisation of amino acid metabolism in *Candida*[39]. However, *C. albicans* is known to use amino acids to replace carbon and other nitrogen sources[48]. *Candida*'s ability to convert arginine to urea allows the neutralisation of an acidic environment triggering the development of hyphae and biofilm formation [32,49,50]. Notably, recent work has shown that *C. albicans* phagocyted by macrophages induces fatty acid β-oxidation and the glyoxylate pathway to induce hypha formation for escape . In a harsh environment that lacks even a nitrogen source, *Candida* can recycle and produce its own proteins and polyamines without host nitrate[51]. Thus, understanding of metabolism and functionality of *Candida* is instrumental in tackling infection and mortality prevalence[52].

Here we developed BioFung – a database focused on functional information and interpretation of biological information. It is based on 128 fungal species using KEGG orthologs (KO). By applying BioFung to *Candida species*, we identified a distinct cluster of *C. albicans, C. glabrata* and *C. auris* referred to as AGAu species. This cluster has a high association with infection and mortality, relative to other *Candida* species[18,53–55]. We applied comparative analysis techniques based on gene, protein, and enzyme-substrate levels and identified metabolic pathways in *Candida* species, such as choline and polyamine pathways. Metabolomics along with experimental validation from gene expression confirmed the AGAu cluster difference. This study (1) provides a novel tool for functional annotation of fungal species, (2) highlights amino acid metabolism importance in AGAu species, and (3) identifies novel potential fungal biomarkers and anti-fungal targets in metabolic pathway.

Results

Development of BioFung database and functional annotation of *Candida* protein-encoding genes

To perform a global functional analysis of *Candida* species, we collected 49 publicly available genomes of different *Candida* strains covering 13 different species (Supplementary Table 1). We selected species based on their clinical importance and abundance within the human mycobiome[56–60]. All 49 *Candida* strains were isolated from different body sites from people in different geographical locations (Fig. 1a and Extended Fig. 1a). The quality of the

genome of each strain was checked at the level of scaffold assembly and genome size (Extended Fig. 1b before reassessing the phylogenetic relationship of these 49 strains based on nucleotide sequence using average nucleotide identity (ANI) (Fig. 1b, Method). We observed strain-specific differences in phylogenetic lineages with 11 distinct branches, including branching of *C. auris*, *C. glabrata* and *C. albicans*, implying genetic diversity that could implicitly be interpreted into functional variances. To elucidate functional details for these strains, we built fungi-specific Hidden Markov Models (HMM) using fungal gene clusters, named BioFung database (Fig. 1c, Method)[61,62]. We analysed 524,288 fungal genes, from 128 fungal species, with a coverage of 4,822 KOs, and 4,430 fungal KO alignments to create BioFung (Supplementary Table 2-4). Comparison of the BioFung database with other eukaryote-specific HMM sources indicated that the BioFung has both higher coverage and specificity of KOs (Extended fig. 1c-e) [61,62].

The collection of *Candida* strains used to integrate functional annotations can be categorised into commonly invasive, non-invasive (requiring a secondary factor to cause infection, such as co-morbidity, immunodeficiency) based on literature (Fig. 1d, Supplementary Table 5, Method). These representative samples of *Candida* were integrated into the functional analysis framework, with a total of 49 *Candida* species annotated with BioFung, Protein families (Pfam)[63] and Carbohydrate-Active enZyme (CAZyme)[64] databases. We applied BioFung using the UCLUST algorithm, to establish core genome features (found in all *Candida* species) and accessory genome features (shared or unique functions) [65]. Intra-strain analysis of *C. albicans* across 24 strains sequenced showed largely conserved metabolic pathways and CAZymes (Extended Figure i-j). In covering KEGG metabolic orthologs, clustering analysis determined a larger number of accessory features of metabolism compared to core characteristics seen in all *Candida* strains (Fig. 1d, Extended Fig. 1f-h, Method).

Identification of global functional annotation profiles in *Candida*

We next determined the CAZyme profile by mapping the 49 *Candida* protein sequences to the dbcan2 database[64]. Doing this allowed us to infer molecular enzyme function[66]. CAZymes are vital enzymes involved in the metabolism of complex carbohydrates. Approximately 205 unique CAZymes were identified in all *Candid*a strains, with various functions (Fig. 1f, Supplementary Table 6, Method). From core *Candida* genome analysis, annotated enzymes were distributed across 6 active families: providing an assortment of enzymatic functions. The glycoside hydrolase (GH) family showed the highest degree of core

coverage (Extended Fig. 2a), with much of the GH family activity in starch and other storage carbohydrates substrate-converters (Fig. 1g).

Additionally, xylan and sugar carbohydrate utilisation were the dominant functions in the accessory genome (Supplementary Table 6). Cell wall composition is a crucial virulence factor and assessing CAZyme components of cell wall substrate converters has been extensively researched[67]. Here, we reveal the presence of pectin lyases, glycan lyases, chitin lyases and mannan lyases (Fig. 1h). Pectin substrate conversion enzyme has been identified as the core feature of *Candidas'* functional cell wall enzyme, though previously only reported in *Candida bodinii*[68]and frequently seen in the fungal plant pathogen, including *Aspergillus Pencillium*[69]. Alongside ß-glucan, mannan and chitin carbohydrate enzyme profiles, *Candida* cell wall activity includes pectin enzyme activity (Supplementary Table 7).

In addition, we identified 1,182 Pfam clans from all *Candida* strains and re-categorised them into 14 functional clans (Supplementary Table 8). Pfam domain annotation indicating genetic information processing, cell machinery, and metabolism was among the most extensive Pfam domains exhibited (Extended fig. 2b). We assessed the diverse functional association of protein domains by analysing core functional clans, and determined similar patterns of dominance for carbohydrate, amino acid and lipid processing-associated domains (Extended Fig. 2c).

The functional and metabolomic activity of clinical AGAu *Candida* strains.

Next, to better explore and understand the link to metabolism and pathogenesis, we clustered the groups based on invasive nature of particular species, linked to contributing to a high percentage of mortality and candidemia (Fig. 2a). *C. albicans, C. glabrata* and the emerging invasive species *C. auris* were grouped together (AGAu cluster). Alternative *Candida* species termed non-AGAu group include opportunistic species that require virulence factors or a defective immune system to cause disease pathology as well as environmental *Candida* species. The AGAu cluster contains those *Candida* species most commonly associated with clinical pathology, contributing to a higher percentage of mortality and candidemia [16–18,53–55,70]. This classification of AGAu is analysed and discussed throughout the rest of this paper.

We compared the CAZyme profile coverage of AGAu and non-AGAu groups (Fig. 2b-c, Supplementary Table S9). The CAZyme GH43_8 (substrate-conversion of α-L-arabinofuranosidase/β-xylosidase [71]) was significantly enriched in AGAu possibly involved

in the breakdown of complex glucans (Wilcoxon signed-rank test, P-value <0.05, Method)[72]. The identification of significant CAZymes in the AGAu cluster showed carbohydrate conversion of xylan (GH43_8), mucin (GH95), cellulose (GH66) and copper oxidase family (AA5) [73–75]. Interestingly, GH66 associated with human oral plaque formation[73] and AA5 enzyme has been reportedly linked to fungal defence [76]. CAZymes seen in non-AGAu species showed carbohydrate-binding module families involved in sugar, polysaccharide and cell wall breakdown, including CBM48.

The AGAu species are morphologically diverse, (*C. albicans* is dimorphic, whilst *C. glabrata* and *C. auris* are not), they are all potential pathogens. Therefore, we assessed metabolic pathway enrichment in KO annotations of both AGAu and non-AGAu *Candida* strains (hypergeometric test and Wilcoxon signed-rank test, p-value < 0.05, Method). We found functional evidence of pathways present in both clusters, indicating a genetic potential for all *Candida* strains to undertake similar metabolic trajectories, as no differences were observed between cluster group in pathway analysis (Fig. 3a). All *Candida* strains notably revealed encoded pathways facilitating carbohydrate catabolism within the system; thus, potential to drive increasing metabolic activity, for example through fructose and mannose metabolism. We identified significant enrichment of amino acid metabolism, including arginine, proline, cysteine and methionine metabolism. We also observed significant levels of fatty acid biosynthesis and glutathione metabolism, which have previously been associated with virulence mechanisms[77–80].

Metabolomics revealed key metabolic pathways assimilated by AGAu group

To elucidate the metabolic trajectory taking place by each cluster group, we performed metabolomics on a collection of 7 clinical *Candida* isolates, representing the diverse pathogenic species (Fig. 3b, Method). These clinical samples were previously isolated from patients infections [53,81–84] and were used to evaluate *in vitro* the critical metabolic activity predicted by our functional analyses (Extended fig. 3a). These *Candida* isolates were representative of both AGAu and non-AGAu grouped species. The metabolomics data were used for partial least squares discriminant analysis (PLS-DA), which resulted in distinct cluster separation of significant analyte classes (Extended Fig. 3b-d, Method). The PLS-DA model identified distinct high metabolites features uniquely in AGAu *Candida* species (Extended Fig. 3e). For example, histidine metabolite production in AGAu species supports this metabolite role in systemic infection and is potential an anti-fungal target[85,86]. We also

identified choline-derived metabolites (choline, phosphatidyl-choline and lysophosphatidylcholine) as increased in certain AGAu species (Fig. 3c). We identified phosphatidylcholines analyte class as contributing the highest number of features across selected clinical *Candida* species (Extended Fig. 3d), which has been observed previously in the hypervirulent *C. albicans* (SC5314) strain[38–41]. Lastly, spermine and spermidine, were found to be significantly associated with AGAu strains, indicating polyamine metabolism could play a functional role in the increased association with disease pathology of these strains.

Integrative global metabolic map of of AGAu *Candida* species

Having identified these pathways in silico, we next determined gene expression levels of essential polyamine (SPE11, SPE3), choline (CKI1, TAZ1) and fatty acid (CEM1) pathways in *C. albicans* from the AGAu cluster to validate these findings (Extended fig. 4d, Supplementary Table 10 and Methods). All 5 genes showed expression substantiating the activity of these pathways. Our conclusions of key pathway associations draws importance of nitrogen sources, specifically the metabolism of amino acids, in this process (Fig. 4), although to date, *Candida* species pathogenesis is better known to be driven by carbon sources[22,37,89,90]. AGAu species exhibited increased levels of metabolites in the choline pathway, polyamine and fatty acid biosynthesis pathways that are primarily propagated through arginine, cysteine and methionine pathways (Extended Fig. 4a-c). Based on integration of computational and experimental data revealed fundamental metabolic pathways applied AGAu species providing a link to the major advantage shown by AGAu species across the human body with considerable contributions to pathogenesis. These important pathways include polyamine, choline and fatty acid biosynthesis. For instance, the polyamine pathway is thought to be involved in *Candida* cell proliferation, and in turn, causes host cellular dysfunction by modulating acetylation levels of aminopropyl groups and inducing autophagy, thus increasing cell life span[89,91,92]. We observed fatty acid biosynthesis production with large numbers of metabolites of triglycerides featuring in PLSDA and family have previously been reported to promote germination and virulence of AGAu *Candida* strains (Extended Fig. 3d)[93,94]. Further, fatty acid biosynthesis is vital in fungal cell membrane viability, energy storage, signalling, and cell proliferation – all functions critical in pathogenesis[95–98].

Discussion

BioFung was used for building metabolic maps of key *Candida* strains and the database provides the mycology community with a resource allowing them to dive more deeply into all fungal species' metabolic capability based on protein encoding genes. Development of data generation technologies development, tools and database for fungal species is currently in its infancy, despite significant advances in these areas for bacteria and archaea. This database enables detailed mechanistic annotations to optimise our understanding of fungal species. to do this, it uses HMM to provide high specificity for fungal annotation. To do this, it is currently the best database available for KEGG-based functional annotation of fungi (Extended fig. 1b-d and Supplementary Table 2-4).

Analysis of annotations allowed us to identify the influential AGAu group of *Candida* strains, highlighting critical metabolic pathways in these strains. In doing so, we developed increased understanding of the metabolism of these strains through integrating multi-omics and experimental data. BioFung can be used extensively to better understand individual fungal species' metabolic pathways but can be extended to explore metabolic interactions between fungi, other organisms, and within the host-mycobiome.

Using this tool in combination with metabolomics validation, the AGAu *Candida* strains appears to be employing specific activity of amino acid metabolism. This shows a degree of metabolic plasticity indicative of fungi, where secretion of these metabolites associated with these pathways aiding in better adaptability to growth, virulence factor production, hyphae and biofilm formation[39,99], enabling more effective adaptation to a wide variety of environments and habitats. Amino acid metabolism has been proposed as an alternative energy source in stress responses and an alternative to carbon sources for growth. We demonstrate here that all Candida species have an amino acid pathways to employ metabolic remodelling (Fig. 3a)[39,42,89]. However, integrating metabolomics from strains grown with abundant nutrient source, shows that the AGAu group as opposed to other species retains significantly more active in the production of polyamine and choline metabolites, which requires the use of amino acid metabolism for production (Fig. 4). The AGAu group employs arginine, methionine and cysteine metabolism and more extensive exploration and experimental data needed to understand the causal effect of amino acid metabolism. For instance, we have identified a confirmed target pathway for anti-fungal drugs with glutathione metabolism (GSH), attributed to fungal mitochondrial maintenance, preservation of membrane integrity, regulation of transcription factors in stress response and protection

against reactive oxygen species. Reducing activity of GSH is under investigation as supplementary aid for anti-fungal drugs (azoles and echinocandin) against resistant strains [100–103]. This finding verifies that pathway enrichment analysis echoes feasibility in clinical relevance within the host.

Among these metabolites identified are the polyamines, including spermine and spermidine. Polyamines play critical roles in normal cell physiology. Spermine is essential for *Candida* hyphal formation, playing a pivotal role in *Candida* invasion[104]. Spermidine, meanwhile, drives genetic modification in fungi by regulating cell cycle and translating the modification of eukaryotic initiation factor (eIF)[105–107]. Excessive polyamines prolong yeast survival via delayed DNA degradation, increasing the likelihood of mutations that could contribute to the development of anti-fungal resistance [108,109]. These mutations are an important consideration, given that *C. glabrata* and *C. auris* are heavily associated with rising anti-fungal resistance[110–112]. Polyamines have also been shown to be anti-inflammatory, depending on the microenvironment, potentially explaining the additional benefits of secondary metabolites to *Candida* species by modulating host immune responses[113]. The use of polyamines is not limited to fungi. Bacteria use polyamines to create and maintain biofilms in order to withstand host defences as well as promoting cancers[114–116]. Viruses use polyamines to promote cell proliferation, thereby promoting their propagation and spread. Intervention in polyamine synthesis has a high degree of potential as a target for antimicrobials. DNA viruses upregulate polyamine synthesis in host during infection and blocking polyamine synthesis is a strategy used in broad-spectrum anti-viral[117,118,119]. These examples along with our findings here indicate that manipulating polyamine secretion from *Candida* species is a realistic target for therapeutic intervention of associated diseases.

Choline metabolism is a critical function for both microbial and host physiology, as demonstrated by the increase seen in AGAu Candida species' related metabolites. Disruption of phospholipid biosynthesis in fungi can occur through inhibition of phosphatidylcholine synthesis, showing preventing virulence within the systemic mice model[87,88,120,121]. Further, acetylcholine is essential in the formation of the chitin wall characteristic of fungi[122,123]. Along with the bacteriome, *Candida* species contribute towards host acquisition of choline. As understanding of choline metabolism is in its infancy, further investigation of host-mycobiome interactions is needed, potentially providing insights for repurposing potential therapeutic intervention. For instance, lack of choline in humans drives liver dysfunction due

304 to the accumulation of lipids within hepatocytes, which can lead to fatty liver diseases and
305 even hepatic liver cancer[124–128].
306
307 Functional analysis indicates that both the AGAu and non-AGAu groups show a great degree
308 of metabolic plasticity (Fig. 3a)[56,59,60,129]. The only functional difference seen between groups
309 was with CAZymes. The AGAu group presented with GH66, which has previously only
310 been associated with the human oral microbiome and as a potential marker for plaque
311 formation[73]. Given that *C. albicans* is constituent of oral plaque, this is constituent with
312 clinical data (Fig. 2b-c) [130]. The finding of GH43_8 remains relatively inconclusive in
313 function but was recently detected in bacteria as β-galactofuranosidase[131]. Although the
314 modes of action for both GH43_8 and AA5 are currently unknown their enrichment in the
315 AGAu group may provide a function-targeted biomarker for *Candida* infection[132,133]. We also
316 highlight fatty acid biosynthesis pathway in AGAu species with significant levels of
317 triglycerides' production detected (Fig. 3c). Fatty acid synthesis has been identified in
318 *Candida* species previously, with focus on OLE1, FASI & FASII genes as key indicators to
319 pathogenesis and virulence[77–80]. This validates the notion of targeting fatty acid biosynthesis
320 pathway within *Candida* species to disrupt *Candida* overgrowth in the host.
321
322 Our study has addressed the need for functional data and tools for fungal species by
323 developing the BioFung resource using the KEGG database. This enables detailed
324 mechanistic pathway analysis of fungi. Our integrative analysis of the AGAu group
325 (associated with the disease pathogenesis) highlighted key pathways that potentially increase
326 virulence and have associating effects in the host. We hypothesise that these markers can aid
327 in identifying routes for intervention in invasive infection. We suggest polyamine, choline
328 and fatty acid biosynthesis metabolism as inception for further investigation. The presence of
329 these metabolites from AGAu *Candida* species potentially directly affects host homeostasis
330 within the mycobiome and adverse effects on humans during infection. As such, the AGAu
331 *Candida* species' metabolic reprogramming may present a novel method of controlling
332 interaction and infection with these fungi. Finally, we focus on fungal metabolism
333 exploration and distinctively towards amino acid metabolism, playing a more significant role
334 in virulence and pathogenicity.
335
336

Code availability

BioFung is public open access database that can be downloaded at: https://www.microbiomeatlas.org/downloads.php. The instruction and the pipeline scripts for BioFung can be found at our GitHub repository https://github.com/sysbiomelab/BioFung.

Acknowledgment

This study was supported by Engineering and Physical Sciences Research Council (EPSRC), EP/S001301/1, Biotechnology Biological Sciences Research Council (BBSRC) BB/S016899/1, Science for Life Laboratory, Swedish National Infrastructure for Computing at SNIC through Uppsala Multidisciplinary Centre for Advanced Computational Science (UPPMAX) under projects SNIC 2020-5-222, SNIC 2019/3-226, SNIC 2020/6-153 and King's College London computational infrastructure facility, Rosalind (https://rosalind.kcl.ac.uk) for high performance computing. We thank Professor Bernhard Hube for kindly sending *C. parapsilosis* strain.

Author contributions

N.B, D.M. and S.S. conceived the project. N.B. performed sample preparation, metabolomics, gene expression data preparation and extraction protocols for the paper. N.B. developed the pipeline, analysis and made the figures. S.L. advised on design, statistical and functional analysis of the data. A.P.and S.S.N. processed qPCR platform. N.B. wrote and drafted the manuscript. J.N. and M.U. provided critical feedback on the data and manuscript. All authors read, edited and reviewed the manuscript.

Conflict of Interest

The authors declare no competing financial interests.

Methods and Materials

Genome sequence collection. Genome sequences of 49 *Candida* species were collected from NCBI database with version release 45 of Ensemble Fungi (date accessed: April 2019)[134,135]. Supplementary information of sample strain, genome ID, ENA ID, Biosample ID, sequence platform, year of collection, sample location of collection, sample tuple and available biological annotation has been provided in Supplementary Table 1. The quality of the sequence was checked to look at GC content, scaffold and genome size (see Supplementary Figure 1b). Nucleotide sequences used with ANI to determine the phylogenetic relationship and determine differences between strains using Pyani package[136].

Construction of fungi-specific functional database (BioFung). Kyoto Encyclopedia of Genes and Genomes (KEGG) database were downloaded for the investigation of all 128 fungal species (3GB file size) from eukaryote database (5GB file size) (downloaded on August 2019)[137]. Around 1,210,746 genes, which are annotated with 4,717 KEGG orthology (KO), were selected among 128 fungal species genes. There were 6071 fungal genes missing sequence to place into KO, and 105 KO failed in multi-sequence alignment due to default settings (minimum of 3 genes sequence required). Those genes per each KO were performed multiple-sequence alignment by ClustalW and generated Hidden Markov Model (HMM) profiles using the hmm-build function of HMMER software (Figure 1A for workflow and supplementary 2a for coverage)[138,139]. Our database of fungal-specific HMM models of 4,722 KOs was freely shared via Github repository (https://github.com/sysbiomelab/BioFung). Missing KO from fungal species was not added due to missing gene sequences from KEGG, or the low number of sequences per KO (<3), thereby failed to perform ClustalW alignment (See details in Supplementary Table 4). Quality check was performed by comparing BioFung HMMs to pre-trained HMMS for eukaryotes (euk90 and euk100 version 91.0) from Raven Toolbox [57,58,] and we observed that BioFung coverage was much higher than both eukaryote profiles (Supplementary figure1c).

Functional annotations of individual fungi species

Fungal KO annotation of each species was performed by HMM scanning of BioFung HMM models by HMMER software. An in-depth exploratory analysis was performed by manually checking KO annotations of individual species deep. Pathway abundance for AGAu and non-AGAu species was performed using KEGG pathway annotations. Hypergeometric testing of pathways with significance confirmed using Wilcoxon signed-rank test (<0.05). CAZymes

annotations were performed by mapping *Candida* protein sequences using HMMs of dbCAN2 database[66]. Substrate conversion of CAZyme families was checked based on literature review[69,140–146].

Candida protein sequences to map against Pfam-A families using HMMs, that are fully annotated and curated above a threshold[63]. Pfam clans' annotations were sub-set into a broader annotation based on a reported standard function of protein domains (please see Supplementary Table 8).

Contrasted functional annotation of *Candida* species. Presence and absence of microbial annotations, i.e., prevalence, was tested for significance based on condition using Chi-squared tests and odd ratio. Percentage coverage of each was also tested between AGAu and non- AGAu *Candida* species. Contrasted functional annotations were checked on individual strains and placed into presence/absence to perform chi-squared for significance (<0.05), and the odds ratio was performed to identify enriched and depleted in AGAu samples. Additional significant functional annotations are seen in AGAu cluster (Supplementary Table 9).

Clustering of protein sequences. Core, shared, and unique proteins were identified based on sequence similarity by a clustering approach called UCLUST algorithm[65]. In short, UCLUST algorithm was applied to identify similarity in protein-encoding gene sequences by clustering and unique protein sequences were identified if included in singleton protein clusters. Core proteins were identified if corresponding proteins from all 49 species were included in the same cluster. Shared proteins were selected if they did not belong to unique and core proteins. Protein sequence clusters were selected based on a threshold 0.5 for representative seed sequence, a default threshold in UCLUST software (Extended fig. 1h). Based on definitions of the core, shared, and unique proteins, we were able to determine the core, shared and unique annotations for KO and CAZymes, accordingly.

Strain growth. 8 strains of *Candida* species (*C. albicans (*SC5314), *C. dubliniensis* (CD36), *C. tropicalis* (CBS94), *C. glabrata* (CBS 138), *C. auris* (47477), *C. parapsilosis* (73/037), and *C. krusei* (CBS573). Strains were grown in liquid sabouraud dextrose broth (Thermo Scientific-Oxoid microbiology, UK)[147]. All strains were grown in 50ml falcon tube in a shaking incubator 95rpm at the temperature of 25°C to encompass all growth rates. Timepoint measurement of growth was taken to measure the exponential and stationary phase

of the optical density of 1 at 600nm absorbance (iEMS Ascent absorbance 96-well plate reader).

Collection and targeted metabolomics on fungal extracellular matrix. Mid-exponential phase indicates bioactive metabolites and time points for the extraction of extracellular metabolites (see Supplementary Figure 1a). 500µl of extracellular medium, proximity to the pellet was removed from growing fungal cells. Samples were placed through 20µm Whatman filter and snap-frozen in liquid nitrogen. Targeted metabolomics performed using the MxP Quant500 kit (Biocrates, Austria). Partial Least Square – Discriminant Analysis (PLS-DA) was performed on targeted metabolomics of fungal extracellular matrices and media as control, using *ropls* package[148]. First, PLS-DA was performed to distinguish between *Candida* samples and control (media). Further, PLS-DA was performed to distinguish between AGAu species and non- AGAu *Candida* samples. PLS-DA indicated a significant difference between AGAu clusters. Further analysis of metabolite concentrations of targeted metabolomics was normalised, and the Wilcoxon rank-sum test was performed to identify critical metabolites and pathways (<0.05).

Validation experiment

RNA was extracted from 3 biological repeats *C. albicans* (SC5139) using RNA Qiagen Powersoil kit adapted with bead beating with interval placement on dry ice and additional 100µl of isopropanol. DNAse cleanup performed using RNA clean-up and concentration kit (NORGEN, Biotek corporation). Primers designed for specific amplification of genes SPE1 targeting Ornithine Carboxylase, SPE3 gene for spermidine synthase, CKI1 specific for bifunctional choline kinase/ethanolamine kinase, TAZ1 gene focused on lysophosphatidylcholine acyltransferase) and CEM1 gene target for fatty acid synthase (Supplementary Table 10 for primer information). These primers are specific for *C. albicans.* *O*ther *Candida* species only predicted gene ontology-based on *C. albicans* and *Saccharomyces* annotation (http://www.candidagenome.org/cgi-bin/GO/goAnnotation.pl?dbid=CAL0000224407&seq_source=C.%20auris%20B8441). Conventional RT-qPCR was performed to identify the expression of these critical pathways for samples, two standard curve analysis with RDN25 which encodes the 25s rRNA subunit.

Figure legends

Figure 1. a, *Candida* strain characterisation. Coverage of *Candida* sample population per species available with the categorisation of species profiled. Numbers around the pie chart signify the number of strain representation in each location. (Supplementary Table 4 for more information about strains and extended figure 1a for the global representation of Candida strains). b, Genome-based phylogenetic tree. The phylogenetic tree was constructed based on average nucleotide identity (ANI) between all strains revealing evolutionary differences across strains (colour coordinated) and indicating distinct metabolic capabilities. See Extended Fig. 1d for quality of sequences. c, BioFung database creation workflow. Eukaryote annotation from KEGG database parsed to extract all fungal species. They were genes parsed, sequences extracted and reassembled to KO. The multi-sequence alignment was performed on each KO with all corresponding sequence available. HMM, profile built based on each KO and assembled to provide a more accurate annotation of fungal species for KO. d, Distribution of *Candida* species based on sample collection and the framework of protein-encoded genes analysis of *Candida* strains. Strains isolated from the various location providing relevant clinical association to host mycobiome and environmental species. * indicates clinical strains used for metabolomics. Functional analysis performed on 49 *Candida* species collected from public repositories. Protein sequences were annotated with Pfam, dbCAN2 and BioFung database for biological information. e, Core and accessory overview of the metabolic pathway across *Candida* strains. Shared genome feature refers to 6-48 species sharing the function and unique genome features is exhibited by less than 5 *Candida* species denoting accessory functions. f, Clustering of carbohydrate-active enzyme profile (CAZyme). Core, shared genome (6-48 strains), and unique genome (<5 strains) illustrates distribution analysis of functions across all *Candida* species. g, Analysis of enzyme function of glycoside hydrolase family across all *Candida* strains. h, Breakdown of cell wall composition of core *Candida* strains with identification of 49 CAZymes.

Figure 2. a, Contribution of individual *Candida* species to candidemia and mortality. The impact of each species in AGAu species' grouping is attributed in this study (literature-based) [16–20]. b, CAZyme analysis across AGAu and non- AGAu groups. Asterisk connotes statistically significant CAZyme. c, CAZyme enrichment and depletion in AGAu strains. Volcano plot showing statistically significant evidence of a relationship between AGAu strains (chi-squared<0.05, in red). Odds ratio applied to distinguish the enriched and depleted effect in invasive strains.

Figure 3. a, Pathway enrichment in 49 *Candida* strains. Coverage of pathways in AGAu and non-AGAu groups indicating genetic pathway potential (hypergeometric testing and Wilcoxon signed-rank test <0.05). b, Framework outline of metabolomic analysis of *Candida* species' metabolism levels of bioactive metabolites in *Candida* strains exhausted media. Computer simulations were performed for pathway analysis, and statistical approach was applied for candidate metabolites that have a potential influence on the host. c, Significantly enriched metabolites detected in metabolomics in AGAu and non- AGAu groups (Wilcoxon rank-sum test <0.05).

Figure 4. Global map metabolism in AGAu *Candida* strains. Enriched pathways observed in functional annotation and further validated by metabolomics with key metabolites markedly significant (Colour in pathway red identify significant metabolites, blue indicate significant pathway enrichment).

Extended Figure 1. a, Map of *Candida* species collection. Indication of global representation of samples. b, Sequence quality assessment. Genome size variation, GC content, number of contig and outcome of quality of the sequence. c, Comparison of BioFung to available euk90 and euk100 KEGG profiles. Euk90_v91 and euk100_v91 are versions of pre-trained hmm for eukaryote database[62]. Coverage of BioFung was higher than of available hmm. d, Venn diagram displays the overlap of annotation. The overlap between BioFung vs euk90_v91, BioFung vs euk100_v91 and comparison of both eukaryote profiles from all candida gene annotation. BioFung still has an enormous scope of unique KO, and small differences between eukaryote profiles can be observed. e, Validation assessing coverage of the yeast-specific pathway. We performed pathway analysis of KO annotations containing yeast-specific cell pathways. Observation of no yeast specific pathways was found in eukaryote annotation. f, Core percentile coverage of functional annotation. First instance, distribution of KEGG module categories reflect functions seen in most microbiota species. Illustrates the extensive functional capability of *Candida* within each species. g, Core KEGG metabolism exploration in *Candida* species. Amino acid, carbohydrate and lipid metabolism contributes the greatest make up of gene function in *Candida* species. This indicates the contribution of different metabolic potential from *Candida*. h, Uclust quality assessment by looking at the frequency of cluster numbers at 0.5 cut-offs. The cut-ff was inclusive of signature clusters capture and a broad range of core clusters. i, Intra-strain metabolism analysis of *C. albicans*

530 indicating genome differentiation. j, Clustering of 24 *C. albicans* strains of Carbohydrate
531 Active enzymes demonstrating indistinguishable changes in genome.
532 **Extended Figure 2** a, Coverage of CAZyme indicates glycoside hydrolase family is the
533 largest contributor in all *Candida* species. b, Pfam coverage indicates that metabolic
534 processes, genetic information processing and cellular process & machinery provide the
535 immense repertoire of protein function in all *Candida* species. c, Core Pfam metabolic
536 process analysis in *Candida* species. Carbohydrate, amino acid, and lipid-associated
537 processes are the largest metabolomic protein function in *Candida* strains.
538
539 **Extended Figure 3** a, Growth curve of strains grown in the laboratory. All the strains were
540 grown, and timepoint measurement of OD1 at 600nm was taken to obtain the mid-
541 exponential phase. b, Targeted metabolomics with PLS-DA analysis differentiating from
542 media. PLS-DA score indicated a precise fitting performance of clustering media distinctly
543 from *Candida* strains. c, Targeted metabolomics PLS-DA analysis based on AGAu and non-
544 AGAu species. PLS-DA score plot indicated a fitting and predictive performance (2 latent
545 variables, $R^2X = 0.501$, $R^2Y = 0.984$, $Q^2 = 0.76$). d, Analyte Class component features in PLS-
546 DA analysis. A significant number of differential components between invasive and non-
547 invasive strains. e, Heatmap of significant metabolites. Identified from PLS-DA, top 20
548 metabolites differentiated in AGAu and non-AGAu based on Wilcoxon rank-sum test
549 (>0.05).
550
551 **Extended Figure 4** a, Choline metabolic pathway. Concise choline metabolism and its affect.
552 b, Fatty acid metabolism. Concise fatty acid biosynthesis metabolism and its affect. c,
553 Polyamine metabolism. Concise polyamine metabolism and its affect. d, Quantitative RT-
554 qPCR validation of pathways with *C. albicans.* Indicating pathway genes (SPE1, SPE3, CK1,
555 TAZ1, CEM1) compared to their relative expression in *Candida albicans*, representative of
556 AGAu species cluster. (n = 3 biological replicates. Display the relative RNA expression level
557 (Ct value)).
558
559 **Supplementary Table** 1: Information on database collection of 49 *Candida* strains. In-depth
560 data collection of strain retrieval from a database, including strain id, genome id, sample
561 type, origin country, year of collection, annotation availability.

562 **Supplementary Table** 2: Fungal species in KEGG database. List of 128 fungal genes in
563 KEGG database, species abbreviation, T number, protein-encoding genes and genes
564 converted to KO.
565 **Supplementary Table 3**: The building of BioFung. List of fungal-specific KO available in
566 KEGG database, genes available per KO, missing genes from KO, missing sequences of
567 genes unable to include and failure of multi-sequence alignment in creating BioFung.
568 **Supplementary Table 4**, Missing information from the KEGG database. List of genes in
569 fungal species had missing sequences and was unable to collate in profile.
570 **Supplementary Table 5**, AGAu strain categorisation. Based on sample names, grouping,
571 location of sample extraction and reference for strain.
572 **Supplementary Table 6**, The categorisation of Pfam clans. Categories were based on clan
573 function to the associated processes; note clans can fit into many functional categories.
574 **Supplementary Table 7**, Significant of functional annotation in enriched or depleted in
575 AGAu group. Some annotations are statistically significant with the odds ratio of falling
576 within the scale of 0 and 1. They have not been included in the figure but remain relevant.
577 These have been captured in this table for CAZyme, KO and Pfam.
578 **Supplementary Table 8**, References for Pectin CAZyme assay confirmation. List of
579 publications confirming the function of pectin. Table 9, References for RT-qPCR validation.
580 List of gene selection, primer sequence and accession numbers for gene expression analysis.

References

1. Kam, A. P. & Xu, J. Diversity of commensal yeasts within and among healthy hosts. *Diagn. Microbiol. Infect. Dis.* **43**, 19–28 (2002).
2. Bougnoux, M.-E. *et al.* Multilocus sequence typing reveals intrafamilial transmission and microevolutions of Candida albicans isolates from the human digestive tract. *J. Clin. Microbiol.* **44**, 1810–1820 (2006).
3. Angebault, C. *et al.* Candida albicans Is Not Always the Preferential Yeast Colonizing Humans: A Study in Wayampi Amerindians. *J. Infect. Dis.* **208**, 1705–1716 (2013).
4. Moyes, D. L., Richardson, J. P. & Naglik, J. R. Candida albicans-epithelial interactions and pathogenicity mechanisms: scratching the surface. *Virulence* **6**, 338–346 (2015).
5. Meeting on Global Surveillance of Antimicrobial Resistance Invasive Candida Infections. https://www.who.int/news-room/events/detail/2018/04/24/default-calendar/meeting-on-global-surveillance-of-antimicrobial-resistance-invasive-candida-infections.
6. Invasive Fungal Infections: A Creeping Public Health Threat. *ASM.org* https://www.asm.org/Articles/2018/September/Invasive-Fungal-Infections-A-Creeping-Public-Healt.
7. Fungal Disease Frequency | Gaffi - Global Action Fund for Fungal Infections. https://www.gaffi.org/why/fungal-disease-frequency/.
8. Nett, J. E. Candida auris: An emerging pathogen “incognito”? *PLOS Pathog.* **15**, e1007638 (2019).
9. Lockhart, S. R. *et al.* Thinking beyond the Common Candida Species: Need for Species-Level Identification of Candida Due to the Emergence of Multidrug-Resistant Candida auris. *J. Clin. Microbiol.* **55**, 3324–3327 (2017).

605 10. Guinea, J. Global trends in the distribution of Candida species causing candidemia. *Clin.*
606 *Microbiol. Infect.* **20**, 5–10 (2014).
607 11. Yapar, N. Epidemiology and risk factors for invasive candidiasis. *Therapeutics and Clinical*
608 *Risk Management* vol. 10 95–105 https://www.dovepress.com/epidemiology-and-risk-
609 factors-for-invasive-candidiasis-peer-reviewed-article-TCRM (2014).
610 12. Spampinato, C. & Leonardi, D. Candida Infections, Causes, Targets, and Resistance
611 Mechanisms: Traditional and Alternative Antifungal Agents. *BioMed Research*
612 *International* vol. 2013 e204237
613 https://www.hindawi.com/journals/bmri/2013/204237/ (2013).
614 13. Jenkinson, H. F. & Douglas, L. J. *Interactions between Candida Species and Bacteria in*
615 *Mixed Infections*. *Polymicrobial Diseases* (ASM Press, 2002).
616 14. Pannanusorn, S., Fernandez, V. & Römling, U. Prevalence of biofilm formation in clinical
617 isolates of Candida species causing bloodstream infection. *Mycoses* **56**, 264–272 (2013).
618 15. Colombo, A. L. *et al.* Global Distribution and Outcomes for Candida Species Causing
619 Invasive Candidiasis: Results from an International Randomized Double-Blind Study of
620 Caspofungin Versus Amphotericin B for the Treatment of Invasive Candidiasis. *Eur. J.*
621 *Clin. Microbiol. Infect. Dis.* **22**, 470–474 (2003).
622 16. Pfaller, M. A. *et al.* Epidemiology and Outcomes of Invasive Candidiasis Due to Non-
623 albicans Species of Candida in 2,496 Patients: Data from the Prospective Antifungal
624 Therapy (PATH) Registry 2004–2008. *PLOS ONE* **9**, e101510 (2014).
625 17. Lindberg, E., Hammarström, H., Ataollahy, N. & Kondori, N. Species distribution and
626 antifungal drug susceptibilities of yeasts isolated from the blood samples of patients
627 with candidemia. *Sci. Rep.* **9**, 3838 (2019).

628 18. Adam, R. *et al.* 378. Candida auris Fungemia: Risk Factors and Outcome. *Open Forum*
629 *Infect. Dis.* **5**, S147–S147 (2018).
630 19. Almeida, J. N. de *et al.* Candida haemulonii Complex Species, Brazil, January 2010–
631 March 2015 - Volume 22, Number 3—March 2016 - Emerging Infectious Diseases
632 journal - CDC. doi:10.3201/eid2203.151610.
633 20. Gomez-Lopez, A. *et al.* Prevalence and Susceptibility Profile of Candida metapsilosis and
634 Candida orthopsilosis: Results from Population-Based Surveillance of Candidemia in
635 Spain. *Antimicrob. Agents Chemother.* **52**, 1506–1509 (2008).
636 21. Aykut, B. *et al.* The fungal mycobiome promotes pancreatic oncogenesis via activation of
637 MBL. *Nature* 1–4 (2019) doi:10.1038/s41586-019-1608-2.
638 22. Brown, A. J. P., Brown, G. D., Netea, M. G. & Gow, N. A. R. Metabolism impacts upon
639 candida immunogenicity and pathogenicity at multiple levels. *Trends Microbiol.* **22**, 614–
640 622 (2014).
641 23. Miramón, P. & Lorenz, M. C. A feast for Candida: Metabolic plasticity confers an edge for
642 virulence. *PLOS Pathog.* **13**, e1006144 (2017).
643 24. Sandai, D. *et al.* The Evolutionary Rewiring of Ubiquitination Targets Has Reprogrammed
644 the Regulation of Carbon Assimilation in the Pathogenic Yeast Candida albicans. *mBio* **3**,
645 (2012).
646 25. Chew, S. Y. *et al.* Physiologically Relevant Alternative Carbon Sources Modulate Biofilm
647 Formation, Cell Wall Architecture, and the Stress and Antifungal Resistance of Candida
648 glabrata. *Int. J. Mol. Sci.* **20**, 3172 (2019).
649 26. Sandai, D., Tabana, Y. & Sandai, R. Carbon Sources Attribute to Pathogenicity in Candida
650 albicans. *Candida Albicans* (2018) doi:10.5772/intechopen.80211.

651 27. Tucey, T. M. *et al.* Glucose Homeostasis Is Important for Immune Cell Viability during
652 Candida Challenge and Host Survival of Systemic Fungal Infection. *Cell Metab.* **27**, 988-
653 1006.e7 (2018).
654 28. Moyes, D. L. *et al.* Candidalysin is a fungal peptide toxin critical for mucosal infection.
655 *Nature* **532**, 64–68 (2016).
656 29. Naglik, J. R., Gaffen, S. L. & Hube, B. Candidalysin: discovery and function in Candida
657 albicans infections. *Curr. Opin. Microbiol.* **52**, 100–109 (2019).
658 30. Rodaki, A. *et al.* Glucose Promotes Stress Resistance in the Fungal Pathogen Candida
659 albicans. *Mol. Biol. Cell* **20**, 4845–4855 (2009).
660 31. Deacon, J. W. *Fungal Biology*. (John Wiley & Sons, 2013).
661 32. Yin, Z. *et al.* Glucose triggers different global responses in yeast, depending on the
662 strength of the signal, and transiently stabilizes ribosomal protein mRNAs. *Mol.*
663 *Microbiol.* **48**, 713–724 (2003).
664 33. Childers, D. S. *et al.* The Rewiring of Ubiquitination Targets in a Pathogenic Yeast
665 Promotes Metabolic Flexibility, Host Colonization and Virulence. *PLOS Pathog.* **12**,
666 e1005566 (2016).
667 34. Lorenz, M. C., Bender, J. A. & Fink, G. R. Transcriptional Response of Candida albicans
668 upon Internalization by Macrophages. *Eukaryot. Cell* **3**, 1076–1087 (2004).
669 35. Ene, I. V. & Brown, A. J. P. Integration of Metabolism with Virulence in Candida albicans.
670 *Fungal Genomics* 349–370 (2014) doi:10.1007/978-3-642-45218-5_14.
671 36. Ene, I. V. *et al.* Carbon source-induced reprogramming of the cell wall proteome and
672 secretome modulates the adherence and drug resistance of the fungal pathogen
673 Candida albicans. *Proteomics* **12**, 3164–3179 (2012).

674 37. Ene, I. V. *et al.* Host carbon sources modulate cell wall architecture, drug resistance and
675 virulence in a fungal pathogen. *Cell. Microbiol.* **14**, 1319–1335 (2012).
676 38. Tudzynski, B. Nitrogen regulation of fungal secondary metabolism in fungi. *Front.*
677 *Microbiol.* **5**, (2014).
678 39. Garbe, E. & Vylkova, S. Role of Amino Acid Metabolism in the Virulence of Human
679 Pathogenic Fungi. *Curr. Clin. Microbiol. Rep.* **6**, 108–119 (2019).
680 40. Ramachandra, S. *et al.* Regulatory Networks Controlling Nitrogen Sensing and Uptake in
681 Candida albicans. *PLoS ONE* **9**, (2014).
682 41. Maidan, M. M., Thevelein, J. M. & Van Dijck, P. Carbon source induced yeast-to-hypha
683 transition in Candida albicans is dependent on the presence of amino acids and on the
684 G-protein-coupled receptor Gpr1. *Biochem. Soc. Trans.* **33**, 291–293 (2005).
685 42. Vylkova, S. *et al.* The Fungal Pathogen Candida albicans Autoinduces Hyphal
686 Morphogenesis by Raising Extracellular pH. *mBio* **2**, (2011).
687 43. Hudson, D. A. *et al.* Identification of the dialysable serum inducer of germ-tube
688 formation in Candida albicans. *Microbiology,* **150**, 3041–3049 (2004).
689 44. Kaur, R., Ma, B. & Cormack, B. P. A family of glycosylphosphatidylinositol-linked aspartyl
690 proteases is required for virulence of Candida glabrata. *Proc. Natl. Acad. Sci.* **104**, 7628–
691 7633 (2007).
692 45. Cheng, S. *et al.* Profiling of Candida albicans Gene Expression During Intra-abdominal
693 Candidiasis Identifies Biologic Processes Involved in Pathogenesis. *J. Infect. Dis.* **208**,
694 1529–1537 (2013).
695 46. Schrevens, S. *et al.* Methionine is required for cAMP-PKA-mediated morphogenesis and
696 virulence of Candida albicans. *Mol. Microbiol.* **108**, 258–275 (2018).

697 47. Silao, F. G. S. *et al.* Mitochondrial proline catabolism activates Ras1/cAMP/PKA-induced
698 filamentation in Candida albicans. (2019) doi:10.1371/journal.pgen.1007976.
699 48. Han, T.-L., Cannon, R. D., Gallo, S. M. & Villas-Bôas, S. G. A metabolomic study of the
700 effect of Candida albicans glutamate dehydrogenase deletion on growth and
701 morphogenesis. *Npj Biofilms Microbiomes* **5**, 1–14 (2019).
702 49. Naglik, J. R., Challacombe, S. J. & Hube, B. Candida albicans Secreted Aspartyl
703 Proteinases in Virulence and Pathogenesis. *Microbiol. Mol. Biol. Rev.* **67**, 400–428
704 (2003).
705 50. Richardson, J. P. & Moyes, D. L. Adaptive immune responses to Candida albicans
706 infection. *Virulence* **6**, 327–337 (2015).
707 51. Mayer, F. L. *et al.* The Novel Candida albicans Transporter Dur31 Is a Multi-Stage
708 Pathogenicity Factor. *PLOS Pathog.* **8**, e1002592 (2012).
709 52. Global and temporal state of the human gut microbiome in health and disease.
710 https://www.researchsquare.com (2021) doi:10.21203/rs.3.rs-339282/v1.
711 53. Moran, C., Grussemeyer, C. A., Spalding, J. R., Benjamin, D. K. J. & Reed, S. D. CANDIDA
712 ALBICANS AND NON-ALBICANS BLOODSTREAM INFECTIONS IN ADULT AND PEDIATRIC
713 PATIENTS: COMPARISON OF MORTALITY AND COSTS. *Pediatr. Infect. Dis. J.* **28**, 433–435
714 (2009).
715 54. Hirano, R., Sakamoto, Y., Kudo, K. & Ohnishi, M. Retrospective analysis of mortality and
716 Candida isolates of 75 patients with candidemia: a single hospital
717 experience. *Infection and Drug Resistance* vol. 8 199–205
718 https://www.dovepress.com/retrospective-analysis-of-mortality-and-candida-isolates-
719 of-75-patient-peer-reviewed-article-IDR (2015).

720 55. Horn, D. L. *et al.* Epidemiology and Outcomes of Candidemia in 2019 Patients: Data from
721 the Prospective Antifungal Therapy Alliance Registry. *Clin. Infect. Dis.* **48**, 1695–1703
722 (2009).
723 56. Nash, A. K. *et al.* The gut mycobiome of the Human Microbiome Project healthy cohort.
724 *Microbiome* **5**, 153–153 (2017).
725 57. Sam, Q., Chang, M. & Chai, L. The Fungal Mycobiome and Its Interaction with Gut
726 Bacteria in the Host. *Int. J. Mol. Sci.* **18**, 330–330 (2017).
727 58. Seed, P. C. The human mycobiome. *Cold Spring Harb. Perspect. Med.* **5**, a019810–
728 a019810 (2015).
729 59. Galloway-Peña, J. R. & Kontoyiannis, D. P. The gut mycobiome: The overlooked
730 constituent of clinical outcomes and treatment complications in patients with cancer
731 and other immunosuppressive conditions. *PLOS Pathog.* **16**, e1008353 (2020).
732 60. Huseyin, C. E., O'Toole, P. W., Cotter, P. D. & Scanlan, P. D. Forgotten fungi-the gut
733 mycobiome in human health and disease. *FEMS Microbiol. Rev.* **41**, 479–511 (2017).
734 61. Agren, R. *et al.* The RAVEN Toolbox and Its Use for Generating a Genome-scale
735 Metabolic Model for Penicillium chrysogenum. *PLoS Comput. Biol.* **9**, (2013).
736 62. SysBioChalmers/RAVEN. *GitHub* https://github.com/SysBioChalmers/RAVEN.
737 63. El-Gebali, S. *et al.* The Pfam protein families database in 2019. *Nucleic Acids Res.* **47**,
738 D427–D432 (2019).
739 64. dbCAN meta server. http://bcb.unl.edu/dbCAN2/.
740 65. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*
741 **26**, 2460–2461 (2010).
742 66. Zhang, H. *et al.* dbCAN2: a meta server for automated carbohydrate-active enzyme
743 annotation. *Nucleic Acids Res.* **46**, W95–W101 (2018).

744 67. Zhao, Z., Liu, H., Wang, C. & Xu, J.-R. Erratum to: Comparative analysis of fungal
745 genomes reveals different plant cell wall degrading capacity in fungi. *BMC Genomics* **15**,
746 6 (2014).
747 68. Nakagawa, T. *et al.* A Methylotrophic Pathway Participates in Pectin Utilization by
748 Candida boidinii. *Appl. Environ. Microbiol.* **66**, 4253–4257 (2000).
749 69. Barrett, K., Jensen, K., Meyer, A. S., Frisvad, J. C. & Lange, L. Fungal secretome profile
750 categorization of CAZymes by function and family corresponds to fungal phylogeny and
751 taxonomy: Example Aspergillus and Penicillium. *Sci. Rep.* **10**, 1–12 (2020).
752 70. Sfeir *et al.* Breakthrough Bloodstream Infections Caused by Echinocandin-Resistant
753 Candida tropicalis: An Emerging Threat to Immunocompromised Patients with
754 Hematological Malignancies. *J. Fungi* **6**, 20–20 (2020).
755 71. CAZy - GH43_8. http://www.cazy.org/GH43_8.html.
756 72. Mewis, K. Functional metagenomic screening for glycoside hydrolases. (University of
757 British Columbia, 2016). doi:10.14288/1.0340610.
758 73. Cantarel, B. L., Lombard, V. & Henrissat, B. Complex Carbohydrate Utilization by the
759 Healthy Human Microbiome. *PLoS ONE* **7**, (2012).
760 74. Katayama, T. *et al.* Molecular Cloning and Characterization of Bifidobacterium bifidum
761 1,2-α-l-Fucosidase (AfcA), a Novel Inverting Glycosidase (Glycoside Hydrolase Family 95).
762 *J. Bacteriol.* **186**, 4885–4893 (2004).
763 75. Manzo, N. *et al.* Carbohydrate-active enzymes from pigmented Bacilli: a genomic
764 approach to assess carbohydrate utilization and degradation. *BMC Microbiol.* **11**, 198
765 (2011).

766 76. Ramzi, A. B., Me, M. L. C., Ruslan, U. S., Baharum, S. N. & Muhammad, N. A. N. Insight
767 into plant cell wall degradation and pathogenesis of Ganoderma boninense via
768 comparative genome analysis. *PeerJ* **7**, e8065 (2019).
769 77. Nguyen, L. N., Gacser, A. & Nosanchuk, J. D. The Stearoyl-Coenzyme A Desaturase 1 Is
770 Essential for Virulence and Membrane Stress in Candida parapsilosis through
771 Unsaturated Fatty Acid Production. *Infect. Immun.* **79**, 136–145 (2011).
772 78. Rodrigues, M. M. Fatty acid synthase 1 in Candida albicans virulence and the in vitro
773 effects of fluconazole, tetracycline in combinatorial therapy. *Masters Theses* (2012).
774 79. Southard, S. B. & Cihlar, R. L. Analysis and expression of the Candida albicans FAS2 gene.
775 *Gene* **156**, 133–138 (1995).
776 80. Zhao, X. & Cihlar, R. L. Isolation and sequence of the Candida albicans FAS1 gene. *Gene*
777 **147**, 119–124 (1994).
778 81. Odds, F. C., Brown, A. J. & Gow, N. A. Candida albicans genome sequence: a platform for
779 genomics in the absence of genetics. *Genome Biol.* **5**, 230 (2004).
780 82. Bader, O. *et al.* Gross Karyotypic and Phenotypic Alterations among Different Progenies
781 of the Candida glabrata CBS138/ATCC2001 Reference Strain. *PLOS ONE* **7**, e52218
782 (2012).
783 83. Douglass, A. P. *et al.* Population genomics shows no distinction between pathogenic
784 Candida krusei and environmental Pichia kudriavzevii: One species, four names. *PLoS*
785 *Pathog.* **14**, e1007138 (2018).
786 84. Barchiesi, F. *et al.* Experimental Induction of Fluconazole Resistance in Candida tropicalis
787 ATCC 750. *Antimicrob. Agents Chemother.* **44**, 1578–1584 (2000).
788 85. Alcazar-Fuoli, L. Amino acid biosynthetic pathways as antifungal targets for fungal
789 infections. *Virulence* **7**, 376–378 (2016).

790 86. Jastrzębowska, K. & Gabriel, I. Inhibitors of amino acids biosynthesis as antifungal
791 agents. *Amino Acids* **47**, 227–249 (2015).
792 87. Tams, R. N. *et al.* Overproduction of Phospholipids by the Kennedy Pathway Leads to
793 Hypervirulence in Candida albicans. *Front. Microbiol.* **10**, (2019).
794 88. Klig, L. S., Friedli, L. & Schmid, E. Phospholipid biosynthesis in Candida albicans:
795 regulation by the precursors inositol and choline. *J. Bacteriol.* **172**, 4407–4414 (1990).
796 89. Mayer, F. L., Wilson, D. & Hube, B. Candida albicans pathogenicity mechanisms.
797 *Virulence* **4**, 119–128 (2013).
798 90. Lorenz, M. C. & Fink, G. R. The glyoxylate cycle is required for fungal virulence. *Nature*
799 **412**, 83–86 (2001).
800 91. Minois, N., Carmona-Gutierrez, D. & Madeo, F. Polyamines in aging and disease. *Aging*
801 **3**, 716–732 (2011).
802 92. Valdés-Santiago, L. & Ruiz-Herrera, J. Stress and polyamine metabolism in fungi. *Front.*
803 *Chem.* **1**, (2014).
804 93. Noverr, M. C. & Huffnagle, G. B. Regulation of Candida albicans Morphogenesis by Fatty
805 Acid Metabolites. *Infect. Immun.* **72**, 6206–6210 (2004).
806 94. Nguyen, L. N., Trofa, D. & Nosanchuk, J. D. Fatty Acid Synthase Impacts the Pathobiology
807 of Candida parapsilosis In Vitro and during Mammalian Infection. *PLOS ONE* **4**, e8421
808 (2009).
809 95. Pan, J., Hu, C. & Yu, J.-H. Lipid Biosynthesis as an Antifungal Target. *J. Fungi* **4**, (2018).
810 96. Gajewski, J., Pavlovic, R., Fischer, M., Boles, E. & Grininger, M. Engineering fungal de
811 novo fatty acid synthesis for short chain fatty acid production. *Nat. Commun.* **8**, 14650
812 (2017).

813 97. Tehlivets, O., Scheuringer, K. & Kohlwein, S. D. Fatty acid synthesis and elongation in
814 yeast. *Biochim. Biophys. Acta BBA - Mol. Cell Biol. Lipids* **1771**, 255–270 (2007).
815 98. Rattray, J. B., Schibeci, A. & Kidby, D. K. Lipids of yeasts. *Bacteriol. Rev.* **39**, 197–231
816 (1975).
817 99. Desai, J. V. Candida albicans Hyphae: From Growth Initiation to Invasion. *J. Fungi* **4**,
818 (2018).
819 100. Miramón, P. *et al.* A family of glutathione peroxidases contributes to oxidative stress
820 resistance in Candida albicans. *Med. Mycol.* **52**, 223–239 (2014).
821 101. Liu, Y. *et al.* Glutathione Reductase Promotes Fungal Clearance and Suppresses
822 Inflammation During Systemic Candida albicans Infection. *J. Immunol.* **200**, 50.1-50.1
823 (2018).
824 102. I, P., Ra, P. & Mj, P. Glutathione, altruistic metabolite in fungi. *Adv. Microb. Physiol.*
825 **49**, 1–76 (2004).
826 103. Maras, B. *et al.* Glutathione Metabolism in Candida albicans Resistant Strains to
827 Fluconazole and Micafungin. *PLOS ONE* **9**, e98387 (2014).
828 104. Dorighetto Cogo, A. J. *et al.* Spermine modulates fungal morphogenesis and activates
829 plasma membrane H^+-ATPase during yeast to hyphae transition. *Biol. Open* **7**,
830 bio029660 (2018).
831 105. Hoyt, M. A. & Davis, R. H. Polyamines in Fungi. in *Biochemistry and Molecular Biology*
832 (eds. Brambl, R. & Marzluf, G. A.) 335–347 (Springer, 2004). doi:10.1007/978-3-662-
833 06064-3_16.
834 106. Tome, M. E. & Gerner, E. W. Cellular eukaryotic initiation factor 5A content as a
835 mediator of polyamine effects on growth and apoptosis. *Biol. Signals* **6**, 150–156 (1997).

836 107. Belda-Palazón, B. *et al.* Biochemical quantitation of the eIF5A hypusination in
837 Arabidopsis thaliana uncovers ABA-dependent regulation. *Front. Plant Sci.* **5**, (2014).
838 108. Yu, Y. *et al.* Position of eukaryotic translation initiation factor eIF1A on the 40S
839 ribosomal subunit mapped by directed hydroxyl radical probing. *Nucleic Acids Res.* **37**,
840 5167–5182 (2009).
841 109. Chattopadhyay, M. K. *et al.* Microarray studies on the genes responsive to the
842 addition of spermidine or spermine to a Saccharomyces cerevisiae spermidine synthase
843 mutant. *Yeast Chichester Engl.* **26**, 531–544 (2009).
844 110. Bhattacharya, S., Sae-Tia, S. & Fries, B. C. Candidiasis and Mechanisms of Antifungal
845 Resistance. *Antibiotics* **9**, 312 (2020).
846 111. Hendrickson, J. A., Hu, C., Aitken, S. L. & Beyda, N. Antifungal Resistance: a
847 Concerning Trend for the Present and Future. *Curr. Infect. Dis. Rep.* **21**, 47 (2019).
848 112. Silva, S. *et al.* Candida glabrata, Candida parapsilosis and Candida tropicalis: biology,
849 epidemiology, pathogenicity and antifungal resistance. *FEMS Microbiol. Rev.* **36**, 288–
850 305 (2012).
851 113. Anti-inflammatory effects of spermidine in lipopolysaccharide-stimulated BV2
852 microglial cells | Journal of Biomedical Science | Full Text.
853 https://jbiomedsci.biomedcentral.com/articles/10.1186/1423-0127-19-31.
854 114. Li, B. *et al.* Polyamine-independent growth and biofilm formation, and functional
855 spermidine/spermine N-acetyltransferases in Staphylococcus aureus and Enterococcus
856 faecalis. *Mol. Microbiol.* **111**, 159–175 (2019).
857 115. Patel, C. N. *et al.* Polyamines Are Essential for the Formation of Plague Biofilm. *J.*
858 *Bacteriol.* **188**, 2355–2363 (2006).

859 116. Villanueva, M. T. Bacterial biofilms may feed colon cancer. *Nat. Rev. Cancer* **15**, 320–
860 320 (2015).
861 117. Li, M. M. H. & MacDonald, M. R. Polyamines: Small Molecules with a Big Role in
862 Promoting Virus Infection. *Cell Host Microbe* **20**, 123–124 (2016).
863 118. Firpo, M. R. & Mounce, B. C. Diverse Functions of Polyamines in Virus Infection.
864 *Biomolecules* **10**, 628 (2020).
865 119. Mounce, B. C., Olsen, M. E., Vignuzzi, M. & Connor, J. H. Polyamines and Their Role in
866 Virus Infection. *Microbiol. Mol. Biol. Rev.* **81**, (2017).
867 120. Wang, J. *et al.* Phospholipid homeostasis plays an important role in fungal
868 development, fungicide resistance and virulence in Fusarium graminearum.
869 *Phytopathol. Res.* **1**, 16 (2019).
870 121. Chen, Y.-L. *et al.* Phosphatidylserine synthase and phosphatidylserine decarboxylase
871 are essential for cell wall integrity and virulence in Candida albicans. *Mol. Microbiol.* **75**,
872 1112–1132 (2010).
873 122. A, A. & Sm, K. Acetylcholine Induces Yeast to Hyphal Form Transition in Candida
874 albicans. *Fungal Genomics Biol.* **08**, (2018).
875 123. Markham, P., Robson, G. D., Bainbridge, B. W. & Trinci, A. P. J. Choline: Its role in the
876 growth of filamentous fungi and the regulation of mycelial morphology. *FEMS Microbiol.*
877 *Rev.* **10**, 287–300 (1993).
878 124. Mehedint, M. G. & Zeisel, S. H. Choline's role in maintaining liver function: new
879 evidence for epigenetic mechanisms. *Curr. Opin. Clin. Nutr. Metab. Care* **16**, 339–345
880 (2013).

881 125. Sherriff, J. L., O'Sullivan, T. A., Properzi, C., Oddo, J.-L. & Adams, L. A. Choline, Its
882 Potential Role in Nonalcoholic Fatty Liver Disease, and the Case for Human and Bacterial
883 Genes. *Adv. Nutr.* **7**, 5–13 (2016).
884 126. Brown, A. L. *et al.* Dietary Choline Supplementation Attenuates High-Fat-Diet–
885 Induced Hepatocellular Carcinoma in Mice. *J. Nutr.* **150**, 775–783 (2020).
886 127. Sun, S. *et al.* Choline and betaine consumption lowers cancer risk: a meta-analysis of
887 epidemiologic studies. *Sci. Rep.* **6**, 35547 (2016).
888 128. Zhou, R. *et al.* Higher dietary intakes of choline and betaine are associated with a
889 lower risk of primary liver cancer: a case-control study. *Sci. Rep.* **7**, 679 (2017).
890 129. Forbes, J. D., Bernstein, C. N., Tremlett, H., Van Domselaar, G. & Knox, N. C. A fungal
891 world: Could the gut mycobiome be involved in neurological disease? *Front. Microbiol.*
892 **10**, 3249–3249 (2019).
893 130. Zijnge, V. *et al.* Oral Biofilm Architecture on Natural Teeth. *PLOS ONE* **5**, e9321
894 (2010).
895 131. Mewis, K., Lenfant, N., Lombard, V. & Henrissat, B. Dividing the Large Glycoside
896 Hydrolase Family 43 into Subfamilies: a Motivation for Detailed Enzyme
897 Characterization. *Appl. Environ. Microbiol.* **82**, 1686–1692 (2016).
898 132. Arfi, Y. *et al.* Characterization of salt-adapted secreted lignocellulolytic enzymes from
899 the mangrove fungus Pestalotiopsis sp . *Nat. Commun.* **4**, 1810 (2013).
900 133. Kumar, V. *et al.* Poplar carbohydrate-active enzymes: whole-genome annotation and
901 functional analyses based on RNA expression data. *Plant J.* **99**, 589–609 (2019).
902 134. National Institutes of Health (U.S.). Office of Communications and Public Liaison. *NIH*
903 *research matters.*
904 135. Ensembl genome browser 100. https://www.ensembl.org/index.html.

905 136. Pritchard, L. *pyani: pyani provides a package and script for calculation of genome-*
906 *scale average nucleotide identity.*
907 137. KEGG: Kyoto Encyclopedia of Genes and Genomes.
908 https://www.genome.jp/kegg/?sess=ebfe2ad23e021e38540f798c803dd061.
909 138. ClustalW2 < Multiple Sequence Alignment < EMBL-EBI.
910 https://www.ebi.ac.uk/Tools/msa/clustalw2/.
911 139. Eddy, S. R. What is a hidden Markov model? *Nat. Biotechnol.* **22**, 1315–1316 (2004).
912 140. Kovatcheva-Datchary, P. *et al.* Simplified Intestinal Microbiota to Study Microbe-
913 Diet-Host Interactions in a Mouse Model. *Cell Rep.* **26**, 3772-3783.e6 (2019).
914 141. Zhao, Z., Liu, H., Wang, C. & Xu, J.-R. Comparative analysis of fungal genomes reveals
915 different plant cell wall degrading capacity in fungi. *BMC Genomics* **14**, 274 (2013).
916 142. Borin, G. P. *et al.* Comparative Secretome Analysis of Trichoderma reesei and
917 Aspergillus niger during Growth on Sugarcane Biomass. *PLOS ONE* **10**, e0129275 (2015).
918 143. Baroncelli, R. *et al.* Gene family expansions and contractions are associated with host
919 range in plant pathogens of the genus Colletotrichum. *BMC Genomics* **17**, 555 (2016).
920 144. Looi, H. K. *et al.* Genomic insight into pathogenicity of dematiaceous fungus
921 Corynespora cassiicola. *PeerJ* **5**, e2841 (2017).
922 145. Wegmann, U. *et al.* Complete genome of a new Firmicutes species belonging to the
923 dominant human colonic microbiota ('Ruminococcus bicirculans') reveals two
924 chromosomes and a selective capacity to utilize plant glucans. *Environ. Microbiol.* **16**,
925 2879–2890 (2014).
926 146. Geisler-Lee, J. *et al.* Poplar carbohydrate-active enzymes. Gene identification and
927 expression analyses. *Plant Physiol.* **140**, 946–962 (2006).

928 147. Oxoid - Product Detail.

929 http://www.oxoid.com/UK/blue/prod_detail/prod_detail.asp?pr=CM0041.

930 148. Analysis of the Human Adult Urinary Metabolome Variations with Age, Body Mass

931 Index, and Gender by Implementing a Comprehensive Workflow for Univariate and

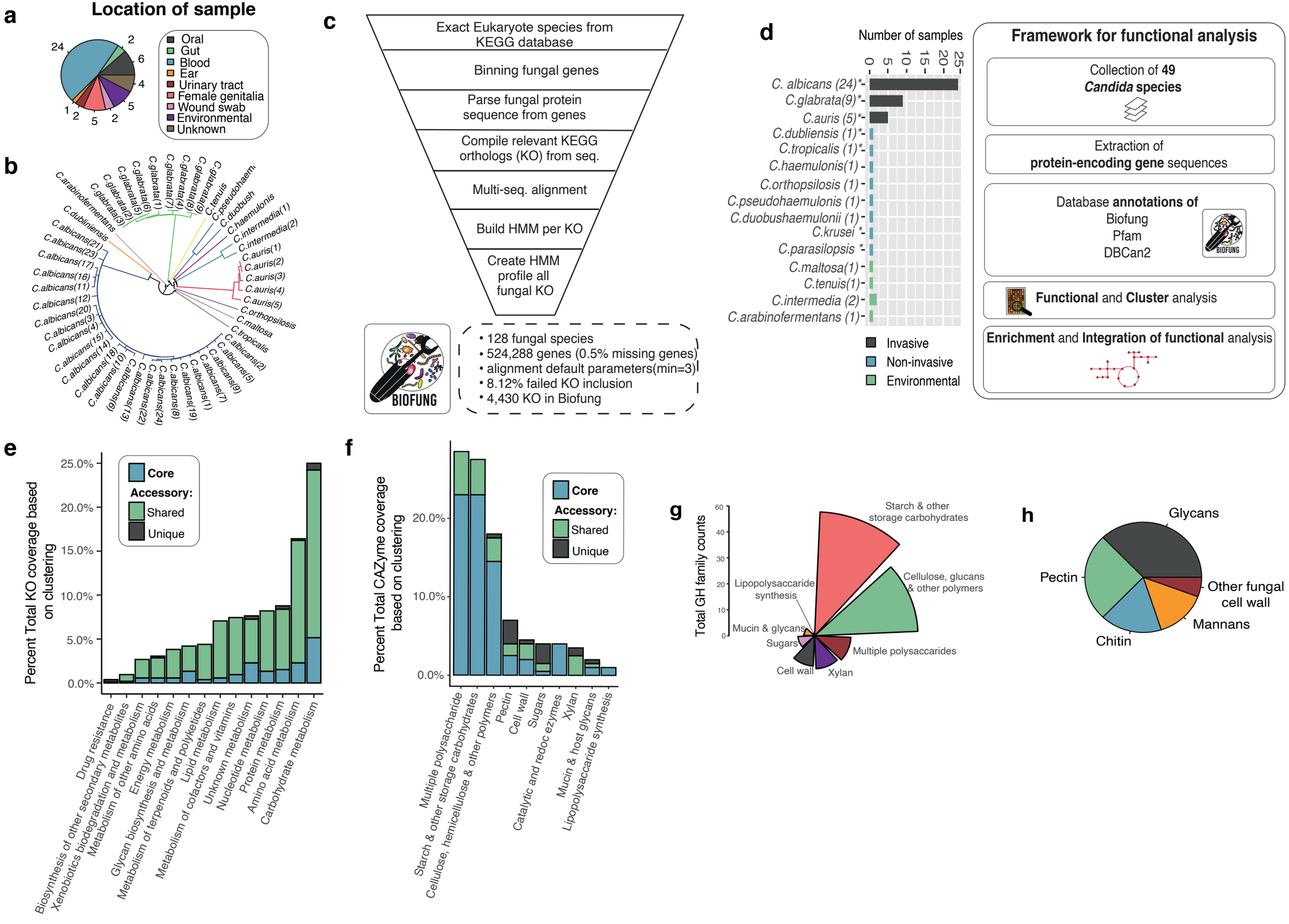
932 OPLS Statistical Analyses | Journal of Proteome Research.

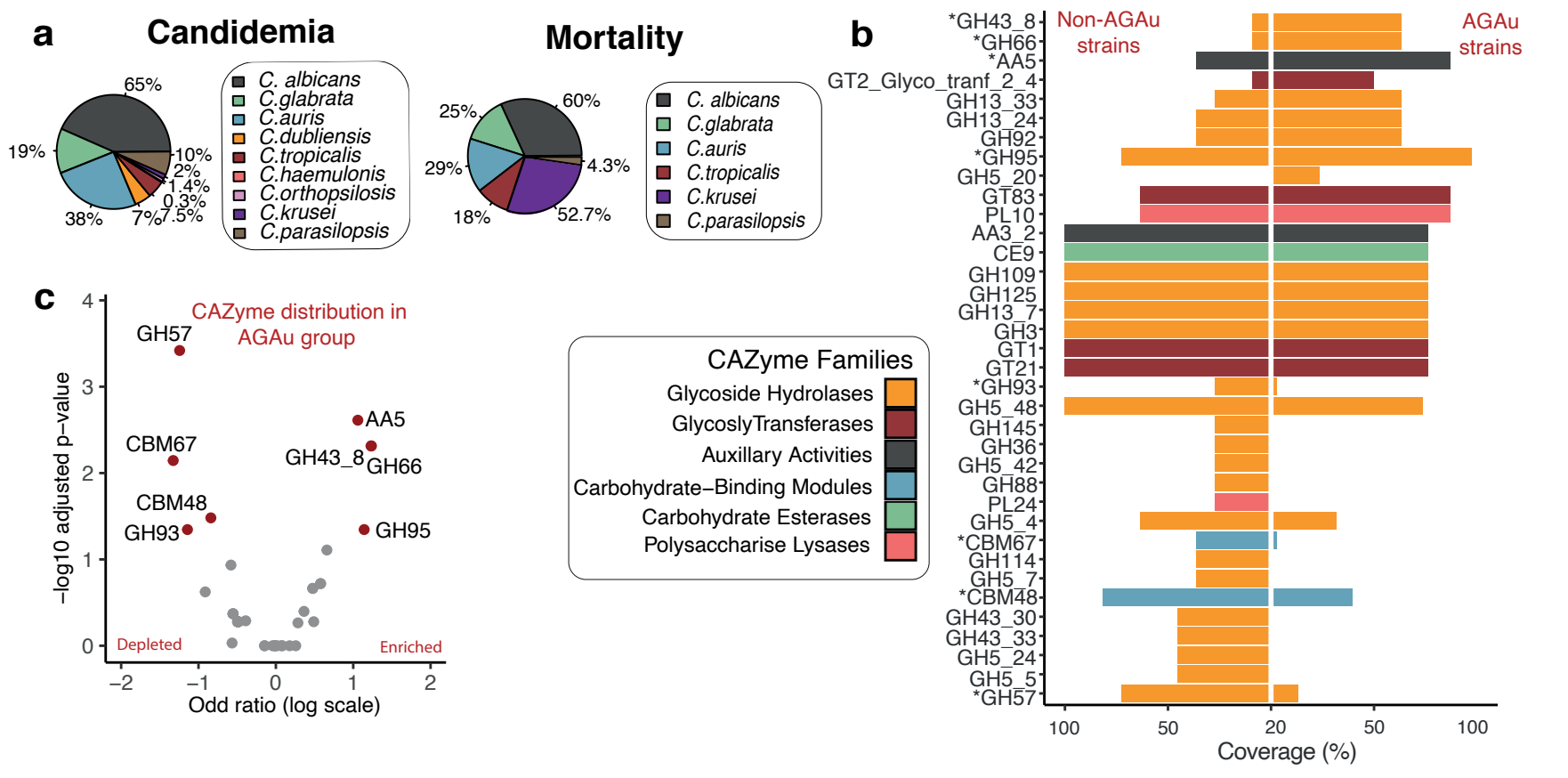
933 https://pubs.acs.org/doi/full/10.1021/acs.jproteome.5b00354.

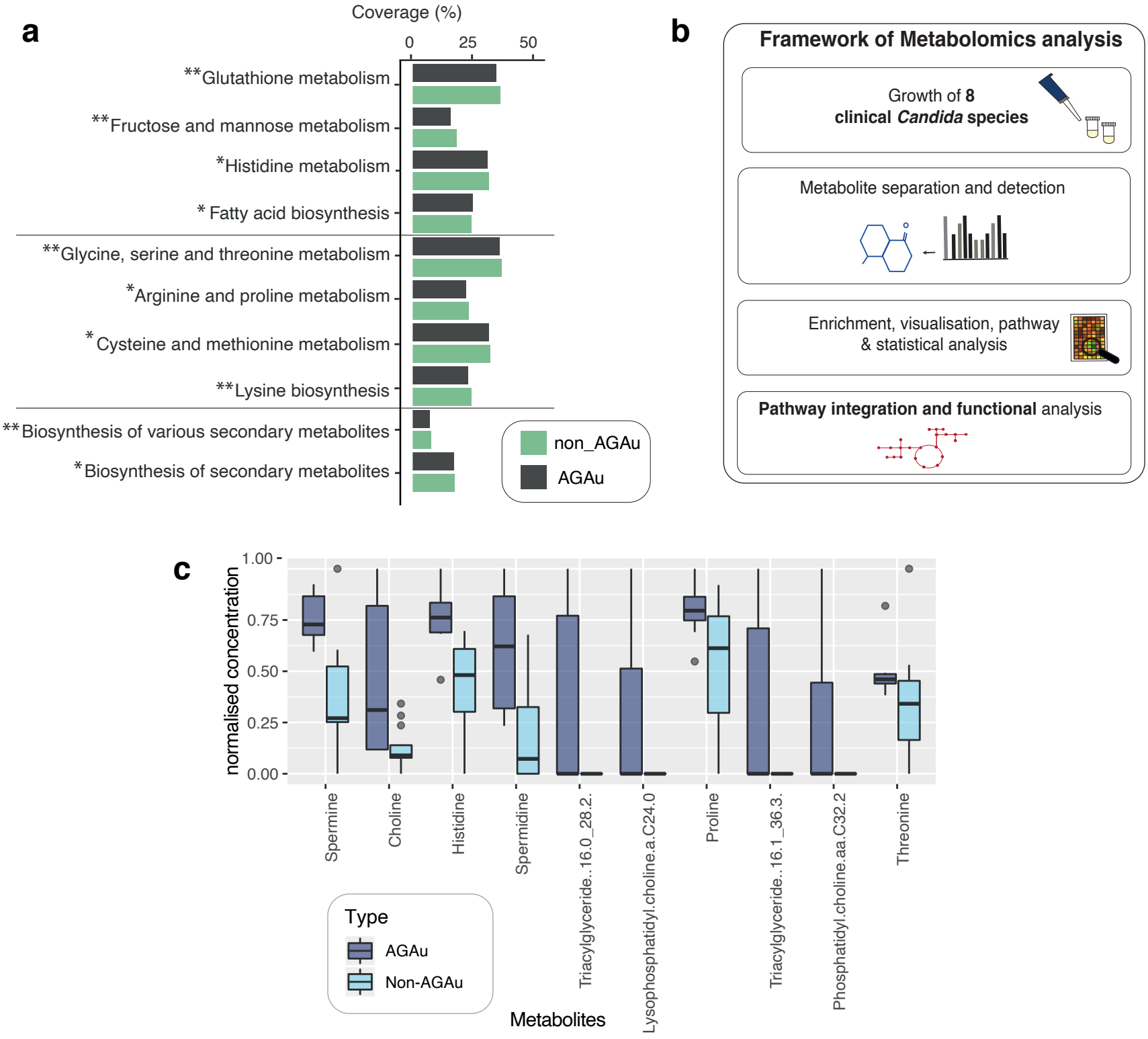
934 935

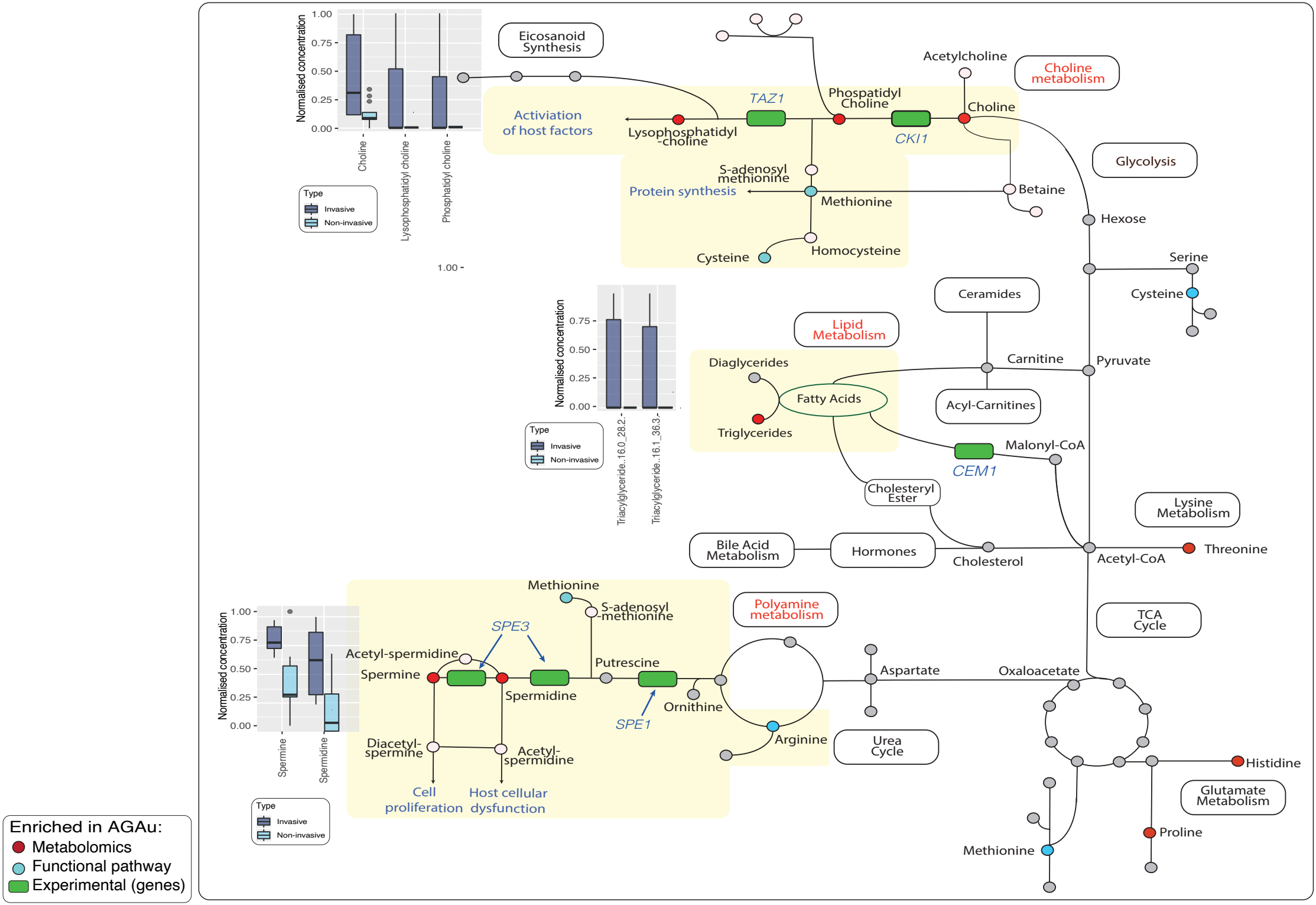
936

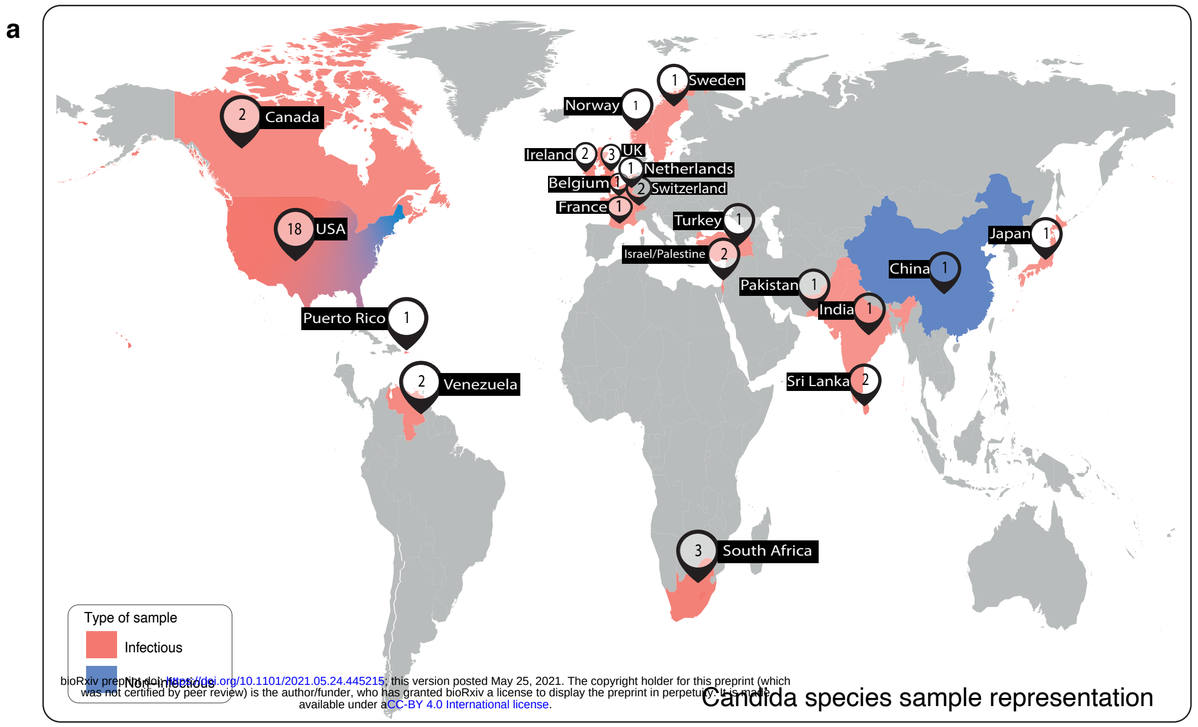
937

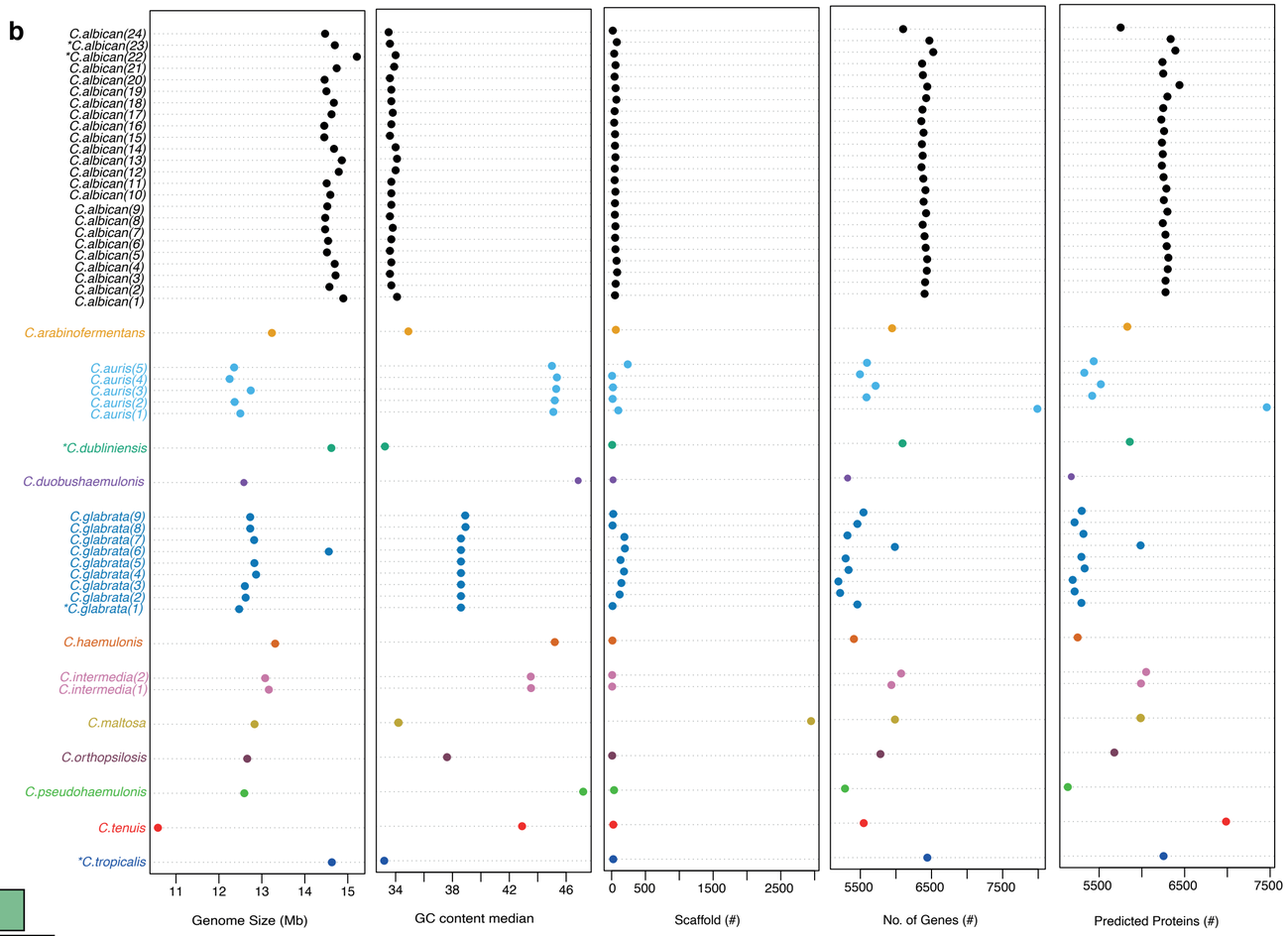

a
Location of sample
24
2
6
4
5
2
5
2
1
Oral
Gut
Blood
Ear
Urinary tract
Female genitalia
Wound swab
Environmental
Unknown
b
C.glabrata(1)
C.glabrata(7)
C.glabrata(4)
C.glabrata(8)
C.glabrata(9)
C.tenuis
C.pseudohaem.
C.duobush
C.haemulonis
C.intermedia(1)
C.intermedia(2)
C.auris(1)
C.auris(2)
C.auris(3)
C.auris(4)
C.auris(5)
C.orthopsilosis
C.maltosa
C.tropicalis
C.albicans(2)
C.albicans(5)
C.albicans(9)
C.albicans(7)
C.albicans(1)
C.albicans(19)
C.albicans(8)
C.albicans(24)
C.albicans(22)
C.albicans(13)
C.albicans(6)
C.albicans(10)
C.albicans(18)
C.albicans(14)
C.albicans(15)
C.albicans(4)
C.albicans(3)
C.albicans(20)
C.albicans(12)
C.albicans(11)
C.albicans(16)
C.albicans(17)
C.albicans(23)
C.albicans(21)
C.dubliniensis
C.arabinofermentans
C.glabrata(3)
C.glabrata(2)
C.glabrata(5)
C.glabrata(6)
c
Exact Eukaryote species from KEGG database
Binning fungal genes
Parse fungal protein sequence from genes
Compile relevant KEGG orthologs (KO) from seq.
Multi-seq. alignment
Build HMM per KO
Create HMM profile all fungal KO
BIOFUNG
• 128 fungal species
• 524,288 genes (0.5% missing genes)
• alignment default parameters(min=3)
• 8.12% failed KO inclusion
• 4,430 KO in Biofung
d
Number of samples
0
5
10
15
20
25
C. albicans (24)*
C.glabrata(9)*
C.auris (5)*
C.dubliensis (1)*
C.tropicalis (1)*
C.haemulonis(1)
C.orthopsilosis (1)
C.pseudohaemulonis (1)
C.duobushaemulonii (1)
C.krusei *
C.parasilopsis *
C.maltosa(1)
C.tenuis(1)
C.intermedia (2)
C.arabinofermentans (1)
Invasive
Non-invasive
Environmental
Framework for functional analysis
Collection of 49 Candida species
Extraction of protein-encoding gene sequences
Database annotations of
Biofung
Pfam
DBCan2
BIOFUNG
Functional and Cluster analysis
Enrichment and Integration of functional analysis
e
Percent Total KO coverage based on clustering
25.0%
20.0%
15.0%
10.0%
5.0%
0.0%
Core
Accessory:
Shared
Unique
Drug resistance
Biosynthesis of other secondary metabolites
Xenobiotics biodegradation and metabolism
Metabolism of other amino acids
Energy metabolism
Glycan biosynthesis and metabolism
Metabolism of terpenoids and polyketides
Lipid metabolism
Metabolism of cofactors and vitamins
Unknown metabolism
Nucleotide metabolism
Protein metabolism
Amino acid metabolism
Carbohydrate metabolism
f
Percent Total CAZyme coverage based on clustering
20.0%
10.0%
0.0%
Core
Accessory:
Shared
Unique
Multiple polysaccharide
Starch & other storage carbohydrates
Cellulose, hemicellulose & other polymers
Pectin
Cell wall
Sugars
Catalytic and redoc ezymes
Xylan
Mucin & host glycans
Lipopolysaccaride synthesis
g
Total GH family counts
60
40
30
20
10
0
Starch & other storage carbohydrates
Cellulose, glucans & other polymers
Multiple polysaccarides
Xylan
Cell wall
Sugars
Mucin & glycans
Lipopolysaccaride synthesis
h
Glycans
Other fungal cell wall
Mannans
Chitin
Pectin

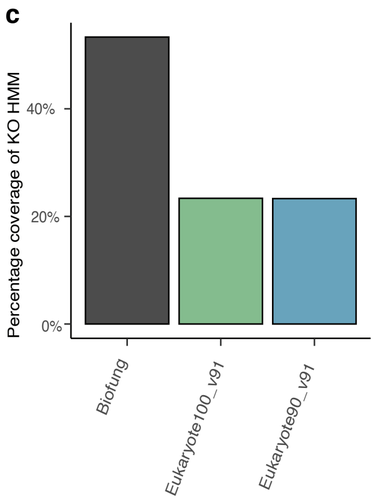

a
Candidemia
65%
19%
10%
2%
1.4%
0.3%
38%
7%
7.5%
C. albicans
C.glabrata
C.auris
C.dubliensis
C.tropicalis
C.haemulonis
C.orthopsilosis
C.krusei
C.parasilopsis
Mortality
25%
60%
29%
4.3%
18%
52.7%
C. albicans
C.glabrata
C.auris
C.tropicalis
C.krusei
C.parasilopsis
b
Non-AGAu strains
AGAu strains
*GH43_8
*GH66
*AA5
GT2_Glyco_tranf_2_4
GH13_33
GH13_24
GH92
*GH95
GH5_20
GT83
PL10
AA3_2
CE9
GH109
GH125
GH13_7
GH3
GT1
GT21
*GH93
GH5_48
GH145
GH36
GH5_42
GH88
PL24
GH5_4
*CBM67
GH114
GH5_7
*CBM48
GH43_30
GH43_33
GH5_24
GH5_5
*GH57
100
50
20
50
100
Coverage (%)
c
CAZyme distribution in AGAu group
GH57
AA5
CBM67
GH43_8
GH66
CBM48
GH93
GH95
Depleted
Enriched
−log10 adjusted p−value
Odd ratio (log scale)
CAZyme Families
Glycoside Hydrolases
GlycoslyTransferases
Auxillary Activities
Carbohydrate−Binding Modules
Carbohydrate Esterases
Polysaccharise Lysases

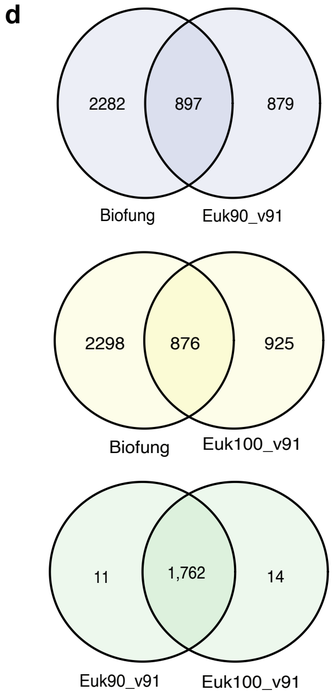

a
Coverage (%)
0
25
50
**Glutathione metabolism
**Fructose and mannose metabolism
*Histidine metabolism
* Fatty acid biosynthesis
**Glycine, serine and threonine metabolism
*Arginine and proline metabolism
* Cysteine and methionine metabolism
**Lysine biosynthesis
**Biosynthesis of various secondary metabolites
*Biosynthesis of secondary metabolites
non_AGAu
AGAu
b
Framework of Metabolomics analysis
Growth of 8
clinical Candida species
Metabolite separation and detection
Enrichment, visualisation, pathway
& statistical analysis
Pathway integration and functional analysis
c
normalised concentration
1.00
0.75
0.50
0.25
0.00
Spermine
Choline
Histidine
Spermidine
Triacylglyceride..16.0_28.2.
Lysophosphatidyl.choline.a.C24.0
Proline
Triacylglyceride..16.1_36.3.
Phosphatidyl.choline.aa.C32.2
Threonine
Type
AGAu
Non-AGAu
Metabolites

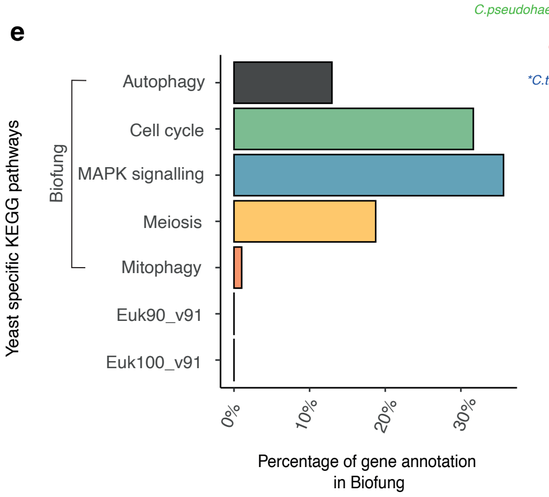

Normalised concentration
1.00
0.75
0.50
0.25
0.00
Choline
Lysophosphatidyl choline
Phosphatidyl choline
Type
Invasive
Non-invasive
Eicosanoid Synthesis
Activiation of host factors
TAZ1
Lysophosphatidyl -choline
Phospatidyl Choline
CKI1
Acetylcholine
Choline
Choline metabolism
Glycolysis
S-adenosyl methionine
Protein synthesis
Methionine
Betaine
Cysteine
Homocysteine
Hexose
Serine
Cysteine
Ceramides
Lipid Metabolism
Diaglycerides
Triglycerides
Fatty Acids
Carnitine
Pyruvate
Acyl-Carnitines
Malonyl-CoA
CEM1
Cholesteryl Ester
Lysine Metabolism
Bile Acid Metabolism
Hormones
Cholesterol
Acetyl-CoA
Threonine
Triacylglyceride..16.0_28.2.
Triacylglyceride..16.1_36.3.
Methionine
S-adenosyl -methionine
Polyamine metabolism
SPE3
Acetyl-spermidine
Spermine
Spermidine
Putrescine
Ornithine
SPE1
Arginine
Urea Cycle
Aspartate
Oxaloacetate
TCA Cycle
Histidine
Glutamate Metabolism
Proline
Methionine
Diacetyl- spermine
Acetyl- spermidine
Cell proliferation
Host cellular dysfunction
Enriched in AGAu:
Metabolomics
Functional pathway
Experimental (genes)

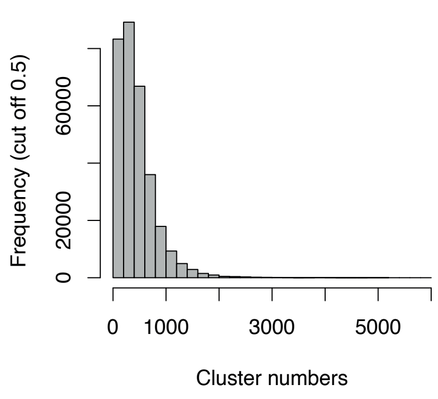


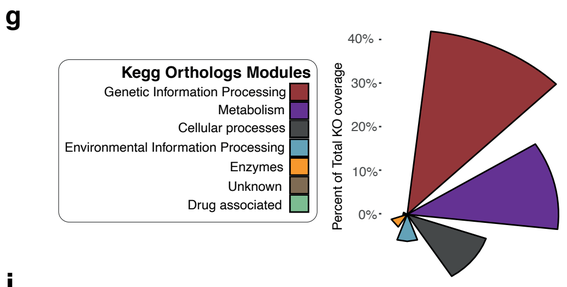


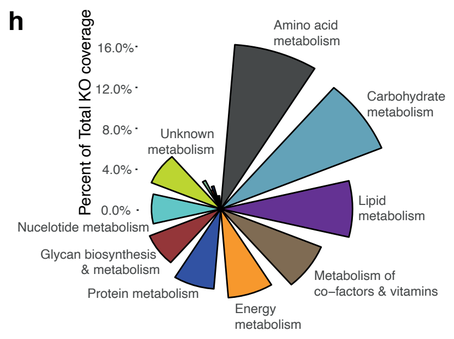


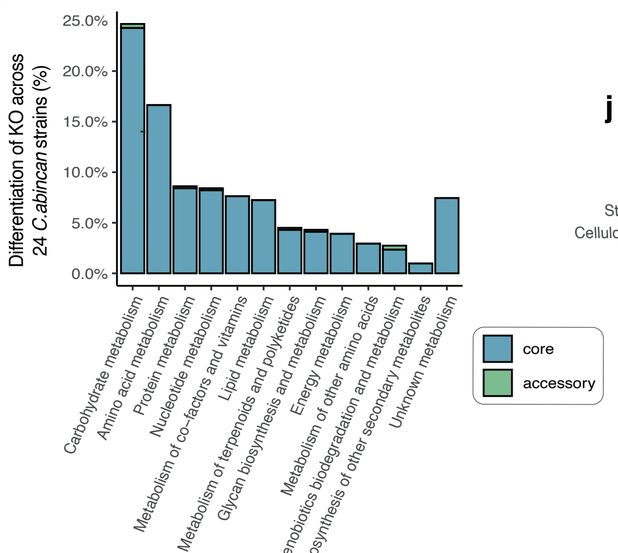


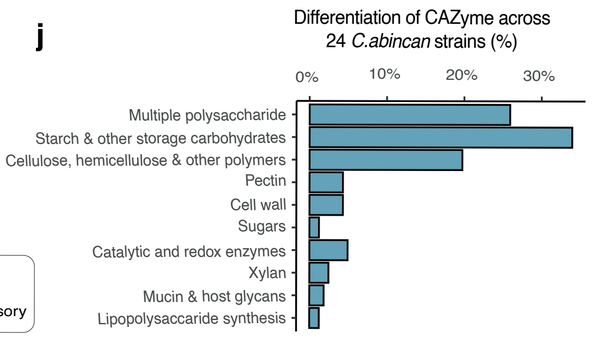


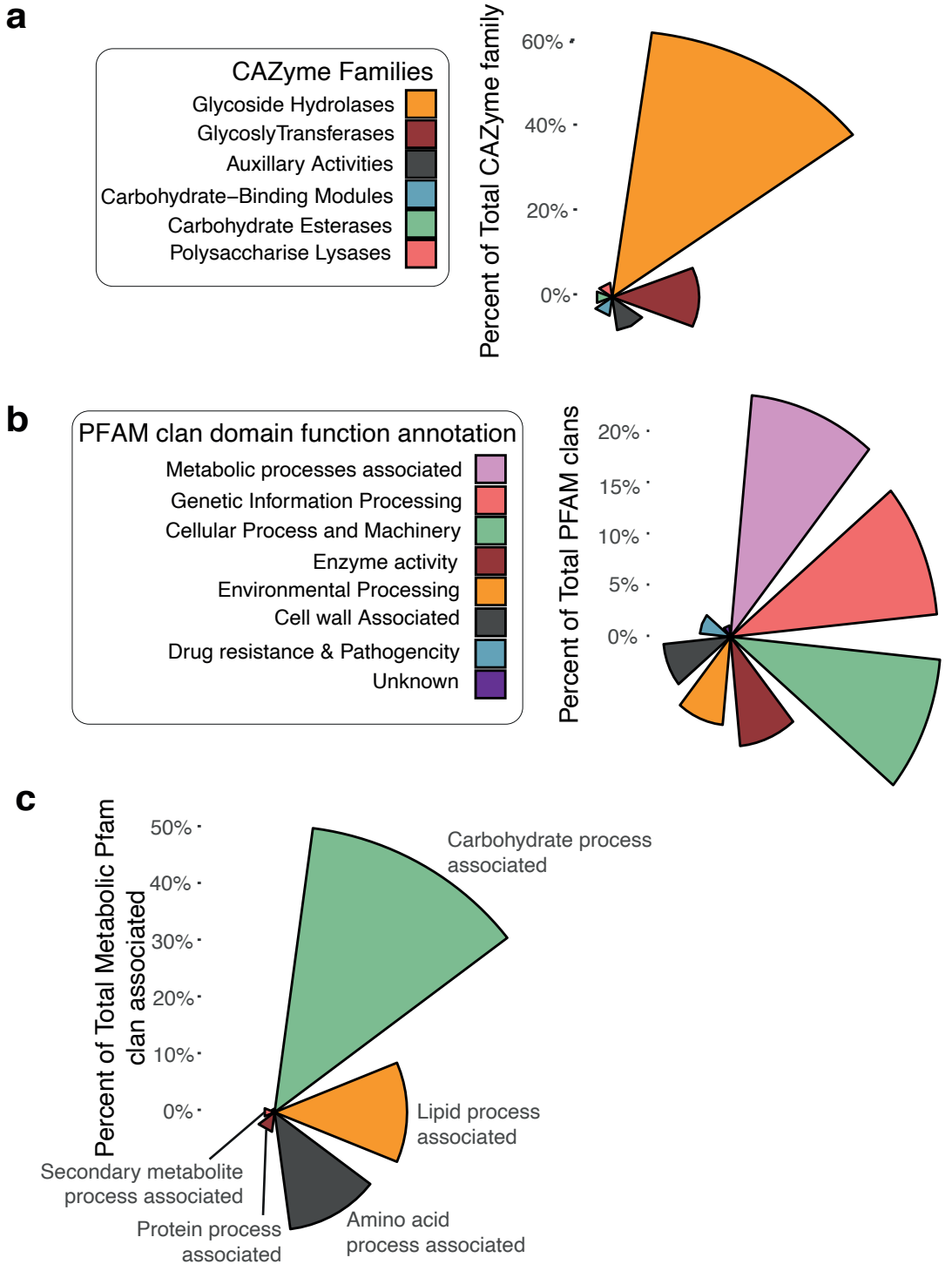


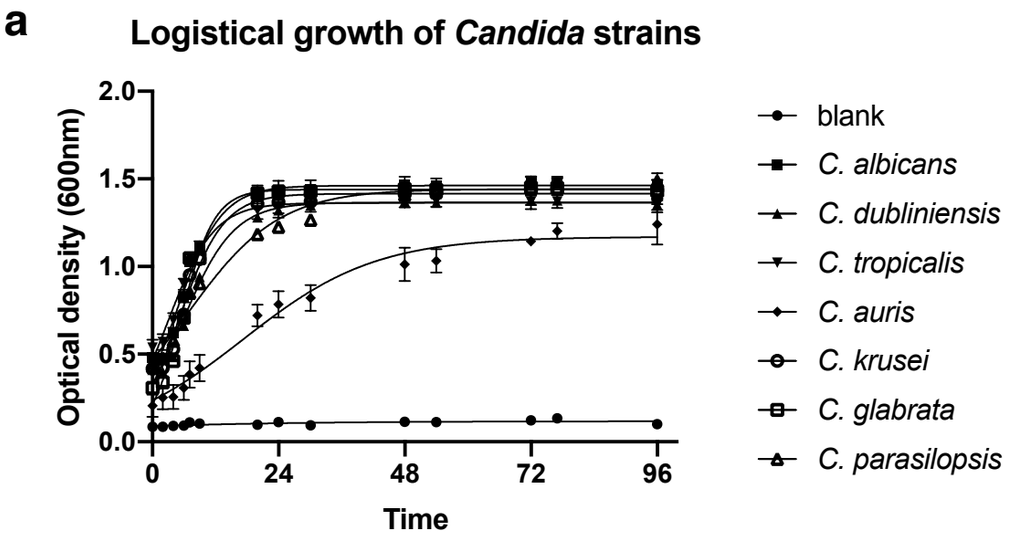


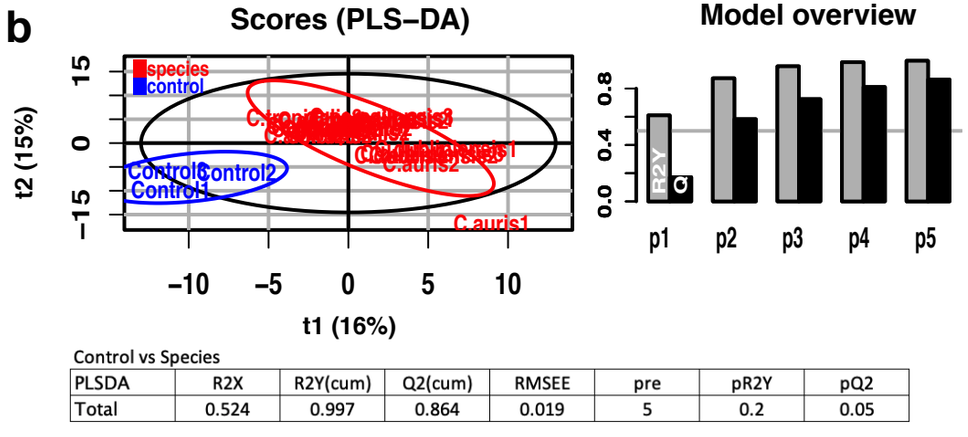






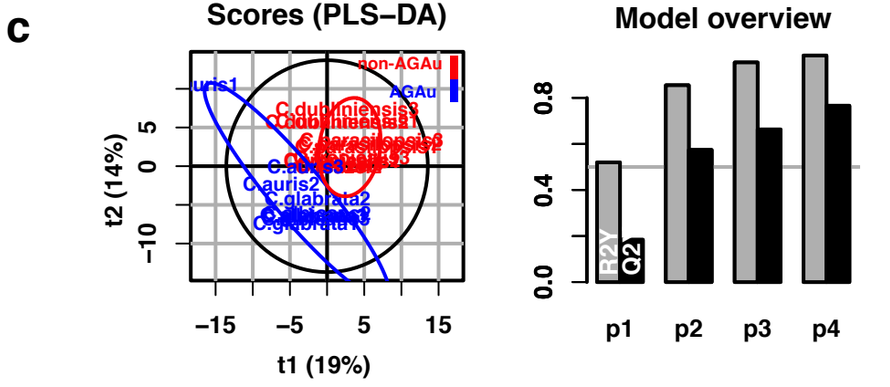

a
CAZyme Families
Glycoside Hydrolases
GlycoslyTransferases
Auxillary Activities
Carbohydrate−Binding Modules
Carbohydrate Esterases
Polysaccharise Lysases
Percent of Total CAZyme family
60%
40%
20%
0%
b
PFAM clan domain function annotation
Metabolic processes associated
Genetic Information Processing
Cellular Process and Machinery
Enzyme activity
Environmental Processing
Cell wall Associated
Drug resistance & Pathogencity
Unknown
Percent of Total PFAM clans
20%
15%
10%
5%
0%
c
Percent of Total Metabolic Pfam clan associated
50%
40%
30%
20%
10%
0%
Carbohydrate process associated
Lipid process associated
Amino acid process associated
Protein process associated
Secondary metabolite process associated





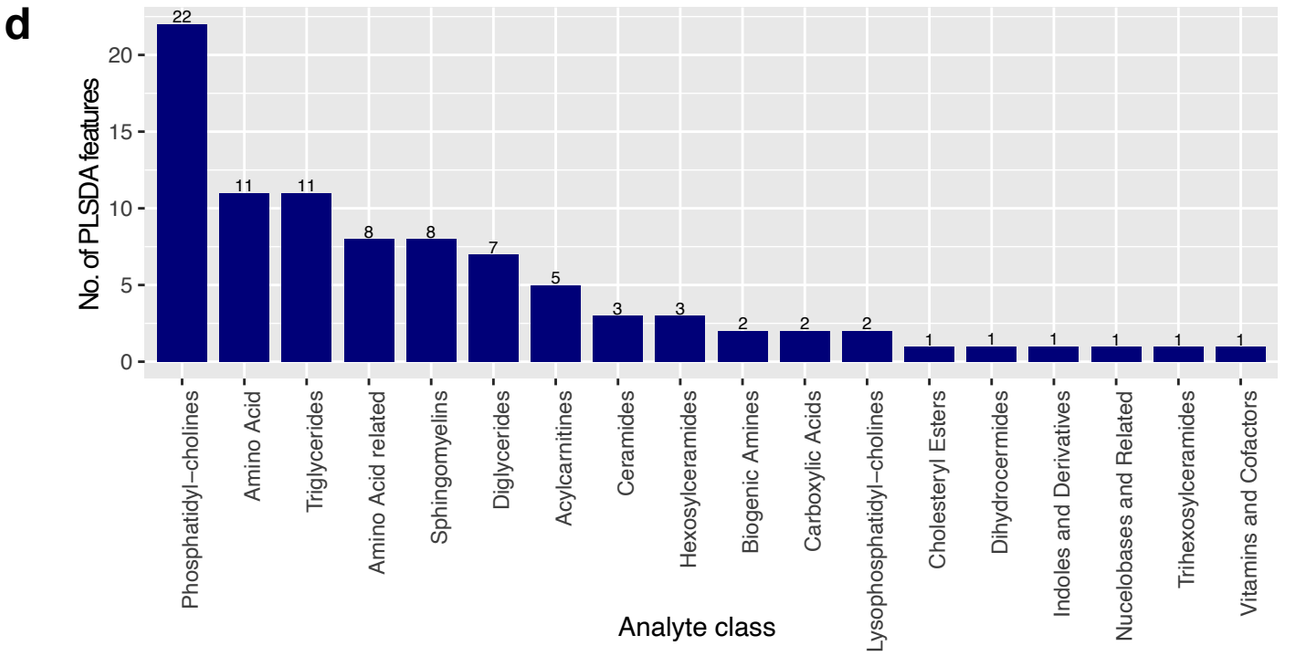
Control vs Species

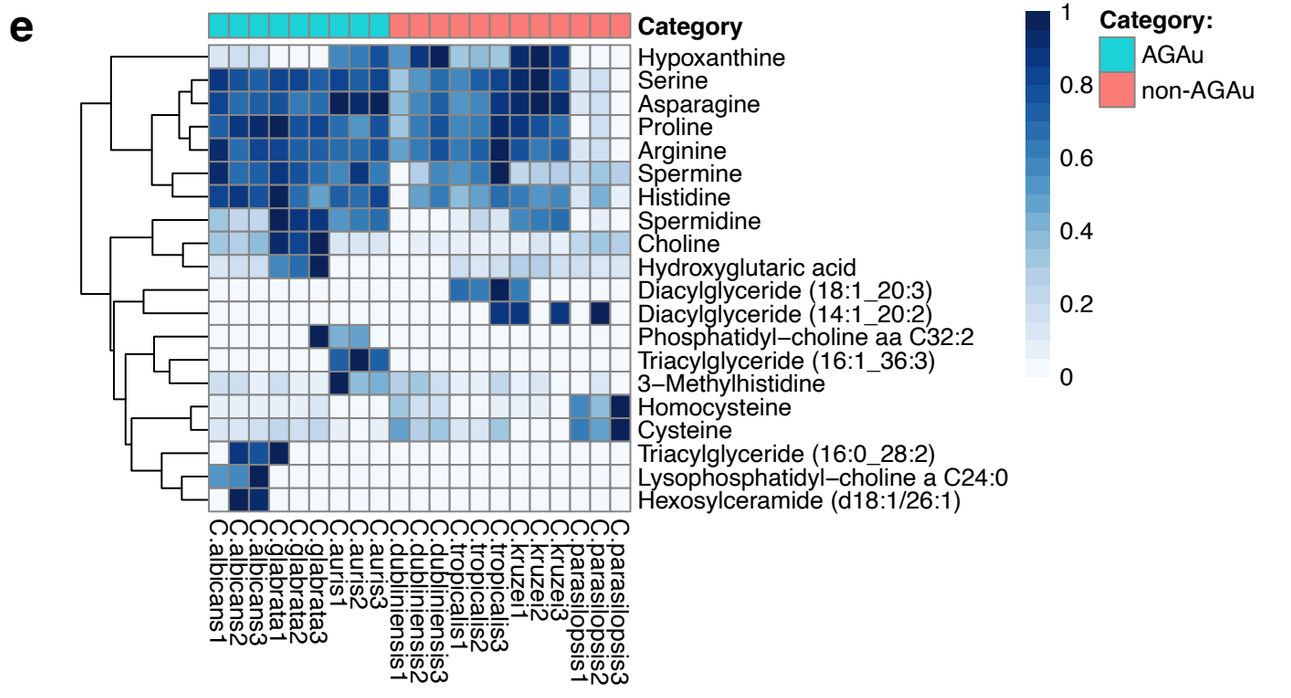
PLSDA	R2X	R2Y(cum)	Q2(cum)	RMSEE	pre	pR2Y	pQ2
Total	0.524	0.997	0.864	0.019	5	0.2	0.05

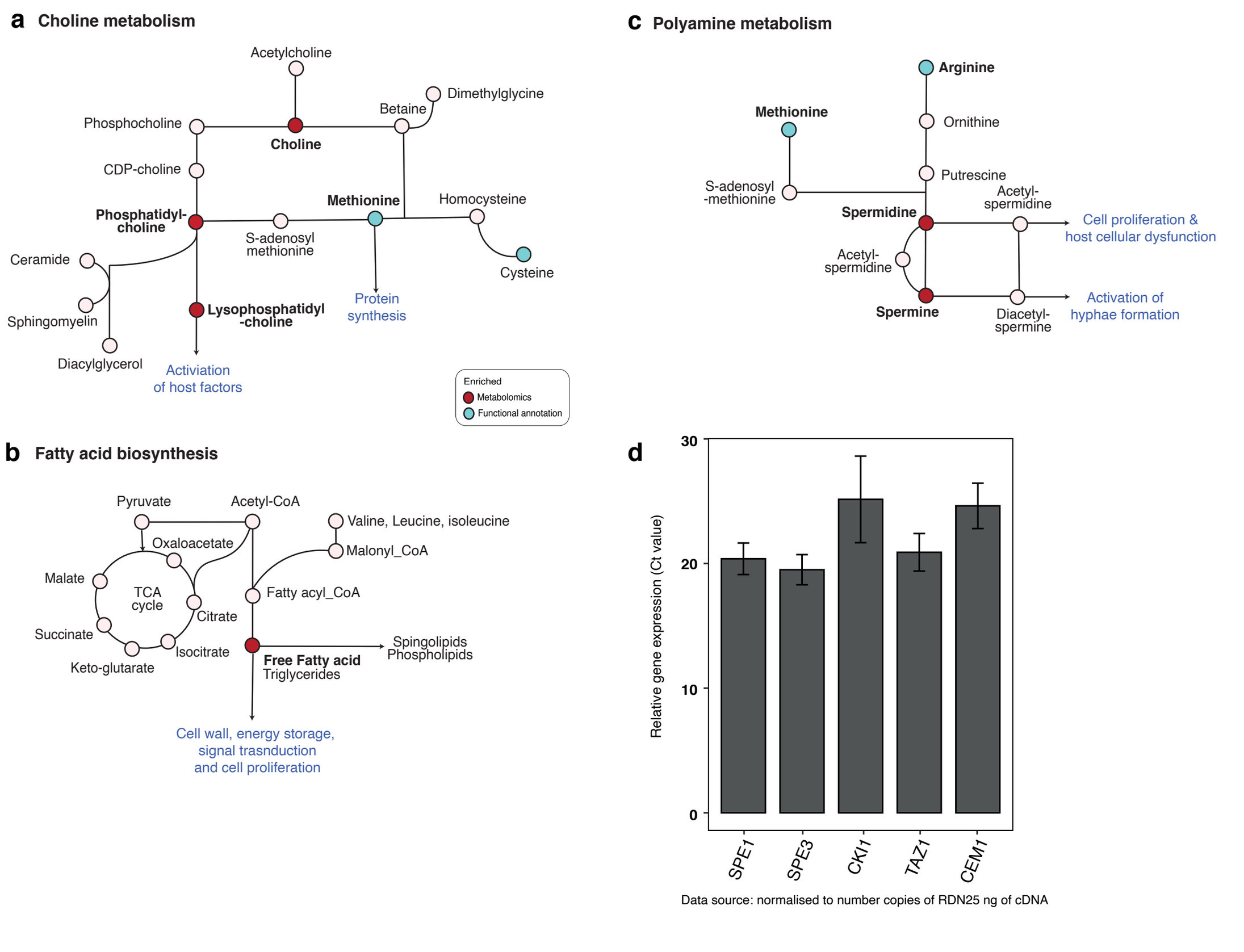


Invasive vs Non-invasive

PLSDA	R2X	R2Y(cum)	Q2(cum)	RMSEE	pre	pR2Y	pQ2
Total	0.501	0.984	0.766	0.071	4	0.3	0.05






a Choline metabolism
Acetylcholine
Dimethylglycine
Betaine
Phosphocholine
Choline
CDP-choline
Methionine
Homocysteine
Phosphatidyl-
choline
S-adenosyl
methionine
Ceramide
Cysteine
Protein
synthesis
Sphingomyelin
Lysophosphatidyl
-choline
Diacylglycerol
Activiation
of host factors
Enriched
Metabolomics
Functional annotation
b Fatty acid biosynthesis
Pyruvate
Acetyl-CoA
Valine, Leucine, isoleucine
Oxaloacetate
Malonyl_CoA
Malate
TCA
cycle
Fatty acyl_CoA
Citrate
Succinate
Isocitrate
Spingolipids
Phospholipids
Keto-glutarate
Free Fatty acid
Triglycerides
Cell wall, energy storage,
signal trasnduction
and cell proliferation
c Polyamine metabolism
Arginine
Methionine
Ornithine
Putrescine
S-adenosyl
-methionine
Acetyl-
spermidine
Spermidine
Cell proliferation &
host cellular dysfunction
Acetyl-
spermidine
Spermine
Diacetyl-
spermine
Activation of
hyphae formation
d
Relative gene expression (Ct value)
30
20
10
0
SPE1
SPE3
CKI1
TAZ1
CEM1
Data source: normalised to number copies of RDN25 ng of cDNA