FungAMR: A comprehensive portrait of antimicrobial resistance mutations in fungi

Camille Bédard[1,2,3,4], Alicia Pageau[1,2,3,4,5], Anna Fijarczyk[1,2,3,4,5], David Mendoza-Salido[14], Alejandro J Alcañiz[14], Philippe C. Després[1,2,3,4], Romain Durand[1,2,3,4,5], Samuel Plante[1,2,3,4], Emilie M. M. Alexander[1,2,3,4], François D. Rouleau[1,2,3,4], Mathieu Giguère[1,2,3,4], Mégane Bernier[1,2,3,4], Jehoshua Sharma[6], Laetitia Maroc[6], Nicholas C. Gervais[6], Anagha C.T. Menon[,12,13], Isabelle Gagnon-Arsenault[1,2,3,4,5], Sibbe Bakker[7], Johanna Rhodes[8], Philippe J. Dufresne[9], Amrita Bharat[15,16], Adnane Sellam[12,13], Domenica G. De Luca[15,16], Aleeza Gerstein[10,11], Rebecca S. Shapiro[6], Narciso M. Quijada[14], Christian R. Landry[1,2,3,4]

1. Département de Biochimie, de Microbiologie et de Bio-informatique, Faculté des Sciences et de Génie, Université Laval, Canada
2. Institut de Biologie Intégrative et des Systèmes, Université Laval, Canada
3. PROTEO, Le regroupement québécois de recherche sur la fonction, l'ingénierie et les applications des protéines, Université Laval, Canada
4. Centre de Recherche sur les Données Massives, Université Laval, Canada
5. Département de Biologie, Faculté des Sciences et de Génie, Université Laval, Canada
6. Department of Molecular and Cellular Biology, University of Guelph, Canada
7. Laboratory of Genetics, Wageningen University and Research, the Netherlands
8. Department of Medical Microbiology, Radboudumc, Nijmegen, the Netherlands
9. Laboratoire de santé publique du Québec, Institut National de Santé Publique du Québec, Canada
10. Department of Microbiology, The University of Manitoba, Canada
11. Department of Statistics, The University of Manitoba, Canada
12. Department of Microbiology, Infectious Diseases and Immunology, Faculty of Medicine, Université de Montréal, Canada
13. Montreal Heart Institute, Université de Montréal, Canada
14. Institute for Agribiotechnology Research (CIALE), Department of Microbiology and Genetics, University of Salamanca, Spain
15. National Microbiology Laboratory, Public Health Agency of Canada, Canada
16. Department of Medical Microbiology and Infectious Diseases, University of Manitoba, Canada

Authors contributions
Design the study: CRL
Data curation: CB, PD, RD, SP, EA, FR, MG, MB, JS, LM, NG, AM, DMS, AG, RS, CRL, DGD, AB
Data analysis: AP, CB, AF, IGA, CRL, SB, NMQ, DMS, JR
Software development and training: NMQ, DMS, AJA
Figures: CB, AP, AF, DMS, NMQ
Drafting the manuscript: CB, CRL
Manuscript editing: All authors
Contributed funding: CRL, ACG, RS, AS, PJD, AB, NMQ.
Contributed data: AB, DD

Corresponding author: christian.landry@bcm.ulaval.ca

Abstract

Antimicrobial resistance (AMR) is a global threat. To optimize the use of our antifungal arsenal, we need rapid detection and monitoring tools that rely on high-quality AMR mutation data. Here, we performed a thorough manual curation of published AMR mutations in fungal pathogens to produce the FungAMR reference dataset. A total of 462 papers were curated, leading to 54,666 mutation entries all classified with the degree of evidence that supports their role in resistance. FungAMR covers 92 species, 202 genes and 184 drugs. We combined variant effect predictors with FungAMR resistance mutations and showed that these tools could be used to help predict the potential impact of mutations on AMR. Additionally, a comparative analysis among species revealed a high level of convergence in the molecular basis of resistance, revealing some potentially universal resistance mutations. The analysis also showed that a significant number of resistance mutations lead to cross-resistance within antifungals of a class, as well as between classes for certain mutated genes. The acquisition of fungal resistance in the clinic and the field is an urging concern. Finally, we provide a computational tool, ChroQueTas, that leverages FungAMR to screen fungal genomes for AMR mutations. These resources are anticipated to have great utility for researchers in the fight against antifungal resistance.

Introduction

Infections caused by fungal pathogens constitute a significant health threat and socio-economic burden. Fungal pathogens have the potential to cause severe diseases, accounting for multiple million deaths annually[1]. The mortality rate of patients with invasive mycosis, an infection spreading within organs, can exceed 50%[2]. Yet, fungal infections are amongst the most overlooked infectious diseases. The World Health Organization (WHO) estimated that less than 1.5% of the funding allocated for infectious disease research is directed towards investigating fungal infections[3]. The burden of pathogenic fungi extends beyond human infections, encompassing challenges to both food security and biodiversity. Fungi are a leading cause of plant diseases, resulting in an estimated 20% worldwide crop yield loss[4]. Additionally, certain fungi have been responsible for significant declines in animal populations, such as amphibian and bat populations across the globe[5,6], and few if any effective drugs are available to mitigate critical loss of animal biodiversity.

Another pressing threat to public health is antimicrobial resistance (AMR)[7]. The widespread use of antifungals in agriculture and clinics, or in other sectors as preservatives, has led to the emergence of resistant strains and populations unresponsive to treatments[8,9]. Fungicide resistance in plant pathogens exacerbates the threat to food security by increasing yield losses and contamination of crops with mycotoxins[9]. AMR in human fungal pathogens poses a particularly daunting challenge due to the limited treatment options available[8,10]. The medical community relies on only a handful major classes of antifungals to treat human systemic infections[11]. Even if other compounds are under development, treatment options are currently rapidly depleted when resistance mutations arise in patients during treatment[8,10]. This is exacerbated by the overlap between the classes of antifungals used to treat infections in humans and the fungicides used on plants, which selects for resistance mutations in environmental fungi that can go on to infect humans[4]. For instance, exposure of *Aspergillus fumigatus* to agricultural azole fungicides leads to cross-resistance to medical azole antifungals[12,13].

One approach to overcome AMR in fungi could be to develop new drugs. However, humans and fungi share a much more recent common ancestor compared to humans and bacteria. This considerable proportion of shared genes and cell morphology limits the number of unique molecular targets that only affect the fungal pathogen and not the human host cells[14]. Despite ongoing efforts to extend our current antifungal arsenal and identify new targets, the rate at which new antimicrobials are produced is far exceeded by the speed at which resistance emerges[14]. Additionally, resistance will eventually likely emerge against these new drugs. While developing and discovering new antifungals is crucial, this alone will not address the problem of antifungal resistance. A better understanding of the different resistance mechanisms is essential in our battle against fungal pathogens to help guide the development of treatment strategies, and hopefully, to better understand and control the evolution of resistance. Having a centralized antifungal mutation repository of high quality will also improve our capacity to detect resistance and paves the way for routine genotypic detection of antifungal resistance (vs long and cumbersome phenotypic assays such as broth microdilution).

Unlike antibiotic resistance, which frequently emerges from the acquisition of genes, antifungal resistance emerges through *de novo* mutations in the genome resulting in a wide array of possible resistance mechanisms[15]. Accordingly, various mechanisms of AMR have been described for all available antifungal classes. Yet, the number of fungal AMR mutations documented in the literature remains relatively limited and dispersed, and in many cases we lack an understanding of the precise impact or causal nature of antifungal resistance mutations. In-depth knowledge of resistance mutations across all available drugs would allow us to better understand which mutations are directly linked with antifungal drug resistance, and further determine the extent of cross-resistance within and between antifungal classes. Similarly, knowledge of resistance mutations across different species and genes would allow us to assess how resistance mechanisms are conserved among fungi. Also, a better understanding of resistance mutations would enable us to streamline the development of antifungal agents that overcome resistance mutations and avoid cross-resistance characteristics by modifying currently available drugs. Developing new antifungals using this strategy would alleviate the pressure to discover new target genes. This knowledge could facilitate the development of genomic tools to monitor and interpret genetic variants of pathogens in real-time during treatment or in the agricultural field to detect and follow resistance. Such information could also be used to build predictive models to foresee resistance mutations before their spread in clinics and agriculture. Databases on antibacterial resistance mutations, such as CARD[16,17], have proved invaluable in addressing the serious threat posed by AMR for instance, by enabling the interpretation of genomic information. However, while some antifungal resistance databases exist, these are unfortunately outdated[18] or contain a large fraction of inaccurate information when done using simple text mining (Supplementary note 1)[19]. Since many studies are reporting potentially spurious associations between mutations and resistance, a careful annotation is needed to build a useful reference dataset (see below). Thus, there is a need to exhaustively and reliably gather all antifungal resistance mutations reported in the literature in one accessible repository.

To address these limitations, we manually curated 462 papers reporting fungal AMR mutations to create the FungAMR compendium (Supplementary table 1). FungAMR contains 54,666 carefully curated entries across 184 drugs (including 104 antifungals) for 92 fungal species and contains missense mutations as well as other genomic changes such as copy-number variations and nonsense mutations. Every mutation is classified with the degree of evidence that supports its role in AMR. We have taken advantage of this new resource to better understand resistance mechanisms among species and antifungals. We have combined variant effect predictions with resistance mutations to confirm that many genes confer resistance through loss of function mutations and show that such analyses could be used to help interpret the impact of mutations on AMR. Furthermore, comparative analysis of resistance mutations among species revealed a high level of convergence in resistance mechanisms at both the gene and the mutation levels. This analysis also confirmed that many resistance mutations provide cross-resistance to antifungals within the same class, but also between classes for mutations present in certain genes. Finally, we present ChroQueTas, a computational pipeline that utilizes the information

contained in FungAMR to easily detect the presence of AMR mutations in fungal genomes and produce an annotated report. Overall, the analysis of the content of FungAMR revealed biases in the study of certain species, proteins and antifungals, and highlighted where more research is needed. This resource will help the scientific community to address the serious threat of antifungal resistance, and has the potential to provide a launchpad for fundamental and applied research.

Methods

Literature review, data curation and development of the FungAMR resource

All papers were read by a scientist in the field and the information was extracted in a systematic manner. In addition to the laboratories involved, some of the curators were recruited among graduate students and postdoctoral fellows from laboratories working on fungal genetics and antifungal resistance, which are part of a pan-Canadian training program on the evolution of fungal pathogens, EvoFunPath. Two undergraduate students were involved in curating papers that more senior scientists also curated. Curation was conducted between April 2023 and May 2024 and included publications from 1988 to 2024. All curators collected papers reporting antimicrobial resistance using standard literature searching and mining references from other papers. We examined human, animal and plant pathogens in addition to model species such as the laboratory model *Saccharomyces cerevisiae*, which is used as a model system to study drug resistance. We also re-curated all papers reported in the MARDy database[18] as these did not come with any measure of confidence levels (see below).

Each paper was read by one of the curators, and mutations reported to be associated with resistance were extracted. Mutations also reported to likely not cause resistance were included, although it is difficult to judge whether this is useful information as assays rarely contain positive controls in parallel. Each entry (Supplementary table 1) is associated with the first author, the journal, the publication year and the PubMedID (PMID) of the paper in which the mutation was reported, the fungal species and its host, the gene name and its protein's accession number, the location of the mutation, its impact on the gene (deletion, insertion, amino acid change, duplication) and the impact on susceptibility to the drugs assayed. We also recorded the origin of the strains (environmental, clinical, laboratory or evolved) as well as the quantitative measures of resistance or susceptibility when provided by the authors. These were often minimum inhibitory concentrations (MIC), but many more measures were reported such as the half maximal effective concentration (EC50), a more common quantitative measure used in the agricultural area. We tried to record these measures as faithfully as possible. Still, we found that the diversity of methods used made any systematic recording of the quantitative impact of the mutations difficult. We therefore relied on the author's assessment to record the mutation as being associated, or not, with resistance. When the authors did not assess if a mutation was associated with resistance, we define strains as resistant when they have a significantly higher growth in the presence of an antifungal than the reference strain or when they present with a minimal inhibitory concentration to an antifungal above the established breakpoint. Each

reported entry concerned a particular strain and its associated phenotype. For instance, if a strain was reported and had multiple amino acid substitutions in the gene or genes sequenced, those multiple substitutions were reported together in a single entry as it was impossible to discern which one could have been causal. When available, we also recorded the identifier of the genes reported so we could track the reference sequences for species of interest. These identifiers came from various databases such as the Candida Genome Database or NCBI. However, this was not systematically done by the authors so some entries are left blank. Not all papers reported the gene identifier for the reference genes the mutations were reported to impact. We therefore performed a quality check by comparing the mutations reported (change from WT genotype to mutant one) with the standard reference genes from the species of interest (Supplementary table 2). A commentary column is included where curators noted information relevant to the entry, such as the experimental method, the culture media or the antifungal susceptibility testing method (e.g EUCAST, CLSI).

All curators annotation sheets were combined into a single table, after refining and editing the nomenclature of species, genes or proteins, mutations, drugs, journal names, and strain origins, to facilitate the comparison of mutations. MIC values were edited and converted to the same units whenever possible. Two additional columns were added, one with an alternative species name and another one with a gene name that corresponds to orthologs or homologs of that gene in other species. In case when the study was curated by multiple curators, entries of both curators were first compared for consistency and corrected if necessary. Redundant entries from the same study (mutations in the same gene, drug, and strain) were combined or one of the duplicate entries was removed.

For each mutation or group of mutations, we assigned a score that assesses the degree of evidence that supports their role in drug resistance. The scoring system was developed based on examples of reports that we had seen in the literature and were agreed upon and discussed among the curators. The list of confidence scores associated with their description is presented in Table 1. A positive confidence score denotes mutations reported to confer resistance, while a negative confidence score relates to mutations reported in susceptible strains. A low positive confidence score indicates that the evidence for the contribution of the mutation to resistance is strong, with 1 being the strongest and 8 being the weakest. Similarly for negative scores, the evidence of susceptibility is stronger for -1 than -8.

Table 1 | Description of the confidence scores associated with curated mutations in FungAMR

Confidence score	Description
1	The mutation was created in the same susceptible species background and effect on resistance was confirmed by measurements.
2	The mutation was created in the same gene but expressed in another species with potentially non endogenous level of expression and effect on resistance was confirmed by experimental measurements.
3	The effect of the mutation was quantified by bulk competition assays such as Deep Mutational Scanning.
4	The mutation was identified by experimental evolution where its high frequency in independent replicates suggests it causes resistance.
5	The deletion of the gene causes resistance.
6	The overexpression or duplication of the gene causes resistance.
7	Significant association between the mutation and resistance in a population as determined by GWAS (or other population-based association such as QTL or classical genetics).
8	The mutation was identified in a 'natural' strain (e.g. a clinical isolate) that is resistant but without any further validations.
Negative scores (-)	Any confidence score where the mutation was found not to cause resistance using the approach of the corresponding positive score.

When a mutation is reported in several independent studies or strains within a study, the strongest resistance evidence (i.e. the best positive confidence score reported) and the strongest sensitivity evidence (i.e. the strongest negative confidence score reported) are provided in two separate columns.

Variant effect predictors

For all protein reference sequences retrieved, we used HHblits[20] from HH-suite3[21] with each sequence as a query to create multiple sequence alignment (MSA) representative of the orthologs diversity. Initial HHblits iterative sequence search filter criteria were a minimum coverage of 80% with query sequence, at least 20% identity to the query and maximum 98% and the e-value cutoff for inclusion was 0.0001. Filter criteria were then gradually relaxed to get a minimum of 200 sequences in the final alignment. 267 alignments were produced with a median of 847 sequences per alignments. With these MSA files as input, we ran GEMME[22] locally from docker desktop to compute the predicted effect of every single amino acid substitution across the proteins.

The protein structures from the proteins contained in FungAMR were obtained from the AlphaFold database (AlphaFold Protein Structure Database, Supplementary table 3). We used

these structures to estimate the effects of amino acid substitutions on protein stability using FoldX[23]. The RepairPDB function was run 10 times with water prediction to repair incorrect torsion angles, VanderWaals clashes and total energy of the structure. We then used the MutateX Python package[24] to perform an in-silico deep mutational scanning of the whole protein running the BuildModel function from FoldX and computing the difference in Gibbs free energy between the mutant and the wild-type.

Tertiary protein structure predictions

When experimentally determined protein structures were available, they were retrieved from the Protein Data Bank (RCSB PDB, PDB). No complete tertiary protein structures were available on the PDB for Erg3, Hmg1, Upc2 and Pdr1, therefore we used the Alphafold 3[25] web server (AlphaFold Server) to predict the entire structures (seed; Erg3: 382113193, Hmg1: 1841159774, Upc2: 601343710 and Pdr1: 264428788). Protein sequences were retrieved from UniProt (UniProt, Erg3: P32353, Hmg1: Q4WHZ1, Upc2: Q59QC7 and Pdr1: P12383). For Hmg1, the NAP – nicotinamide-adenine-dinucleotide-phosphate ligand was included in the prediction. For the transcription factors Upc2 and Pdr1, a DNA sequence containing one of the described DNA binding site of the enzymes was included in the prediction (Upc2: AATATCGTACCCGATTATGTCGTATATT from *C. albicans ERG11* promoter[26,27] and Pdr1: GAGAAATGTCTCCGCGGAACTCTTCTAC from *S. cerevisiae PDR5* promoter[28]).

Overall data description

We constructed tables of correspondence of amino acid sequences among the orthologs of the proteins reported in the data. Reference sequences for all proteins were retrieved from the Uniprot or NCBI database and MSA were performed on orthologous sequences using Muscle5[29]. All of the alignments for each family of orthologs are available in Fasta format as supplementary material (Supplementary data 1). These MSA files were used to assign a group number to orthologous mutations in different species, which can be found in an additional column. Alignment positions were also used to produce figures and to map resistance mutations on protein structures. The classification of the host of the fungal species, drug classes and gene types used in the figures is available in Supplementary table 4.

To display phylogenetic relationships among the species present in FungAMR, we used a phylogenetic tree constructed by Li et al.[30] which includes 67 species from the database. To include approximate positions of the 18 remaining taxa at the species level, subphylogenies of species subsets were built using ITS sequences from either ITS RefSeq (BioProject number PRJNA177353), UNITE database v9, or if these were missing, from GenBank. Specific host forms (forma specialis) and hybrids were omitted from the species tree. Any ambiguous bases were resolved randomly, then sequences were aligned with mafft v7.525[31] and trimmed using trimAl v1.4[32] with “-gappyout” option. The maximum likelihood trees were built with iq-tree v2.2.2.7[33]. Branches with missing species were added to the main phylogenetic tree retaining the proportion of the branch lengths and position of the nodes with respect to related species from the ITS tree. The visualization of the tree was made using ggtree package[34] in R.

The visualization of protein structures was done with ChimeraX[35]. Figures and analyses were done using the following Python packages: Matplotlib[36], NumPy[37], SciPy[38], pandas[39] and seaborn[40], and the R packages: dplyr[41], tidyr[42], stringr[43], igraph[44,45], ape[46], castor[47], RRphylo[48], and ggplot2[49].

Computational pipeline to detect mutations and InDels causing resistance

In addition to the FungAMR repository, a bioinformatic tool named *Chromosome Query Targets* (ChroQueTas) was developed in order to rapidly screen for AMR mutations in fungal genomes. ChroQueTas was devised to work CLI-based on UNIX environments (tested on Debian and Ubuntu-based OS) in a user-friendly manner so it can be run by following straightforward instructions and commands. ChroQueTas is open-source and publicly available at https://github.com/nmquijada/ChroQueTas, together with instructions for its installation and usage.

ChroQueTas requires an assembled fungal genome (either complete or draft assembly in FASTA format, no need for annotation) and a "scheme" (to be provided by the user under the *--scheme* flag) corresponding to the species of the fungal genome (thus corresponding to the species reported in the FungAMR repository). With that information, ChroQueTas will: i) extract from the fungal genome the CDS and protein where a point mutation is known to cause AMR in that particular species by using miniprot v0.12-r237[50] and the information contained in FungAMR; ii) evaluate sequence similarity against the reference by using BLASTP v2.14.1+ [51] and discard low confidence hits (to be specified by the user and the *--min-id*, *--min-cov* flags); iii) deal with potential introns, exons and InDels; iv) evaluate amino acid positions between the query and the reference proteins accounting for FungAMR information; and v) report amino acid changes and InDels that could lead to AMR according to the confidence score in FungAMR. The output from ChroQueTas will consist of different text files with the information resulting from the AMR screening and further metadata contained in FungAMR. All temporary files could also be stored for user inspection. Additionally, multiple ChroQueTas outputs can be combined into a single and interactive HTML report that could be visualized in any browser to ease user investigation and sharing of the results, including MSA and interactive 3D reconstructions of the proteins target for fungicides with the AMR mutations highlighted for deeper investigation (an example is available at https://github.com/nmquijada/ChroQueTas). ChroQueTas is intended to be a "living project", being hosted in a public repository with enabled discussion forums where the users could ask for further improvements that will be considered and implemented on a regular basis.

In order to illustrate the power of the software, the Illumina raw sequencing data from 46 *Candida albicans* and 144 *Zymoseptoria tritici* isolates was downloaded using SRA Toolkit v3.1.0 (https://github.com/ncbi/sra-tools) (Supplementary Table 5). The sequencing data quality was assessed using FastQC v0.11.9[52]. Residual adapters and barcodes were discarded and quality filtering was performed by using FastP v0.23.2[53] by setting a minimum Phred score of 25 and a minimum length of at least 75% of the length of the raw FASTQ files. Genomes were assembled

by using SPAdes v3.15.5[54] and contigs below 1,000 bp were discarded. The quality and completeness of each draft genome was assessed with QUAST v5.0.2[55], BUSCO v5.4.3[56] and by aligning the high-QC reads against the draft genomes by using Bowtie2 v2.4.2[57]. Once the quality of the genomes was assessed, ChroQueTas was used for the inference of AMR by using the following command:

```
ChroQueTas --input <genome>.fasta --scheme <Calbicans/Ztritici> --threads 32 --output <genome>_ChroQueTas.txt --report
```

The average nucleotide identity (ANI) of the fungal genomes was calculated using fastANI v1.32[58] and the distance matrix was imported into R environment[59] to build a phylogenetic tree by using *ggplot2*[49], *ggtree*[34] and *reshape*[60] packages. The interactive HTML is rendered in R by using *rmarkdown*[61] and *bio3d*[62], *dplyr*[63], *msa*[64], *seqinr*[64] and *r3dmol*[65].

Results

Content of FungAMR

A total of 462 papers were curated, leading to 54,666 mutation entries (Supplementary table 1). All entries are associated with information on the publication from which it was curated, as well as the gene, the location of the mutation, and the impact on drug susceptibility as experimentally assessed. Since the definition of a resistance mutation varies between studies, we defined a resistant strain as a strain presenting with a minimal inhibitory concentration to an antifungal above the established breakpoint when it was available or as a strain with significantly higher growth in the presence of an antifungal than the reference strain. Also, the level of evidence for the causal link between reported mutations and drug resistance varies considerably from one study to another. Some reports demonstrate a causal link between the mutation and the phenotype, for instance, by creating the mutation in a control genetic background. Other reports simply identify a mutation in a resistant strain, without further analysis, and often focus on sequencing only known resistance genes. Therefore, we associated all mutations or groups of mutations with a confidence score assessing the degree of evidence that supports their role in drug resistance (Table 1). A positive confidence score denotes mutations reported to confer resistance, while a negative confidence score relates to reported mutations that were not found to confer resistance. A low positive confidence score indicates that the evidence for the contribution of the mutation to resistance is strong, with 1 being the strongest and 8 being the weakest. We do not use the negative score as extensively as those may depend on many factors, including the lack of sensitivity of the assays. Nevertheless, we think this information is useful to have recorded. When a mutation is associated with both positive and negative scores, we take the lowest positive one for our analyses. Ultimately, it is important to bear in mind that confidence scores do not translate into a probability of treatment failure, but only reflect the level of evidence we have that a mutation confers resistance in laboratory conditions.

We gathered more than 10,000 unique mutations for 184 drugs across 202 genes and 92 fungal species. Mutations assayed by Deep Mutational Scanning (DMS, confidence score of 3) come

from studies where complete or nearly complete libraries of gene mutants were subjected to resistance testing. We have only four studies reporting mutations assayed by DMS,but they represent 86% of the entries (3: 19% and -3: 67%, Fig. 1a), showing how powerful systematic approaches can be at characterizing mutations. We set aside these entries for some analyses, since they are not representative of the overall content. Although DMS studies contribute to a large fraction of the data, 459 studies report non-DMS data for 7,690 entries and over 2,500 unique mutations covering the same set of species. The drugs listed include 31 antifungals used in the clinic (80% of entries), 73 used in agriculture (14% of entries) and 79 other molecules (6% of entries, e.g. anticancer drugs and immunosuppressants). However, most reported entries are of low confidence, with 60% having a confidence score of 8 (8: 44% and -8: 16%, Fig. 1a). Around 20% of the entries are mutations validated directly in the susceptible pathogen (1: 12% and -1: 4%, Fig. 1a) or by heterologous expression in the model yeast *Saccharomyces cerevisiae* (2: 5% and -2: 1%, Fig. 1a). This highlights the fact that although many AMR mutations have been reported in the literature, most have very little to no experimental support.

Nakaseomyces glabratus (formerly *Candida glabrata*), *Aspergillus fumigatus and Candida albicans* have the largest number of entries with more than 1,200 each. They are followed by the model yeast *Saccharomyces cerevisiae* (1,054 entries) and the emergent human fungal pathogen *Candida auris* (465 entries) (Fig. 1b). Roughly 75% of the listed species have less than 25 entries, reflecting a strong bias towards the study of certain fungi. Antifungals used in clinics are the most studied, contributing to more than 80% of entries (Fig. 1b). The drug targets of azoles and echinocandins, Cyp51 (also named Erg11 in yeast) and Fks, are unsurprisingly the most studied genes, being listed in nearly 40% and 25% of the entries, respectively. Although azoles and echinocandins are the most commonly used antifungals to treat invasive infection, this representation probably overestimates Cyp51 and Fks1 contribution to resistance as many studies focus only on these genes as the potential cause of resistance. The other genes are much less studied as more than 70% of them are listed fewer than 25 times. Most of the biases in the entries can be explained by the fact that AMR in human fungal pathogens are overall better studied or more often reported than in plant pathogens. However, the diversity of species and antimicrobials reported for plant pathogens is larger than for human pathogens (Fig. 1b, c). Also, azoles are widely used in both clinics and agriculture, making them and their target Cyp51 the subject of many studies (Fig. 1b).

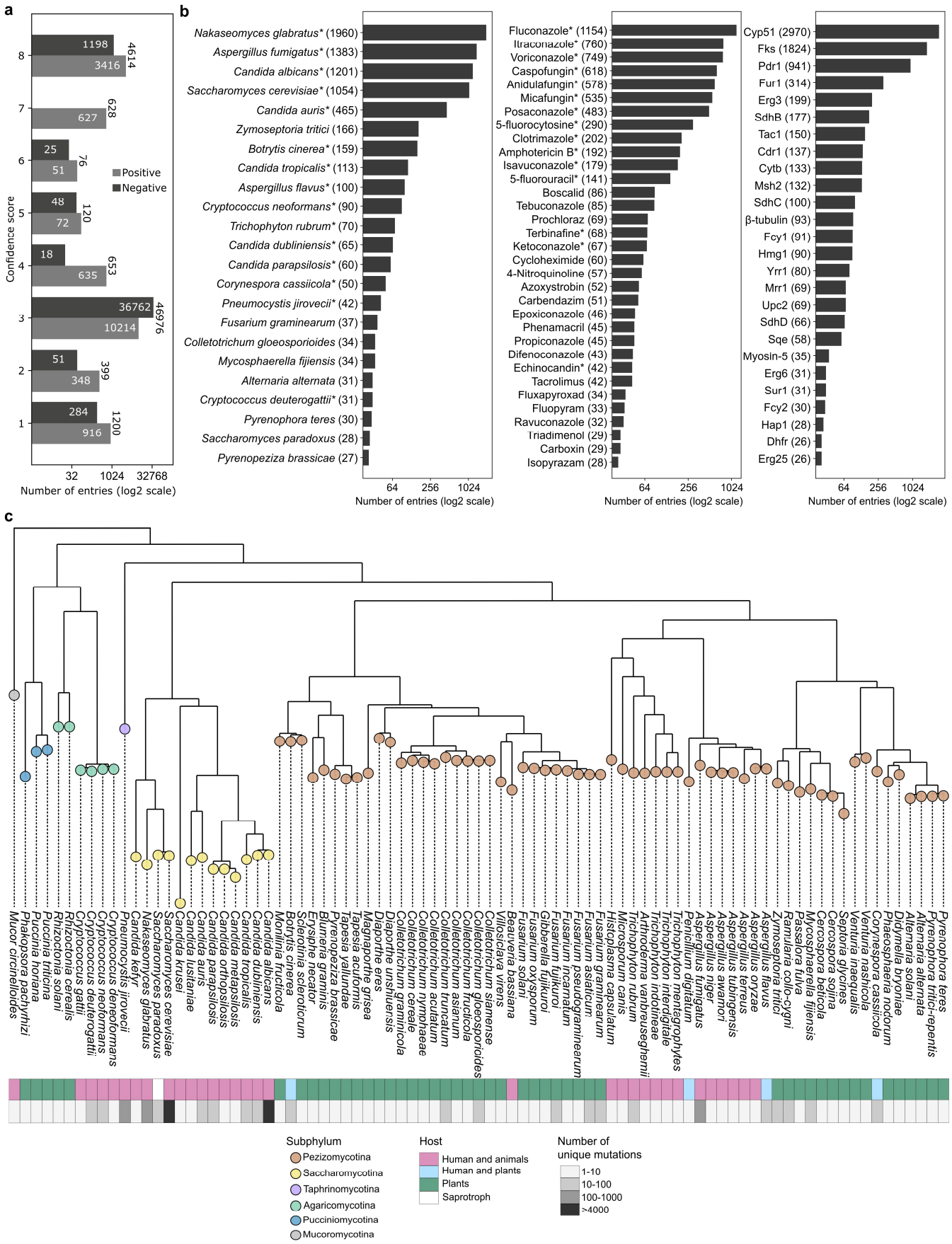


Fig. 1 | Content of FungAMR per confidence scores, species, antifungals and proteins. a, Distribution of positive and negative confidence scores associated with mutations reported in the literature (n = 54, 666 entries).

A positive confidence score indicates mutations that have been reported to confer resistance, while a negative confidence score indicates mutations that have been reported but have not been found to confer resistance. A low positive confidence score means that the evidence for the effect of the mutation on resistance is strong, with 1 being the strongest and 8 being the weakest. Most entries come from high throughput experiments and have a confidence score of 3. Most remaining reports have limited experimental support. **b,** Number of observations per species, per drugs and per proteins. Species with * can cause human diseases (e.g. asthma, dermatosis, invasive mycosis) and drugs with * are used in clinics. Entries with a confidence score of 3 were excluded and cases with fewer than 25 reports are not shown to simplify the figure. **c,** Phylogenetic relationship of the species represented. On the right of species names are indicated which hosts these species can be isolated from and the number of unique mutations reported in the dataset in each case. The number of unique mutations can be lower than the number of mutation entries since many mutations have been reported multiple times.

Identifying AMR mutations and their impact on protein function

Antifungal resistance frequently arises through loss-of-function mutations (LOF). LOF mutations disrupt gene function, for instance, by introducing premature stop codons, InDels causing frameshift or missense mutations at critical sites in proteins. For some molecules, how LOF leads to resistance can be deduced from the antifungal mode of action. For instance, deletions of genes other than *CYP51* in the ergosterol biosynthesis pathway have been shown to confer resistance to azoles. These drugs bind to Cyp51, which blocks ergosterol biosynthesis and leads to the accumulation of a toxic dienol intermediate by the sterol desaturase Erg3[14]. The deletion of *ERG3* prevents the synthesis of the toxic sterol[66], causing resistance. Upstream of the pathway, the deletion of *HMG1*, an enzyme with sterol sensing functions, increases sterol production which causes resistance[66]. Another example is resistance to flucytosine (5-FC), a prodrug that must be converted by the cytosine deaminase to gain potency. A LOF in the gene encoding this enzyme (*FCY1* in *S. cerevisiae*) or other downstream genes that further process the molecule (e.g. *FUR1*), for instance, through premature stop codons, leads to resistance[67,68].

We examined all instances of resistance associated with premature stop codons and InDels causing frameshifts to identify genes whose contribution to resistance may occur through LOF. LOF in *FCY1* and *FUR1* are well-established mechanisms of resistance in *S. cerevisiae* and unsurprisingly, they are also observed in pathogens such as *C. neoformans*[69] and *N. glabratus*[70,71]. Experimental evolution of *S. cerevisiae* in nystatin led to the identification of LOF causing resistance in *ERG3* and *ERG6*[72]. Frameshifts and stop codons in *ERG3* and *ERG6* also appear to cause resistance to Amphotericin B in *C. auris*[73], revealing that such LOF in these genes could be a conserved mechanism of resistance to polyenes. LOF in other genes of the ergosterol pathway is also associated with resistance, including frameshifts in *ERG2* in *C. albicans* causing resistance to azoles[74]. LOF of *ERG4* has also been associated with azole and echinocandin resistance in *N. glabratus*[75] but with a confidence score of 8, which means that this requires further examination. Interestingly, all of these cases occur in species that are mainly haploid (or haploid strains of *S. cerevisiae*), which is anticipated as most LOF mutations are expected to be recessive.

Amino acid substitutions in Fcy1 or Fur1, typically at highly conserved sites, are often destabilizing for the protein and sufficiently disruptive to their function so that they also confer

resistance[67,68]. It is therefore possible that conservation level and estimates of destabilization could help to better understand whether resistance emerges from LOF, without the need to observe premature stop codons or frameshift mutations in resistant isolates. To examine this and see if other mutations may be indicative that LOF causes AMR, we integrated into the data two predictions of the impact of mutations based on conservation analysis[22 76] and one based on the prediction of the destabilization by the mutation using protein structure[23,24] (Fig. 2). For genes that lead to resistance upon LOF, such as *FCY1* and *FUR1*, we indeed clearly see that resistance mutations are predicted to have a larger impact on protein function and stability in comparison to other mutations in the same genes . Other extreme cases are Sur1 and Csg2, proteins involved in sphingolipid synthesis. Amino acid substitutions in these proteins that cause resistance are particularly destabilizing and occur at conserved sites, suggesting LOF. This observation is validated by recent experimental evolution results showing that stop codons in these two enzymes in *S. cerevisiae* cause resistance to clotrimazole[77]. For *ERG3* and *HMG1*, mutations are predicted to have a greater impact on the protein function but to have a low or no impact on protein stability (*ERG6* did not have enough entries to perform this analysis). Nonetheless, this suggests that resistance mutations in *ERG3* and *HMG1* likely arose through LOF which can be predicted through their potential effects on protein function.

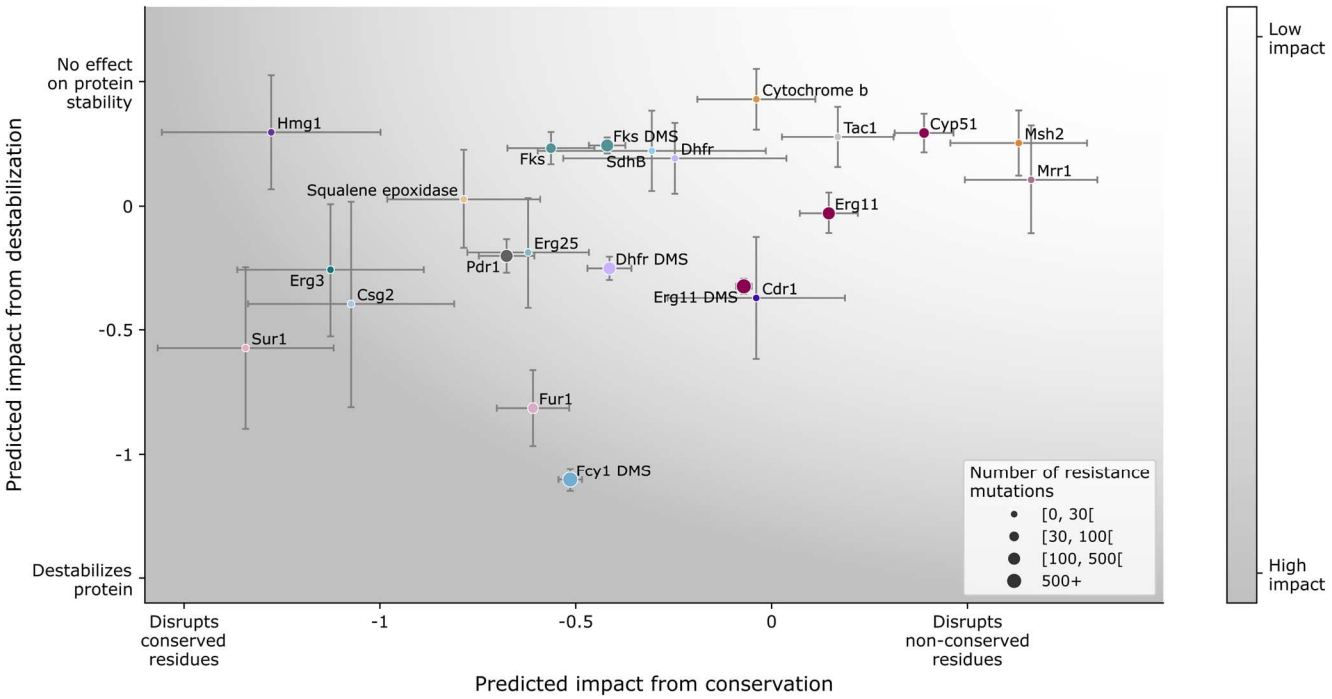


Fig. 2 | Prediction of mutations' effect on protein function and stability can aid differentiation between loss and gain of function resistance mutations. Relationship between the predicted impact of resistance mutations inferred from long-term evolutionary constraints (GEMME[22]) and protein stability (FoldX[23,24]). Each dot represents the mean deviation of the predicted impact for resistance mutations in a given protein relative to all possible amino acid changes. The y-axis presents the z-score of the normalized predicted change in the Gibbs free energy (ΔΔG) upon mutation predicted by FoldX. Negative values predict destabilization of the protein. The x-axis represents the z-score of the predicted impact of the mutation computed by GEMME, which uses patterns of protein conservation to predict the tolerance of mutations. Negative values predict a higher impact on the protein function based on evolutionary conservation. For each gene, when mutations were available from DMS experiments (confidence score

of 3), the DMS data is shown separately. The size of the dots shows the sample size and the colors show the different proteins. Error bars represent one standard error. The distribution of the data for each gene is presented in Extended data Fig. 1.

As most resistance genes for which we have enough data to perform such analysis do not show patterns consistent with LOF mutations, this suggests that a large fraction of resistance genes do not evolve by simple inactivation. In fact, antifungal resistance can also arise through gain of function (GOF) mutations. The mechanisms by which a gene gains a function by mutations can be diverse. One way GOF mutations can lead to resistance is through the production of an hyperactive protein. For instance, mutations in the transcription factor Upc2 increase the expression of the azole drug target, which compensates for the inhibition by azoles[78]. Alternatively, GOF mutations in efflux pumps such as Cdr1/Pdr5 and Mdr1[79–81] or in their transcription factors such as Pdr1, Tac1 and Mrr1[81] increase drug efflux outside the cells, and are often responsible for multi-drug resistance. Another more ambiguous type of GOF are mutations in some drug target proteins, such as Cyp51 for azoles or Fks1 for echinocandins, that can prevent drug binding while maintaining the regular activity of the enzymes[82]. As expected for GOF resistance mechanisms, the impact of resistance mutations on protein function and stability are predicted to be lower than for LOF resistance mutations (Fig. 2). Consequently, GOF mechanisms are more difficult to identify directly from the predicted impact of mutations on the proteins.

Mapping AMR mutations on protein structures

Mutations leading to LOF through destabilization need to be located in regions of proteins important for stability. Alternatively, they could occur at the active sites of the proteins. We mapped resistance mutations on the protein structures to detect potential clustering of mutations. Fur1 and Fcy1 are tetrameric and dimeric proteins, respectively. In both cases, many of the mutations occur in structured and core residues, including at the interfaces, potentially disrupting complex assembly (Fig 3a, b). Erg3 and Hmg1 structures were predicted with AlphaFold 3[25]. For Erg3 (Fig. 3c, pTM = 0.91), reported mutations are in transmembrane helices. For Hmg1, Alphafold 3 predictions have a lower confidence level (Fig. 3d, ipTM = 0.64 and pTM = 0.47), but prior knowledge on the protein allows us to deduce that mutations are located in the conserved sterol sensing domain of the protein[83]. This could explain the results above (Fig. 2) where resistance mutations occur at conserved sites but are not predicted to destabilize the protein structure. Structural mapping therefore also helps interpret the impact of mutations.

Mutations involved in preventing drug binding are expected to cluster in regions of proteins where the drug and the protein interact. We mapped resistance mutations on the protein structures for proteins that are the most represented in the literature. For Fks, echinocandin resistance mutations are clustered in three different hotspots (Fig. 3e) as it has been previously observed[84,85]. However, there are also reported resistance mutations outside of the described hotspots, suggesting they might be larger than previously proposed (although we note some of these mutations have low confidence scores). The proximity of the hotspots on the enzyme

structure suggests a putative binding site of echinocandin drugs[85,86]. For Cyp51, azole resistance mutations are more dispersed. However, some regions of the protein appear to be enriched with such mutations (Fig. 3f). A DMSof *C. albicans* Cyp51 (Erg11) has been conducted, and the authors similarly concluded that resistance mutations are dispersed and not limited to distinct clusters[87].

One potential way to identify GOF mutations in transcription factors would be to examine whether mutations cluster in space in specific domains or regions associated with particular functions. No complete experimentally determined structures were available for any lead transcription factors. Therefore, we predicted two of them using Alphafold 3[25] in order to map resistance mutations. The prediction scores are low (Upc2 : ipTM = 0.37 and pTM = 0.45, Pdr1 : ipTM = 0.4 and pTM = 0.69), meaning we cannot be confident of the predicted structure. For Upc2, the structure of the ligand-binding domain in complex with ergosterol has been determined experimentally. Therefore, we superposed the crystalized ligand-binding domain with the complete Alphafold 3 predicted structure (Fig. 3g). We found that the few reported resistance mutations cluster in Upc2 but not in the DNA binding region. Mutations in Upc2 have already been found to increase the protein activity through the interruption of ergosterol binding in the sterol-binding pocket[88]. For Pdr1, although the predicted structure is of low confidence, we found that the resistance mutations are located across the structure, but not at the DNA binding site (Fig. 3h). Interestingly, Pdr1 mutations are in the middle range in terms of predicted effects based on conservation and destabilization (Fig. 2). Thus mutations could modify Pdr1 activity through minor destabilization. These observations suggest that resistance mutations in transcription factors often do not directly alter DNA binding but rather other facets of their activity.

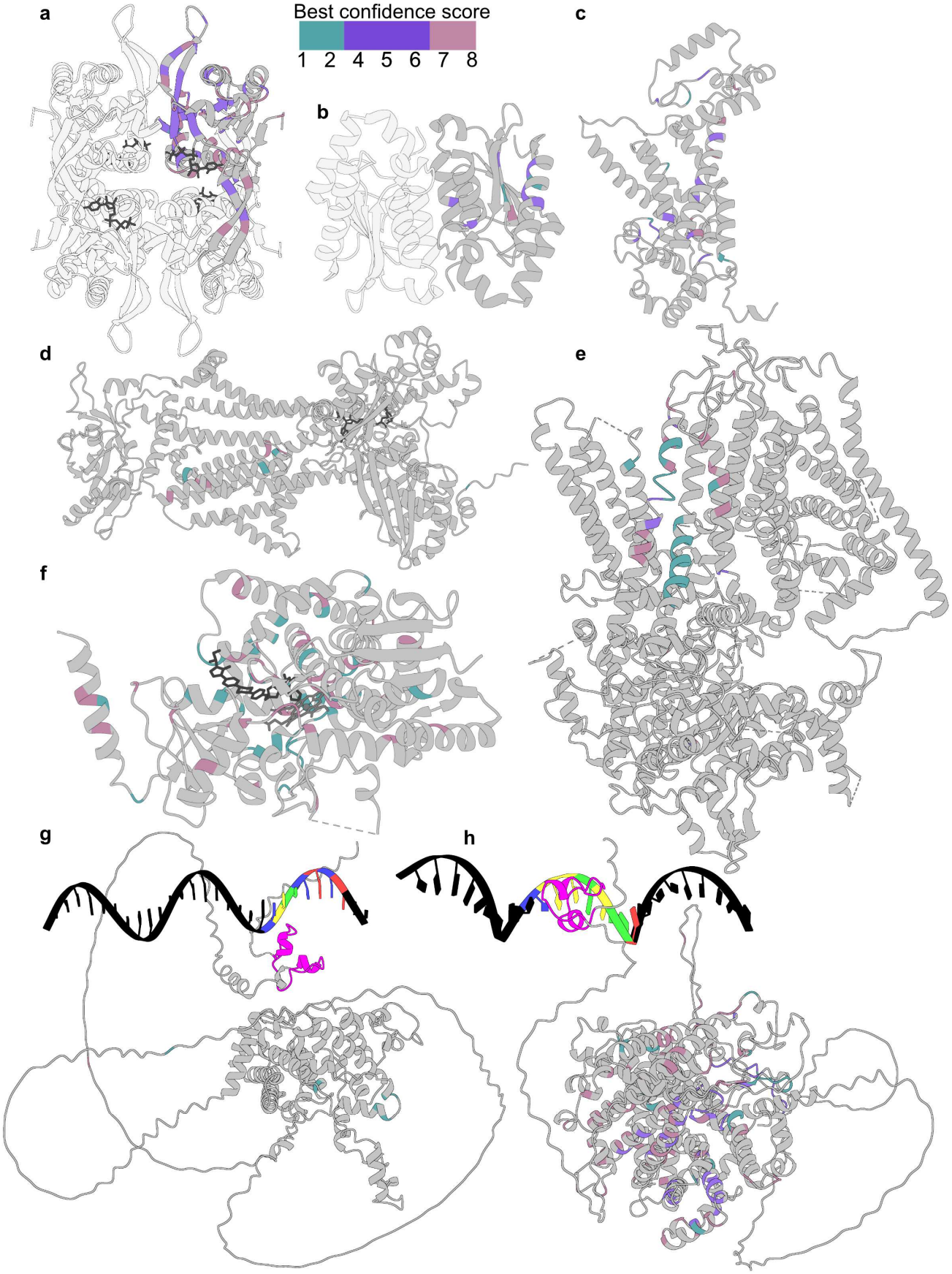


Fig. 3 | Mapping of resistance mutations on protein tertiary structures. a,b,c,d,e,f,g,h For each protein structure, residues are colored by the best positive confidence score for all reported mutations at this residue.

Multiple sequence alignments were used to map resistance mutations of orthologs. Confidence scores of 3 (DMS assays) were excluded. *C. albicans* Fur1 in complex with UTP (PDB ID 7RH8) (**a**), *S. cerevisiae* Fcy1 (PDB ID 1P6O) (**b**), *S. cerevisiae* Erg3 (AlphaFold 3 prediction with Uniprot P32353 protein sequence, pTM = 0.91) (**c**), *A. fumigatus* Hmg1 (AlphaFold 3 prediction with Uniprot Q4WHZ1 protein sequence and NADP, ipTM = 0.64 and pTM = 0.47) (**d**), *S. cerevisiae* Fks1 (PDB ID 7XE4) (**e**), *C. albicans* Erg11 bound to itraconazole (PDB ID 5V5Z) (**f**), *C. albicans* Upc2 (AlphaFold 3 prediction with Uniprot Q59QC7 protein sequence and AATATCGTACCCGATTATG**TCGTATA**TT DNA sequence[26], ipTM = 0.37 and pTM = 0.45, grey) superposed with *N. glabrata* Upc2 ligand-binding domain in complex with ergosterol (PDB ID 7VPR, white with orthologous resistance mutations in dark yellow) (**g**)**,** *S. cerevisiae* Pdr1 (AlphaFold 3 prediction with Uniprot P12383 protein sequence and GAGAAATGTC**TCCGCGGA**ACTCTTCTAC DNA sequence[28], ipTM = 0.4 and pTM = 0.69) (**h**). **g,h** In bright pink are the Zn2-C6 fungal-type DNA binding domains. On the DNA strands, colors represent the nucleotides (T:Blue, C:Yellow, G:Green, A:Red) of the core Upc2 DNA binding site **TCGTATA** of *C. albicans ERG11* promoter[26,27] (**g**) and of the pleiotropic drug response element (PDRE) **TCCGCGGA** present in *S. cerevisiae PDR5* promoter[28] (**h**).

Extreme convergence of AMR mutations among species

The use of the same antifungals to treat infections by different species allows us to examine evolutionary convergence in the mechanisms of resistance. Convergence is often used as evidence that a trait has been under selection, which in the case of AMR helps interpret and prioritize mutations for validation and for modifying drugs to overcome frequently encountered resistance mutations[89]. Experimental evolution studies have shown that the diversity of mutated genes to adapt to antifungals was narrow within a species and often overlapped with mechanisms described in other fungi[72,90,91]. We first examined the level of convergence in the resistance mechanisms among fungi by comparing resistance genes between species (Fig. 4a, b, Extended data Fig. 2). Outside of the specific drug targets such as Erg11 and Fks, which are expected to be convergent, we found that, generally, fungi evolve resistance through mutations in homologous genes or in genes with similar functions. One example of this is efflux pumps of the ABC (e.g. Cdr1 and Mdr2) and MFS (e.g. Mfs1) transporter families[81] and the transcription factors that regulate their expression (e.g. Tac1 and Pdr1). The absence of convergence, or species-specificity in AMR mechanisms, is more difficult to assess due to incomplete data resulting from varying research efforts across species.

Beyond gene-level convergence, we can also examine convergence at the level of individual mutations. Some proteins, such as Cyp51 and Fks, are well conserved across fungi, and it is therefore possible that the same mutations would confer resistance across species. Using multiple sequence alignments (MSA), we compared the identity of resistance mutations across orthologs. We found several cases where the same resistance mutations occur in two or more species (Fig. 4c). The most impressive examples are amino acid substitutions in Cyp51. For instance, Y132H and F in Cyp51 (*C. albicans* aa numbering) have been found to confer azole resistance in more than 20 species, including the distantly related species *C. albicans, A. fumigatus and C. neoformans*. Amino acid substitutions at position Y132 have been found to disrupt azole binding[92]. Other striking examples are mutations at position S645 of Fks (*C. albicans* aa numbering) and G143 of the cytochrome b (*A. fumigatus* aa numbering), which are associated with resistance in 10 species or more. In Fks, S645 is located in the echinocandin resistance hotspot 1 and is most likely involved in drug binding[85,86]. For cytochrome b, it is

hypothesized that the G143A mutation alters the binding of Qo inhibitors, which are a class of fungicides that target the quinol oxidation (Qo) site of the cytochrome bc1 complex, through steric interactions[93]. Mutations at residues where resistance mutations with a high level of convergence are observed can confidently be interpreted as resistance mutations across species. Such convergence could be explained by mutation bias, but given the breadth of species concerned, it is likely that convergence is rather due to the fact that these mutations confer high resistance levels with little to no evolutionary tradeoff in comparison to other possible resistance mutations[89]. We provide the MSA of the most frequently reported genes across multiple species as a resource (Supplementary data 1).

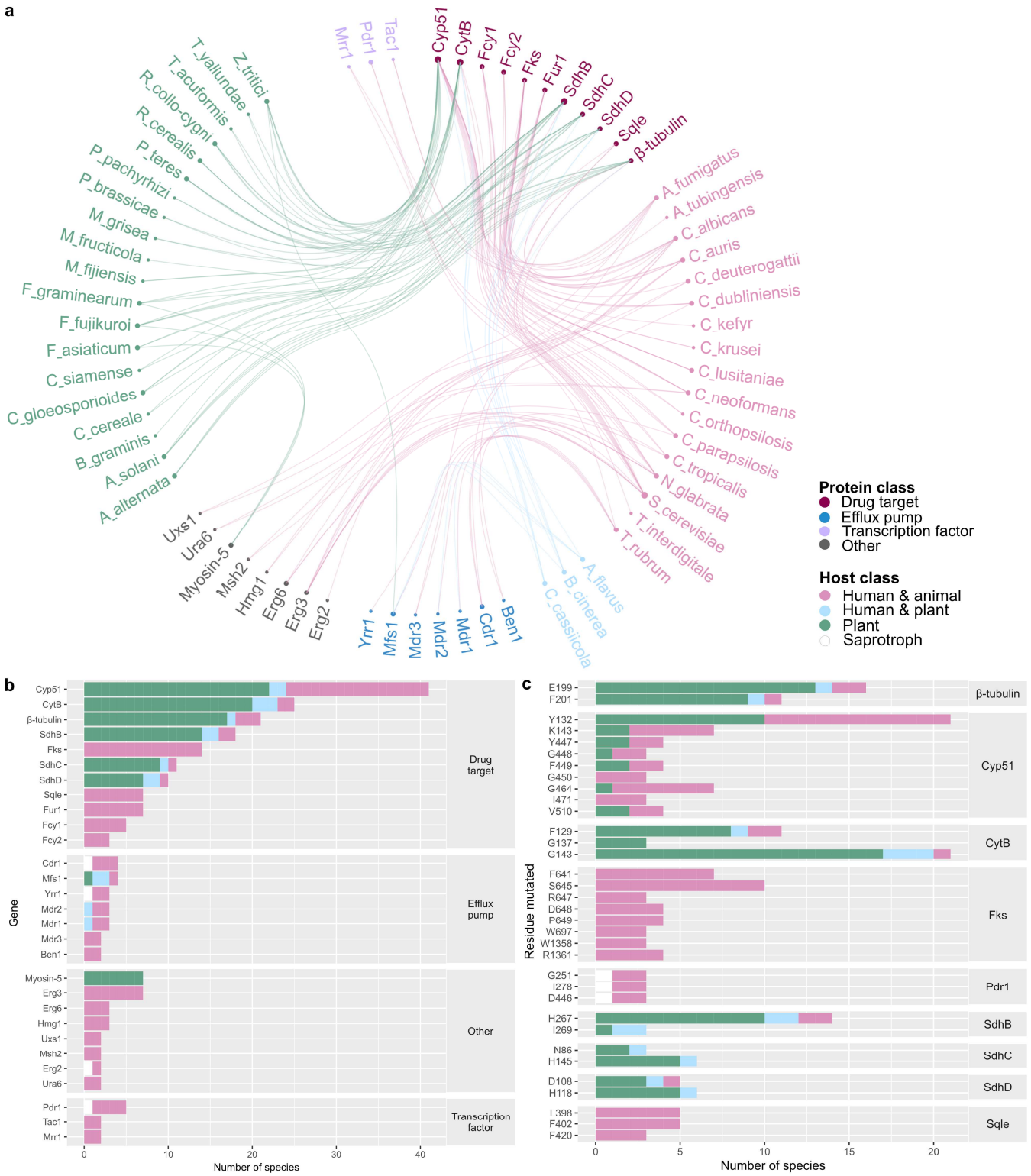


Fig. 4 | Massive convergence in the mechanisms of resistance among fungi at the gene and mutation level. **a,** Edgebundle plot[94] of mutated proteins associated with resistance in more than one fungal pathogen. Only proteins and species with more than five entries were included in this analysis. The interactive html file is available in Supplementary data 2. Genes classes were colored based on their function. **b,** Number of species with the same mutated protein associated with resistance. Only instances of proteins associated with more than two species are shown for clarity. **c,** Orthologous mutations conferring resistance in several species show a high level of convergence at the mutation level. Cyp51, Sqle (Squalene epoxidase) and Fks use *C. albicans* numbering in the MSA. SdhB, SdhC, SdhD, Beta-tubulin and Cytb (Cytochrome b) use *Z.tritici* numbering in the MSA. Pdr1 uses

S.cerevisiae numbering. Only instances of mutations associated with more than two species are shown for clarity. Entries with multiple proteins were excluded.

Cross-resistance is common between and within classes of antifungals

One major question regarding AMR mutations is whether they confer cross-resistance to more than one drug, either of the same class or different classes. FungAMR allows quantifying the extent of cross-resistance. Not all resistant strains present in our dataset were systematically assayed for multiple drugs, so we cannot identify clear cases of resistance mutations that are specific to one drug. Our extensive data nevertheless provides a global view of which mutations in which genes confer cross-resistance within and between drug classes. We found that if a mutation in a gene provides resistance to one antifungal, it will likely confer cross-resistance to others within a class compared to between classes (Fig. 5a, b, Extended data Fig. 2). As expected, we also found extensive cross-resistance between clinical and agricultural azole antifungals. This is expected given the similarities of molecules within a class[95].

Although within-class cross-resistance is more frequent, between-class cross-resistance also occurs[96,97]. Mutations found in clinical isolates in proteins implicated in ergosterol biosynthesis such as Cyp51, Upc2, Hmg1, Erg3 and Erg6 can confer resistance to both azoles and polyenes, even if these antifungals target different elements of the ergosterol biosynthesis pathway (Fig. 5b). Additionally, some mutations in Fks or Erg proteins seem to be involved in both azole and echinocandin resistance (Fig. 5b). This connection is more surprising since echinocandins do not target cell membrane proteins but target a beta-glucan synthase, and beta-glucan is a component of the fungal cell wall. We can hypothesize that changes in the cell wall or membrane composition can provide an advantage for cells under drug-induced stress, making resistance to the different clinical antifungal classes more interconnected than anticipated from the drugs' mode of action. Cross-resistance caused by interconnection of transcriptional pathways has been reported before[98] but our observations suggest that cross-resistance between classes could also happen from mutations in the drug targets. On the side of agricultural antifungals, cross-resistance between the different classes is also present, but seems to be less frequent than among human pathogens. Such a result could also be due to lack of data, as resistance has been less frequently tested systematically among classes of plant antifungals. Although incomplete, these analyses are sufficient to further strengthen important concerns. First, cross-resistance between medical and agricultural azoles is frequent. Second, azoles, polyenes and echinocandins are connected through a core set of genes, meaning that cross-resistance to clinical drugs can evolve through single mutations. The full extent of antifungal cross-resistance remains to be assessed but this analysis reveals that three out of the four classes of drugs used to treat fungal infections are not entirely orthogonal. This interdependence stresses the need for finding new molecular targets and drugs for treatment and most importantly, increase stewardship between clinical and agricultural drugs when new drugs are developed and new targets identified[97].

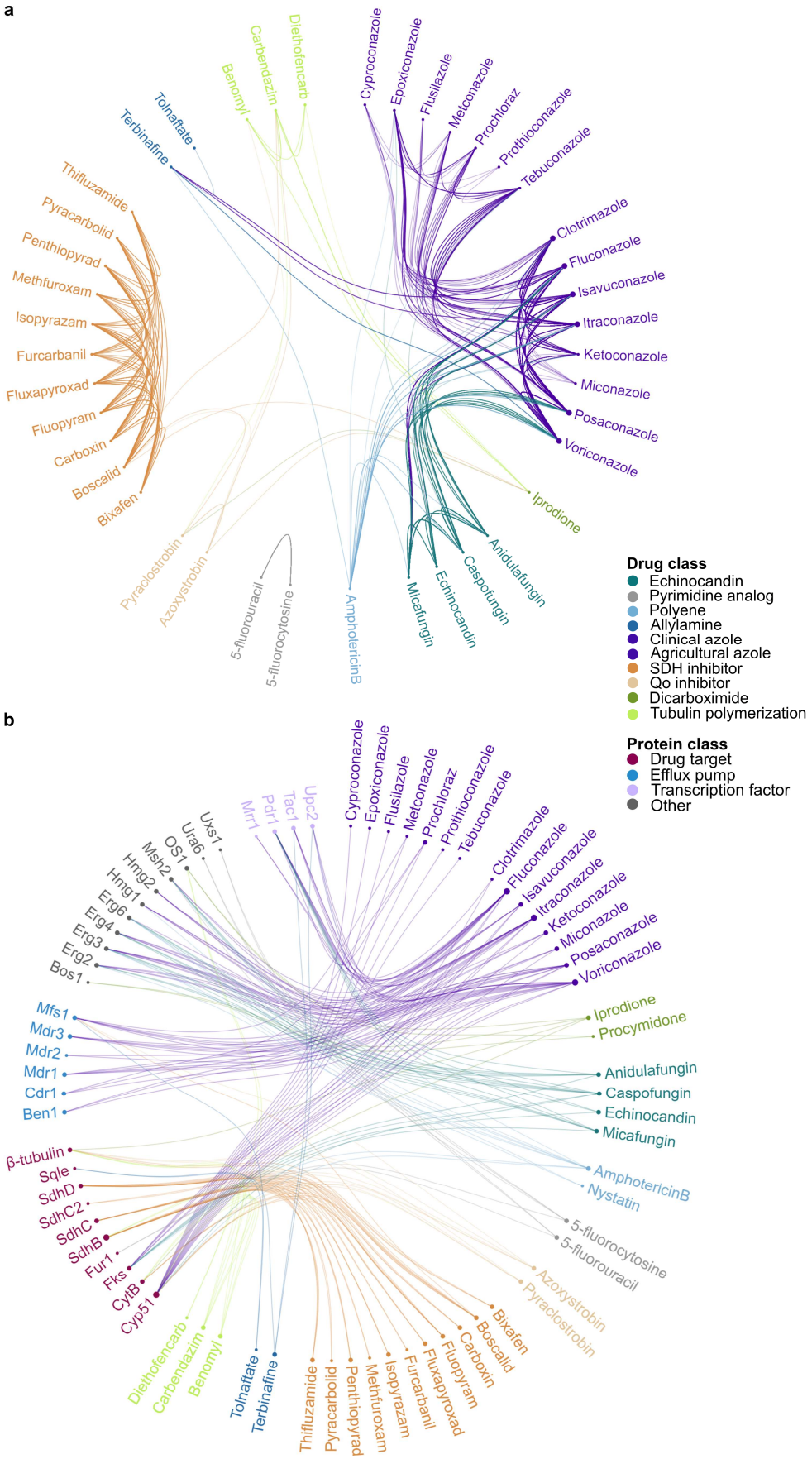


Fig. 5 | Rampant cross-resistance between and within classes of antifungals. a, Edgebundle plot[94] of cross-resistance between antifungals. Drugs are linked when a unique mutation in a single protein can confer resistance

to multiple drugs. Only proteins and drugs with more than five entries were included. The interactive html file is available in Supplementary data 3. **b,** Edgebundle plot[94] of mutated proteins associated with at least one mutation conferring resistance to more than one clinical or agricultural antifungal. Only proteins and drugs with more than five entries were included. Entries with multiple proteins mutated in the same strain were excluded. The interactive html file is available in Supplementary data 4.

ChroQueTas efficiently screen AMR in fungal genomes

To facilitate the analysis of AMR in fungal genomes from the FungAMR data, we developed ChroQueTas, a user-friendly bioinformatic tool that only necessitates a fungal genome in FASTA format. As a practical example of how ChroQueTas can be used to identify AMR mutations using FungAMR, we downloaded the raw sequencing data (Illumina) from 46 *C. albicans* and 144 *Zymoseptoria tritici* isolates[99–101] (Supplementary table 5). These two species were selected to represent human and agricultural pathogens, respectively. The *C. albicans* dataset included isolates from two countries (China and the United States of America). The *Z. tritici* dataset included isolates recovered from wheat from four countries (Australia, Israel, Switzerland and the United States of America). After quality control of the sequencing data and further assembly, the genomes were submitted to ChroQueTas by choosing their species-associated scheme and Cyp51 (Erg11) protein as target.

The screening of the *C. albicans* and *Z. tritici* genomes performed with ChroQueTas was based on FungAMR data, which contains the largest number of AMR mutations reported to date, and thus, we were able to identify all the AMR mutations reported by the authors. As a summary, we found that 83.3% of the selected *C. albicans* genomes harbored at least one mutation in Erg11 known to cause AMR (Fig. 6a). The most commonly identified mutated positions were A114 (24% of genomes, n=11) and Y132 (24% of genomes). Nine of the *C. albicans* that harbored the mutation A114S also harbored Y257H, a combination that has been previously reported[102] to cause AMR against different azoles, including fluconazole and itraconazole. The other two genomes (SRR13587870 and SRR13587591) harbored mutation A114V (previously reported to not cause AMR alone[87]) together with F145L, a combination that have shown to confer resistance against fluconazole[103]. Regarding the position Y132, the nine Chinese isolates harbored the mutation Y132H, whereas the two U.S. isolates harbored the mutation Y132F, also associated with azole resistance. For the analysis of the *Z. tritici* genomes, 84.7% of the isolates showed at least one mutation in Cyp51 known to confer resistance (Fig. 6b). Additionally, 30.5% of the genomes (n=44) harbored Y459Del and/or G460Del deletions, associated with AMR against different azoles[104,105]. Additionally, the combination of ChroQueTas and FungAMR allowed us to identify eight additional AMR mutations in *C. albicans* and one in *Z. tritici* that were not reported by the authors in their corresponding studies (highlighted in green in Fig. 6). Additionally, we identified Y461D in SRR3740350 (*Z. tritici* isolated in Switzerland) which occurs at a position previously associated with AMR due to another mutation, suggesting this mutation may also contribute to resistance. The results summarized here show the strength of ChroQueTas and FungAMR to efficiently screen AMR in fungal genomes and the potential to identify novel mutations involved in AMR.

Fig. 6 | Practical example of AMR screening by using ChroQueTas and FungAMR. a, b, Phylogenetic comparison of *C. albicans* (**a**) and *Z. tritici* (**b**) datasets based on average nucleotide identity (ANI)**.** The trees' tips are coloured or shaped according to the country in which they were isolated. Each heatmap represents the position within the query protein sequences (Erg11 for *C. albicans* and Cyp51 for *Z. tritici*) and are coloured when a mutation

was identified. AA stands for amino acids. “Missing information” in (**a**) stands for proteins that could only be resolved partially due to genome fragmentation.

Discussion

Despite the threat that fungal pathogens and AMR represents, we were still missing a high quality centralized antifungal mutations repository. Here, we present FungAMR, a compendium of 54,666 carefully curated entries of clinical and agricultural importance. One major finding is that many resistance mutations reported in the literature have little experimental support beyond identifying a mutation in a known resistance gene in a resistant isolate, which can be coincidental (confidence scores of 7 and 8, Fig. 1a). Such low quality reports create an overconfidence in our knowledge of the identity of resistance mutations. Our observations call for more experiments and for the development of computational tools focusing on the direct association of mutations to resistance. The reconstruction of individual mutations in pathogens will always remain more challenging than in model organisms but recent progress with genome editing directly in the pathogens makes such experiments possible on an appreciable scale (confidence score of 1, Fig. 1a). Heterologous expression of resistance genes in a model organism such as *S. cerevisiae* has been proven to be a powerful solution in a large number of studies (confidence score of 2, Fig. 1a). Since human and *E. coli* genes [106,107] can complement many of *S. cerevisiae*'s orthologs, it is likely that most genes present in FungAMR can be expressed and be studied in model species. Furthermore, high throughput experiments such as deep mutational scanning are powerful ways to characterize a large number of genotype-resistance phenotype associations[67,82,87,108,109](confidence score of 3, Fig. 1a). As we show, the fact that AMR mutations share specific features and locations in proteins also opens the possibility of using computational approaches to further increase confidence into the likely causal role of mutations that could otherwise not be validated experimentally.

FungAMR also highlights how research is extremely skewed toward a few fungal species, genes and drugs (Fig. 1). The scientific community would benefit from a diversification of studies. For instance, by focusing analysis on only a few genes that are *a priori* known to be associated with resistance, we are limiting our comprehension of the full suite of different resistance mechanisms. Our ability to screen for resistant isolates through whole genome sequencing should greatly facilitate such an approach, yet sometimes even such datasets are only mined for variants in a subset of genes. Also, for most genes, our knowledge of resistance mutations across different species is too limited to assess how convergent they are. The same applies for the impact of mutations on AMR within and between classes in order to assess the level of cross-resistance. Although limited, the information we currently have is crucial in our fight against fungal pathogens and suggest that there is a high level of convergence in the mechanisms of resistance among fungi (Fig. 4) and that cross-resistance seems to be common within and between classes of antifungals (Fig. 5). FungAMR has the potential to help predict resistance mutations and develop new drugs, but more studies incorporating different species and different antifungals are needed to determine how universal our observations are. In the meantime, there is still an urgent need to identify new molecular targets and drugs for treatment, and the dual

use of some antifungal classes in the clinic and in agriculture should be considered a critical threat to our ability to treat human fungal infections.

In addition to providing an AMR mutation catalog, we show that variant effect predictors based on protein stability and sequence conservation can be used to differentiate between LOF and GOF resistance mechanisms. Along with structural mapping, these tools can be used to interpret the impact of mutations in proteins involved in antifungal resistance. However, we are missing structural models for most proteins listed in FungAMR. Even though advances are made in protein structure predictions, algorithms such as AlphaFold are dependent on training data availability. Even if predictions are extremely valuable to interpret the impact of mutations in a protein, accurate knowledge of the 3D structure is crucial for protein structure-based drug design[110]. The scientific community would greatly benefit from experimentally determined structures of proteins associated with resistance.

Finally, we developed ChroQueTas, a user-friendly bioinformatic tool based on the FungAMR resource, to facilitate the screening of AMR mutations in fungal genomes. We show that ChroQueTas and FungAMR together allow efficient screening of AMR in published fungal genomes by identifying all resistance mutations reported by the authors in addition to new resistance mutations that were not reported initially.

Acknowledgements

Funding

This work was supported by a Genome Québec and Genome Canada grant 6569, a CRL Canadian Institutes of Health Research Foundation grant (number 387697), and a NSERC CREATE grant (EvoFunPath) to CRL. CRL Holds the Canada Research Chair in Cellular Systems and Synthetic Biology. ACG and RSS acknowledge the support of the CIFAR Azrieli Global Scholars Program. NSERC DG and Research Manitoba grants to ACG. NMQ is currently funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 101034371. CB was supported by fellowships from the Vanier Canada Graduate Scholarship agency, Fonds de Recherche du Québec Santé (FRQS), EvoFunPath NSERC CREATE program and Université Laval. PCD was supported by a Vanier Canada Graduate Fellowship and a FRQS doctoral fellowship. RD was supported by a NSERC postdoctoral fellowship and a FRQS postdoctoral fellowship. NCG was supported by an Ontario Graduate Scholarship and is an EvoFunPath Fellow (NSERC CREATE). D.M-S. is currently funded by the State Research Agency (AEI) of the Ministry of Science, Innovation and Universities and the European Social Fund (MCIN/AEI/10.13039/501100011033 grant agreement No PRE2022-105868). A.J.A is currently funded by "Programa Investigo" and "Next Generation EU" funds. AB = Genome Research and Development Initiative

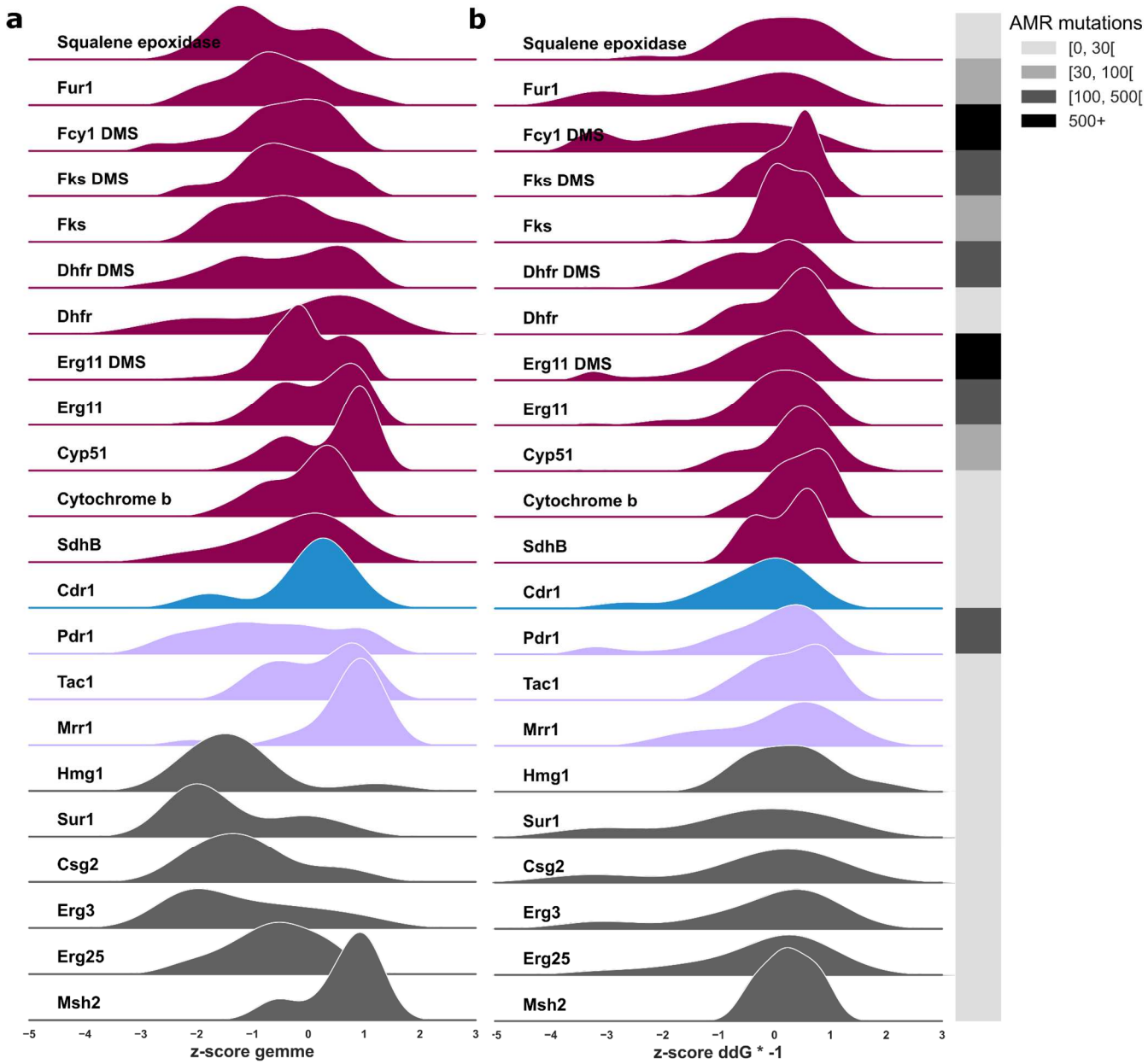
Data availability

The FungAMR resource is available in Supplementary table 1. The multiple sequence alignment files for the genes present in FungAMR are available in Supplementary data 1. Whole genome Illumina reads of *C. albicans* are deposited in NCBI SRA under Project number PRJNA592373. The dataset used for the Supplementary note 1 is available at 10.5281/zenodo.12583470. The Protein Data Bank is available at https://www.rcsb.org/. UniProt is available at UniProt. The AlphaFold database is available at AlphaFold Protein Structure Database.

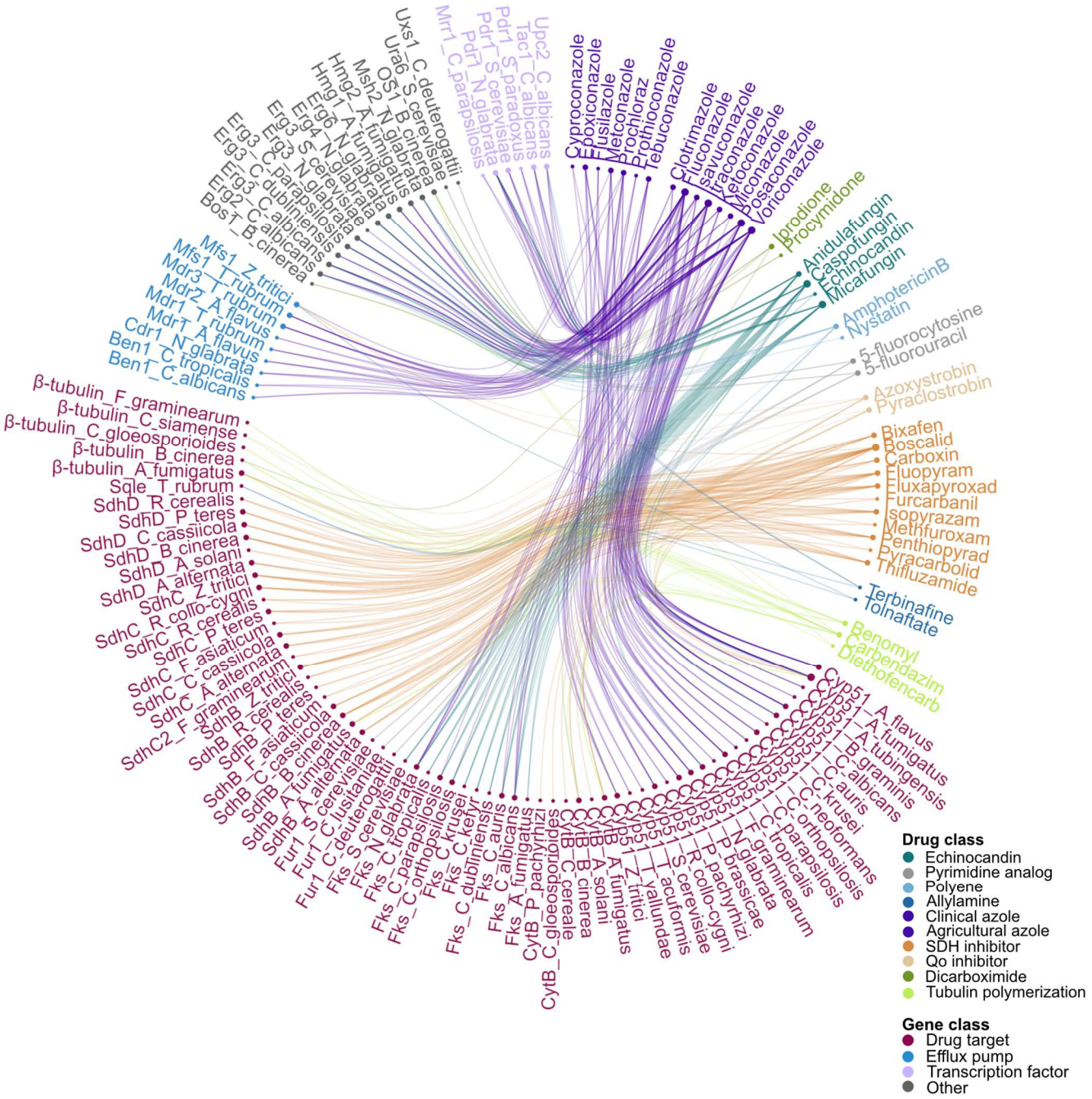
Code availability

All scripts for figures 1 to 5 and for FungAMR content analysis are available on GitHub at https://github.com/Landrylab/FungAMR. ChroQueTas is available on GitHub at https://github.com/nmquijada/ChroQueTas.

Extended data and figures



Extended data Fig. 1 | Distribution of predicted effect of AMR mutations on protein function and stability. a, Distribution of the standardized GEMME score for all unique resistance mutations across proteins. Negative values predict a higher impact on the protein function based on evolutionary conservation. **b,** Distribution of the standardized negated predicted change in the Gibbs free energy (ΔΔG) upon mutation predicted by FoldX for all unique resistance mutations across proteins. Negative values predict destabilization of the protein. For each protein, when mutations were available from DMS experiments (confidence score of 3), the DMS data is shown separately. The grayscale shows the number of unique resistance mutations per protein.



Extended data Fig. 2 | Convergence of resistance mechanisms and cross-resistance between antifungals. Edgebundle plot[94] linking species specific proteins to antifungal resistance. Orthologous proteins are associated with drug resistance across multiple fungal pathogens, demonstrating evolutionary convergence. Furthermore, mutations in the same protein are linked to resistance against multiple drugs and drug families, indicating cross-resistance. Only proteins and drugs with more than five entries were included. The interactive html file is available in Supplementary data 5.

References

1. Denning, D. W. Global incidence and mortality of severe fungal disease. *Lancet Infect. Dis.* (2024) doi:10.1016/S1473-3099(23)00692-8.
2. Brown, G. D. *et al.* Hidden Killers: Human Fungal Infections. *Sci. Transl. Med.* **4**, 165rv13–165rv13 (2012).
3. World Health Organisation. WHO fungal priority pathogens list to guide research, development and public health action. at https://www.who.int/publications-detail-redirect/9789240060241 (2022).
4. Fisher, M. C., Hawkins, N. J., Sanglard, D. & Gurr, S. J. Worldwide emergence of resistance to antifungal drugs challenges human health and food security. *Science* **360**, 739–742 (2018).
5. Fisher, M. C. & Garner, T. W. J. Chytrid fungi and global amphibian declines. *Nat. Rev. Microbiol.* **18**, 332–343 (2020).
6. Hoyt, J. R., Kilpatrick, A. M. & Langwig, K. E. Ecology and impacts of white-nose syndrome on bats. *Nat. Rev. Microbiol.* **19**, 196–210 (2021).
7. World Health Organization. Antimicrobial resistance. (2021).
8. Arastehfar, A. *et al.* The Quiet and Underappreciated Rise of Drug-Resistant Invasive Fungal Pathogens. *J Fungi (Basel)* **6**, (2020).
9. Yin, Y. *et al.* Fungicide Resistance: Progress in Understanding Mechanism, Monitoring, and Management. *Phytopathology* **113**, 707–718 (2023).
10. Gow, N. A. R. *et al.* The importance of antimicrobial resistance in medical mycology. *Nat. Commun.* **13**, 5352 (2022).
11. Perfect, J. R. The antifungal pipeline: a reality check. *Nat. Rev. Drug Discov.* **16**, 603–616 (2017).
12. Shelton, J. M. G. *et al.* Citizen science reveals landscape-scale exposures to multiazole-resistant Aspergillus fumigatus bioaerosols. *Science Advances* **9**, eadh8839 (2023).
13. Rhodes, J. *et al.* Population genomics confirms acquisition of drug-resistant Aspergillus fumigatus infection by humans from the environment. *Nat Microbiol* **7**, 663–674 (2022).
14. Lee, Y., Robbins, N. & Cowen, L. E. Molecular mechanisms governing antifungal drug resistance. *npj Antimicrob Resist* **1**, 1–9 (2023).

15. Fisher, M. C. *et al.* Tackling the emerging threat of antifungal resistance to human health. *Nat. Rev. Microbiol.* **20**, 557–571 (2022).

16. McArthur, A. G. *et al.* The comprehensive antibiotic resistance database. *Antimicrob. Agents Chemother.* **57**, 3348–3357 (2013).

17. Alcock, B. P. *et al.* CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Res.* **51**, D690–D699 (2023).

18. Nash, A. *et al.* MARDy: Mycology Antifungal Resistance Database. *Bioinformatics* **34**, 3233–3234 (2018).

19. Jain, A., Singhal, N. & Kumar, M. AFRbase: a database of protein mutations responsible for antifungal resistance. *Bioinformatics* **39**, (2023).

20. Remmert, M., Biegert, A., Hauser, A. & Söding, J. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nat. Methods* **9**, 173–175 (2011).

21. Steinegger, M. *et al.* HH-suite3 for fast remote homology detection and deep protein annotation. *BMC Bioinformatics* **20**, 473 (2019).

22. Laine, E., Karami, Y. & Carbone, A. GEMME: a simple and fast global epistatic model predicting mutational effects. *Mol. Biol. Evol.* **36**, 2604–2619 (2019).

23. Delgado, J., Radusky, L. G., Cianferoni, D. & Serrano, L. FoldX 5.0: working with RNA, small molecules and a new graphical interface. *Bioinformatics* **35**, 4168–4169 (2019).

24. Tiberti, M. *et al.* MutateX: an automated pipeline for in silico saturation mutagenesis of protein structures and structural ensembles. *Brief. Bioinform.* **23**, (2022).

25. Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* (2024) doi:10.1038/s41586-024-07487-w.

26. Vandeputte, P., Ferrari, S. & Coste, A. T. Antifungal resistance and new strategies to control fungal infections. *Int. J. Microbiol.* **2012**, 713687 (2012).

27. Hoot, S. J., Brown, R. P., Oliver, B. G. & White, T. C. The UPC2 promoter in Candida albicans contains two cis-acting elements that bind directly to Upc2p, resulting in transcriptional

autoregulation. *Eukaryot. Cell* **9**, 1354–1362 (2010).

28. Gao, C., Wang, L., Milgrom, E. & Shen, W.-C. W. On the mechanism of constitutive Pdr1 activator-mediated PDR5 transcription in Saccharomyces cerevisiae: evidence for enhanced recruitment of coactivators and altered nucleosome structures. *J. Biol. Chem.* **279**, 42677–42686 (2004).

29. Edgar, R. C. Muscle5: High-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. *Nat. Commun.* **13**, 6968 (2022).

30. Li, Y. *et al.* A genome-scale phylogeny of the kingdom Fungi. *Curr. Biol.* **31**, 1653–1665.e5 (2021).

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

32. Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).

33. Nguyen, L.-T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).

34. Yu, G., Smith, D. K., Zhu, H., Guan, Y. & Lam, T. T.-Y. Ggtree: An r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* **8**, 28–36 (2017).

35. Meng, E. C. *et al.* UCSF ChimeraX: Tools for structure building and analysis. *Protein Sci.* **32**, e4792 (2023).

36. Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).

37. Harris, C. R. *et al.* Array programming with NumPy. *Nature* **585**, 357–362 (2020).

38. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).

39. Team, T. P. D. *Pandas-Dev/pandas: Pandas*. (Zenodo, 2022).

40. Waskom, M. L. seaborn: statistical data visualization. *Journal of Open Source Software* **6**, 3021 (2021).

41. Wickham, H., François, R., Henry, L., Müller, K. & Vaughan, D. *Dplyr: A Grammar of Data*

Manipulation. (2023).

42. Wickham, H., Vaughan, D. & Girlich, M. *Tidyr: Tidy Messy Data*. (2024).

43. Wickham, H. *Stringr: Simple, Consistent Wrappers for Common String Operations*. (2022).

44. Csárdi, G. *et al. igraph: Network Analysis and Visualization in R*. (2024). doi:10.5281/zenodo.7682609.

45. Csardi, G. & Nepusz, T. The igraph software package for complex network research. *InterJournal* **Complex Systems**, 1695 (2006).

46. Paradis, E. & Schliep, K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528 (2019).

47. Louca, S. & Doebeli, M. Efficient comparative phylogenetics on large trees. *Bioinformatics* **34**, 1053–1055 (2018).

48. Castiglione, S. *et al.* A new method for testing evolutionary rate variation and shifts in phenotypic evolution. *Methods in Ecology and Evolution* **9**, 974–983 (2018).

49. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer New York, 2010).

50. Li, H. Protein-to-genome alignment with miniprot. *Bioinformatics* **39**, (2023).

51. Camacho, C. *et al.* BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).

52. Babraham bioinformatics - FastQC A quality control tool for high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc.

53. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).

54. Bankevich, A. *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).

55. Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* **29**, 1072–1075 (2013).

56. Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).

57. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

58. Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T. & Aluru, S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* **9**, 5114 (2018).

59. Ripley, B. D. The R project in statistical computing. *MSOR Connect.* **1**, 23–25 (2001).

60. Wickham, H. Reshaping Data with the reshape Package. *J. Stat. Softw.* **21**, 1–20 (2007).

61. *Rmarkdown: Dynamic Documents for R*. (Github).

62. Grant, B. J., Rodrigues, A. P. C., ElSawy, K. M., McCammon, J. A. & Caves, L. S. D. Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics* **22**, 2695–2696 (2006).

63. Wickham, H., François, R., Henry, L. & Müller, K. A Grammar of Data Manipulation [R package dplyr version 1.0.2]. (2020).

64. Bodenhofer, U., Bonatesta, E., Horejš-Kainrath, C. & Hochreiter, S. msa: an R package for multiple sequence alignment. *Bioinformatics* **31**, 3997–3999 (2015).

65. Create Interactive 3D Visualizations of Molecular Data [R package r3dmol version 0.1.2]. https://CRAN.R-project.org/package=r3dmol.

66. Bhattacharya, S., Esquivel, B. D. & White, T. C. Overexpression or Deletion of Ergosterol Biosynthesis Genes Alters Doubling Time, Response to Stress Agents, and Drug Susceptibility in Saccharomyces cerevisiae. *MBio* **9**, (2018).

67. Després, P. C. *et al.* Asymmetrical dose responses shape the evolutionary trade-off between antifungal resistance and nutrient use. *Nat Ecol Evol* **6**, 1501–1515 (2022).

68. Durand, R., Jalbert-Ross, J., Fijarczyk, A., Dubé, A. K. & Landry, C. R. Cross-feeding affects the target of resistance evolution to an antifungal drug. *PLoS Genet.* **19**, e1011002 (2023).

69. Chang, Y. C. *et al.* Moderate levels of 5-fluorocytosine cause the emergence of high frequency resistance in cryptococci. *Nat. Commun.* **12**, 3418 (2021).

70. Edlind, T. D. & Katiyar, S. K. Mutational analysis of flucytosine resistance in Candida glabrata. *Antimicrob. Agents Chemother.* **54**, 4733–4738 (2010).

71. Charlier, C. *et al.* Acquired Flucytosine Resistance during Combination Therapy with Caspofungin and Flucytosine for Candida glabrata Cystitis. *Antimicrob. Agents Chemother.* **60**, 662–665 (2016).

72. Gerstein, A. C., Lo, D. S. & Otto, S. P. Parallel Genetic Changes and Nonparallel Gene–Environment Interactions Characterize the Evolution of Drug Resistance in Yeast. *Genetics* **192**, 241–252 (2012-9).

73. Carolus, H. *et al.* Acquired amphotericin B resistance and fitness trade-off compensation in Candida auris. *Research Square* (2023) doi:10.21203/rs.3.rs-3621420/v1.

74. Jensen, R. H. *et al.* Stepwise emergence of azole, echinocandin and amphotericin B multidrug resistance in vivo in Candida albicans orchestrated by multiple genetic alterations. *J. Antimicrob. Chemother.* **70**, 2551–2555 (2015).

75. Lim, H. J. *et al.* Whole-Genome Sequence Analysis of Candida glabrata Isolates from a Patient with Persistent Fungemia and Determination of the Molecular Mechanisms of Multidrug Resistance. *J Fungi (Basel)* **9**, (2023).

76. Vaser, R., Adusumalli, S., Leng, S. N., Sikic, M. & Ng, P. C. SIFT missense predictions for genomes. *Nat. Protoc.* **11**, 1–9 (2016).

77. Taylor, M. B. *et al.* yEvo: experimental evolution in high school classrooms selects for novel mutations that impact clotrimazole resistance in Saccharomyces cerevisiae. *G3* **12**, (2022).

78. MacPherson, S. *et al.* Candida albicans zinc cluster protein Upc2p confers resistance to antifungal drugs and is an activator of ergosterol biosynthetic genes. *Antimicrob. Agents Chemother.* **49**, 1745–1752 (2005).

79. Harris, A. *et al.* Structure and efflux mechanism of the yeast pleiotropic drug resistance transporter Pdr5. *Nat. Commun.* **12**, 5254 (2021).

80. Prasad, R., Banerjee, A., Khandelwal, N. K. & Dhamgaye, S. The ABCs of Candida albicans Multidrug Transporter Cdr1. *Eukaryot. Cell* **14**, 1154–1164 (2015).

81. Cannon, R. D. *et al.* Efflux-mediated antifungal drug resistance. *Clin. Microbiol. Rev.* **22**, 291–321, Table of Contents (2009).

82. Durand, R. *et al.* Mutational landscape and molecular bases of echinocandin resistance. *bioRxiv*

2024.07.21.604487 (2024) doi:10.1101/2024.07.21.604487.

83. Rybak, J. M. *et al.* Mutations in hmg1, Challenging the Paradigm of Clinical Triazole Resistance in Aspergillus fumigatus. *MBio* **10**, (2019).

84. Johnson, M. E., Katiyar, S. K. & Edlind, T. D. New Fks hot spot for acquired echinocandin resistance in Saccharomyces cerevisiae and its contribution to intrinsic resistance of Scedosporium species. *Antimicrob. Agents Chemother.* **55**, 3774–3781 (2011).

85. Hu, X. *et al.* Structural and mechanistic insights into fungal β-1,3-glucan synthase FKS1. *Nature* **616**, 190–198 (2023).

86. Jospe-Kaufman, M. *et al.* Reshaping Echinocandin Antifungal Drugs To Circumvent Glucan Synthase Point-Mutation-Mediated Resistance. *Angew. Chem. Int. Ed Engl.* **63**, e202314728 (2024).

87. Bédard, C. *et al.* Most azole resistance mutations in the Candida albicans drug target confer cross-resistance without intrinsic fitness cost. *Nat. Microbiol.* (2024) doi: 10.1038/s41564-024-01819-2.

88. Tan, L. *et al.* Structural basis for activation of fungal sterol receptor Upc2 and azole resistance. *Nat. Chem. Biol.* **18**, 1253–1262 (2022).

89. Storz, J. F. Causes of molecular convergence and parallelism in protein evolution. *Nat. Rev. Genet.* **17**, 239–250 (2016).

90. Carolus, H. *et al.* Genome-Wide Analysis of Experimentally Evolved Candida auris Reveals Multiple Novel Mechanisms of Multidrug Resistance. *MBio* **12**, e03333–20.

91. Ksiezopolska, E. *et al.* Narrow mutational signatures drive acquisition of multidrug resistance in the fungal pathogen Candida glabrata. *Curr. Biol.* **31**, 5314–5326.e10 (2021).

92. Sagatova, A. A. *et al.* Triazole resistance mediated by mutations of a conserved active site tyrosine in fungal lanosterol 14α-demethylase. *Sci. Rep.* **6**, 26213 (2016).

93. Fisher, N., Meunier, B. & Biagini, G. A. The cytochrome bc1 complex as an antipathogenic target. *FEBS Lett.* **594**, 2935–2952 (2020).

94. Tarr, G. *edgebundleR: Circle Plot with Bundled Edges*. (Github).

95. Robbins, N., Wright, G. D. & Cowen, L. E. Antifungal Drugs: The Current Armamentarium and

Development of New Agents. *Microbiol Spectr* **4**, (10 2016).

96. Gerstein, A. C. & Sethi, P. Experimental evolution of drug resistance in human fungal pathogens. *Curr. Opin. Genet. Dev.* **76**, 101965 (2022).

97. van Rhijn, N. *et al.* Aspergillus fumigatus strains that evolve resistance to the agrochemical fungicide ipflufenoquin in vitro are also resistant to olorofim. *Nat Microbiol* **9**, 29–34 (2024).

98. van Rhijn, N. *et al.* Antagonism of the Azoles to Olorofim and Cross-Resistance Are Governed by Linked Transcriptional Networks in Aspergillus fumigatus. *MBio* **13**, e0221522 (2022).

99. Ford, C. B. *et al.* The evolution of drug resistance in clinical isolates of Candida albicans. *Elife* **4**, e00662 (2015).

100. McDonald, M. C. *et al.* Rapid Parallel Evolution of Azole Fungicide Resistance in Australian Populations of the Wheat Pathogen Zymoseptoria tritici. *Appl. Environ. Microbiol.* **85**, (2019).

101. Gong, J. *et al.* Emergence of Antifungal Resistant Subclades in the Global Predominant Phylogenetic Population of Candida albicans. *Microbiol Spectr* **11**, e0380722 (2023).

102. Xiang, M.-J. *et al.* Erg11 mutations associated with azole resistance in clinical isolates of Candida albicans. *FEMS Yeast Res.* **13**, 386–393 (2013).

103. Kumar, A. *et al.* Assessment of antifungal resistance and associated molecular mechanism in Candida albicans isolates from different cohorts of patients in North Indian state of Haryana. *Folia Microbiol.* **65**, 747–754 (2020).

104. Cools, H. J. *et al.* Heterologous Expression of Mutated Eburicol 14α-Demethylase (CYP51) Proteins of Mycosphaerella graminicola To Assess Effects on Azole Fungicide Sensitivity and Intrinsic Protein Function. *Appl. Environ. Microbiol.* **76**, 2866–2872 (2010).

105. Leroux, P., Albertini, C., Gautier, A., Gredt, M. & Walker, A.-S. Mutations in the CYP51 gene correlated with changes in sensitivity to sterol 14 alpha-demethylation inhibitors in field isolates of Mycosphaerella graminicola. *Pest Manag. Sci.* **63**, 688–698 (2007).

106. Kachroo, A. H. *et al.* Evolution. Systematic humanization of yeast genes reveals conserved functions and genetic modularity. *Science* **348**, 921–925 (2015).

107. Kachroo, A. H. *et al.* Systematic bacterialization of yeast genes identifies a near-universally

swappable pathway. *Elife* **6**, (2017).

108. Fowler, D. M. & Fields, S. Deep mutational scanning: a new style of protein science. *Nat. Methods* **11**, 801–807 (2014).

109. Rouleau, F. D. *et al.* Deep mutational scanning of Pneumocystis jirovecii dihydrofolate reductase reveals allosteric mechanism of resistance to an antifolate. *PLoS Genet.* **20**, e1011252 (2024).

110. Karelina, M., Noh, J. & Dror, R. How accurately can one predict drug binding modes using AlphaFold models? at https://doi.org/10.1101/2023.05.18.541346 (2023).