Full Title: Analysis of Cyp51 protein sequences shows 4 major Cyp51 gene family groups across Fungi

Running Title: 4 Fungal groups of Cyp51 proteins

Authors

Brandi N Celia-Sanchez[1], Brandon Mangum[1], Marin Brewer[2], Michelle Momany[1*]

Affiliations

[1] Department of Plant Biology, University of Georgia, Athens, Georgia, United States of America

[2] Department of Plant Pathology, University of Georgia, Athens, Georgia, United States of America

*Corresponding Author (0000-0001-7110-063X)

Abstract

Azole drugs target fungal sterol biosynthesis and are used to treat millions of human fungal infections each year. Resistance to azole drugs has emerged in multiple fungal pathogens including *Candida albicans*, *Cryptococcus neoformans*, *Histoplasma capsulatum*, and *Aspergillus fumigatus*. The most well-studied resistance mechanism in *A. fumigatus* arises from missense mutations in the coding sequence combined with a tandem repeat in the promoter of *cyp51A*, which encodes a cytochrome P450 enzyme in the fungal sterol biosynthesis pathway. Filamentous members of Ascomycota such as *A. fumigatus* have either one or two of three Cyp51 paralogs (Cyp51A, Cyp51B, and Cyp51C). Most previous research in *A. fumigatus* has focused on Cyp51A due to its role in azole resistance. We used the *A. fumigatus* Cyp51A protein sequence as the query in database searches to identify Cyp51 proteins across Fungi. We found 435 Cyp51 proteins in 301 species spanning from early-diverging fungi (Blastocladiomycota, Chytridiomycota, Zoopagomycota and Mucormycota) to late-diverging fungi (Ascomycota and Basidiomycota). We found these sequences formed 4 major Cyp51 groups: Cyp51, Cyp51A, Cyp51B, and Cyp51C. Surprisingly, we found all filamentous Ascomycota had a Cyp51B paralog, while only 50% had a Cyp51A paralog. We created maximum likelihood trees to investigate the evolution of Cyp51 in Fungi. Our results suggest Cyp51 is present in all fungi with three paralogs emerging in Pezizomycotina, including Cyp51C which appears to have diverged from the progenitor of the Cyp51A and Cyp51B groups.

Author Summary

Each year millions of people are infected by a fungal pathogen and receive antifungal treatment with azole drugs. Resistance to azole drugs is becoming increasingly prevalent and is mostly caused by mutations in the azole drug target, Cyp51. *Aspergillus fumigatus* is an airborne fungal pathogen that causes more than 600,000 deaths every year. Azole resistance in *A. fumigatus* is primarily driven by a promoter repeat coupled with mutations in *cyp51A*. In our study, we found 435 Cyp51 proteins in 4 major groups across Fungi, with some species having multiple Cyp51 proteins (Cyp51, Cyp51A, Cyp51B, and Cyp51C). Although most research in *A. fumigatus* has focused on Cyp51A, we found Cyp51B in all filamentous Ascomycota fungi showing it is more conserved than Cyp51A and likely plays a vital role in these fungi.

Author Summary (Shortened)

Resistance to azole drugs is becoming increasingly prevalent and is mostly caused by mutations in the azole drug target, Cyp51. Azole resistance in Aspergillus fumigatus is primarily driven by a promoter repeat coupled with mutations in cyp51A. We found 435 Cyp51 proteins in 4 major groups across Fungi, with some species having multiple Cyp51 proteins (Cyp51, Cyp51A, Cyp51B, and Cyp51C). Although most research focuses on Cyp51A, we found Cyp51B in all filamentous Ascomycota fungi showing it's more conserved than Cyp51A.

Introduction

Fungal pathogens caused over 9 million diagnosed infections in 2017 in the United States, but the true fungal burden is hard to estimate since many cases are likely undiagnosed [1, 2]. The infections caused by fungal pathogens include severe chronic conditions, complex chronic respiratory conditions, recurrent infections, and many life-threatening invasive diseases [3]. Invasive fungal infections generally occur in individuals with suppressed or compromised immune systems [4]. These infections have a high mortality rate if not treated early with appropriate antifungal drugs [4]. Major drugs used to treat invasive fungal infections are echinocandins, polyenes, flucytosine, and azole drugs [5]. Azoles, which target synthesis of the fungal-specific membrane component ergosterol, are among the most highly used antifungal drugs.

Cyp51 proteins, also known as Erg11 in Ascomycota yeast in the Saccharomycotina and Taphrinomycotina subphyla, are in all biological kingdoms and are highly conserved [6]. Cyp51 proteins have 6 substrate recognition sites (SRS), an oxygen-binding motif (AGXDTT), PER and EXXR motifs that create an E-R-R triad within the heme pocket, and a conserved heme-binding motif (FXXGXXXCXG) (Fig. 1) [7, 8]. Azole drugs competitively bind to sterol 14 alpha-demethylase (Cyp51, Erg11), a cytochrome P450 in the ergosterol biosynthesis pathway in fungi. Azole drugs consist of a heterocyclic ring with either two (imidazoles) or three (triazoles) nitrogens and a sidechain. The side chain of azoles interacts with the Cyp51 polypeptide while the nitrogen in the azole heterocyclic ring interacts directly with the sixth ligand of the heme ferric ion, a cofactor of the Cyp51 protein [9]. Cytochrome P450 proteins conduct a three-step reaction within the sterol biosynthesis pathway leading to the production of cholesterol in animals, sitosterol in plants, and ergosterol in fungi [10, 11]. Sterols are integrated into the cell membrane where they aid in membrane fluidity and permeability [10]. The binding of azoles to Cyp51 depletes intracellular ergosterol and causes accumulation of methylated sterols and toxic intermediate sterols within the fungal cell membrane causing arrested growth and cell membrane stress [12].

Many fungi have acquired mutations in *cyp51*that alter the ability of azoles to bind and inhibit Cyp51 [13]. Many missense mutations that have been shown to decrease sensitivity to azoles as determined by minimum inhibitory concentration (MIC) in the human fungal pathogens *Candida albicans*, *Cryptococcus neoformans, Histoplasma capsulatum*, and *A. fumigatus* [13-17] occur in substrate recognition sites causing azoles to interact and bind differently within Cyp51. Increased expression levels of *cyp51A* due to 34-, 46-, 53- and 120-bp tandem repeats in the promoter have occurred in *A. fumigatus* leading to high levels of pan-azole resistance, resistance to more than one azole drug [18-21]. Tandem repeats in the *cyp51A* promoter reduce the affinity of the promoter and the CGAAT binding complex (CBC), which binds to CGAAT in the promoter and

downregulates *cyp51A* expression [22]. Although Cyp51A has been the focus of most studies in *A. fumigatus*, a second paralog (Cyp51B) has also been documented to cause resistance through upregulation and missense mutations [23, 24]. Like human pathogens, plant pathogens (such as *Erysiphe necator, Mycosphaerella* spp., *Penicillium digitatum, Venturia inaequalis*) undergo changes in the *cyp51A* promoter (substitutions, insertions, duplications) and/or mutations in substrate recognition sites to alter expression and binding of Cyp51A [25-34].

Filamentous Ascomycota have multiple Cyp51 paralogs that may play different roles in the cell under normal conditions or under azole stress. The goal of our study was to understand the evolutionary relationships among fungal Cyp51 proteins.

Results and Discussion

435 fungal Cyp51 proteins were analyzed. Cyp51 proteins were previously defined as having 6 substrate recognition sites (SRS), an oxygen-binding motif (AGXDTT), PER and EXXR motifs that create an E-R-R triad within the heme pocket, and a conserved heme-binding motif (FXXGXXXCXG) [7, 8] (Fig. 1). To understand Cyp51 genes across Fungi, the *A. fumigatus* Cyp51A protein was used as a reference in a protein BLAST (Basic Local Alignment Search Tool) [35]. A total of 4404 protein sequences resulted and were filtered to retain those with greater than 50% coverage and greater than 30% percent identity to the reference sequence, resulting in 480 sequences (Supplemental Table 1). The resulting protein sequences were analyzed for the presence of full length SRS1-6 domains and the four Cyp51 motifs. Of these, 439 proteins had SRS1-6 domains, the oxygen-binding motif AGXDTT, the PER and EXXR motifs, and the conserved heme-binding motif FXXGXXXCXG (Supplemental Table 1) and were considered to be functional Cyp51 proteins. Within the 439 Cyp51 proteins, we found 19 fusion proteins in various Ascomycota in which a kinase immediately upstream and Cyp51B were mistakenly fused in genome processing. According to the Joint Genome Institute (JGI), a frameshift mutation occurred *in silico* causing the stop codon of the kinase to appear to be deleted (https://img.jgi.doe.gov/data-processing.html). The Cyp51B portions of the fusion proteins were extracted based on their Cyp51 domains and kept in the analyses. Four Cyp51B proteins (XP_007688940.1, XP_014555703.1, XP_033384229.1, XP_018700143.1) were not included in the analyses due to missing amino acids in SRS1-6 or missing amino acids in conserved motifs resulting in a total of 435 Cyp51 proteins retained for further analysis.

Figure 1: Typical organization of Cyp51 domains. Cyp51 proteins contain six substrate recognition sites (SRS 1-6), an oxygen binding motif (AGXDTT), PER and EXXR motifs, and a conserved heme-binding motif (FXXGXXXCXG). Black boxes represent SRS domains. Gray shading represents other motifs. Diagram is based on *A. fumigatus* Cyp51A (XP_752137.1) and is shown to scale.

Fungal Cyp51 proteins fall into 4 groups. To understand presence of Cyp51 homologs across Fungi, we analyzed 435 Cyp51 proteins (Supplemental Table 2). Most fungal species in our study had one or two Cyp51 paralogs (180/295 and 100/295, respectively) (Table 1). Fewer had three (14/295) and only one fungus, *Basidiobolus meristosporus*, had four copies of Cyp51 (Table 1). Fungi with one or two Cyp51 proteins were found across all taxonomic groups (Table 1). Fungi with three Cyp51 proteins were found in Basidiomycota and Ascomycota. One species in Zoopagomycota had four Cyp51 proteins (Table 1). Members in Pezizomycotina had different combinations of Cyp51 paralogs (Table 2). Most species had Cyp51A and Cyp51B paralogs or only Cyp51B (69/171 and 63/171, respectively) (Table 2). All members of Pezizomycotina contained a Cyp51B paralog (Table 2). As shown in Supplemental Table 2, we named fungal Cyp51 proteins without a designation in NCBI based on the group assignment in Supplemental Figure 1.

Major Fungal Taxonomic Clades[1]			Number of species with various copy numbers of Cyp51 Proteins[2]			
			1	2	3	4
Blastocladiomycota			-	1	-	-
Chytridiomycota			4	-	-	-
Zoopagomycota			9	1	-	1
Mucoromycota			3	1	-	-
Basidiomycota	Ustilaginomycotina		15	-	-	-
	Pucciniomycotina		3	-	-	-
	Agaricomycotina		25	5	1	-
	Wallemiomycotina		1	1	-	-
Ascomycota	Taphrinomycotina		7	-	-	-
	Saccharomycotina		48	2	-	-
	Pezizomycotina	Orbiliomycetes	-	1	-	-
		Xylonomycetes	-	1	-	-
		Eurotiomycetes	15	61	4	-
		Lecanoromycetes	-	1	-	-
		Dothideomycetes	19	5	1	-
		Leotiomycetes	9	1	-	
		Sordariomycetes	21	19	8	-
# Species/Total # Species			180/295 (61%)	100/295 (33.9%)	14/295 (4.7%)	1/295 (0.3%)

Table 1. Species of Fungi have different numbers of Cyp51 proteins. Orange, red, black, and blue boxes represent the presence of Cyp51, Cyp51A, Cyp51B, and Cyp51C within fungal groups, respectively.

[1]Major fungal taxonomic clades based on James et al 2020 [36].
[2]Based on all fungal Cyp51 proteins identified in NCBI databases (435) as described in methods and shown in Supplemental Table 2.

Table 2. Species in Pezizomycotina have different combinations of Cyp51 paralogs.

Cyp51 Paralog	B	A+B	B+B	B+C	A+B+C	A+A+B	A+B+B	B+C+C
# Species/Total # Species	63/171	69/171	4/171	18/171	9/171	5/171	1/171	1/171

To investigate how fungal Cyp51 proteins are related to each other, we used RAxML to create a maximum likelihood tree with 435 fungal Cyp51 proteins and 2 human Cyp51 proteins to root the tree (Supplemental Figure 1). To aid in visualization, we collapsed branches based on phyla, subphyla or classes (Figure 2). We found fungal Cyp51 proteins fell into four major groups which we designated as Cyp51, Cyp51A, Cyp51B, and Cyp51C based on naming in previous literature (Figure 2, Supplemental Figure 1). Cyp51 in members of Saccharomycotina and Taphrinomycotina are also known as "Erg11" in the literature. The topology of our Cyp51 protein tree largely followed the topology of the fungal tree of life (Figure 2, Supplemental Figure 1, [36]). Proteins from early-diverging fungi (Blastocladiomycota, Chytridiomycota, Zoopagomycota and Mucormycota), Basidiomycota, Saccharomycotina, and Taphrinomycotina fell into group Cyp51. Our phylogenetic analyses show a divergence of three Cyp51 paralogs in filamentous Ascomycota (Figure 2, Supplemental Figure 1). Those from Pezizomycotina fell into Cyp51A, Cyp51B, and Cyp51C. Divergence of Cyp51 from the common ancestor of paralogs Cyp51A, Cyp51B and Cyp51C has strong support (100% bootstrap support), but divergence of paralogs Cyp51A, Cyp51B and Cyp51C from each other does not have strong support (41%, 36%, and 41% bootstrap support, respectively).

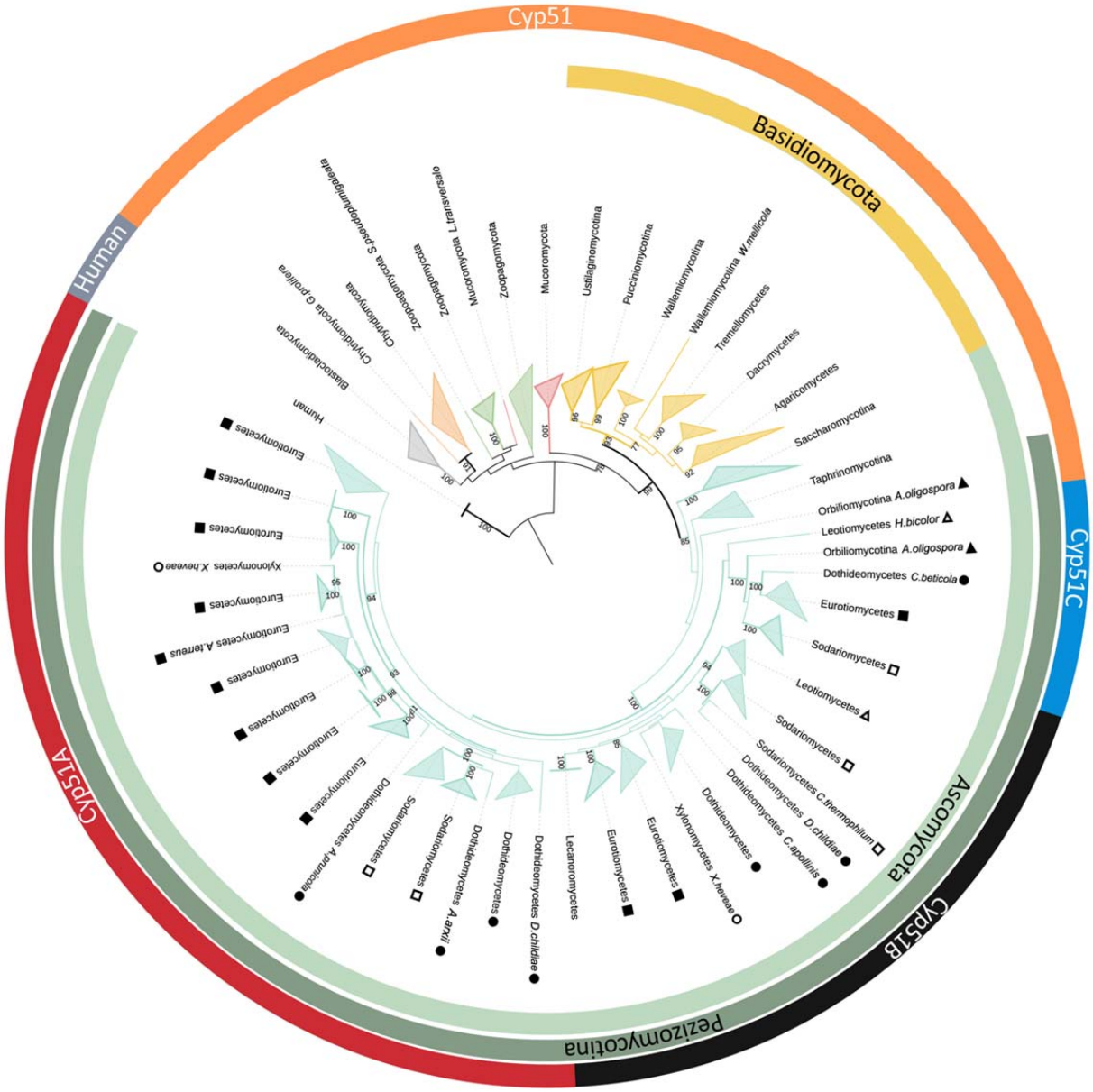

Figure 2: Collapsed Cyp51 protein tree for Fungi. Branch colors match the colors used for taxonomic clades in the Fungal Tree of Life [36]. Branches for Blastocladiomycota, Chytridiomycota, Zoopagomycota, Mucoromycota, Basidiomycota, and Ascomycota are represented by gray, orange, green, red, yellow, and teal, respectively. Branches with bootstrap support of at least 90 are in bold. Collapsed branches represent phyla, subphyla and classes and are named accordingly. Shapes represent subphyla and classes in Ascomycota. Filled triangles represent subphylum Orbiliomycotina. Empty triangles, filled circles, empty circles, filled squares, and empty squares represent classes Leotiomycetes, Dothideomycetes, Xylonomycetes, Eurotiomycetes, and Sodariomycetes, respectively. The most inner ring shows phyla Basidiomycota (yellow) and Ascomycota (teal). The second ring shows subphylum Pezizomycotina (dark teal). The outer ring shows human Cyp51 proteins used as the outgroup (grey) and 4 groups of fungal Cyp51 proteins – Cyp51, Cyp51C, Cyp51B, and Cyp51A represented by orange, blue, black and red, respectively.

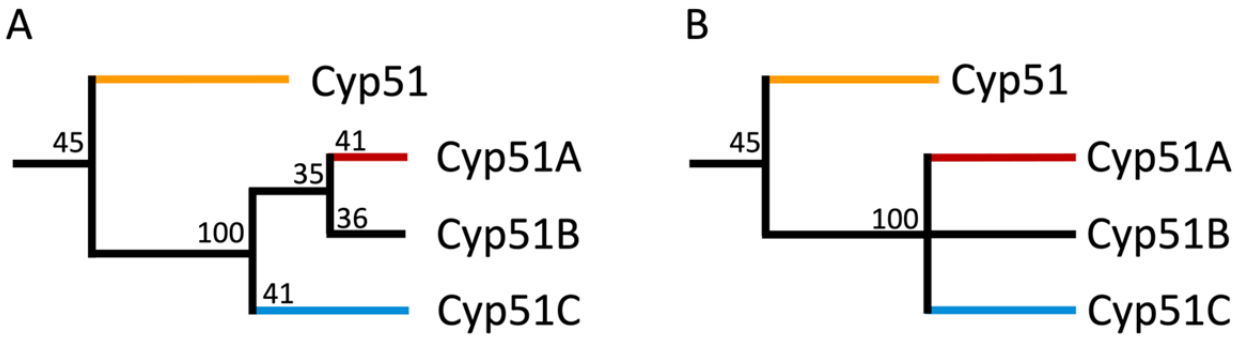

Figure 3. Possible Cyp51 homolog evolutionary paths. Simplified diagrams of possible Cyp51 evolutionary paths based on Figure 2. The Cyp51 branch in early diverging fungi, Basidiomycota, Saccharomycotina and Taphrinomycotina is colored orange. Filamentous Ascomycota (Pezizomycotina) Cyp51A, Cyp51B, and Cyp51C branches are colored red, black and blue, respectively. Numbers represent bootstrap support.

We postulate two possible evolutionary paths for Cyp51 paralogs as shown in Figure 3. In the first possible evolutionary path, shown in 3A, after an initial Cyp51 duplication paralog C diverged followed by another duplication event and divergence of paralogs A and B. The divergence of paralogs A (41%), B (36%), and C (41%) has low support. In the second possible evolutionary path, shown in 3B, the poorly supported nodes are removed so that the three paralogs diverged after two unresolved duplication events or a triplication event placing them on the same branch. In either scenario it is possible that subsequent gene loss(es) or duplication(s) led to species with the different combinations of paralogs shown in Table 2.

Distinguishing between these possible evolutionary paths is complicated by the small number of Cyp51C sequences and the relatively low number of characters (414-624 amino acids in Cyp51 genes) resulting in low bootstrap support for some nodes. Our analysis only had 29 Cyp51C sequences compared to Cyp51A and B with 87 and 171 sequences, respectively. To see if we could better resolve the relationships among Cyp51A, Cyp51B, and Cyp51C, we analyzed pairwise conservation of all Cyp51 protein sequences using Geneious Prime (Supplemental Table 3). We found that individual members of the Cyp51 group varied the most from each other, with only 46% similarity. Similarity within the Cyp51A, Cyp51B, and Cyp51C groups was much higher (64.7%-68.8%). Comparing between groups, members of the Cyp51 group were 45-50% similar to members of Cyp51A, Cyp51B, or Cyp51C groups while members of Cyp51A, Cyp51B, and Cyp51C groups were roughly 60% similar to each other. , (Supplemental Table 3). We then examined conservation within the highly conserved motifs (SRS 1-6, AGXDTT, PER, EXXR and FXXGXXXCXG) (Supplemental Table 4). Once more the general trend was that motifs within the Cyp51 group were more variable that those in other groups. .

Consensus sequences from the four Cyp51 groups were compared to each other to create a visual representation of differences in motifs across groups (Figure 4, Supplemental Table 4). The most conserved regions (>95% across all Cyp51 proteins) are in the EXXR, PER, and FXXGXXXCIG motifs (Figure 4, Supplemental Table 4). These three motifs are presumably highly conserved due to their roles in Cyp51 structure and function. The **E**XX**R** and PE**R** motif form the **E**-**R**-**R** triad that stabilizes the core structure of Cyp51, while the FXXGXXXCIG motif

is a heme-binding domain that is essential for Cyp51 function [37, 38]. Out of the 17 amino acids within these three motifs, 13 are more than 95% conserved (Figure 4, Supplemental Table 4). Looking at all motifs, there are 59 amino acids shared between the four consensus sequences with 39 having greater than 95% conservation (Figure 4, Supplemental Table 4). Cyp51A and Cyp51B have the highest amount of shared amino acids in motifs (88%) (Figure 4, Supplemental Table 4). Cyp51C has 23 amino acids across motifs that are unique compared to 15 in Cyp51, 8 in Cyp51A, and 5 in Cyp51B, suggesting Cyp51C may be a specialized group, though the low number of Cyp51C sequences included in the analysis might also explain the pattern (Figure 4).

Figure 4. Conservation of motifs in Cyp51 proteins across Fungi. Consensus sequences for each of the four Cyp51A groups are shown arranged from the N- to C-terminus. Amino acids conserved across all 435 Cyp51 protein sequences are denoted by an *. Amino acids found in

more than 95% of all Cyp51s are denoted by >. Lower than 95% conservation in all Cyp51s are denoted by #. Numbers represent amino acid position within motifs.

Research on Cyp51 in filamentous Ascomycota tends to be focused on the Cyp51A paralog because of its role in azole resistance in fungal pathogens of plants and animals. To our surprise, we found Cyp51A in only half of the species of filamentous Ascomycota we analyzed (86/171) and Cyp51B in all species of filamentous Ascomycota (Table 2, Supplemental Table 2). The low bootstrap support we saw for divergence of Cyp51A from Cyp51B suggests they may play very similar roles with Cyp51B being the essential paralog and the possible Cyp51 ortholog (Figure 3A). Indeed, in *A. fumigatus*, Cyp51A and Cyp51B act in a compensatory manner; when one paralog is knocked out expression of the other paralog increases [39]. Our results confirm and expand previous work by Hawkins et al [28] that compared 86 fungal Cyp51 proteins in 50 different species and found all species of filamentous Ascomycota retained a Cyp51B paralog, but Cyp51A had been lost in multiple lineages, and Cyp51C was only found in *Fusarium* spp. [40-42]. [40, 41].). Cyp51C was subsequently reported in *Fusarium spp.*, *Gibberella zeae*, and *Nectria haematoccoa* [40, 41]. We found Cyp51C in 9 other genera (Supplemental Table 2). Interestingly, all are pathogens of plants or animals (Supplemental Table 2). Cyp51C has been shown to be necessary for invasion of plant host tissues in *Fusarium* [43]. This raises the interesting possibility that it could play a similar role in other genera, though further functional studies are needed to test the role of Cyp51C in invasion and virulence.

Maximum likelihood analysis suggested two possible evolutionary paths for Cyp51 paralogs:1) Cyp51C diverged before Cyp51A and Cyp51B (Figure 3A; 2) Cyp51A, Cyp51B, and Cyp51C diverged from each other at the same time (Figure 3B). Based on the higher amino acid conservation between Cyp51A and Cyp51B, the evolutionary path shown in Figure 3A seems more likely; that is to say Cyp51C is a specialized group of proteins that diverged first (Figure 3A, Figure 4, Supplemental Table 4). More functional studies are needed to understand shared and unique roles Cyp51A, Cyp51B, and Cyp51C paralogs play in fungi.

Materials and Methods

NCBI Protein Blast

Cyp51 protein sequences (XP_006955920.1, XP_016612382.1, XP_025190100.1, XP_01661282.1, XP_018263285.1, XP_752137.1, XP_018249823.1, XP_018226934.1, XP_015469227.1) were used in an NCBI Protein Blast to search the refence sequence database for other Cyp51 proteins in Fungi (Supplemental Table 1). The following setting were used for the Protein BLAST: Database: Reference Proteins, Exclude: uncultured/environmental sample sequences, Algorithm: blastp (protein-protein BLAST), Max Target Sequences: 1000, Expect Threshold: 0.001, Word size: 6, Max matches in a query range: 0, Matrix: BLOSUM62, Gap Costs: Existence: 11 Extension: 1, and Compositional adjustments: Conditional compositional score matrix adjustment. Some clades were not represented in the reference sequence database, so identical settings were used to search the non-redundant proteins sequences database. The

unfiltered searches resulted in a total of 4404 sequence hits. Sequences with less than 50% coverage and less than 30% percent identity were eliminated. FASTA files of the 480 sequences resulting from filtering were downloaded and opened in Geneious Prime 2019.1.1. To confirm the sequences were a Cyp51, sequences were checked for the presence of SRS1-6, the oxygen-binding motif AGXDTT, PER and EXXR motifs, and the conserved FXXGXXXCXG heme-binding domain.
The following sequences were eliminated for missing amino acids in SRS1-6 or missing amino acids in conserved motifs: EPZ30787.1, EPZ31936.1, KXN68292.1, RKP07181.1, RKP16598.1, RKP17874.1, RKP18653.1, RKP18926.1, XP_001218650.1, XP_002563403.1, XP_002583031.1, XP_002842283.1, XP_003005233.1, XP_003325369.2, XP_007375289.1, XP_007756389.1, XP_007802603.1, XP_008039623.1, XP_009649122.1, XP_013258864.1, XP_015404015.1, XP_018230821.1, XP_018249826.1, XP_018270027.1, XP_018712692.1, XP_020066776.1, XP_022578172.1, XP_025599710.1, XP_027619241.1, XP_027619242.1, XP_031034290.1, XP_031059536.1, ORZ32486.1, XP_017991977.1, XP_003017020.1, XP_003019064.1, and XP_033461214.1. The following 19 fusion proteins were found in various members in Ascomycota: XP_022511803.1, XP_013278994.1, XP_024670717.1, XP_031899359.1, XP_031927262.1, XP_026621438.1, XP_025554838.1, XP_018192447.1, XP_031935516.1, XP_025433272.1, XP_015404994.1, XP_024709601.1, XP_007688940.1, XP_014073145.1, XP_025394250.1, XP_014555703.1, XP_007712864.1, XP_033384229.1, XP_018700143.1. Four fusions (XP_007688940.1, XP_014555703.1, XP_033384229.1, XP_018700143.1) were not included in the analyses due to missing amino acids in SRS1-6 or missing amino acids in conserved motifs. *Aspergillus flavus* contained three Cyp51 proteins, but were later identified to be contaminated with foreign sequence by NCBI (https://www.ncbi.nlm.nih.gov/protein/XP_002375123.1/) and were removed (XP_002375123.1 , XP_002379130.1, XP_002383931.1).

Phylogenetic analyses

Protein sequences were aligned once with MAFFT version 7.407 then once with PASTA version 1.8.5 [44, 45]. Maximum likelihood trees were constructed with RAxML version 8.2.11 with a PROTGAMMAAUTO or GTRGAMMA substitution model and 1000 bootstraps [46]. Interactive Tree of Life (iTOL) was used for visualization and annotation of the trees [47].

Amino Acid Analyses

Geneious Prime (version 2021.2.2) was used to generate pairwise identities and consensus sequences. Similarity tables for the whole protein and motifs are based on pairwise identity in each group, number of similar amino acids divided by the total number of amino acids in the protein or motif. “Weblogo-like” diagrams were created manually for visualization of conservation across groups. The height of one letter amino acid designation was based on frequency across all four consensus sequences. Colors and symbols were used as described in Figure 4 legend to denote conservation within groups.

Data availability

All Cyp51 sequences used are listed in Supplementary Table 2 and are publicly available through NCBI (https://www.ncbi.nlm.nih.gov/).

Funding

This work was supported by the Centers for Disease Control and Prevention (CDC; contract 200-2017-96199 to M.M. and M.T.B.) and United States Department of Agriculture, National Institute of Food and Agriculture (USDA NIFA AFRI grant 2019-67017-29113 to M.T.B. and M.M.). B.N.C-S. was also supported by the National Science Foundation under Grant No. DGE-1545433.

Conflicts of interest

The authors declare that there is no conflict of interest.

References

1. Benedict K, Jackson BR, Chiller T, Beer KD. Estimation of Direct Healthcare Costs of Fungal Diseases in the United States. Clin Infect Dis. 2019;68(11):1791-7. Epub 2018/09/12. doi: 10.1093/cid/ciy776. PubMed PMID: 30204844; PubMed Central PMCID: PMCPMC6409199.
2. Vallabhaneni S, Mody RK, Walker T, Chiller T. The Global Burden of Fungal Diseases. Infect Dis Clin North Am. 2016;30(1):1-11. Epub 2016/01/08. doi: 10.1016/j.idc.2015.10.004. PubMed PMID: 26739604.
3. Perlin DS, Rautemaa-Richardson R, Alastruey-Izquierdo A. The global problem of antifungal resistance: prevalence, mechanisms, and management. Lancet Infect Dis. 2017;17(12):e383-e92. Epub 2017/08/05. doi: 10.1016/S1473-3099(17)30316-X. PubMed PMID: 28774698.
4. Brown GD, Denning DW, Gow NA, Levitz SM, Netea MG, White TC. Hidden killers: human fungal infections. Sci Transl Med. 2012;4(165):165rv13. doi: 10.1126/scitranslmed.3004404. PubMed PMID: 23253612.
5. Pound MW, Townsend ML, Dimondi V, Wilson D, Drew RH. Overview of treatment options for invasive fungal infections. Med Mycol. 2011;49(6):561-80. Epub 2011/03/04. doi: 10.3109/13693786.2011.560197. PubMed PMID: 21366509.
6. Lepesheva GI, Waterman MR. Structural basis for conservation in the CYP51 family. Biochim Biophys Acta. 2011;1814(1):88-93. Epub 2010/06/16. doi: 10.1016/j.bbapap.2010.06.006. PubMed PMID: 20547249; PubMed Central PMCID: PMCPMC2962772.
7. Lepesheva GI, Waterman MR. Sterol 14alpha-demethylase cytochrome P450 (CYP51), a P450 in all biological kingdoms. Biochim Biophys Acta. 2007;1770(3):467-77. Epub 2006/09/12. doi: 10.1016/j.bbagen.2006.07.018. PubMed PMID: 16963187; PubMed Central PMCID: PMCPMC2324071.
8. Cordova P, Gonzalez AM, Nelson DR, Gutierrez MS, Baeza M, Cifuentes V, et al. Characterization of the cytochrome P450 monooxygenase genes (P450ome) from the carotenogenic yeast Xanthophyllomyces dendrorhous. BMC Genomics. 2017;18(1):540. Epub 2017/07/21. doi: 10.1186/s12864-017-3942-9. PubMed PMID: 28724407; PubMed Central PMCID: PMCPMC5516332.
9. Jefcoate CR, Gaylor JL, Calabrese RL. Ligand interactions with cytochrome P-450. I. Binding of primary amines. Biochemistry. 1969;8(8):3455-63. Epub 1969/08/01. doi: 10.1021/bi00836a049. PubMed PMID: 5809236.

373 10. Dufourc EJ. Sterols and membrane dynamics. J Chem Biol. 2008;1(1-4):63-77. Epub
374 2009/07/02. doi: 10.1007/s12154-008-0010-6. PubMed PMID: 19568799; PubMed Central
375 PMCID: PMCPMC2698314.
376 11. Schaller H. The role of sterols in plant growth and development. Prog Lipid Res.
377 2003;42(3):163-75. Epub 2003/04/12. doi: 10.1016/s0163-7827(02)00047-4. PubMed PMID:
378 12689617.
379 12. Cowen LE, Steinbach WJ. Stress, drugs, and evolution: the role of cellular signaling in
380 fungal drug resistance. Eukaryot Cell. 2008;7(5):747-64. Epub 2008/04/01. doi:
381 10.1128/EC.00041-08. PubMed PMID: 18375617; PubMed Central PMCID:
382 PMCPMC2394973.
383 13. Flowers SA, Colon B, Whaley SG, Schuler MA, Rogers PD. Contribution of clinically
384 derived mutations in ERG11 to azole resistance in Candida albicans. Antimicrob Agents
385 Chemother. 2015;59(1):450-60. Epub 2014/11/12. doi: 10.1128/AAC.03470-14. PubMed PMID:
386 25385095; PubMed Central PMCID: PMCPMC4291385.
387 14. Rodero L, Mellado E, Rodriguez AC, Salve A, Guelfand L, Cahn P, et al. G484S amino
388 acid substitution in lanosterol 14-alpha demethylase (ERG11) is related to fluconazole resistance
389 in a recurrent Cryptococcus neoformans clinical isolate. Antimicrob Agents Chemother.
390 2003;47(11):3653-6. Epub 2003/10/25. doi: 10.1128/AAC.47.11.3653-3656.2003. PubMed
391 PMID: 14576140; PubMed Central PMCID: PMCPMC253790.
392 15. Sionov E, Chang YC, Garraffo HM, Dolan MA, Ghannoum MA, Kwon-Chung KJ.
393 Identification of a Cryptococcus neoformans cytochrome P450 lanosterol 14alpha-demethylase
394 (Erg11) residue critical for differential susceptibility between fluconazole/voriconazole and
395 itraconazole/posaconazole. Antimicrob Agents Chemother. 2012;56(3):1162-9. Epub
396 2011/12/14. doi: 10.1128/AAC.05502-11. PubMed PMID: 22155829; PubMed Central PMCID:
397 PMCPMC3294891.
398 16. Wheat LJ, Connolly P, Smedema M, Durkin M, Brizendine E, Mann P, et al. Activity of
399 newer triazoles against Histoplasma capsulatum from patients with AIDS who failed
400 fluconazole. J Antimicrob Chemother. 2006;57(6):1235-9. Epub 2006/04/22. doi:
401 10.1093/jac/dkl133. PubMed PMID: 16627592.
402 17. Lockhart SR, Frade JP, Etienne KA, Pfaller MA, Diekema DJ, Balajee SA. Azole
403 resistance in Aspergillus fumigatus isolates from the ARTEMIS global surveillance study is
404 primarily due to the TR/L98H mutation in the cyp51A gene. Antimicrob Agents Chemother.
405 2011;55(9):4465-8. Epub 2011/06/22. doi: 10.1128/AAC.00185-11. PubMed PMID: 21690285;
406 PubMed Central PMCID: PMCPMC3165364.
407 18. Snelders E, van der Lee HA, Kuijpers J, Rijs AJ, Varga J, Samson RA, et al. Emergence
408 of azole resistance in Aspergillus fumigatus and spread of a single resistance mechanism. PLoS
409 Med. 2008;5(11):e219. Epub 2008/11/13. doi: 10.1371/journal.pmed.0050219. PubMed PMID:
410 18998768; PubMed Central PMCID: PMCPMC2581623.
411 19. Snelders E, Karawajczyk A, Schaftenaar G, Verweij PE, Melchers WJ. Azole resistance
412 profile of amino acid changes in Aspergillus fumigatus CYP51A based on protein homology
413 modeling. Antimicrob Agents Chemother. 2010;54(6):2425-30. Epub 2010/04/14. doi:
414 10.1128/AAC.01599-09. PubMed PMID: 20385860; PubMed Central PMCID:
415 PMCPMC2876375.
416 20. Hodiamont CJ, Dolman KM, Ten Berge IJ, Melchers WJ, Verweij PE, Pajkrt D.
417 Multiple-azole-resistant Aspergillus fumigatus osteomyelitis in a patient with chronic
418 granulomatous disease successfully treated with long-term oral posaconazole and surgery. Med

Mycol. 2009;47(2):217-20. Epub 2008/12/23. doi: 10.1080/13693780802545600. PubMed PMID: 19101840.

21. Hare RK, Gertsen JB, Astvad KMT, Degn KB, Lokke A, Stegger M, et al. In Vivo Selection of a Unique Tandem Repeat Mediated Azole Resistance Mechanism (TR120) in Aspergillus fumigatus cyp51A, Denmark. Emerg Infect Dis. 2019;25(3):577-80. Epub 2019/02/23. doi: 10.3201/eid2503.180297. PubMed PMID: 30789127; PubMed Central PMCID: PMCPMC6390761.

22. Gsaller F, Hortschansky P, Furukawa T, Carr PD, Rash B, Capilla J, et al. Sterol Biosynthesis and Azole Tolerance Is Governed by the Opposing Actions of SrbA and the CCAAT Binding Complex. PLoS Pathog. 2016;12(7):e1005775. Epub 2016/07/22. doi: 10.1371/journal.ppat.1005775. PubMed PMID: 27438727; PubMed Central PMCID: PMCPMC4954732.

23. Buied A, Moore CB, Denning DW, Bowyer P. High-level expression of cyp51B in azole-resistant clinical Aspergillus fumigatus isolates. J Antimicrob Chemother. 2013;68(3):512-4. Epub 2012/12/05. doi: 10.1093/jac/dks451. PubMed PMID: 23208831.

24. Gonzalez-Jimenez I, Lucio J, Amich J, Cuesta I, Sanchez Arroyo R, Alcazar-Fuoli L, et al. A Cyp51B Mutation Contributes to Azole Resistance in Aspergillus fumigatus. J Fungi (Basel). 2020;6(4). Epub 2020/12/02. doi: 10.3390/jof6040315. PubMed PMID: 33255951; PubMed Central PMCID: PMCPMC7712412.

25. Ghosoph JM, Schmidt LS, Margosan DA, Smilanick JL. Imazalil resistance linked to a unique insertion sequence in the PdCYP51 promoter region of Penicillium digitatum. Postharvest Biol Tec. 2007;44(1):9-18. doi: 10.1016/j.postharvbio.2006.11.008. PubMed PMID: WOS:000245490200002.

26. Delye C, Laigret F, CorioCostet MF. A mutation in the 14 alpha-demethylase gene of Uncinula necator that correlates with resistance to a sterol biosynthesis inhibitor. Appl Environ Microb. 1997;63(8):2966-70. PubMed PMID: WOS:A1997XP06700002.

27. Gadoury DM, Cadle-Davidson L, Wilcox WF, Dry IB, Seem RC, Milgroom MG. Grapevine powdery mildew (Erysiphe necator): a fascinating system for the study of the biology, ecology and epidemiology of an obligate biotroph. Mol Plant Pathol. 2012;13(1):1-16. doi: 10.1111/j.1364-3703.2011.00728.x. PubMed PMID: WOS:000298088200001.

28. Stammler G, Cordero J, Koch A, Semar M, Schlehuber S. Role of the Y134F mutation in cyp51 and overexpression of cyp51 in the sensitivity response of Puccinia triticina to epoxiconazole. Crop Prot. 2009;28(10):891-7. doi: 10.1016/j.cropro.2009.05.007. PubMed PMID: WOS:000270123900015.

29. Canas-Gutierrez GP, Angarita-Velasquez MJ, Restrepo-Florez JM, Rodriguez P, Moreno CX, Arango R. Analysis of the CYP51 gene and encoded protein in propiconazole-resistant isolates of Mycosphaerella fijiensis. Pest Manag Sci. 2009;65(8):892-9. doi: 10.1002/ps.1770. PubMed PMID: WOS:000268442600010.

30. Cools HJ, Fraaije BA. Update on mechanisms of azole resistance in Mycosphaerella graminicola and implications for future control. Pest Manag Sci. 2013;69(2):150-5. Epub 2012/06/26. doi: 10.1002/ps.3348. PubMed PMID: 22730104.

31. Cools HJ, Bayon C, Atkins S, Lucas JA, Fraaije BA. Overexpression of the sterol 14alpha-demethylase gene (MgCYP51) in Mycosphaerella graminicola isolates confers a novel azole fungicide sensitivity phenotype. Pest Manag Sci. 2012;68(7):1034-40. Epub 2012/03/14. doi: 10.1002/ps.3263. PubMed PMID: 22411894.

464 32. Sun XP, Xu Q, Ruan RX, Zhang TY, Zhu CY, Li HY. PdMLE1, a specific and active
465 transposon acts as a promoter and confers Penicillium digitatum with DMI resistance. Env
466 Microbiol Rep. 2013;5(1):135-42. doi: 10.1111/1758-2229.12012. PubMed PMID:
467 WOS:000314474500016.
468 33. Schnabel G, Jones AL. The 14 alpha-demethylase (CYP51A1) gene is overexpressed in
469 Venturia inaequalis strains resistant to myclobutanil. Phytopathology. 2001;91(1):102-10. doi:
470 Doi 10.1094/Phyto.2001.91.1.102. PubMed PMID: WOS:000166027300011.
471 34. Pfeufer EE, Ngugi HK. Orchard Factors Associated with Resistance and Cross
472 Resistance to Sterol Demethylation Inhibitor Fungicides in Populations of Venturia inaequalis
473 from Pennsylvania. Phytopathology. 2012;102(3):272-82. doi: 10.1094/Phyto-04-11-0117.
474 PubMed PMID: WOS:000300044100005.
475 35. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search
476 tool. J Mol Biol. 1990;215(3):403-10. Epub 1990/10/05. doi: 10.1016/S0022-2836(05)80360-2.
477 PubMed PMID: 2231712.
478 36. James TY, Stajich JE, Hittinger CT, Rokas A. Toward a Fully Resolved Fungal Tree of
479 Life. Annu Rev Microbiol. 2020;74:291-313. Epub 2020/07/15. doi: 10.1146/annurev-micro-
480 022020-051835. PubMed PMID: 32660385.
481 37. Sezutsu H, Le Goff G, Feyereisen R. Origins of P450 diversity. Philos T R Soc B.
482 2013;368(1612). doi: ARTN 20120428
483 10.1098/rstb.2012.0428. PubMed PMID: WOS:000313343600002.
484 38. Deng JX, Carbone I, Dean RA. The evolutionary history of Cytochrome P450 genes in
485 four filamentous Ascomycetes. Bmc Evol Biol. 2007;7. doi: Artn 30
486 10.1186/1471-2148-7-30. PubMed PMID: WOS:000244937100001.
487 39. Roundtree MT, Juvvadi PR, Shwab EK, Cole DC, Steinbach WJ. Aspergillus fumigatus
488 Cyp51A and Cyp51B Proteins Are Compensatory in Function and Localize Differentially in
489 Response to Antifungals and Cell Wall Inhibitors. Antimicrob Agents Chemother. 2020;64(10).
490 Epub 2020/07/15. doi: 10.1128/AAC.00735-20. PubMed PMID: 32660997; PubMed Central
491 PMCID: PMCPMC7508596.
492 40. Becher R, Weihmann F, Deising HB, Wirsel SG. Development of a novel multiplex
493 DNA microarray for Fusarium graminearum and analysis of azole fungicide responses. BMC
494 Genomics. 2011;12:52. Epub 2011/01/25. doi: 10.1186/1471-2164-12-52. PubMed PMID:
495 21255412; PubMed Central PMCID: PMCPMC3037902.
496 41. Hawkins NJ, Cools HJ, Sierotzki H, Shaw MW, Knogge W, Kelly SL, et al. Paralog re-
497 emergence: a novel, historically contingent mechanism in the evolution of antimicrobial
498 resistance. Mol Biol Evol. 2014;31(7):1793-802. Epub 2014/04/16. doi:
499 10.1093/molbev/msu134. PubMed PMID: 24732957; PubMed Central PMCID:
500 PMCPMC4069618.
501 42. James JE, Lamping E, Santhanam J, Milne TJ, Abd Razak MF, Zakaria L, et al. A 23 bp
502 cyp51A Promoter Deletion Associated With Voriconazole Resistance in Clinical and
503 Environmental Isolates of Neocosmospora keratoplastica. Front Microbiol. 2020;11:272. Epub
504 2020/04/17. doi: 10.3389/fmicb.2020.00272. PubMed PMID: 32296397; PubMed Central
505 PMCID: PMCPMC7136401.
506 43. Fan J, Urban M, Parker JE, Brewer HC, Kelly SL, Hammond-Kosack KE, et al.
507 Characterization of the sterol 14alpha-demethylases of Fusarium graminearum identifies a novel
508 genus-specific CYP51 function. New Phytol. 2013;198(3):821-35. Epub 2013/02/28. doi:
509 10.1111/nph.12193. PubMed PMID: 23442154.

510 44. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7:
511 improvements in performance and usability. Mol Biol Evol. 2013;30(4):772-80. Epub
512 2013/01/19. doi: 10.1093/molbev/mst010. PubMed PMID: 23329690; PubMed Central PMCID:
513 PMCPMC3603318.
514 45. Mirarab S, Nguyen N, Guo S, Wang LS, Kim J, Warnow T. PASTA: Ultra-Large
515 Multiple Sequence Alignment for Nucleotide and Amino-Acid Sequences. J Comput Biol.
516 2015;22(5):377-86. Epub 2014/12/31. doi: 10.1089/cmb.2014.0156. PubMed PMID: 25549288;
517 PubMed Central PMCID: PMCPMC4424971.
518 46. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of
519 large phylogenies. Bioinformatics. 2014;30(9):1312-3. Epub 2014/01/24. doi:
520 10.1093/bioinformatics/btu033. PubMed PMID: 24451623; PubMed Central PMCID:
521 PMCPMC3998144.
522 47. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new
523 developments. Nucleic Acids Res. 2019;47(W1):W256-W9. Epub 2019/04/02. doi:
524 10.1093/nar/gkz239. PubMed PMID: 30931475; PubMed Central PMCID: PMCPMC6602468.
525