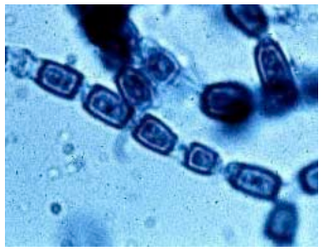
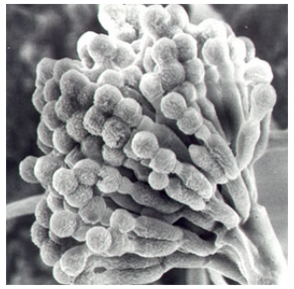
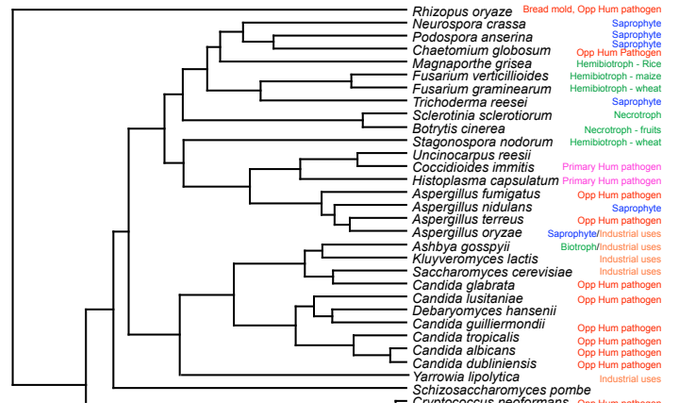
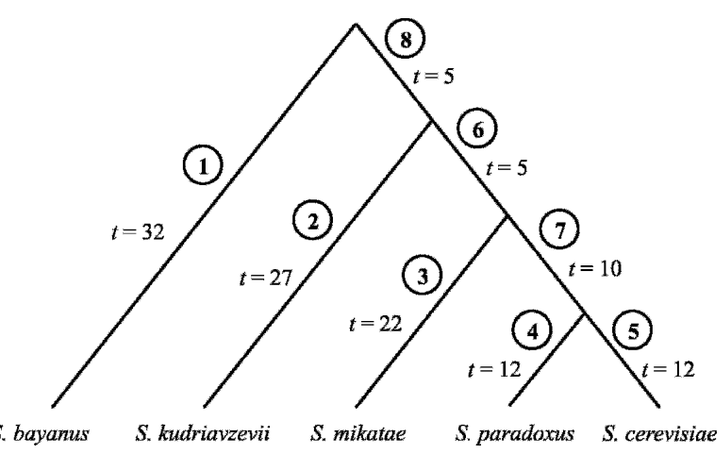


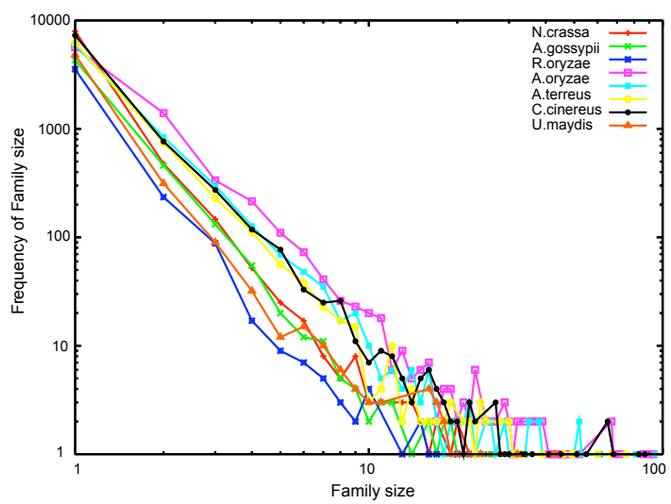





Evolution of gene family size change in fungi

Jason Stajich
University of California, Berkeley



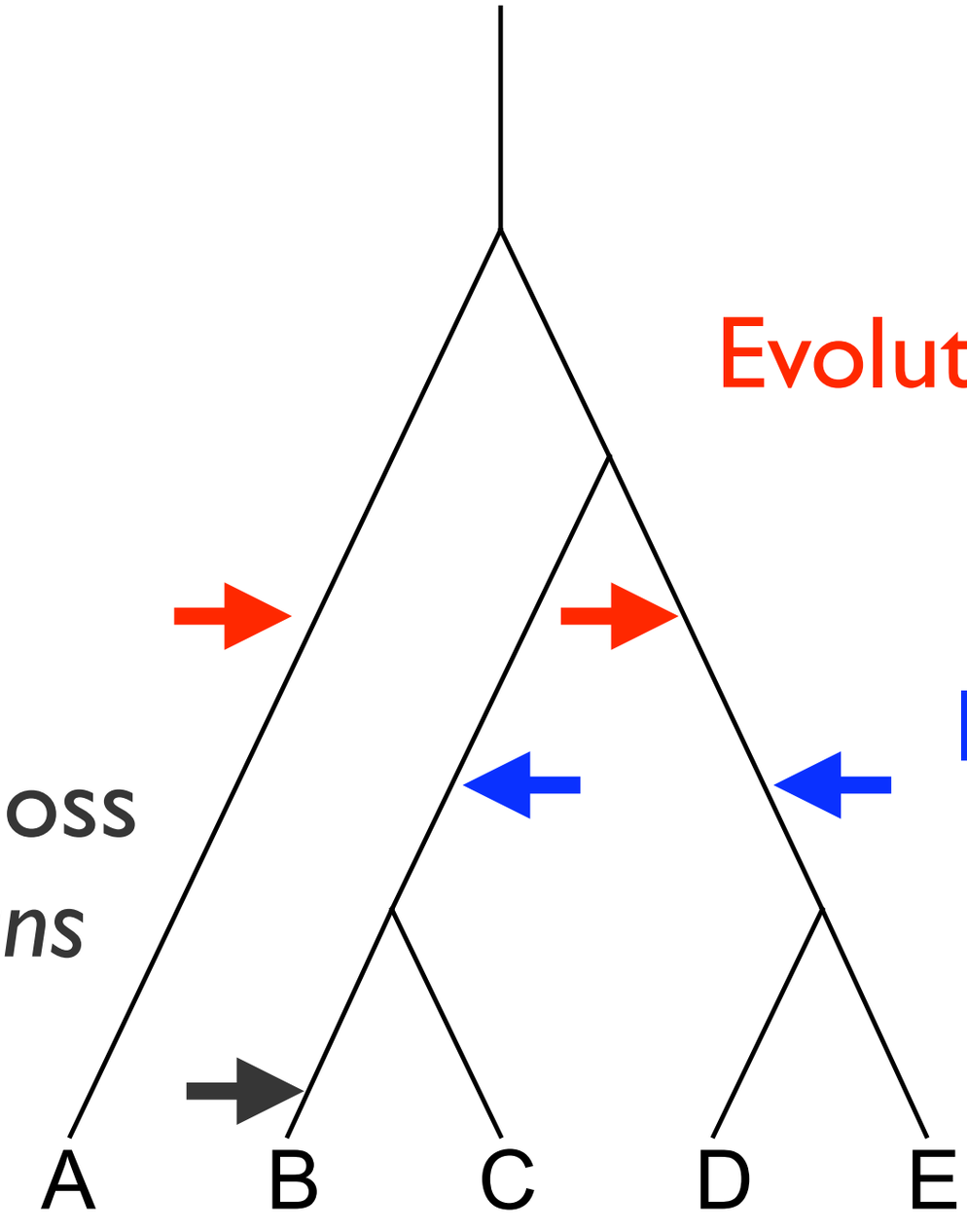


Outline

- Motivation
- Model of Gene family size change
- Cornucopia of fungal genomes
- Methodology for comparing family size
- Lineage specific expansions in several groups of fungi

Fungal comparative genomics

Evolution of fungal introns



Recent intron loss in *C. neoformans*

Fungal gene family evolution

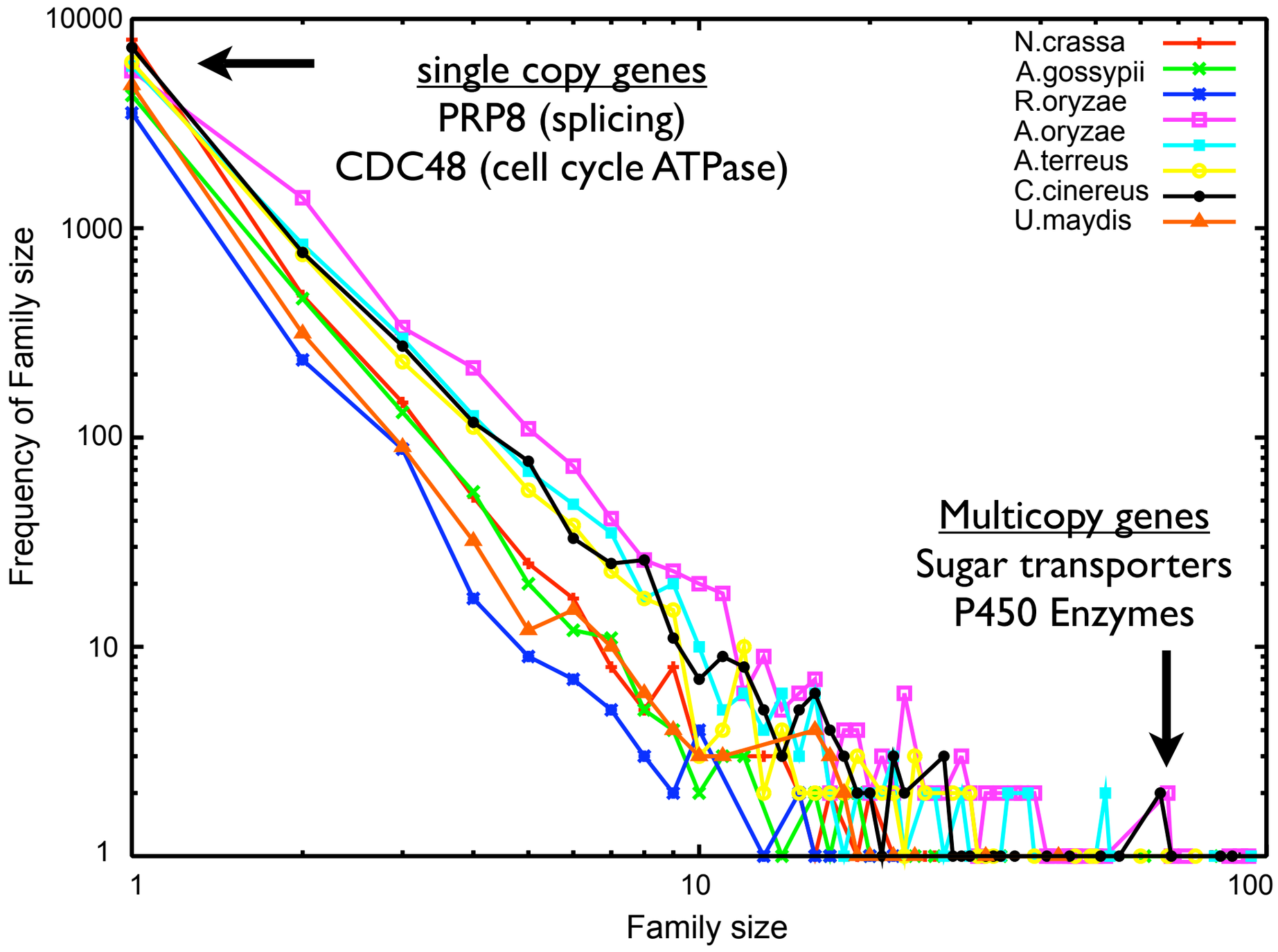
Why study family size change?

- Gene duplicates are fixed because one copy aquires a novel function (Ohno 1970)
  - Gene duplications are the crucible of new genes and thus new function.
- Not all duplicates are preserved because of directional selection (Lynch and Katju, TIG 2004)
- Are some lineage-specific expansions of gene families the result of adaptive evolution?

Identifying family expansions

- Previous work primarily considered pairwise comparisons
- *Ad hoc* comparison of gene family sizes
  - *C.elegans*-*C.briggsae* - GPCR family expansions (Stein et al, *PLOS Biology* 2004).
  - *A. gambiae*-*D. melanogaster* - Mosquito specific family expansions related to symbiotic bacteria (Holt et el, *Science* 2002).
  - Loss of olfactory receptors corr with increased vision capabilities in Humans (Gilad, PNAS 2003, PLOS Bio 2004)
- Need a null model

Gene family sizes follow power law distribution



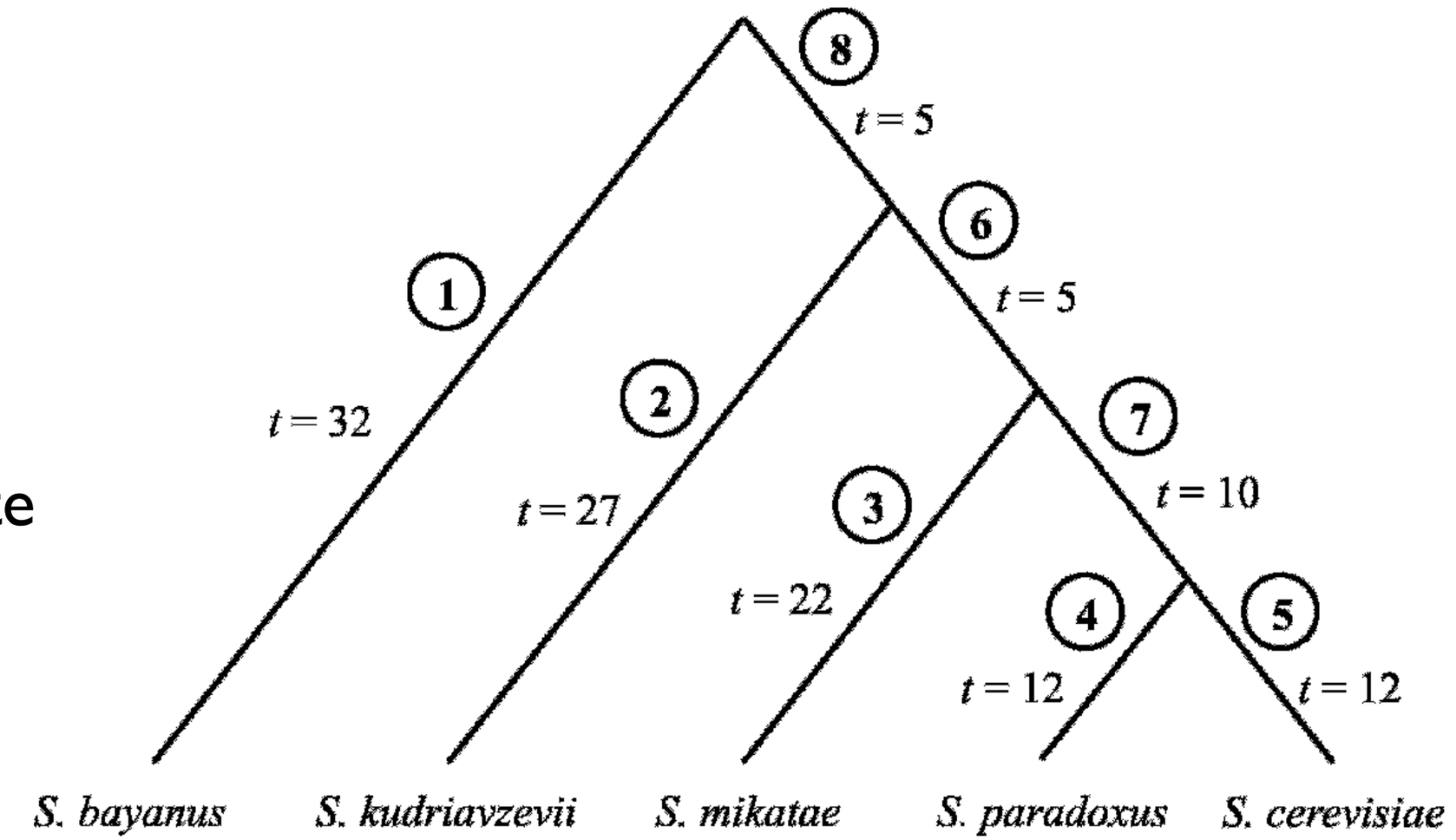
Phylogenetic evaluation of gene family size change

- Previous methods only *ad hoc* statistics
- Explicit model for gene family size change according to a Birth-Death models
- Apply BD to family size along phylogeny using probabilistic graph models
- CAFE - Computational Analysis of gene Family Evolution

Hahn et al, *Genome Res* 2005
De Bie, et al *Bioinformatics* 2006

CAFE

- Use a Probabilistic Graph Model for:
- Ancestral states
- Birth and Death rate (lamda)
- Per branch changes
- P-values



Hahn et al, *Genome Res* 2005

Phylogeny of 37 fungi with fully sequenced genomes

Zygomycota

Euascomycota

Hemiascomycota

Archiascomycota

Basidiomycota

Rhizopus oryaze Bread mold, Opp Hum pathogen

Neurospora crassa

Podospora anserina

Chaetomium globosum

Magnaporthe grisea

Fusarium verticillioides

Fusarium graminearum

Trichoderma reesei

Sclerotinia sclerotiorum

Botrytis cinerea

Stagonospora nodo

Uncinocarpus reesii

Coccidioides immitis

Histoplasma capsula

Aspergillus fumigatu

Aspergillus nidulans

Aspergillus terreus

Aspergillus oryzae

Ashbya gosspyii

Kluyveromyces lactis

Saccharomyces cerevisia

Candida glabrata

Candida lusitaniae

Debaryomyces hansenii

Candida guilliermondii

Candida tropicalis

Candida albicans

Candida dubliniensis

Yarrowia lipolytica

Schizosaccharomyces pomb

Cryptococcus neoformans

Cryptococcus neoformans H

Cryptococcus gattii WM276

Cryptococcus gattii R265

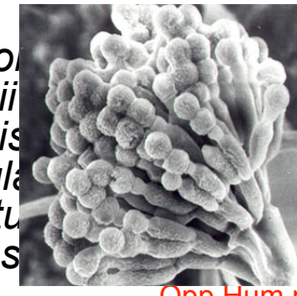
Phanerochaete chrysosporiu

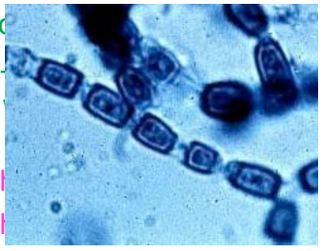
Coprinus cinereus

Ustilago maydis

Million years ago

900 800 700 600 500 400 300 200 100 0





Saprophyte

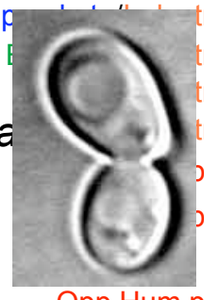
Necr

Opp Hum pathogen

Saprophyte

Opp Hum pathogen

Industrial uses



Saprophyte

Industrial uses

Industrial uses

Industrial uses

Opp Hum pathogen

Opp Hum pathogen

Opp Hum pathogen

Opp Hum pathogen

Opp Hum pathogen

Industrial uses

Saprophyte

Biotroph - maize

50+ More funded and in progress world-wide

Sequencing In-Progress*

Species	Clade	Sequencing center	
Schizosaccharomyces japonicus	Archaeascomycta	Broad-FGI	
Schizosaccharomyces octosporus	Archaeascomycta	Broad-FGI	
Pneumocystis carinii	Archaeascomycta	Sanger, UC, Broad-FGI	
Pneumocystis carinii hominis	Archaeascomycta	UC, Broad-FGI, UC	
Amanita bisporigera	Basidiomycota: Homobasidiomycota	MSU	
Crinipellis perniciosa	Basidiomycota: Homobasidiomycota	Univ Campinas	
Ganoderma lucidum	Basidiomycota: Homobasidiomycota	Yang-Ming Univ	
Hebeloma cylindrosporum	Basidiomycota: Homobasidiomycota	INRA	
Laccaria bicolor	Basidiomycota: Homobasidiomycota	JGI-DOE	R
Phakopsora pachyrhizi	Basidiomycota: Homobasidiomycota	JGI-DOE	
Postia placenta	Basidiomycota: Homobasidiomycota	JGI-DOE	
Schizophyllum commune	Basidiomycota: Homobasidiomycota	JGI-DOE	
Sporobolomyces roseus	Basidiomycota: Urediniomycota	JGI-DOE	
Phakopsora meibomiae	Basidiomycota: Urediniomycota	JGI-DOE	
Batrachochytrium dendrobatidis	Chytridiomycota	Broad-FGI & JGI-DOE	
Piromyces sp.	Chytridiomycota	JGI-DOE	
Glomus intraradices	Glomeromycota	JGI-DOE	
Phycomyces blakesleeanus	Zygomycota	JGI-DOE	R
Brachiola algerae	Microsporidia	Genoscope	
Nosema (Antonospora) locustae	Microsporidia	MBL	R
Enterocytozoon bieneusi	Microsporidia	Tufts Univ	

Sequencing In-Progress*

Species	Clade	Sequencing center	
Aspergillus niger	Euascomycota: Eurotiomycota	DOE-JGI	R
Aspergillus flavus	Euascomycota: Eurotiomycota	NCSU	
Aspergillus clavatus	Euascomycota: Eurotiomycota	OU	
Neosartorya fischeri	Euascomycota: Eurotiomycetes	TIGR	
Histoplasma capsulatum WU24	Euascomycota: Eurotiomycota	Broad-FGI	R
Histoplasma capsulatum 186R,217B	Euascomycota: Eurotiomycota	WUSTL	
Coccidioides posadasii	Euascomycota: Eurotiomycota	TIGR	
Coccidioides immitis 10 strains	Euascomycota: Eurotiomycota	Broad-FGI & TIGR	
Paracoccidioides brasiliensis	Euascomycota: Eurotiomycota	Univ of Brazil	
Ascosphaera apis	Euascomycota: Eurotiomycota	BCM	
Epichloe festucae	Euascomycota: Sordariomycetes	UK	
Podospora anserina	Euascomycota: Sordariomycetes	Broad-FGI	
Trichoderma atroviride	Euascomycota: Sordariomycetes	DOE-JGI	
Trichoderma virens	Euascomycota: Sordariomycetes	DOE-JGI	
Leptosphaeria maculans	Euascomycota: Dothideomycetes	Genoscope	
Alternaria brassicicola	Euascomycota: Dothideomycetes	VPI & WUSTL	R
Xanthoria parietina (lichen)	Euascomycota: Lecanoromycetes	DOE-JGI	
Candida albicans WO-1	Hemiascomycota	Broad-FGI	
Lodderomyces elongisporus	Hemiascomycota	Broad-FGI	R
Pichia stipitis	Hemiascomycota	JGI-DOE	
Saccharomces bayanus	Hemiascomycota	(49, 167)	
Saccharomces castellii	Hemiascomycota	(49)	
Saccharomces cerevevisiae RM11-1A	Hemiascomycota	Broad-FGI	
Saccharomces cerevevisiae YJM789	Hemiascomycota	(113) +++	
Saccharomyces kluyeri	Hemiascomycota	WUSTL (finishing)	
Saccharomces kudriavzevii	Hemiascomycota	(49)	
Saccharomces mikatae	Hemiascomycota	(49, 167)	
Saccharomces paradoxus	Hemiascomycota	(167)	
Saccharomyces pastorianus	Hemiascomycota	Kitasato Univ	
Zygosaccharomyces rouxii	Hemiascomycota	CNRS-Genoscope	

25 strains

Genome annotation

- Many of the fungal genomes were only assembled genomic sequence.
- Automated annotation pipeline was built to generate to get systematic gene prediction.
- Several gene prediction programs were trained and results were combined with GLEAN (Liu, Mackey, Roo, et al unpublished) to produce composite gene calls.

Zygomycota

Euascomycota

Hemiascomycota

Ascomycota

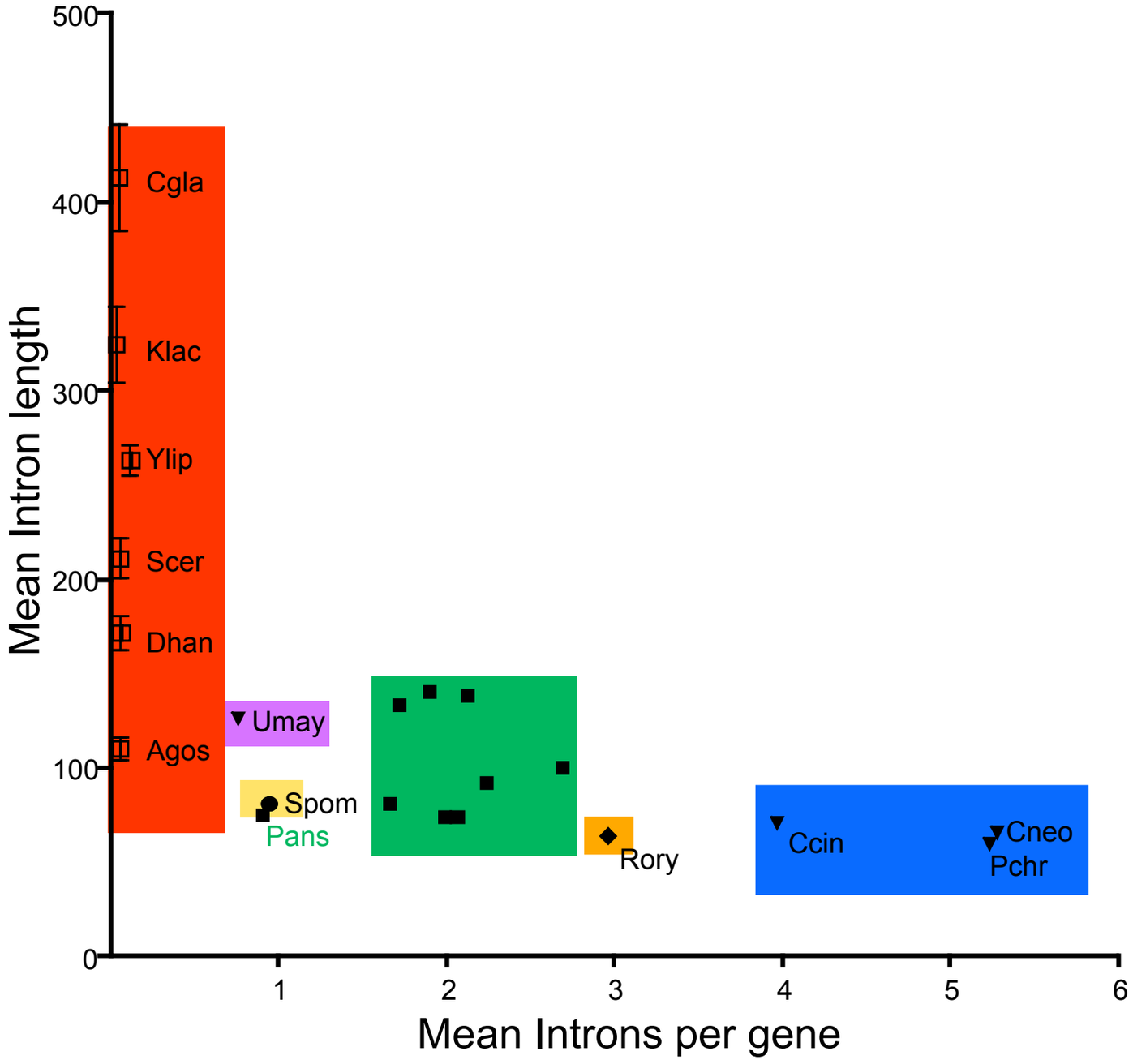
Basidiomycota

Rhizopus oryaze	Bread mold, Opp Hum pathogen
Neurospora crassa	Saprophyte
Podospora anserina	Saprophyte
Chaetomium globosum	Saprophyte
Magnaporthe grisea	Opp Hum Pathogen
Fusarium verticillioides	Hemibiotroph - Rice
Fusarium graminearum	Hemibiotroph - maize
Trichoderma reesei	Hemibiotroph - wheat
Sclerotinia sclerotiorum	Saprophyte
Botrytis cinerea	Necrotroph
Stagonospora nodorum	Necrotroph - fruits
Uncinocarpus reesii	Hemibiotroph - wheat
Coccidioides immitis	Primary Hum pathogen
Histoplasma capsulatum	Primary Hum pathogen
Aspergillus fumigatus	Opp Hum pathogen
Aspergillus nidulans	Saprophyte
Aspergillus terreus	Opp Hum pathogen
Aspergillus oryzae	Saprophyte/Industrial uses
Ashbya gosspyii	Biotroph/Industrial uses
Kluyveromyces lactis	Industrial uses
Saccharomyces cerevisiae	Industrial uses
Candida glabrata	Opp Hum pathogen
Candida lusitaniae	Opp Hum pathogen
Debaryomyces hansenii	
Candida guilliermondii	Opp Hum pathogen
Candida tropicalis	Opp Hum pathogen
Candida albicans	Opp Hum pathogen
Candida dubliniensis	Opp Hum pathogen
Yarrowia lipolytica	Industrial uses
Schizosaccharomyces pombe	
Cryptococcus neoformans	Opp Hum pathogen
Cryptococcus neoformans H99	Opp Hum pathogen
Cryptococcus gattii WM276	Opp Hum pathogen
Cryptococcus gattii R265	Opp Hum pathogen
Phanerochaete chrysosporium	Saprophyte
Coprinus cinereus	Saprophyte
Ustilago maydis	Biotroph - maize

Million years ago

900 800 700 600 500 400 300 200 100 0

Intron frequency varies among the fungi



Methodology

Analysis

Rfam
tRNAscan

SNAP
Twinscan

ZFF to GFF3

GFF2 to GFF3

Glimmer
Genscan

Tools::Glimmer

Tools::Genscan

BLASTZ
BLASTN

SearchIO

Genome

Proteins

exonerate
protein2genome

GFF2 to GFF3

ESTs

exonerate
est2genome

Bio::DB::GFF

GFF to AA

protein to
genome
coordinates

SearchIO

Tools::GFF

predicted
proteins

FASTA
all-vs-all

Proteins

SearchIO

HMMER

Find
orthologs

MCL

GLEAN
(combiner)

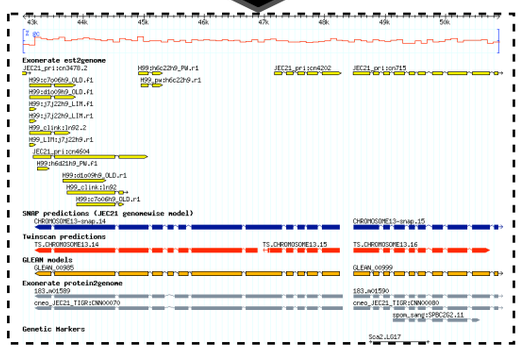
Multiple
sequence
alignment

Gene families

Bio::AlignIO
aa2cds alignment

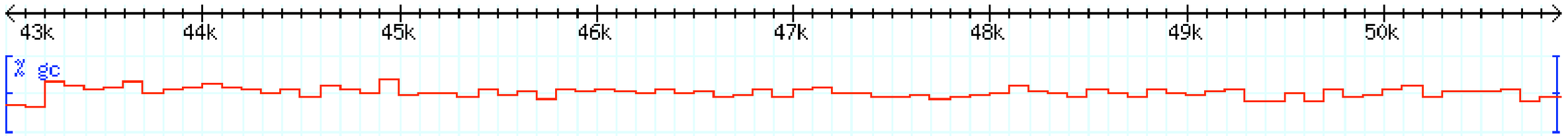
Intron
mapping into
alignment

Intron
analysis



http://fungal.genome.duke.edu

Generic Genome Browser



Exonerate est2genome

JEC21_pri:cn3478.2
H99:c7o06h9_OLD.f1
H99:d1o09h9_OLD.f1
H99:j7j22h9_LIM.f1
H99:j7j22h9_LIM.r1
H99_clink:ln92.2
H99_LIM:j7j22h9.r1
JEC21_pri:cn4604
H99:h6d21h9_PW.f1
H99:d1o09h9_OLD.r1
H99_clink:ln92
H99:c7o06h9_OLD.r1
H99:h6c22h9_PW.r1
H99_pw:h6c22h9.r1
JEC21_pri:cn4202
JEC21_pri:cn715

SNAP predictions (JEC21 genomewise model)

CHROMOSOME13-snap.14
CHROMOSOME13-snap.15

Twinscan predictions

TS.CHROMOSOME13.14
TS.CHROMOSOME13.15
TS.CHROMOSOME13.16

GLEAN models

GLEAN_00985
GLEAN_00999

Exonerate protein2genome

183.m01589
cneo_JEC21_TIGR:CNN00070
183.m01590
cneo_JEC21_TIGR:CNN00080
spom_sang:SPBC2G2.11

Genetic Markers

Sca2.LG17

365k

366k

% gc

Named Genes CDS

trp1

Twinscan predictions

TS.ccin_1.103.134

GLEANmax modes (Best set)

Jan06m300_GLEAN_00356

Probability 0.904599

Exonerate est2genome

gi|61244574|gb|DN593801.1|DN593801

Genezilla CDS (C.neoformans trained)

gz.5222

gz.5223

SNAP CDS (gbk trained)

gbk_ccin_1.103-snap.149

SNAP CDS (Genomewise trained)

gw_ccin_1.103-snap.106

SNAP CDS

ccin_1.103-snap.104

AUGUSTUS (Genomewise trained)

ccin_1.103-g116.1

Exonerate protein2genome

spom_sang:SPAC19A8.15

ctro_SNAP:ctro_1.6-snap.238

hcap_SNAP:186R_hcap-186R_33.38-snap.1

fgra_brd:FG11345.1

sklu_SNAP:sklu_c1623-snap.2

Pfam Genewise domains

Trp_syntA

Bits=450.84

Trp_syntA-1

Bits=450.84

PALP

Bits=403.13

Coprinus cinereus TRP1

Curated model

Named Genes CDS
STE20_H99_alpha

Exonerate est2genome
JEC21_pri:cn1598
H99_sing:AF315638.2
B3501:k5f12j2_OLD.f1
B3501:p5i20b3_LIM.r1
B3501:p5i20b3_LIM.f1
JEC21_pri:cn1597
B3501:k1i04j2_OLD.r1
B3501:k1i04j2_OLD.f1
B3501_pri:cn730
H99_sing:AF315638
184A:b1i16a8.f1

SNAP predictions (JEC21 genomewise model)
CHROMOSOME5-snap.75

Twinscan predictions
TS.CHROMOSOME5.82

GLEAN models
GLEAN_02118
Probability 0.999994
GLEAN_gz_02107
Probability 0.999996

GlimmerHMM (Cryptococcus trained)
CHROMOSOME5-glimmerhmm.69
CHROMOSOME5-glimmerhmm.68

AUGUSTUS (JEC21 trained)
CHROMOSOME5-g65.1

Genezilla CDS (JEC21 trained)
gz.2664

Exonerate protein2genome
cneo_JEC21_TIGR:CND05790
pchr_SNAP:pchr_198-snap.7
rory_SNAP:rory_1.21-snap.157
rory_SNAP:rory_1.34-snap.115
rory_SNAP:rory_1.44-snap.84
rory_SNAP:rory_1.3-snap.303
pchr_SNAP:pchr_110-snap.41
rory_SNAP:rory_1.56-snap.48
rory_SNAP:rory_1.19-snap.236
rory_SNAP:rory_1.3-snap.302
fgra_brd:FG09492.1.2

Pfam domains
PH
PH domain E-value: 2.2e-11
Pkinase
Protein kinase domain E-value: 0

Final call

C.neoformans
STE20

Methods: gene family identification

- All-vs-All pairwise sequence searches (FASTP)
- Cluster genes by similarity using Markov CLustering (MCL) algorithm
- Identify families with unusually large size changes along phylogeny with CAFE
- Use 37 fungal genomes from 5 major clades

FASTA
all-vs-all

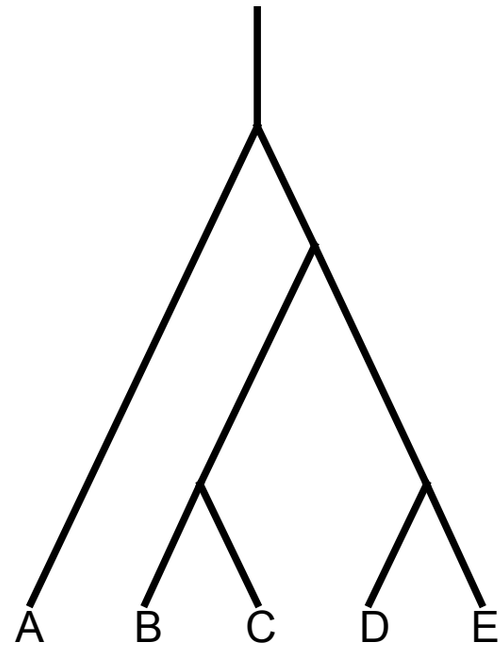
MCL

Gene
families

Family count

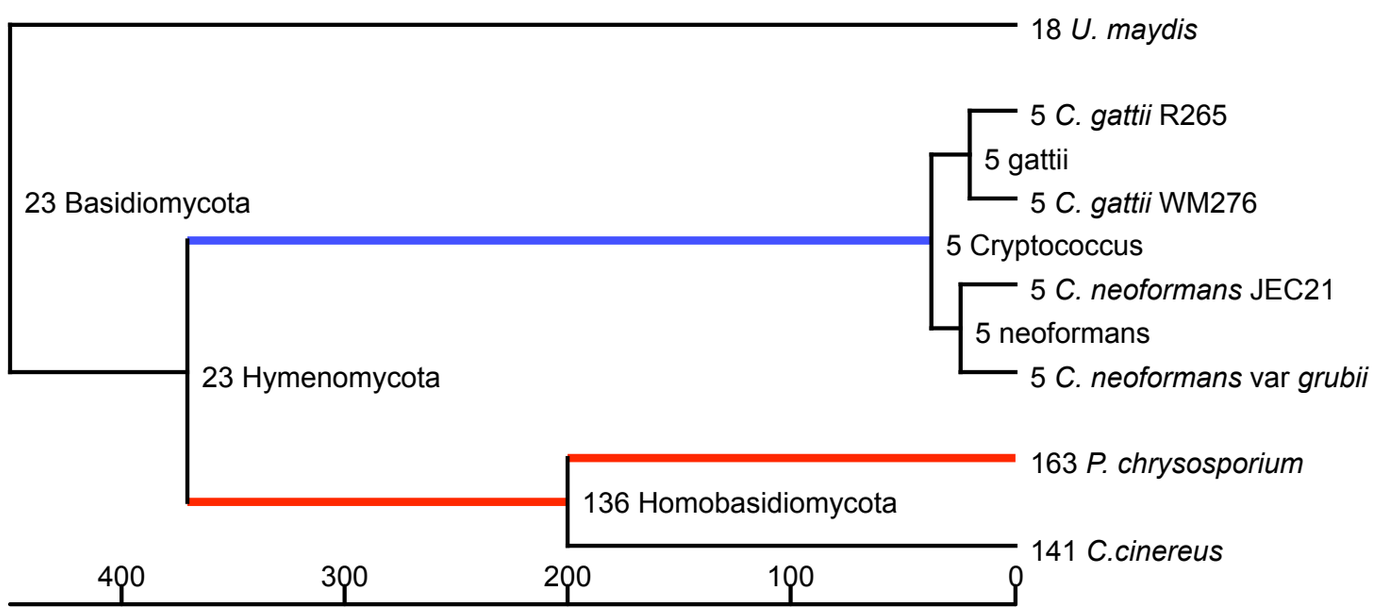
Species			
	10	1	2
	14	18	2
	7	1	1
	6	1	12
	6	1	8
	3	1	1

+



CAFE

Family 1	P < 0.001	Branch A
Family 2	P < 0.001	Branch B
Family 3	P=0.02	Branch C,E
Family 4	P=0.03	Branch D



Families with significant expansions

49 significant families

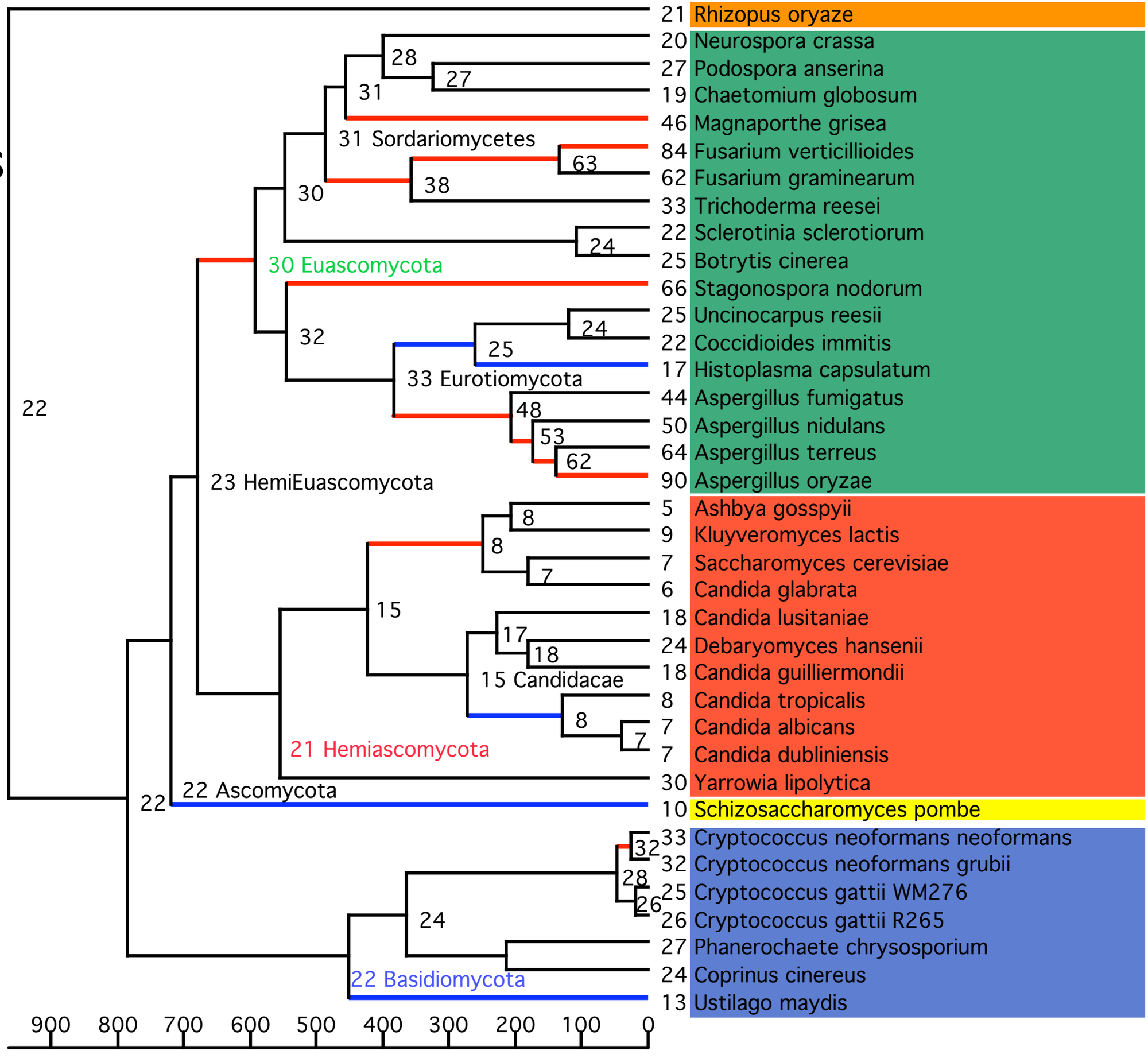
Transporters
Kinases
P450
Oxidation

Vitamin & Cofactor transport	Methytransferase
Lactose & sugar transport	Cytochrome P450: CYP64
Amine transport	Cytochrome P450: CYP53,57A
Myo-instol, quinate, and glucose transport	Cytochrome P450
Oligopeptide transport	Kinase
ABC transporter	Subtilase family
MFS, drug pump, & sugar transport	NADH flavin oxidoreductase
Transport	Aldehyde dehydrogenase
Monocarboxylate & sugar transport	Aldo/kedo reductase
ABC transport	Multicopper oxidase
Amino acid permease	AMP-binding enzyme

Transporters

- Of 45 significant families, 22 were related to transport
  - Vitamin and amino acid transport
  - Sugar and sugar-like transporters
  - Multidrug and efflux pumps
  - ABC transporters (ATP Binding Cassette)

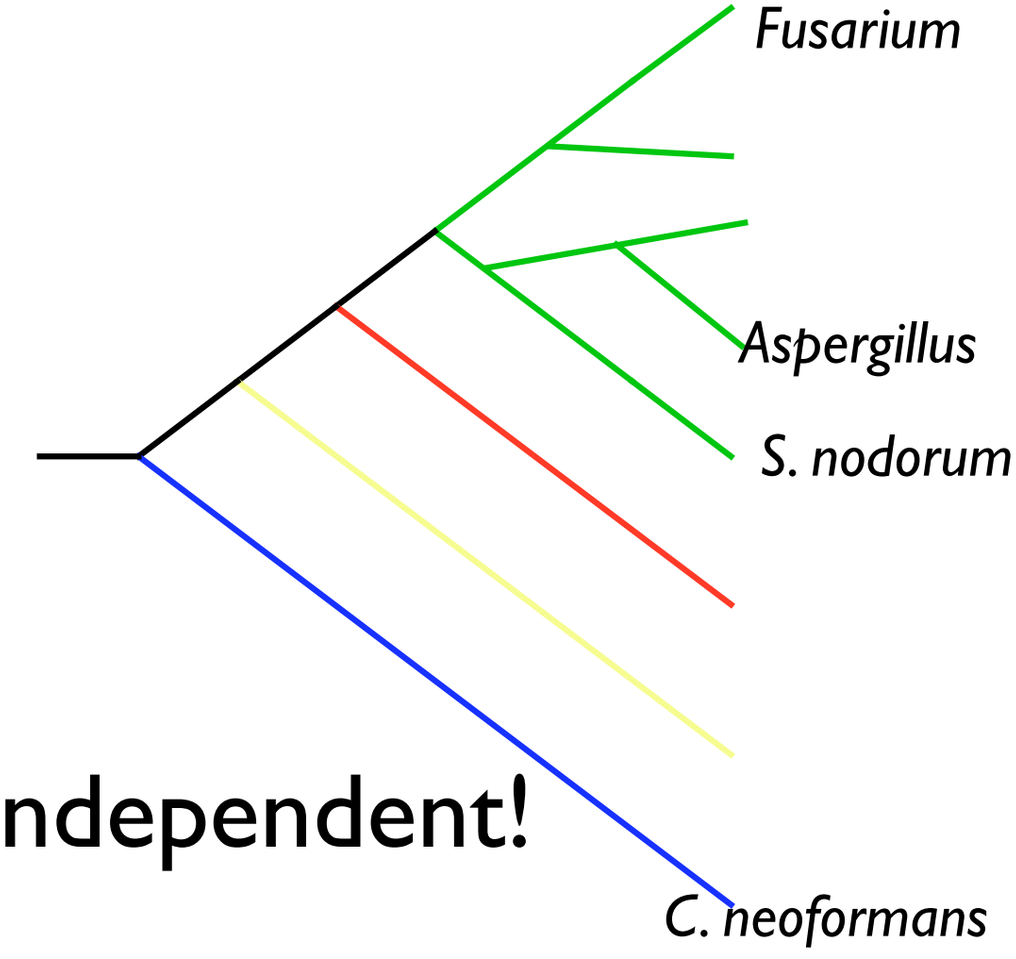
Vitamin & Cofactor Transporters



Marked branches with significant (P<0.05) expansions or contractions

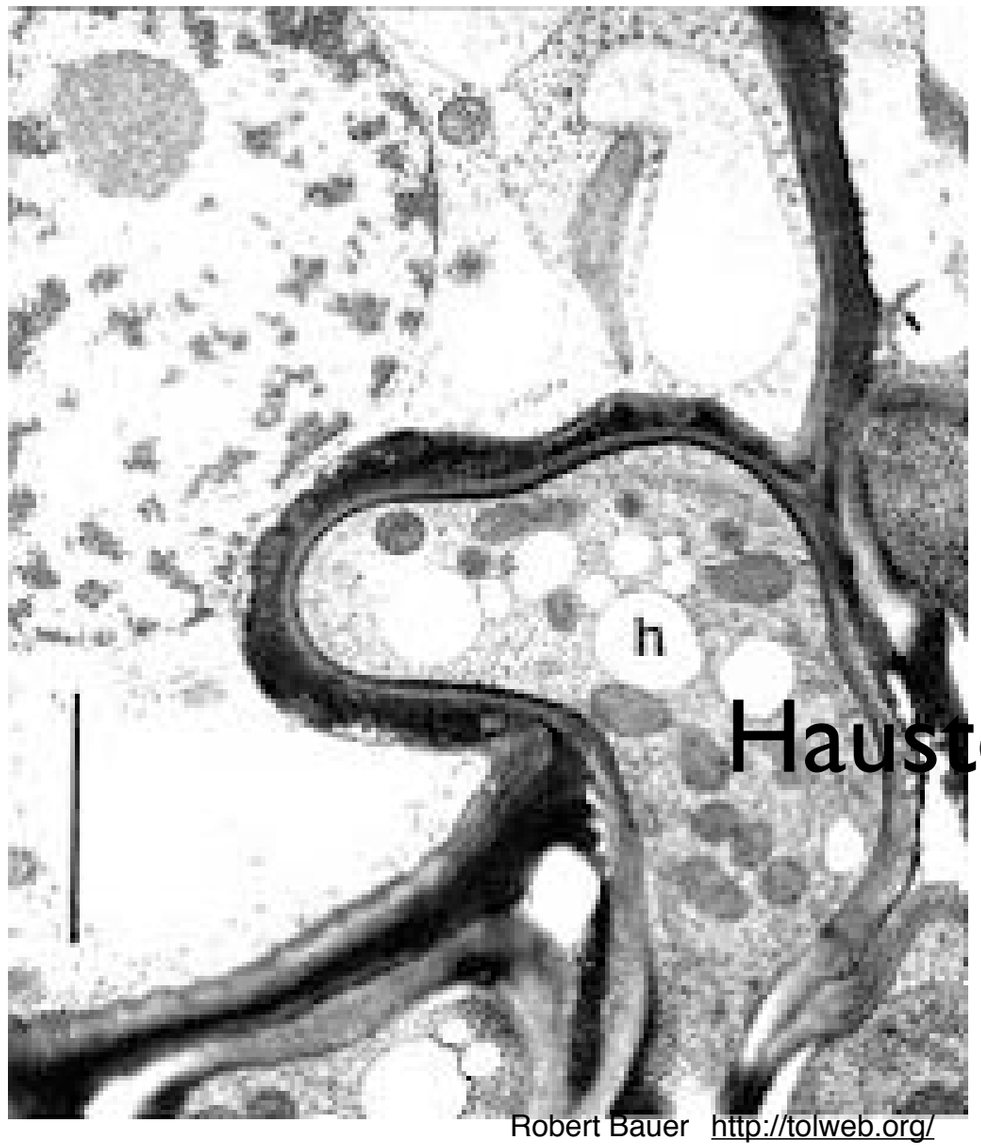
Transporter expansions

- Sugar related, Drug pump, and Major Facilitator Superfamily
  - *Aspergillus spp*, *Fusarium spp*, *S. nodorum*
    - Euascomycota
- Vitamin transport
  - *C. neoformans*, *Fusarium*
  - *A. nidulans* (Biotin)
- Saccharomyces expansions independent!



Sugar transporter use in phytopathogens

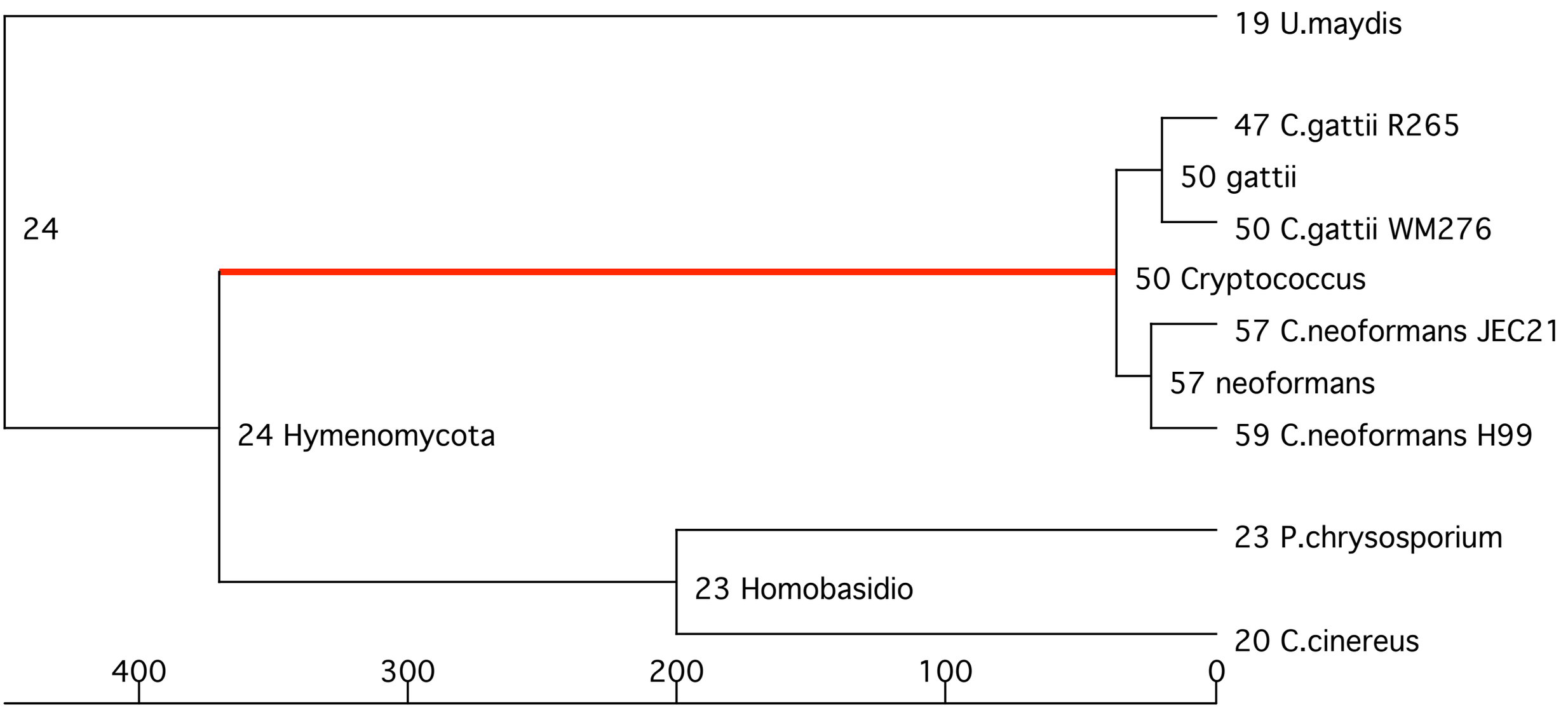
- Sugar transporters are used to extract nutrients from host
- Haustorium: specialized structure for plant parasitism
- Many sugar transporters highly and specifically expressed in haustoria



Haustorium

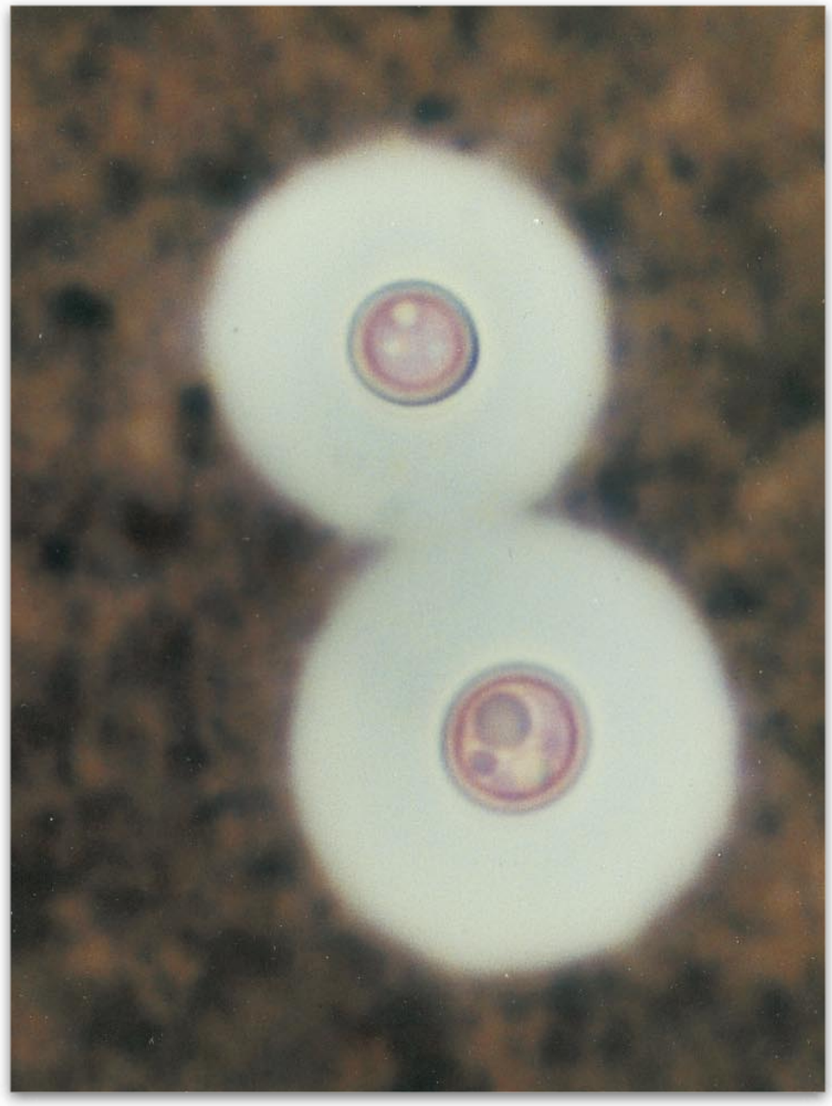
Robert Bauer http://tolweb.org/

Cryptococcus sugar transporters expansion



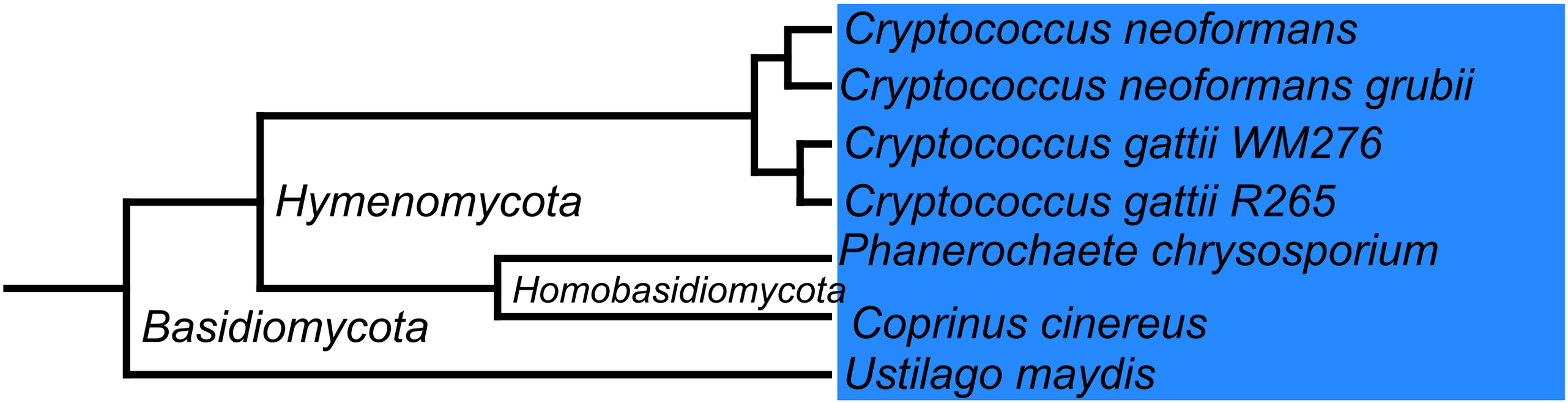
Cryptococcus sugar transporters

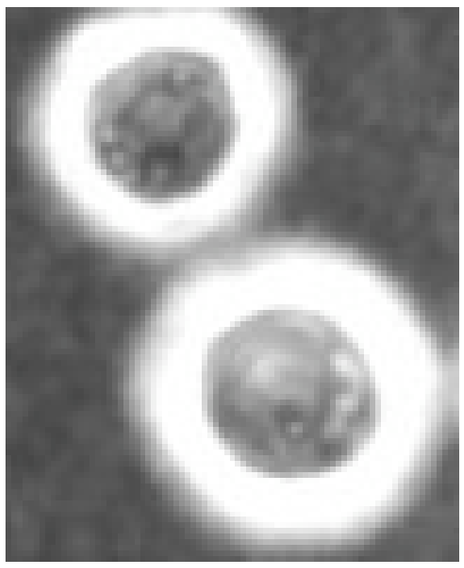
- 3x as many sugar transporters in *C. neoformans* (~50) than other basidiomycetes
- "sugar coated killer"
- Capsule is a mixture of glucose, xylose, and mannose.
- Transporters could be important in capsule synthesis



Zerpa et al, 1996

Basidiomycota changes





C.neoformans

P.chrysosporium

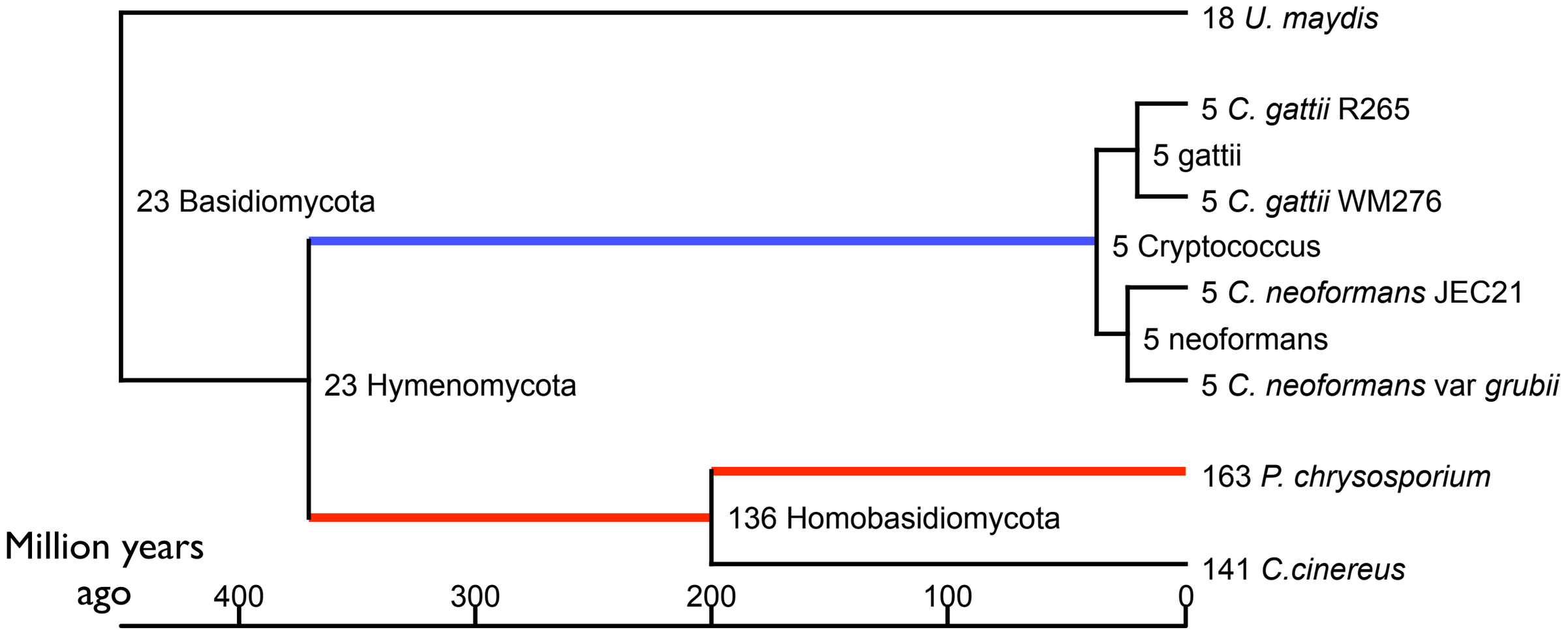
C.cinereus

U.maydis

P450 CYP64

P450 enzymes involved in synthesis and cleavage of chemical bonds. Drug metabolism in animals.

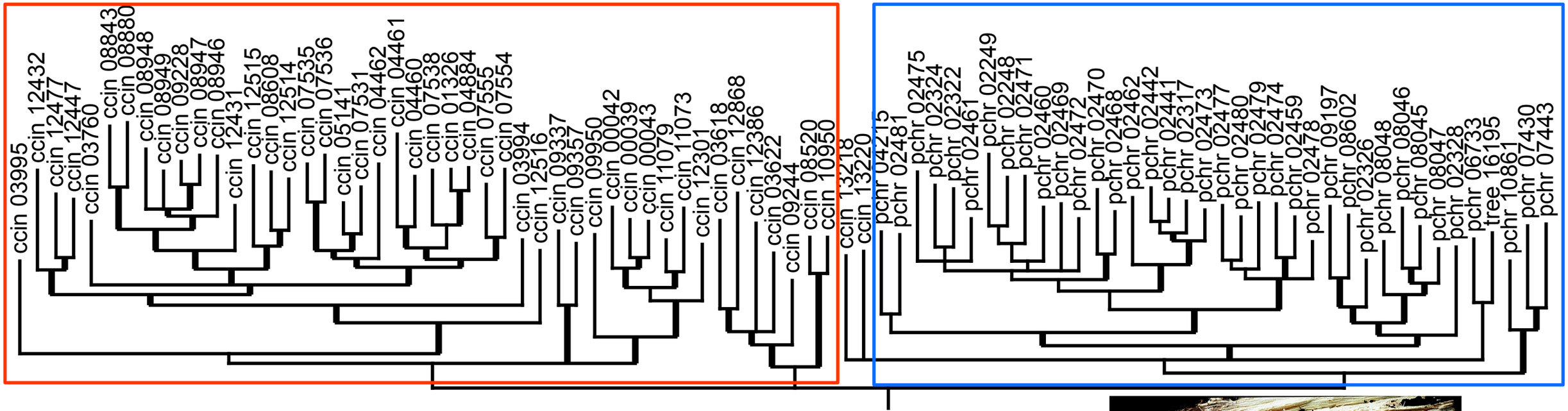
CYP64: Step in *Aspergillus* spp aflatoxin pathway
P. chrysosporium implicated in lignin and hydrocarbon degradation.



CYP64 was from independent duplication

C. cinereus expansion

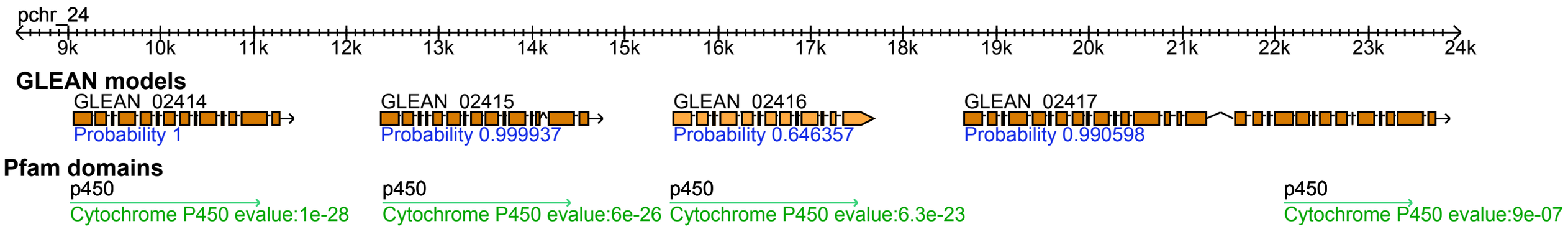
P. chrysosporium expansion



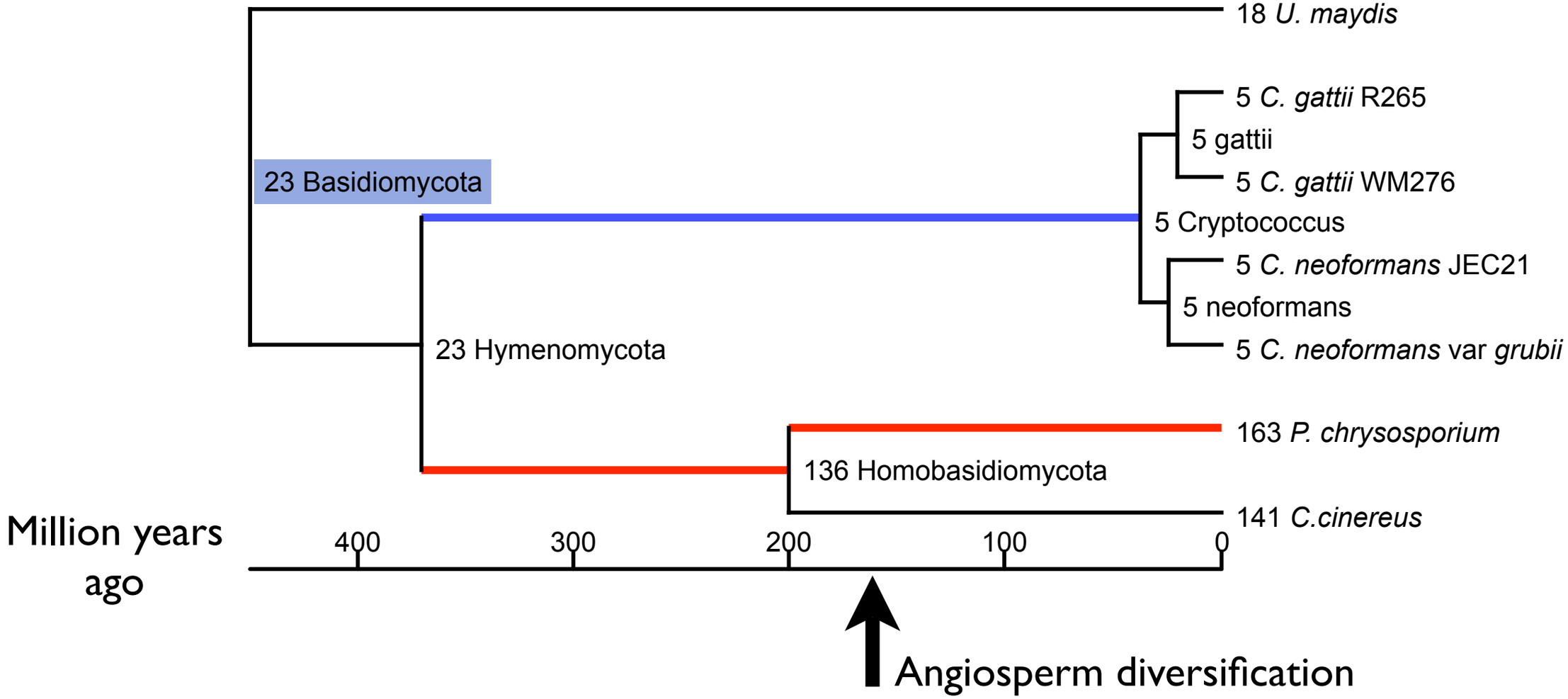
Mario Cervini

Tom Volk

Local duplications created CYP64 expansion



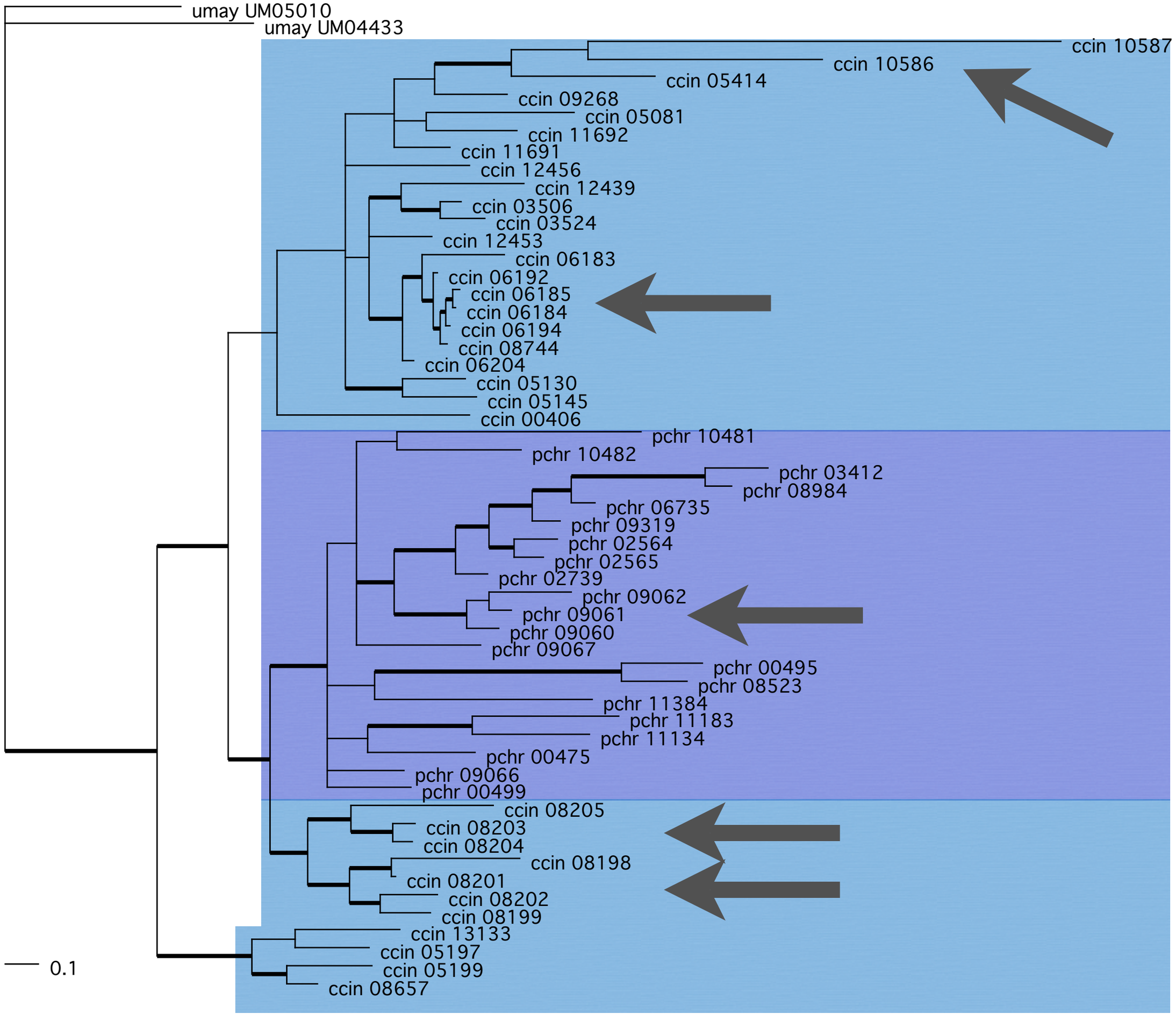
Interpretation of CYP64 expansion



Hydrophobin Family

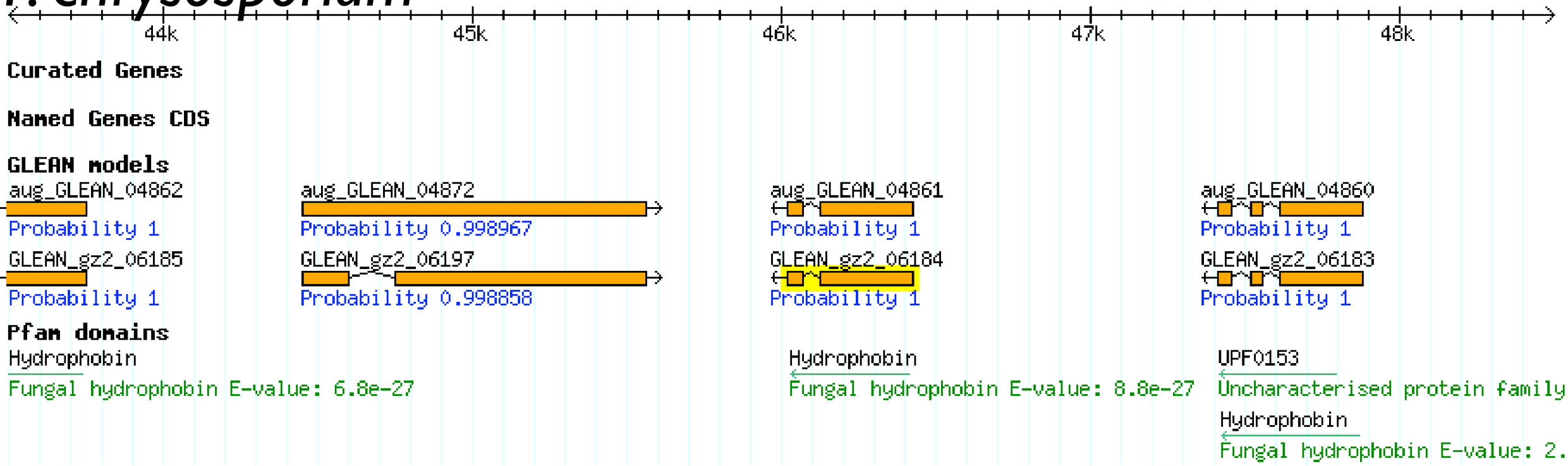
P.chr	*C.cin*	*C.neo*	*U.may*
21	33	0	2

- Self assembling proteins involved in fungal cell wall
- Part of what makes a mushroom
- 8 Cysteine residues critical to function
- Help spores stay airborne resisting water

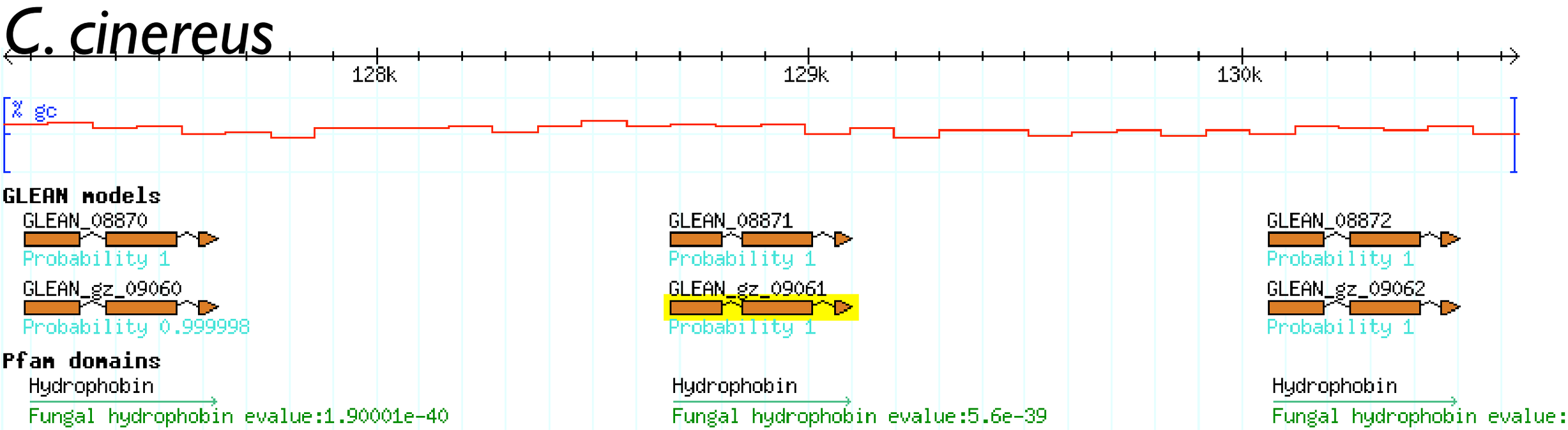

umay UM05010
umay UM04433
ccin 10587
ccin 10586
ccin 05414
ccin 09268
ccin 05081
ccin 11692
ccin 11691
ccin 12456
ccin 12439
ccin 03506
ccin 03524
ccin 12453
ccin 06183
ccin 06192
ccin 06185
ccin 06184
ccin 06194
ccin 08744
ccin 06204
ccin 05130
ccin 05145
ccin 00406
pchr 10481
pchr 10482
pchr 03412
pchr 08984
pchr 06735
pchr 09319
pchr 02564
pchr 02565
pchr 02739
pchr 09062
pchr 09061
pchr 09060
pchr 09067
pchr 00495
pchr 08523
pchr 11384
pchr 11183
pchr 11134
pchr 00475
pchr 09066
pchr 00499
ccin 08205
ccin 08203
ccin 08204
ccin 08198
ccin 08201
ccin 08202
ccin 08199
ccin 13133
ccin 05197
ccin 05199
ccin 08657
0.1

Local Duplications

P. chrysosporium



C. cinereus



Family size contractions

- Many families contracted in *Onygenales*
  - *Histoplasma, Coccidioides*
  - Several Human pathogenic fungi are found in this clade

Coccidioides

- Human pathogen - one of very few fungi that infect immno-competent individuals.
- Found in deserts of Southwest US and in Mexico
- Spores are infectious propogules
- *Onygenales* genomes
  - 2 species have sequenced genomes (*C. immitis* and *C. posadasii*)
  - 1 non-pathogenic (*U. reesii*)
  - 3 *Histoplasma* strains.

Zygomycota

Euascomycota

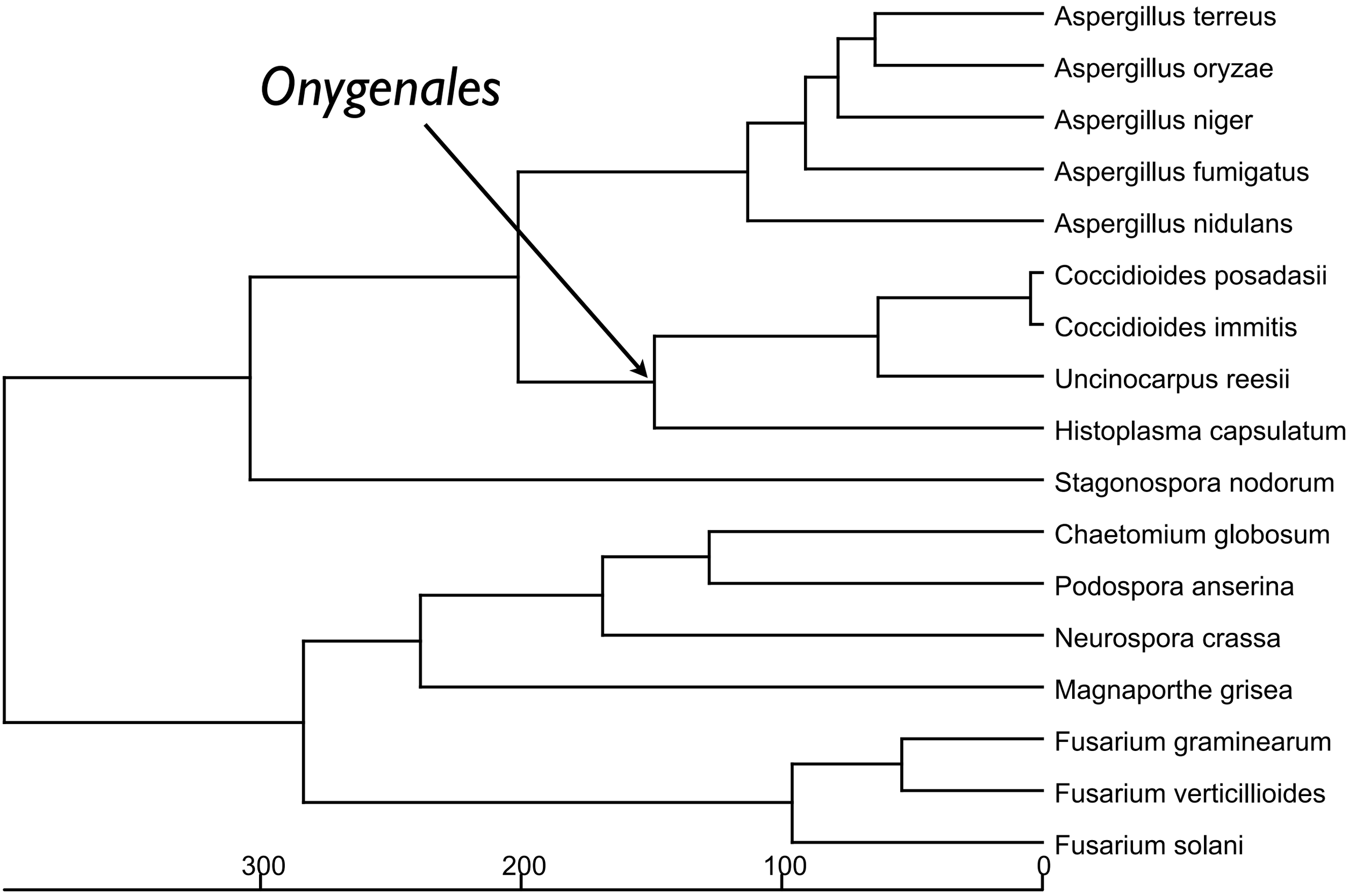
Hemiascomycota

Ascomycota

Basidiomycota

Rhizopus oryaze	Bread mold, Opp Hum pathogen
Neurospora crassa	Saprophyte
Podospora anserina	Saprophyte
Chaetomium globosum	Saprophyte
Magnaporthe grisea	Opp Hum Pathogen
Fusarium verticillioides	Hemibiotroph - Rice
Fusarium graminearum	Hemibiotroph - maize
Trichoderma reesei	Hemibiotroph - wheat
Sclerotinia sclerotiorum	Saprophyte
Botrytis cinerea	Necrotroph
Stagonospora nodorum	Necrotroph - fruits
Uncinocarpus reesii	Hemibiotroph - wheat
Coccidioides immitis	Primary Hum pathogen
Histoplasma capsulatum	Primary Hum pathogen
Aspergillus fumigatus	Opp Hum pathogen
Aspergillus nidulans	Saprophyte
Aspergillus terreus	Opp Hum pathogen
Aspergillus oryzae	Saprophyte/Industrial uses
Ashbya gosspyii	Biotroph/Industrial uses
Kluyveromyces lactis	Industrial uses
Saccharomyces cerevisiae	Industrial uses
Candida glabrata	Opp Hum pathogen
Candida lusitaniae	Opp Hum pathogen
Debaryomyces hansenii	
Candida guilliermondii	Opp Hum pathogen
Candida tropicalis	Opp Hum pathogen
Candida albicans	Opp Hum pathogen
Candida dubliniensis	Opp Hum pathogen
Yarrowia lipolytica	Industrial uses
Schizosaccharomyces pombe	
Cryptococcus neoformans	Opp Hum pathogen
Cryptococcus neoformans H99	Opp Hum pathogen
Cryptococcus gattii WM276	Opp Hum pathogen
Cryptococcus gattii R265	Opp Hum pathogen
Phanerochaete chrysosporium	Saprophyte
Coprinus cinereus	Saprophyte
Ustilago maydis	Biotroph - maize

Million years ago: 900 800 700 600 500 400 300 200 100 0


Onygenales
Aspergillus terreus
Aspergillus oryzae
Aspergillus niger
Aspergillus fumigatus
Aspergillus nidulans
Coccidioides posadasii
Coccidioides immitis
Uncinocarpus reesii
Histoplasma capsulatum
Stagonospora nodorum
Chaetomium globosum
Podospora anserina
Neurospora crassa
Magnaporthe grisea
Fusarium graminearum
Fusarium verticillioides
Fusarium solani
300
200
100
0

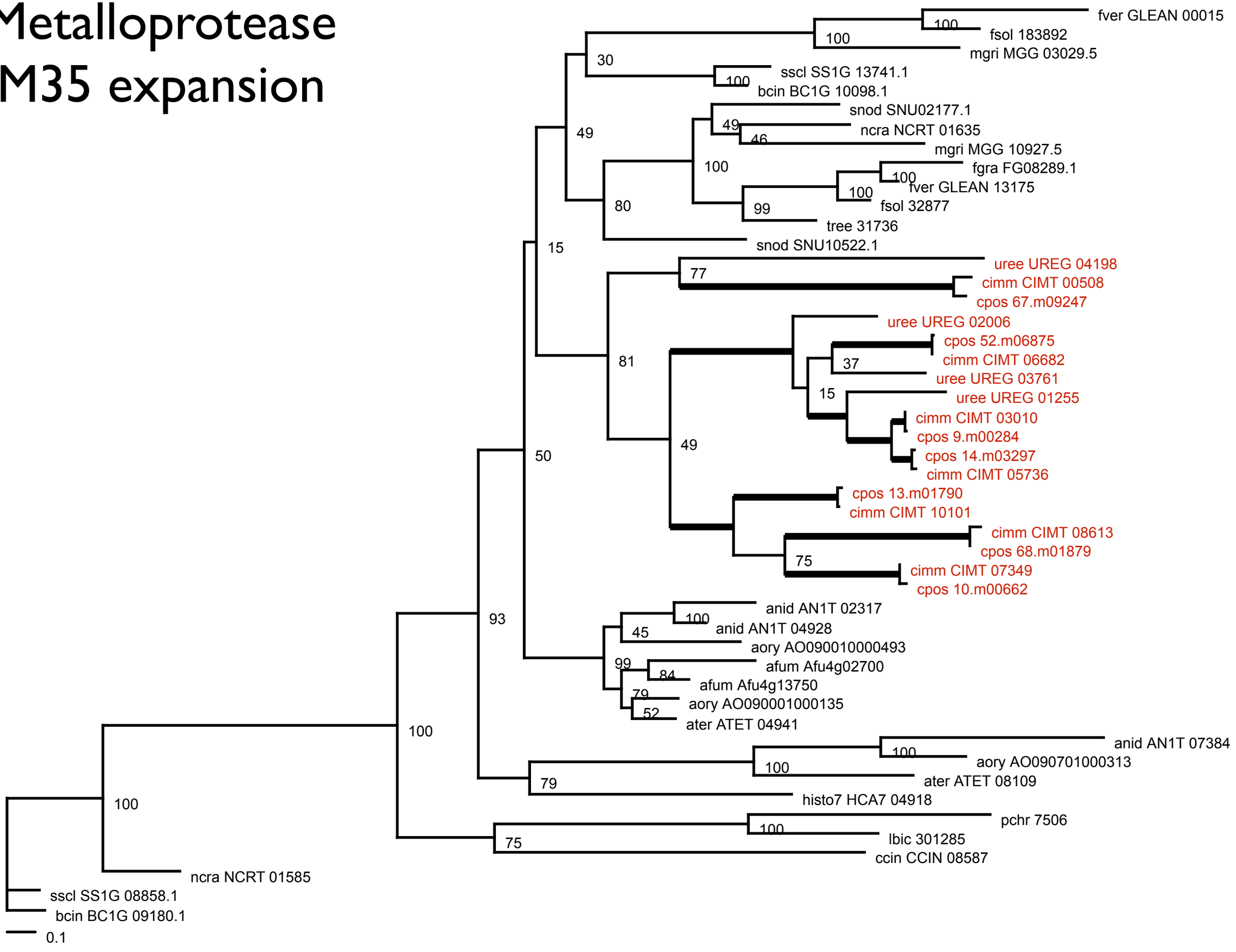
Analysis of protein domain content

- Pfam database is semi-curated set of well-conserved protein domains
- Alignments of the domains can be searched against proteins of each fungal genome
- Identify copy number in each species
- Map values onto phylogenetic tree to identify significant differences
- Only evaluated filamentous euascomycete fungi

Coccidioides expansions

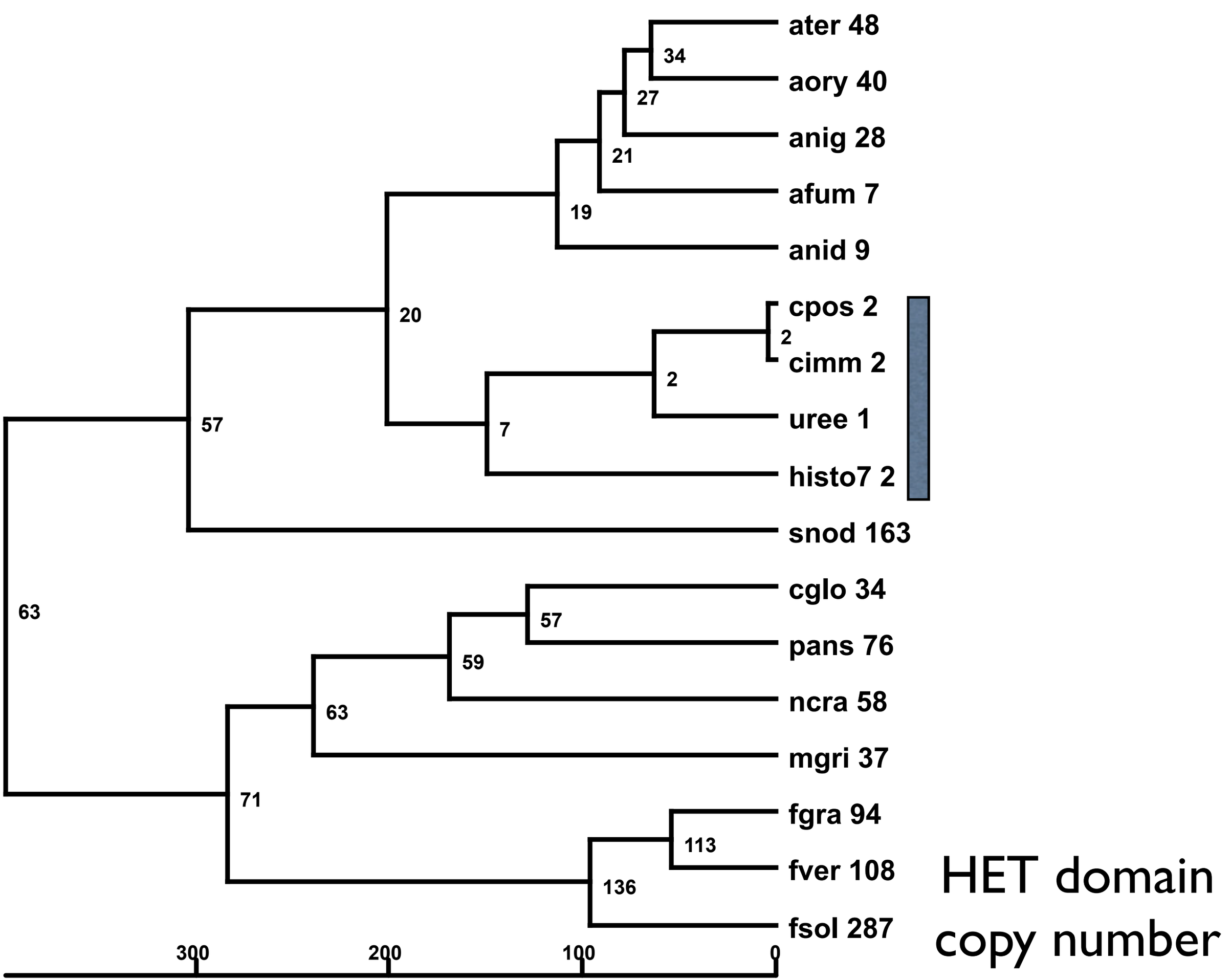
- Metalloprotease (Peptidase M35)
- Subtilisin - peptide proteinase inhibitor

Metalloprotease M35 expansion



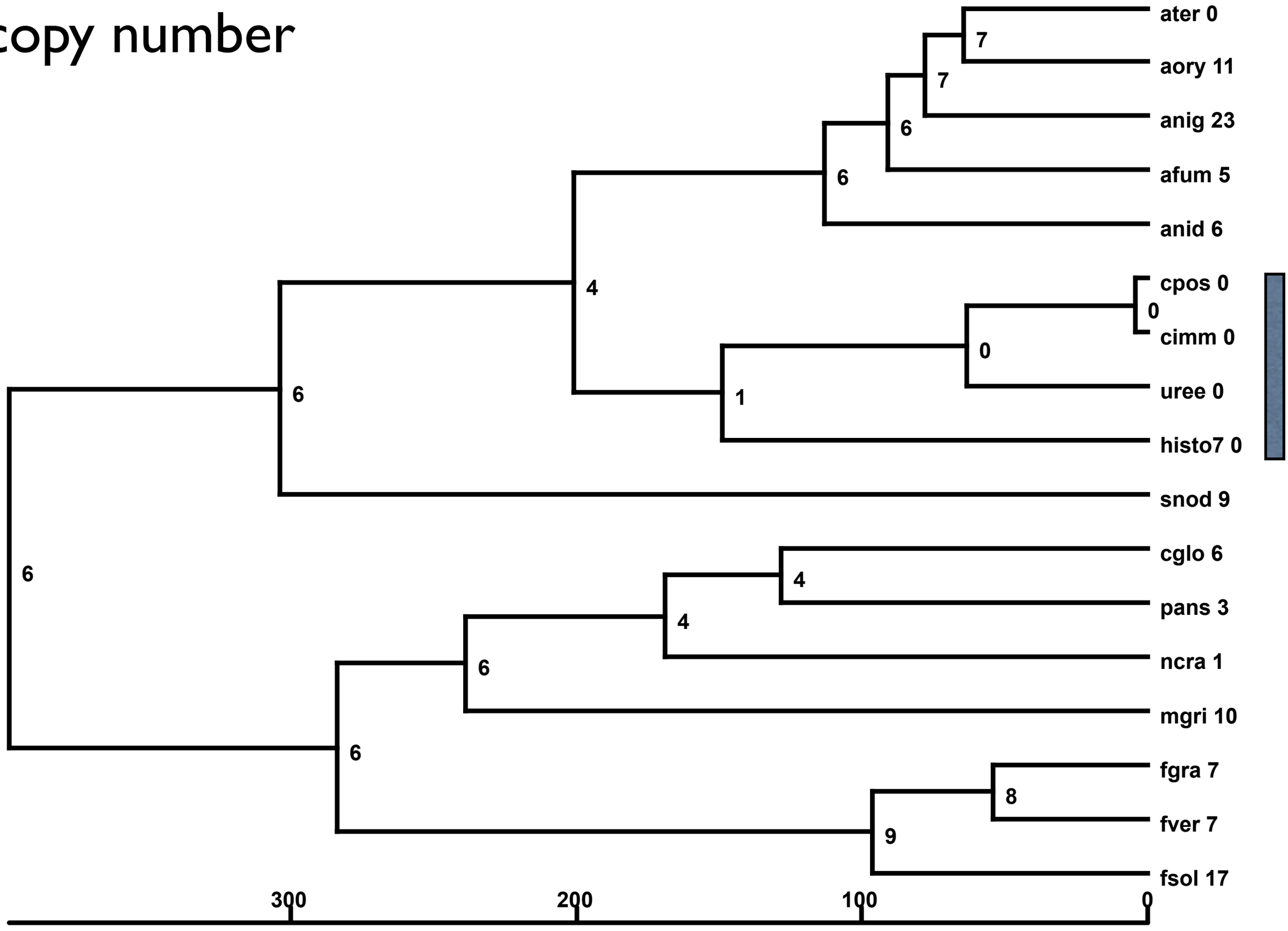
Onygenales contractions

- Heterokaryon incompatibility domain (HET)
- Tannase, Cellulose Binding Domain, Cellulase, Cutinase, NPP1 (necrosis inducing protein)
- Pectin lyase and Pectinesterase - rotting of soft tissue
- Several families of peptidases and hydrolases



HET domain copy number

Tannase domain copy number



Trends

- Lack of many putative plant degrading enzymes
- Are *Onygenales* losing genes related to being saprophytic?
- Genome streamlining as part of pathogenic lifestyles?

Conclusions

- Transporters are highly expanded in independent lineages
  - Saprophytic and phytopathogenic lifestyles
- Homobasidiomycete (mushroom) expansions
  - Lignin degradation - saprophytic lifestyles
  - Hydrophobins - cell wall structures
  - Convergent evolution to generate similar complements of a gene family

Are lineage-specific size changes adaptive?

- Some promising candidates can be identified by these methods
- May need functional data to interpret the changes
- Additional methods to look at timing of duplication and speciation with good sampling

Acknowledgments

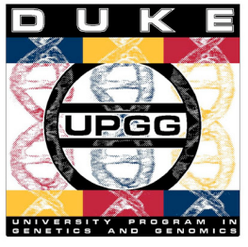
Matthew Hahn (Indiana)
Jeff Demuth
Sang-Gook Han

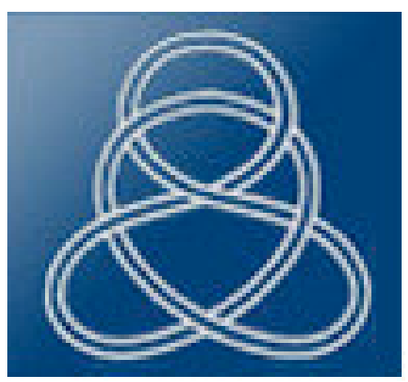
Tijl De Bie
Nello Cristianini

Aaron Mackey
Ian Korf
Mario Stanke

Fred Dietrich

John Taylor
Thomas Sharpton





Sequencing centers
Broad Institute
Joint Genome Institute
Génolevures
Stanford University
TIGR
Welcome Trust Sanger Centre
(NIH and NSF)