Gene family evolution across the Fungi

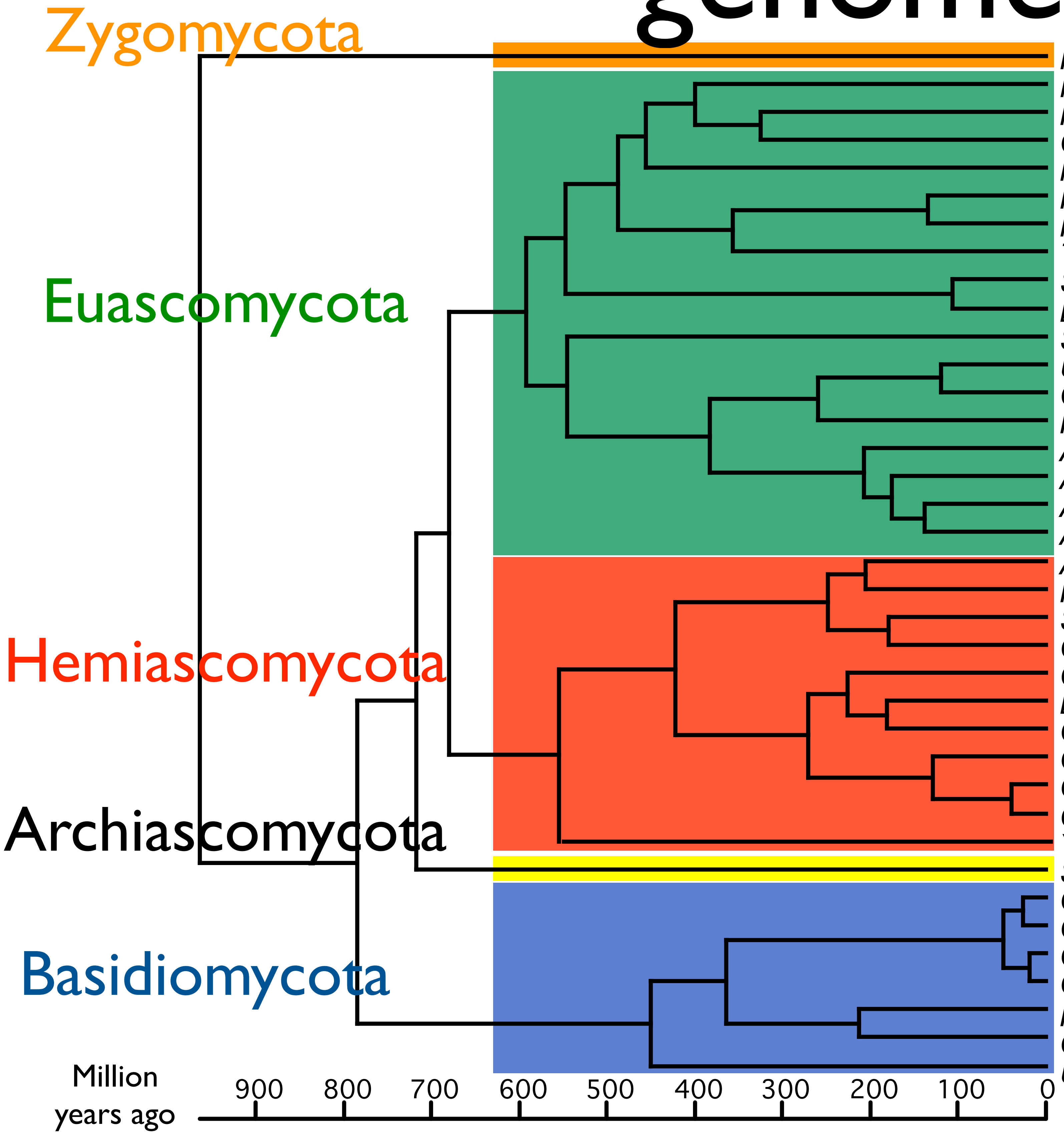
Lineage-specific expansions and contractions

Jason Stajich
Duke University

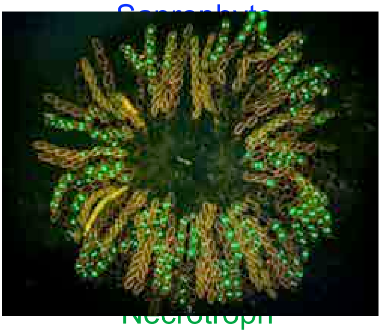
Fungal genome evolution

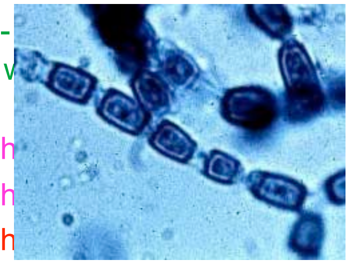
- How have fungi adapted to niches?
- Are family expansions or contractions the result of adaptive evolution?

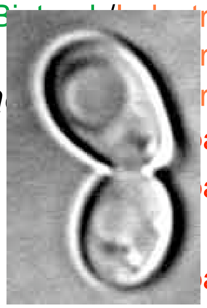
37 Fully sequenced fungal genomes



Rhizopus oryaze Bread mold, Opp Hum pathogen
Neurospora crassa
Podospora anserina
Chaetomium globosum
Magnaporthe grisea
Fusarium verticillioides H
Fusarium graminearum H
Trichoderma reesei
Sclerotinia sclerotiorum Necrotroph
Botrytis cinerea Necrotroph -
Stagonospora nodor - v
Uncinocarpus reesii
Coccidioides immitis ath
Histoplasma capsula ath
Aspergillus fumigatu ath
Aspergillus nidulans rophyte
Aspergillus terreus athogen
Aspergillus oryzae Saprophyte/Industrial uses
Ashbya gosspyii rial uses
Kluyveromyces lactis rial uses
Saccharomyces cerevisia rial uses
Candida glabrata athogen
Candida lusitaniae athogen
Debaryomyces hansenii
Candida guilliermondii athogen
Candida tropicalis Opp Hum pathogen
Candida albicans Opp Hum pathogen
Candida dubliniensis Opp Hum pathogen
Yarrowia lipolytica Industrial uses
Schizosaccharomyces pombe
Cryptococcus neoformans C
Cryptococcus neoformans H
Cryptococcus gattii WM276 C
Cryptococcus gattii R265 C
Phanerochaete chrysosporiu
Coprinus cinereus
Ustilago maydis Biotroph - maize







51+ 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
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
Brachiola algerae	Microsporidia	Genoscope
Nosema (Antonospora) locustae	Microsporidia	MBL
Enterocytozoon bieneusi	Microsporidia	Tufts Univ

Sequencing In-Progress

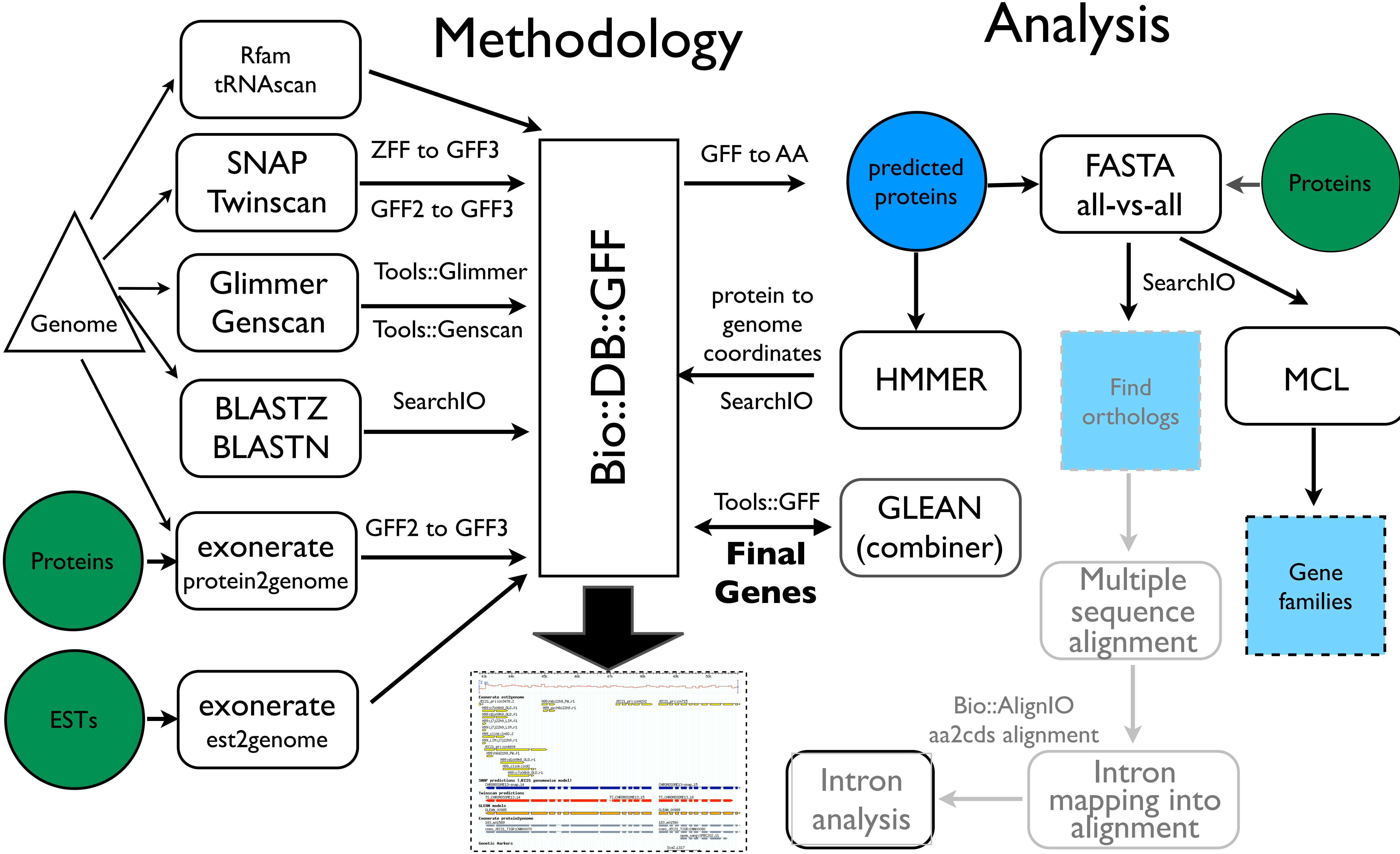
Species	**Clade**	**Sequencing center**
Aspergillus niger	Euascomycota: Eurotiomycota	DOE-JGI
Aspergillus flavus	Euascomycota: Eurotiomycota	NCSU
Aspergillus clavatus	Euascomycota: Eurotiomycota	OU
Neosartorya fischeri	Euascomycota: Eurotiomycetes	TIGR
Histoplasma capsulatum WU24	Euascomycota: Eurotiomycota	Broad-FGI
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
Xanthoria parietina (lichen)	Euascomycota: Lecanoromycetes	DOE-JGI
Candida albicans WO-1	Hemiascomycota	Broad-FGI
Lodderomyces elongisporus	Hemiascomycota	Broad-FGI
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

Genome Annotation

- Genome assembly only for half of genomes
- Developed automated annotation pipeline for fungal genome
- Combine evidence based and *ab initio* gene prediction and species-specific training

Methodology

Analysis



Genome

Rfam
tRNAscan

SNAP
Twinscan

ZFF to GFF3

GFF2 to GFF3

Glimmer
Genscan

Tools::Glimmer

Tools::Genscan

BLASTZ
BLASTN

SearchIO

Proteins

exonerate
protein2genome

GFF2 to GFF3

ESTs

exonerate
est2genome

Bio::DB::GFF

GFF to AA

protein to
genome
coordinates

SearchIO

Tools::GFF

**Final
Genes**

predicted
proteins

HMMER

GLEAN
(combiner)

FASTA
all-vs-all

Proteins

SearchIO

Find
orthologs

MCL

Gene
families

Multiple
sequence
alignment

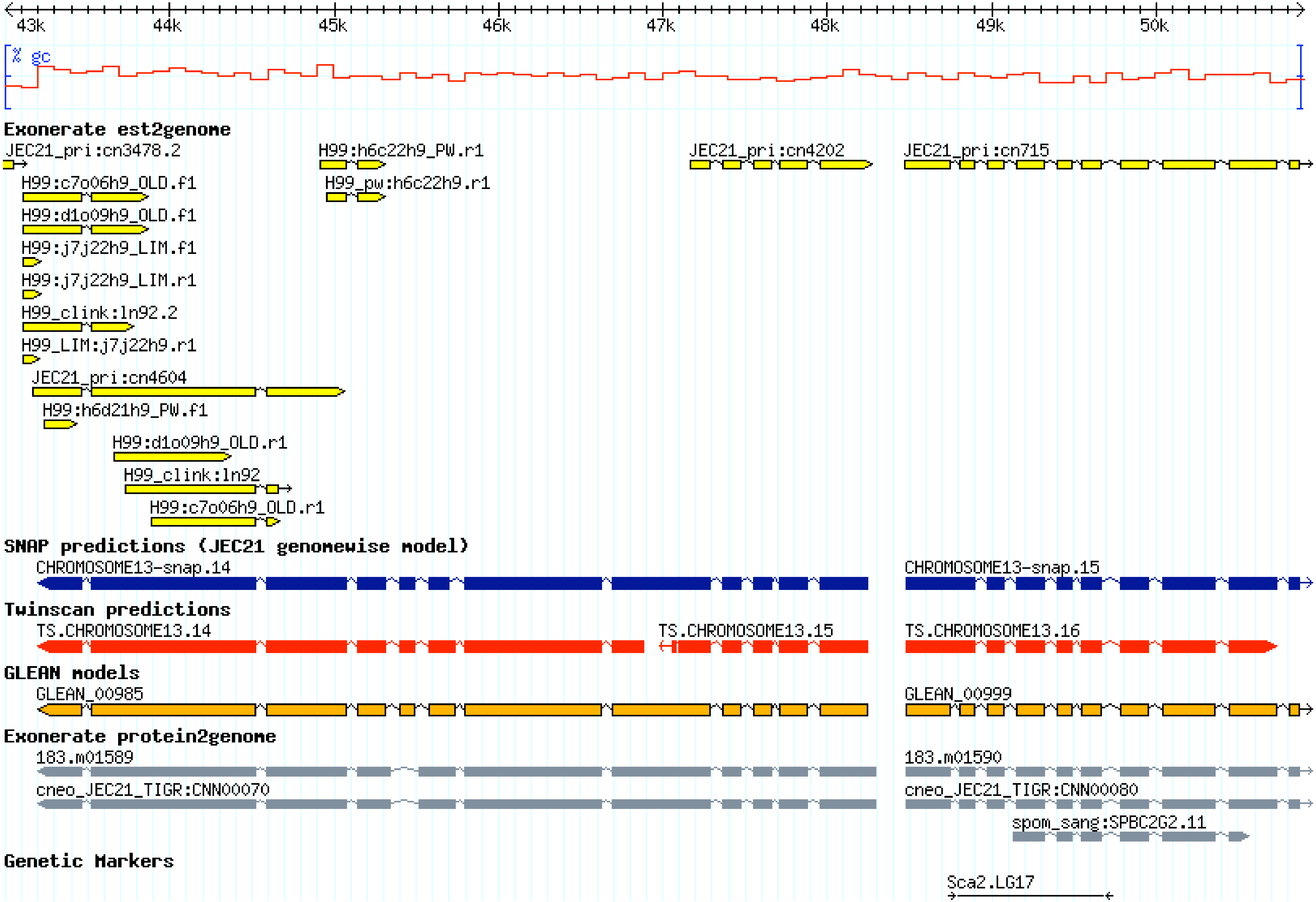
Bio::AlignIO
aa2cds alignment

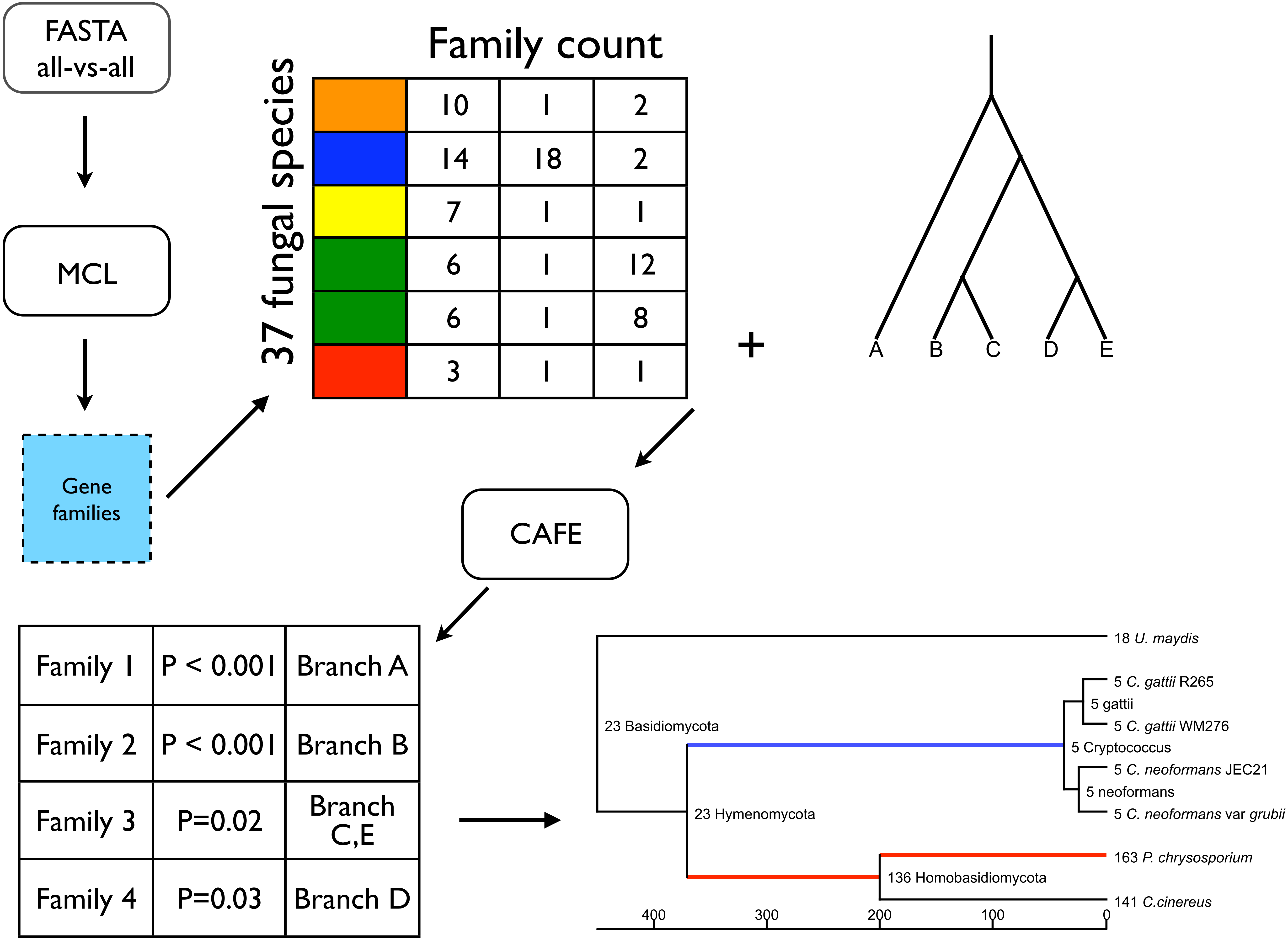
Intron
mapping into
alignment

Intron
analysis

http://fungal.genome.duke.edu

Generic Genome Browser





FASTA all-vs-all

MCL

Gene families

Family count

37 fungal species

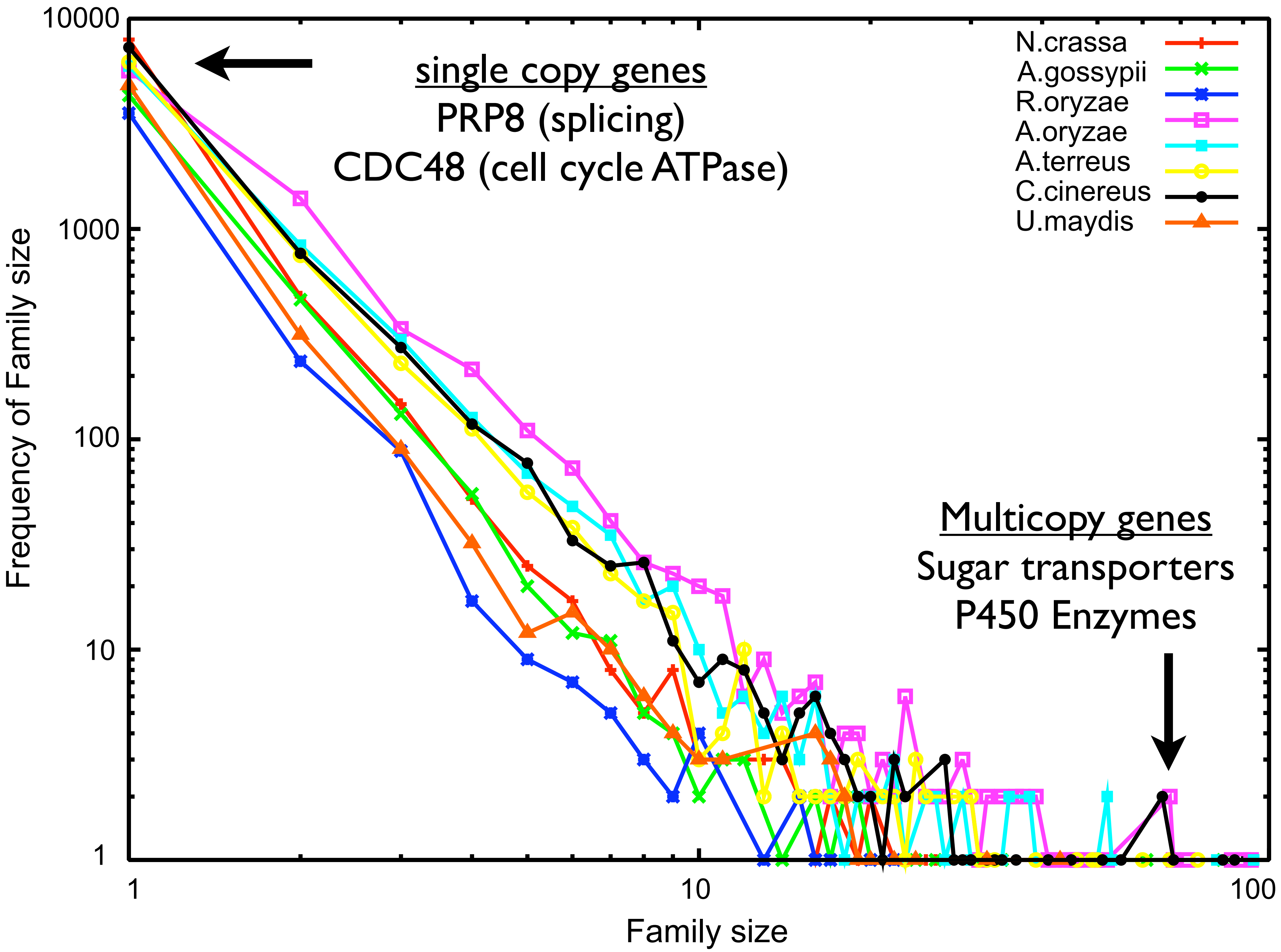
🟧	10	1	2
🟦	14	18	2
🟨	7	1	1
🟩	6	1	12
🟩	6	1	8
🟥	3	1	1

+

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

Gene family sizes follow power law distribution



single copy genes
PRP8 (splicing)
CDC48 (cell cycle ATPase)

Multicopy genes
Sugar transporters
P450 Enzymes

Phylogenetic evaluation of gene family size change

- Previous methods only used *ad hoc* statistics
- Explicit model for gene family size change according to a Birth-Death (BD) 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

- Ancestral states
- Birth and Death rate
- Per branch changes
- P-values

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
- 37 fungal genomes from 5 major clades

Summary of fungal family analysis

- 917 families have at least one member in outgroup present in > 75% of ingroup.
- 47 families had significant expansions and contractions

Families with significant expansions

49 significant families

Transporters
Kinases
P450
Oxidation

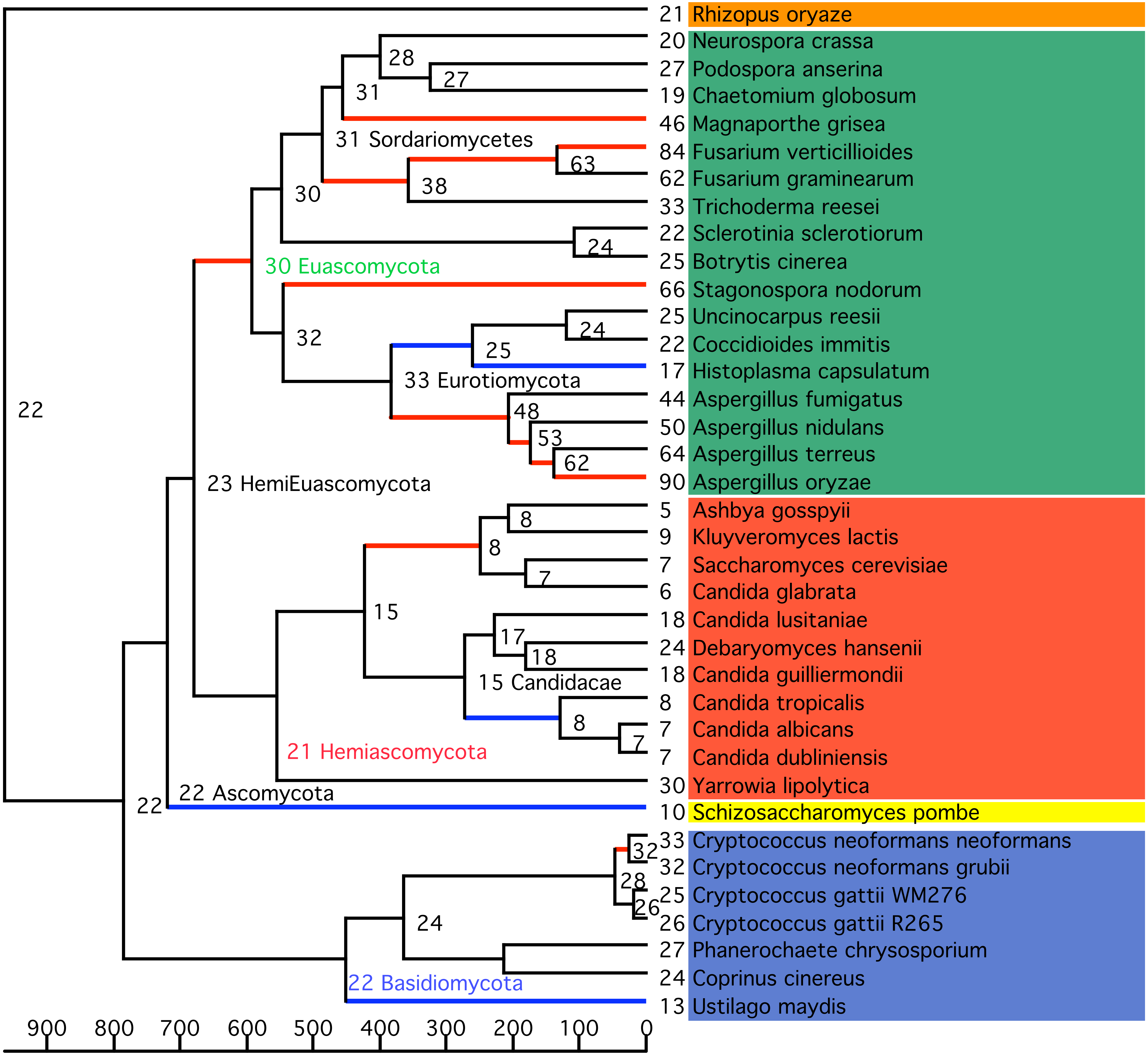
- Vitamin & Cofactor transport
- Lactose & sugar transport
- Amine transport
- Myo-instol, quinate, and glucose transport
- Oligopeptide transport
- ABC transporter
- MFS, drug pump, & sugar transport
- Transport
- Monocarboxylate & sugar transport
- ABC transport
- Amino acid permease

- Methytransferase
- Cytochrome P450: CYP64
- Cytochrome P450: CYP53,57A
- Cytochrome P450
- Kinase
- Subtilase family
- NADH flavin oxidoreductase
- Aldehyde dehydrogenase
- Aldo/kedo reductase
- Multicopper oxidase
- 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)

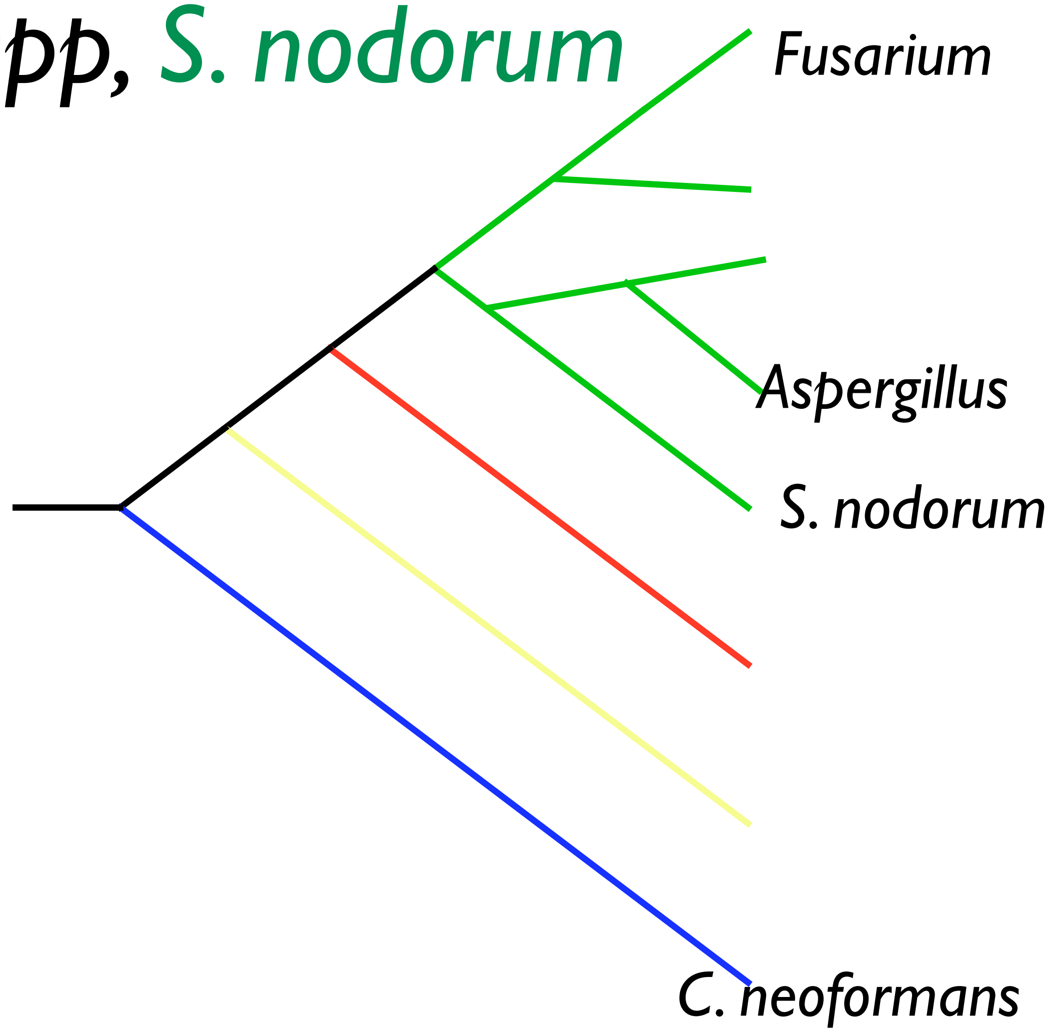
Vitamin & Cofactor Transporters



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

Branches with transporter expansions

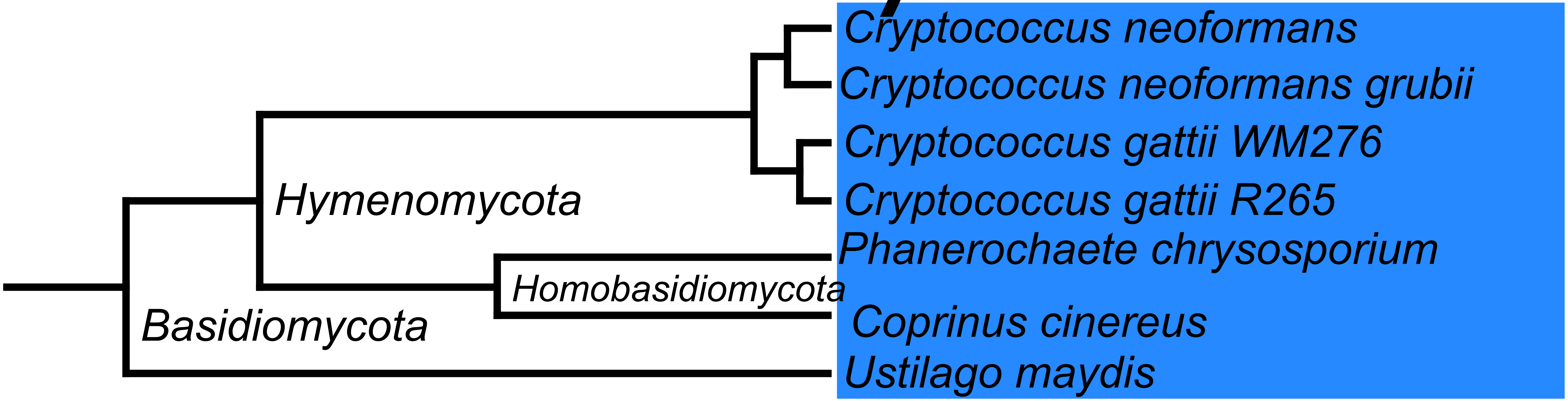
- Sugar related, Drug pump, and Major Facilitator Superfamily
    - *Aspergillus* spp, *Fusarium* spp, *S. nodorum*
    - Euascomycota
- Vitamin transport
    - *C. neoformans*, *Fusarium*
    - *A. nidulans* (Biotin)



Family size contractions

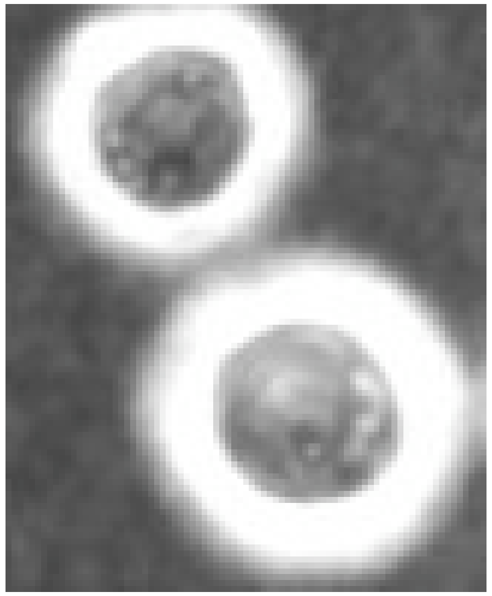
- *Histoplasma, Coccidioides* many families
- Hemiascomycetes - P450
- *C. neoformans* - P450
- *U. maydis* - Lactose transport

Focus on Basidiomycota



Cryptococcus neoformans
Cryptococcus neoformans grubii
Cryptococcus gattii WM276
Cryptococcus gattii R265
Phanerochaete chrysosporium
Coprinus cinereus
Ustilago maydis

Hymenomycota
Homobasidiomycota
Basidiomycota



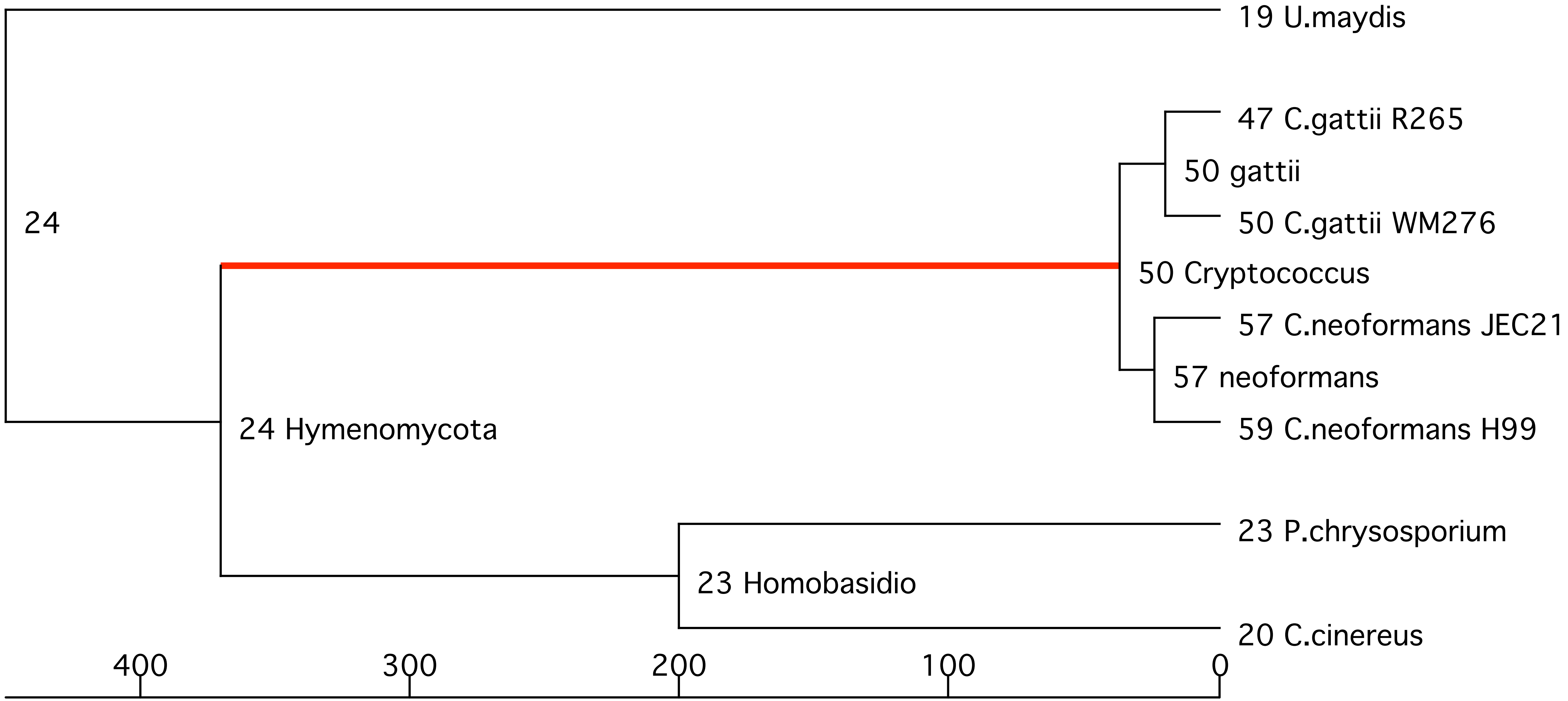
C.neoformans

P.chrysosporium

C.cinereus

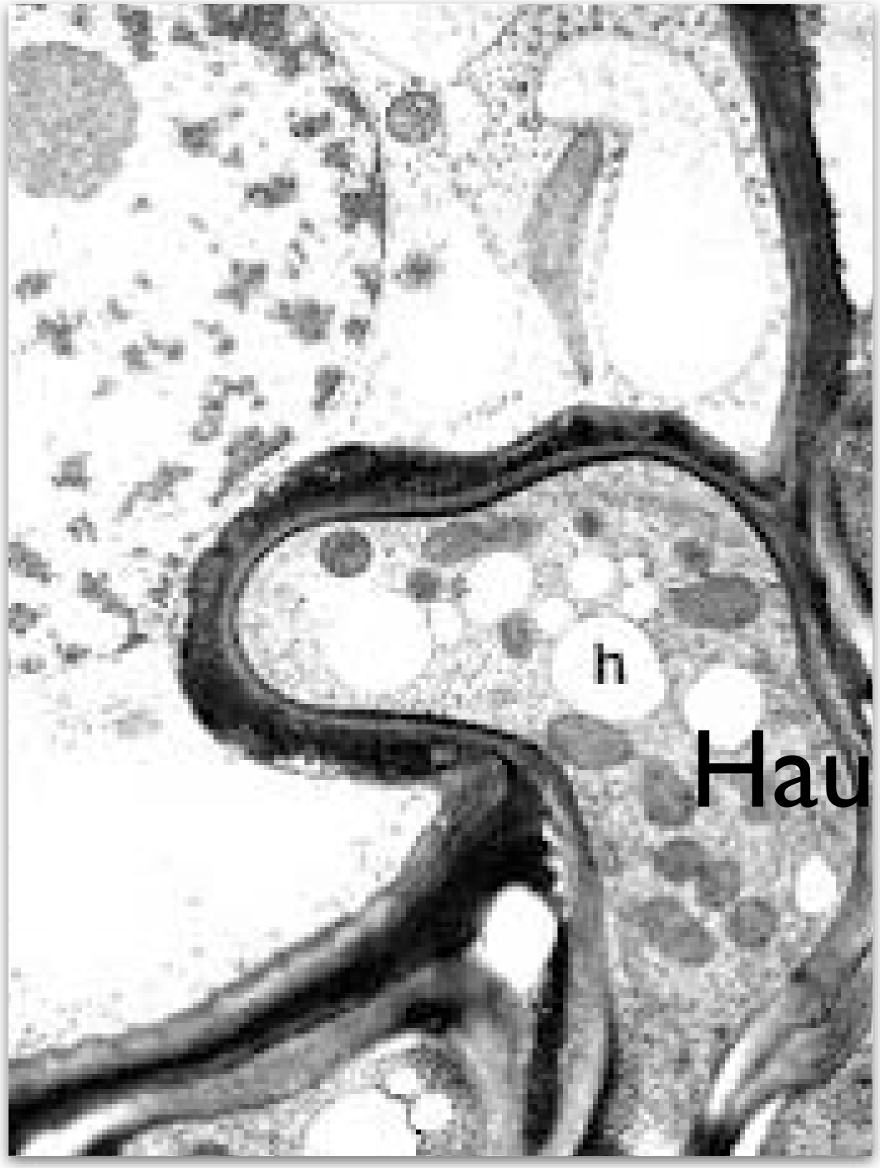
U.maydis

Cryptococcus sugar transporters expansion



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



Robert Bauer http://tolweb.org/

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

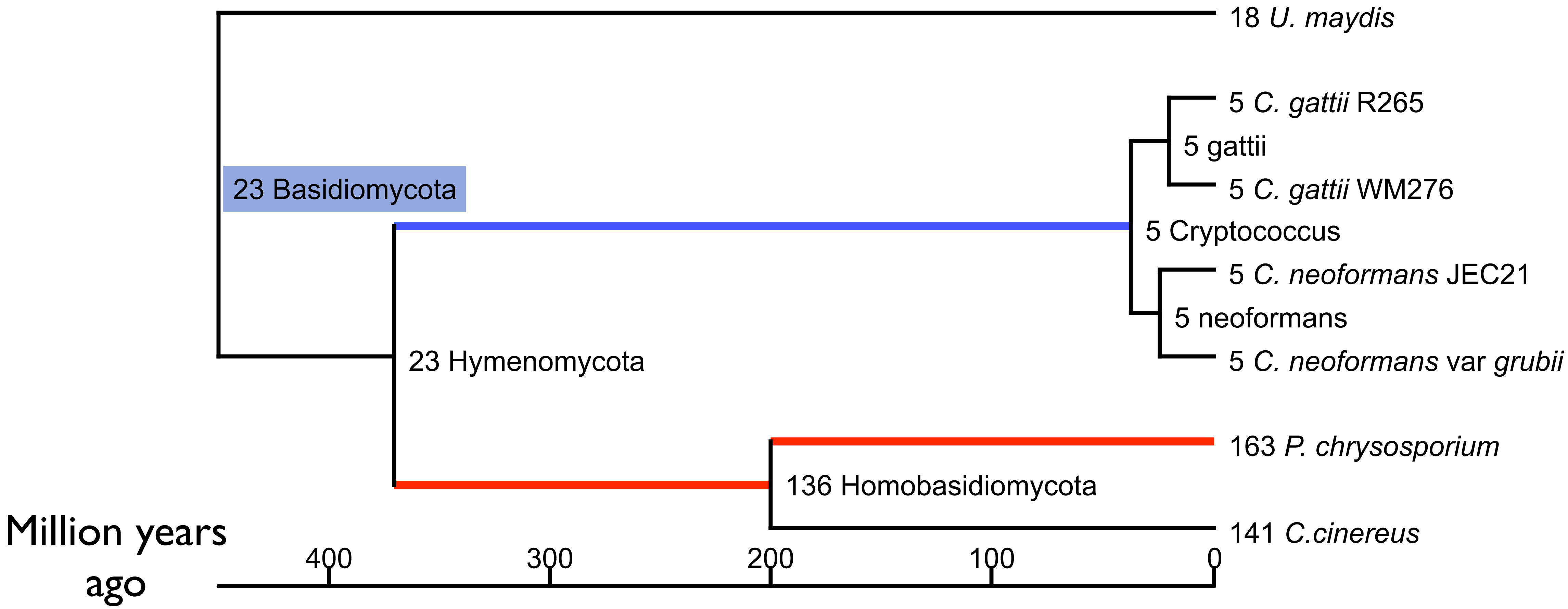
Gene family changes

- P450 (CYP64) - [Homobasidio]
- Hydrophobins - [Homobasidio]
- Monosaccharide metabolism - [Homobasidio]
- Oxidoreductase - [Cryptococcus]

P450 CYP64

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

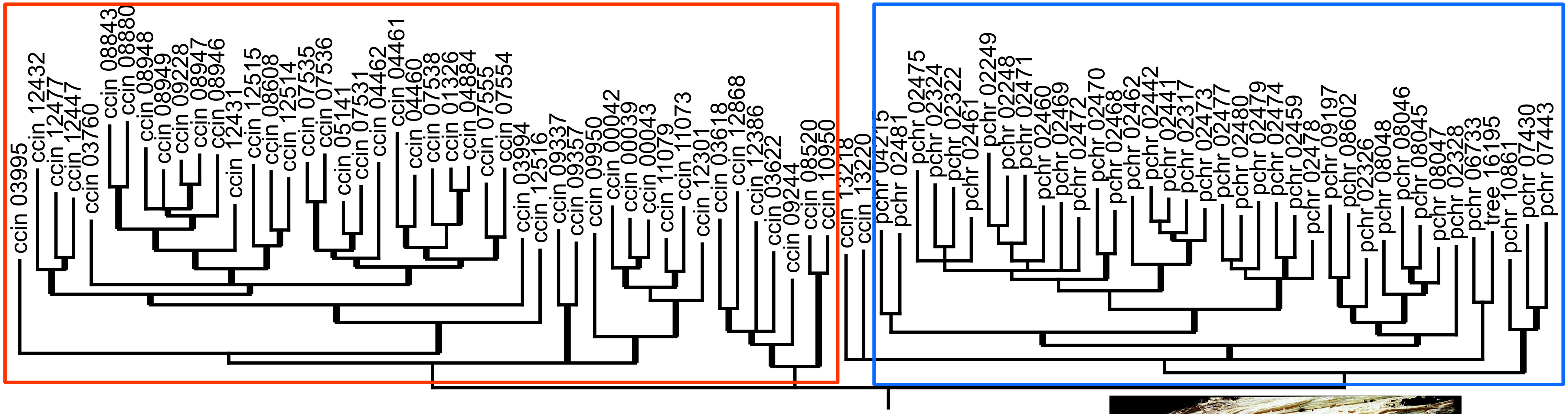
CYP64: Step in *Aspergillus* spp aflatoxin pathway
P. chrysosporium implicated in lignin and hydrocarbon degradation.



CYP64 was from independent duplication

C. cinereus expansion

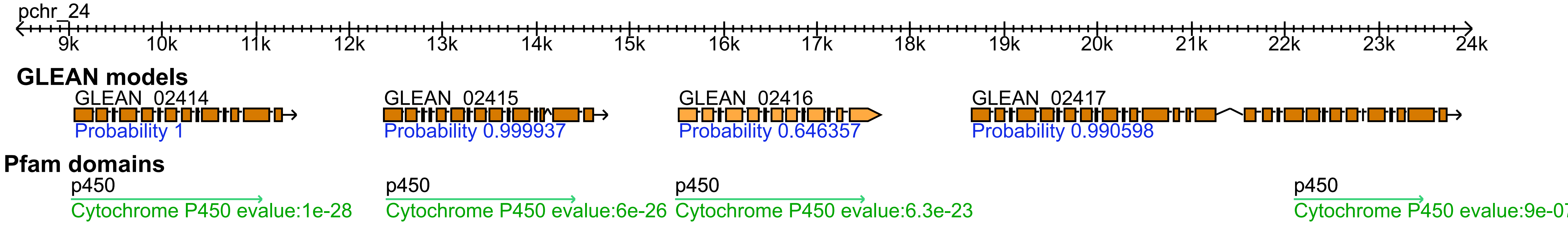
P. chrysosporium expansion



Mario Cervini

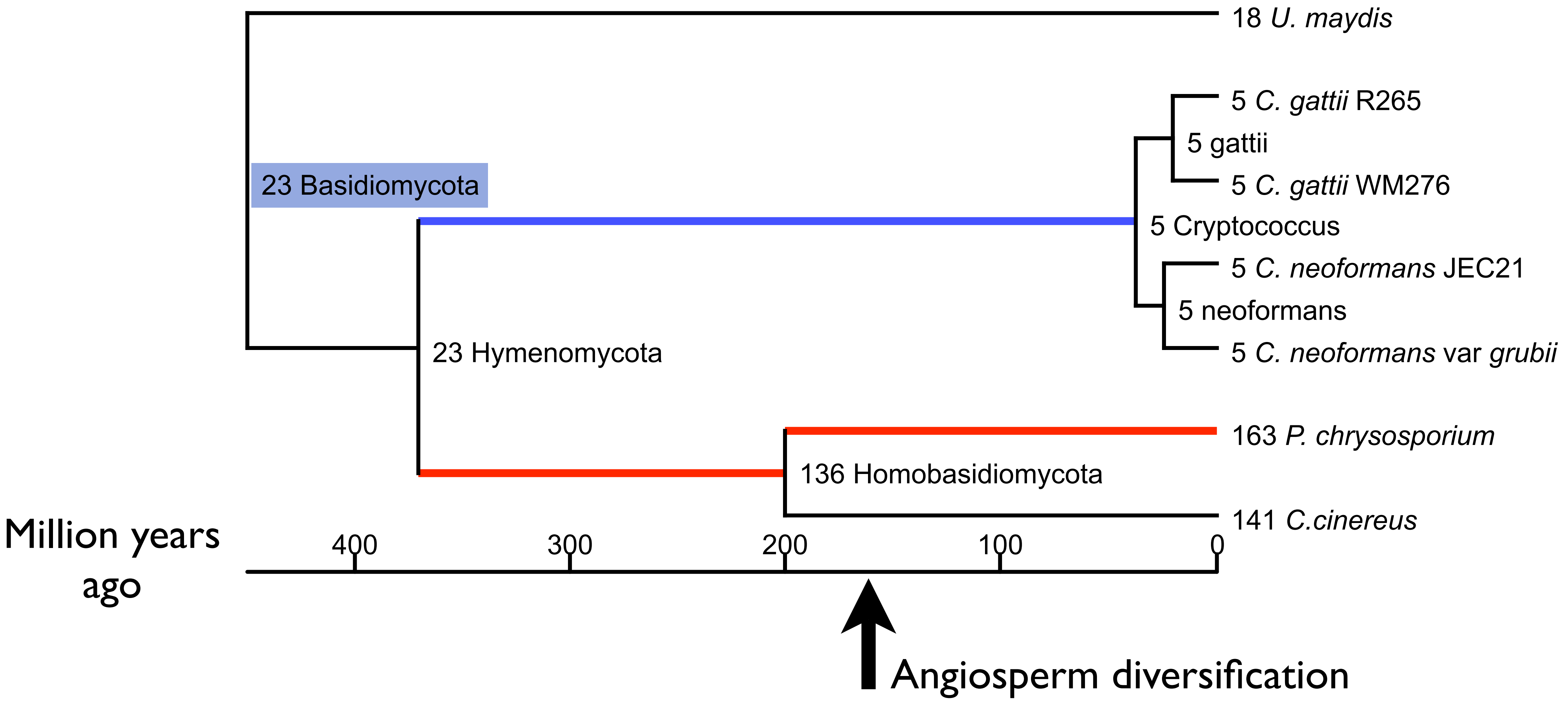
Tom Volk

Local duplications created CYP64 expansion



http://fungal.genome.duke.edu

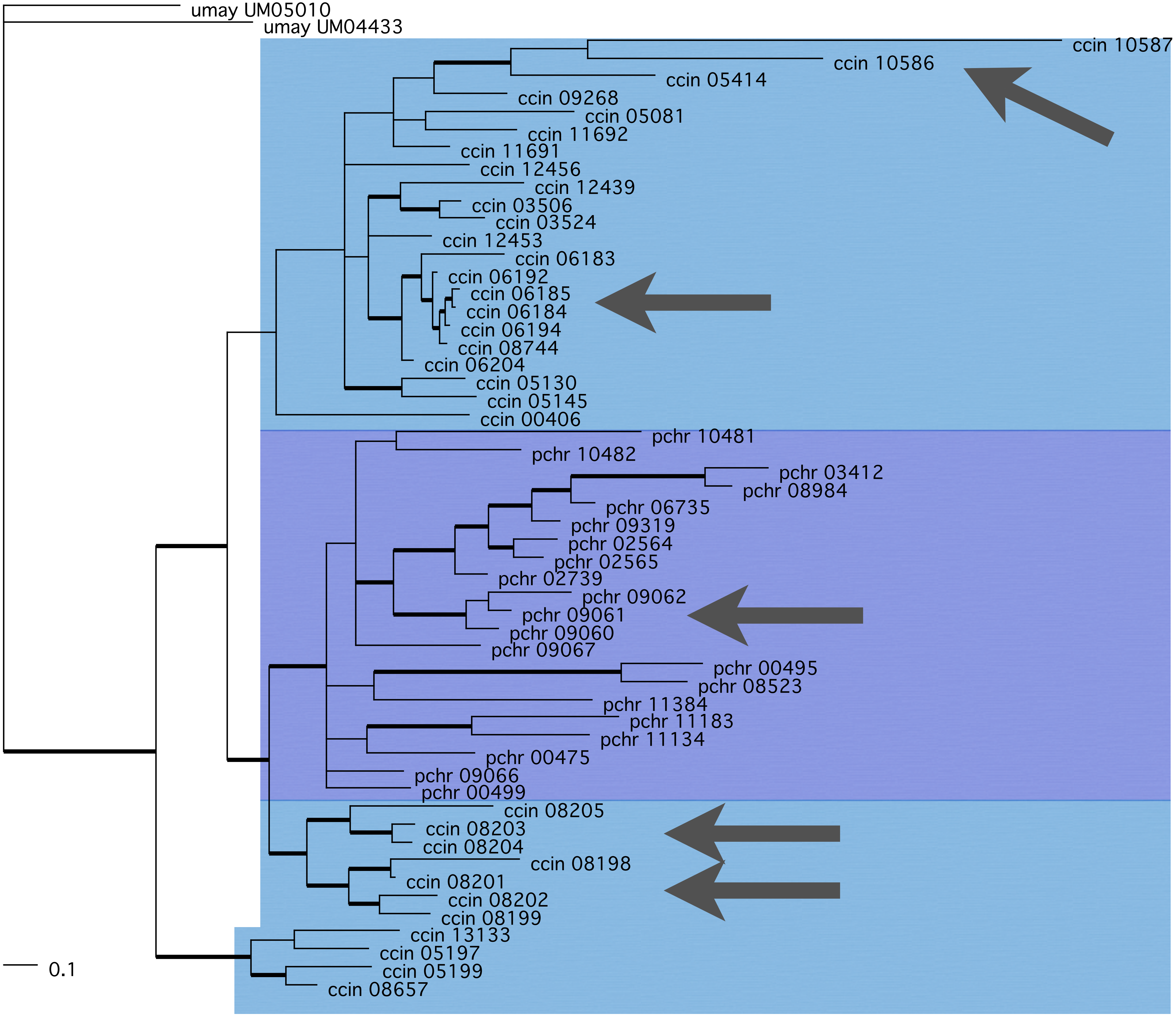
Interpretation of CYP64 expansion



Hydrophobin Family

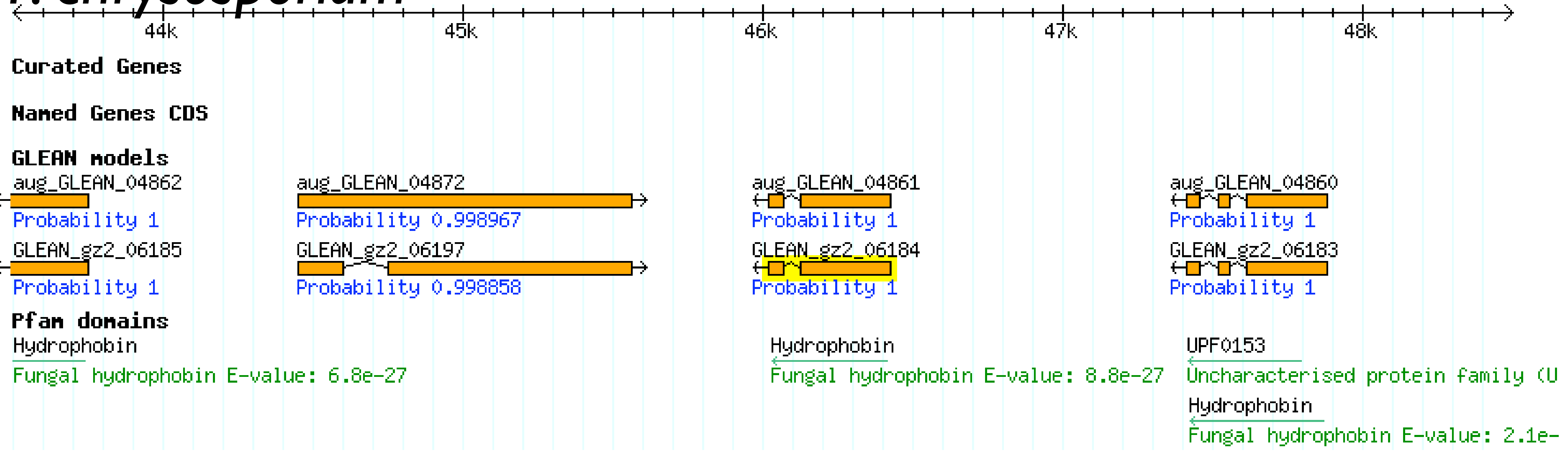
P.chr	*C.cin*	*C.neo*	*U.may*
21	33	0	2

- 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

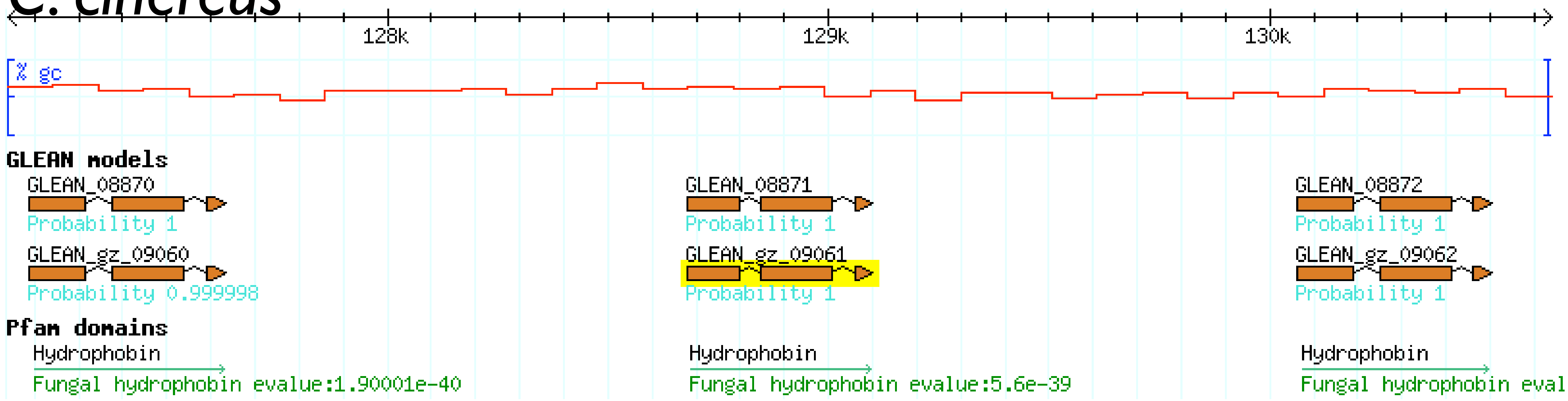

umay UM05010
umay UM04433
ccin 10587
ccin 10586
ccin 05414
ccin 09268
ccin 05081
ccin 11692
ccin 11691
ccin 12456
ccin 12439
ccin 03506
ccin 03524
ccin 12453
ccin 06183
ccin 06192
ccin 06185
ccin 06184
ccin 06194
ccin 08744
ccin 06204
ccin 05130
ccin 05145
ccin 00406
pchr 10481
pchr 10482
pchr 03412
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



Hydrophobin Expansion

- Due to several local duplications
- Expansion is lineage specific
- Important in cell wall construction - mushroom formation

Conclusions

- Sugar transporters are highly expanded in independent lineages
  - Saprophytic and phytopathogenic lifestyles
- P450 CYP64 independent expansions in Homobasidiomycetes
  - Lignin degradation and saprophytic lifestyles
- Family size contractions among lineages containing primary pathogens
  - Genome streamlining?

Future directions

- Confirm contractions in clade of highly pathogenic fungi
- Focus on more clade specific families

- Can we classify fungal lifestyles by family composition?

Acknowledgements

Matthew Hahn

Fred Dietrich

Mario Stanke
Ian Korf
Aaron Mackey

NSF

CSEM

Duke
UPGG

Sequencing centers
Broad Institute
Duke University
Joint Genome Institute
Génolevures
Stanford University
Washington University