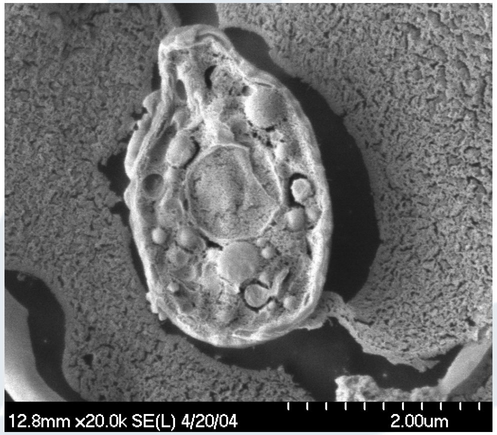
Genus-Wide Comparative Genomics of Malassezia

Phylogeny, Physiology and Niche Adaptation on Human Skin

Niranjan Nagarajan
Associate Director and Group Leader,
Computational & Systems Biology

Skin Mycobiome: *Malassezia*

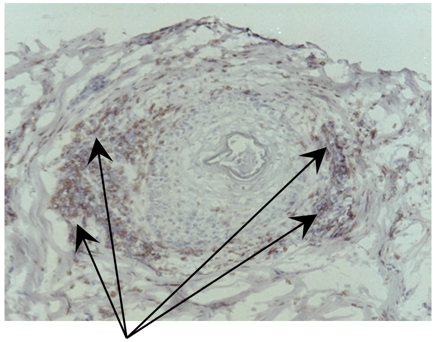


- Belong to a group of largely **plant pathogens** (Ustilagomycotina)
- Strongly **lipophilic** fungi (hard to grow)

Malassezia and Skin Diseases

Dandruff

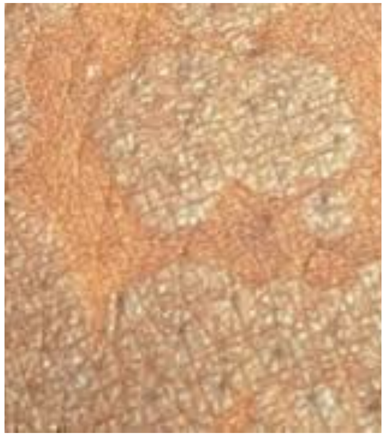
- Antifungal treatments



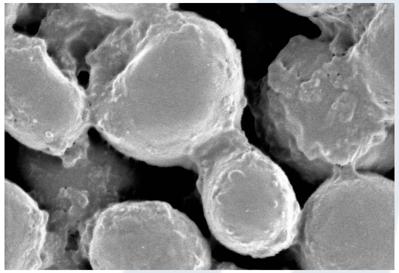
staining for CD4$^+$ immune cells

Seborrhoeic dermatitis

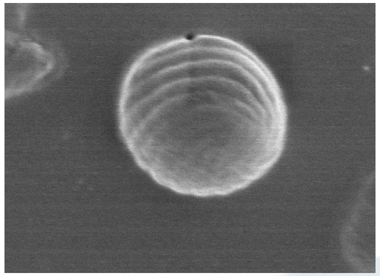
Tinea versicolor



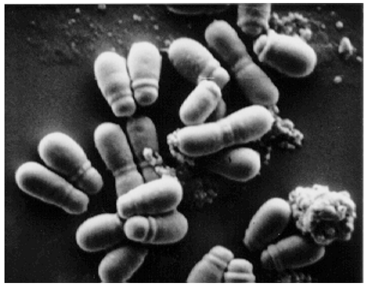
Malassezia Genus



Malassezia globosa



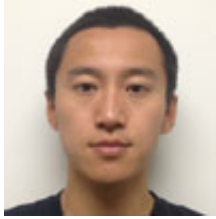
Malassezia restricta



Malassezia furfur

14 *Malassezia* Species

- What are their unique genetic features?
- Which *Malassezia* are associated with disease?



Guangxi

Malassezia Genome Database

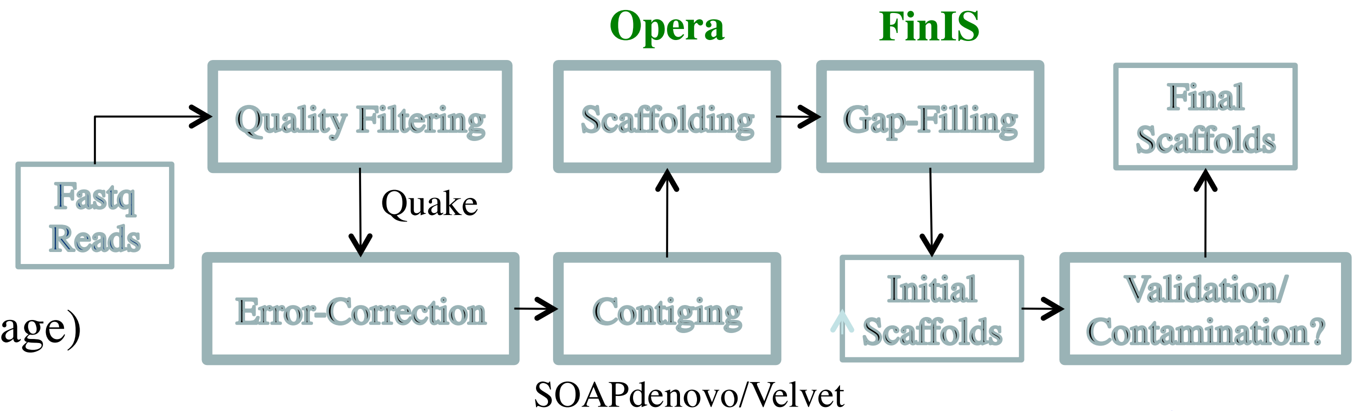
Sample set

(24 strains, all 14 species)

M. caprae 10434, *M. cuniculi* 11721, *M. dermatis* 9169, *M. equina* 9969, *M. furfur* 1878, 4172, 7019, 7710, JPLK23, 7982, *M. globosa* 7990, 7966, 7874, *M. japonica* 9431, *M. nana* 9557, *M. obtusa* 7876, *M. pachydermatis* 1879, *M. restricta* 7877, 8742, *M. slooffiae* 7956, *M. sympodialis* 42132, 44340, 96806, *M. yamatoenis* 9725

Sequencing & Assembly

(>300x median coverage)



Fastq Reads

Quality Filtering

Quake

Error-Correction

Contiging

SOAPdenovo/Velvet

Opera

Scaffolding

FinIS

Gap-Filling

Initial Scaffolds

Validation/ Contamination?

Final Scaffolds

GERMS Pipeline

Assembly Results

99.92% identity, <5 breakpoints per 100 kbp, median N50 of 54 kbp and max. N50 of 1.4 Mbp

Distribution of *Malassezia*

Environmental samples (PLoS Pathogens, 2014)

(ocean water, marine sediments, soil and rhizosphere samples)

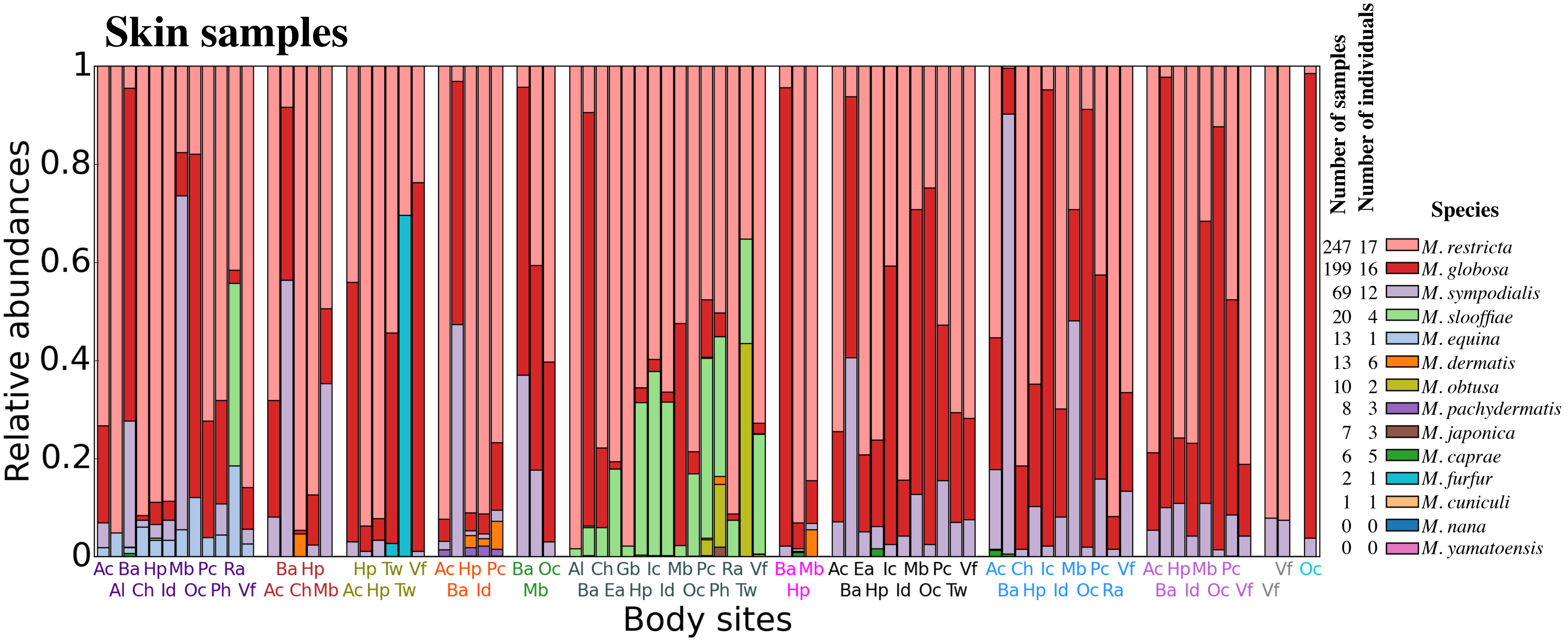
Not Detectable

Oral samples (PLoS ONE 2014)

(keratinized gingiva, buccal mucosa, palatine tonsil, saliva, subgingival plaque, supragingival plaque, throat, tongue dorsum, hard palate)

Not Detectable

Skin samples



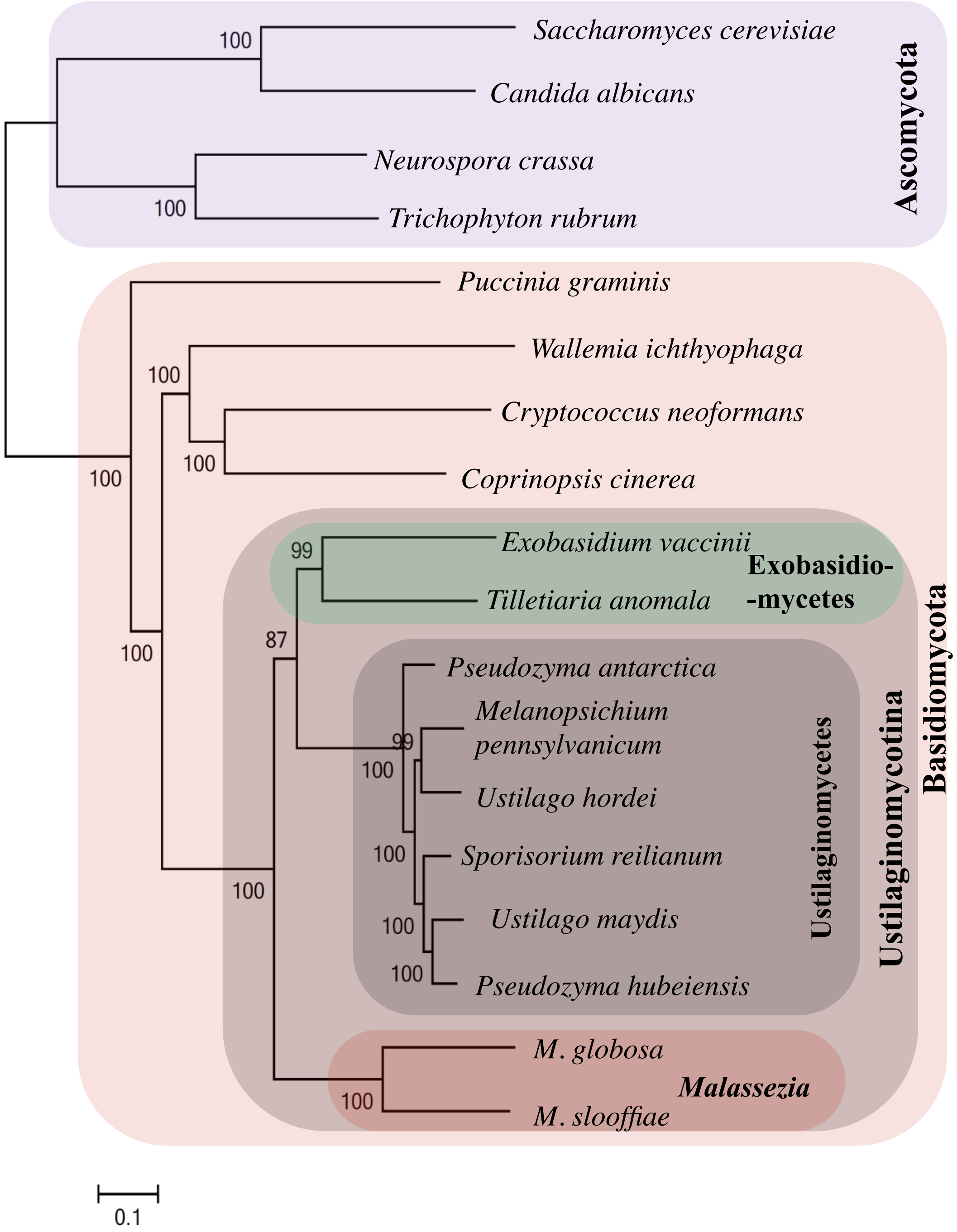
Ac: Antecubital Crease	**Al**: Alar Crease	**Ba**: Back	**Ch**: Cheek	**Id**: Interdigital Web Space	**Mb**: Manubrium	**Oc**: Occiput	**Pc**: Popliteal Crease
Ea: External Auditory Canal	**Gb**: Glabella	**Hp**: Hypothenar Palm	**Ic**: Inguinal Crease	**Ph**: Plantar Heel	**Ra**: Retroauricular Crease	**Tw**: Toeweb Space	**Vf**: Volar Forearm

What makes *Malassezia* unique?

Compact genome: 4,000-5,000 genes

Except for 4 *M. furfur* strains: ~10,000 genes

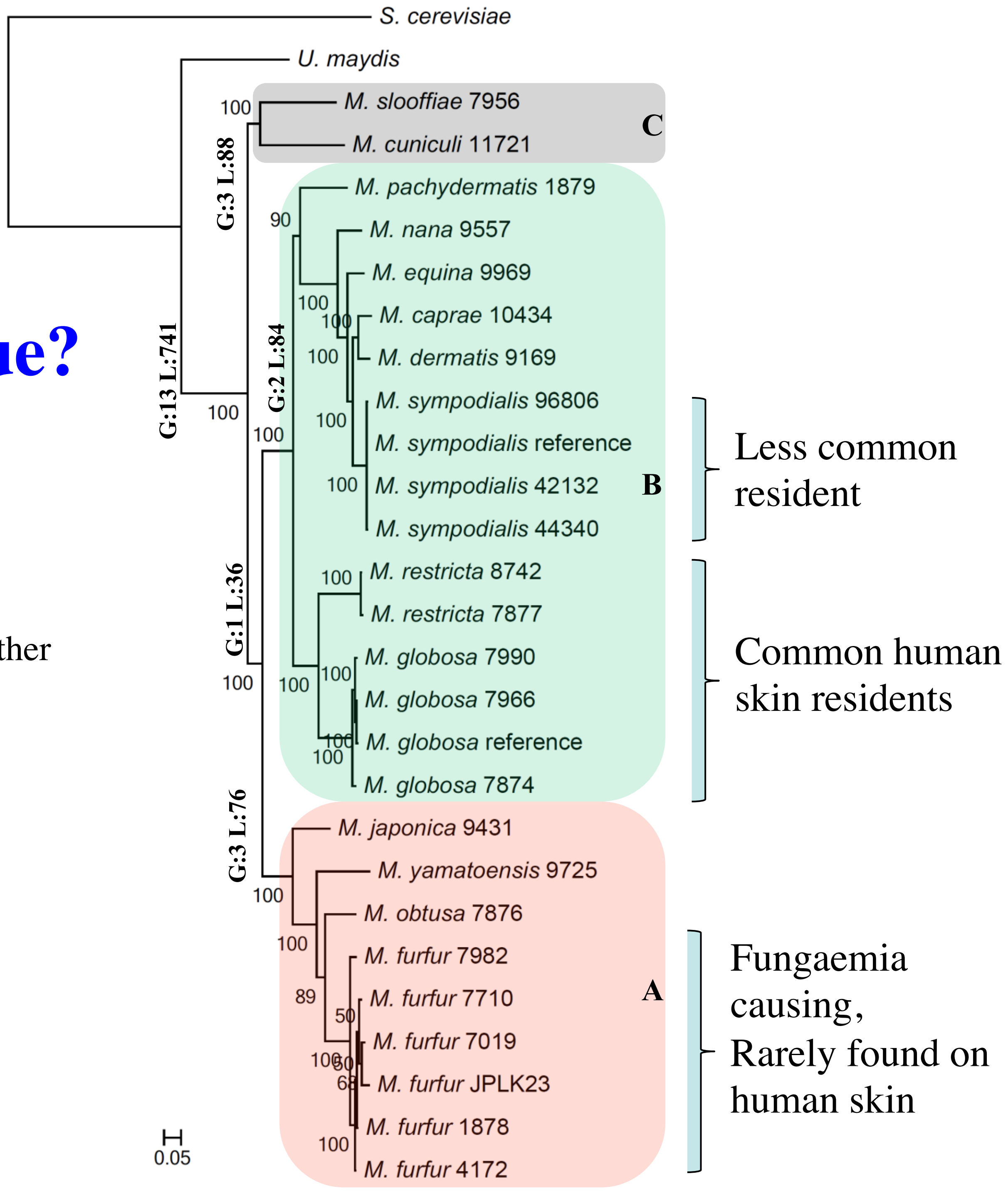
- potential hybrids



What makes *Malassezia* unique?

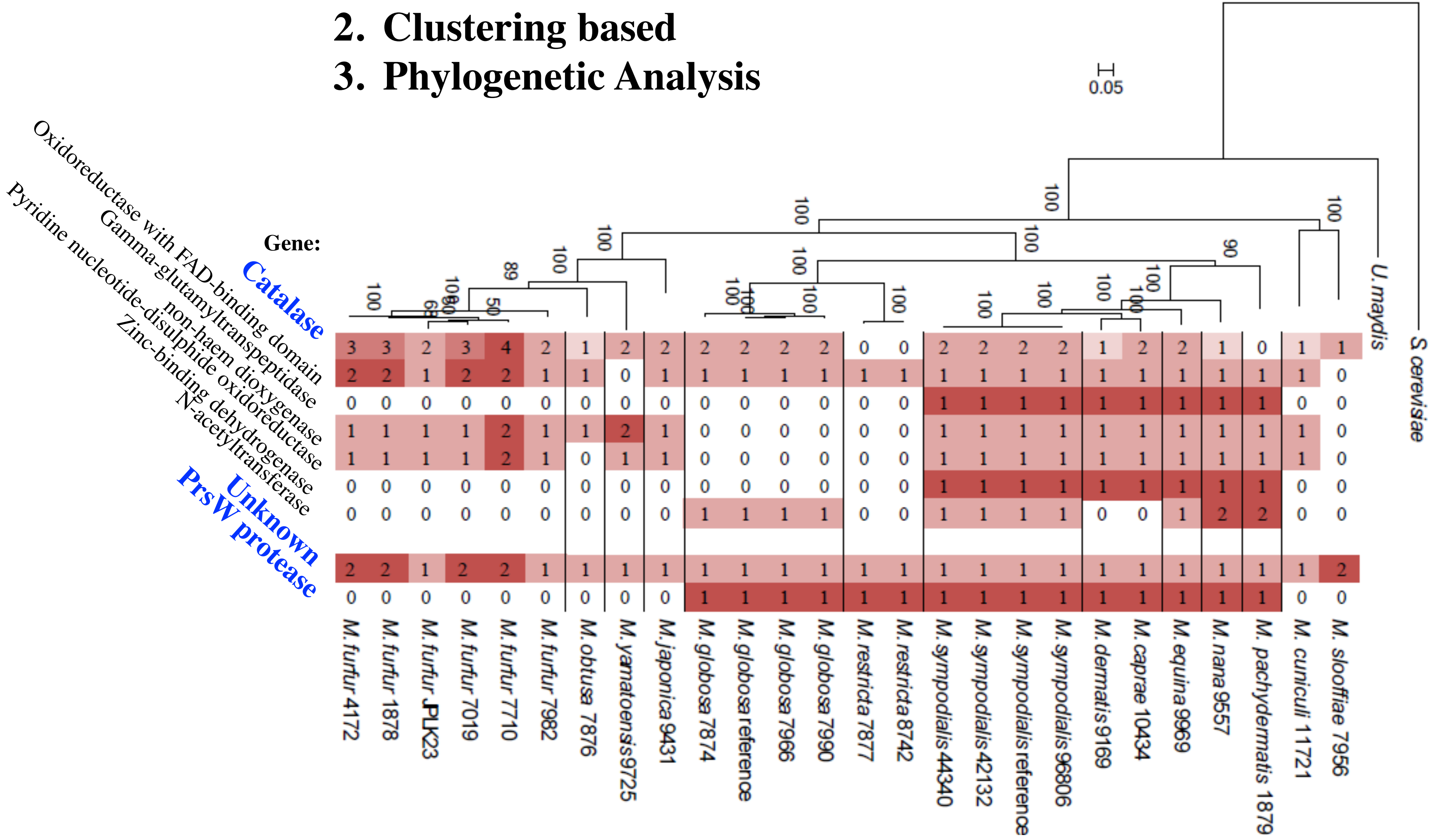
Comparison to a panel of 16 fungal genomes

- All Ustilagomycotina, other Basidiomycetes and Ascomycetes
- **64 gene gains, >700 genes lost**



Horizontal Gene Transfers

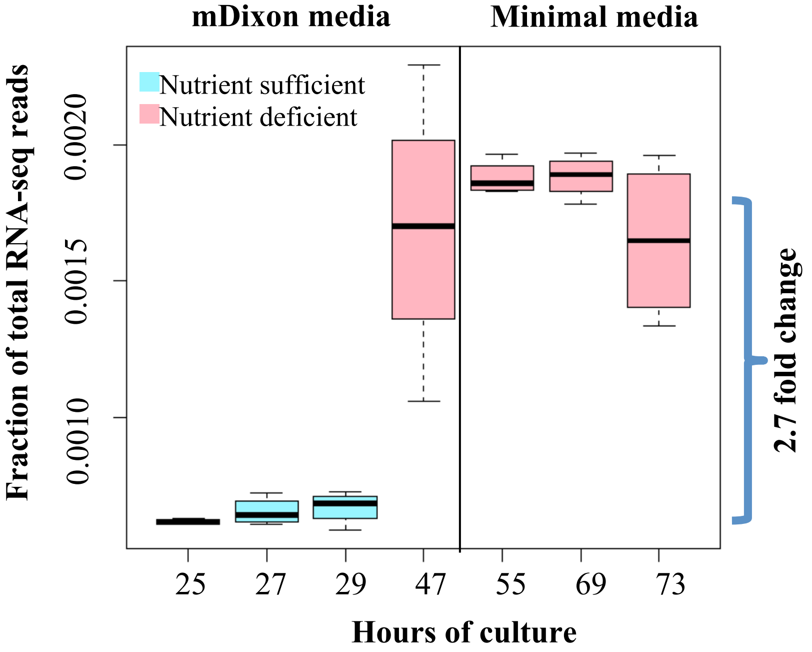
1. **PFam Analysis**
2. **Clustering based**
3. **Phylogenetic Analysis**

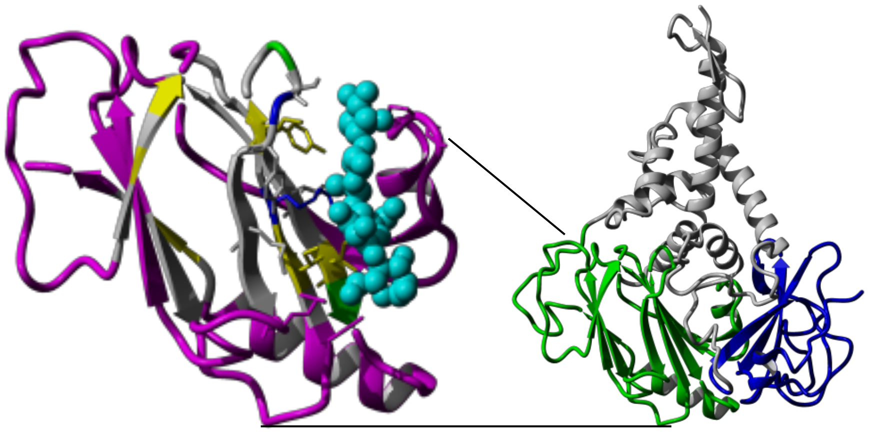


Gene:	M.furfur 4172	M.furfur 1878	M.furfur JPLK23	M.furfur 7019	M.furfur 7710	M.furfur 7982	M.obtusa 7876	M.yamatoensis 9725	M.japonica 9431	M.globosa 7874	M.globosa reference	M.globosa 7966	M.globosa 7990	M.restricta 7877	M.restricta 8742	M.sympodialis 44340	M.sympodialis 42132	M.sympodialis reference	M.sympodialis 96806	M.dermatis 9169	M.caprae 10434	M.equina 9969	M.nana 9557	M. pachydermatis 1879	M. cuniculi 11721	M. slooffiae 7956
Catalase	3	3	2	3	4	2	1	2	2	2	2	2	2	0	0	2	2	2	2	1	2	2	1	0	1	1
Oxidoreductase with FAD-binding domain	2	2	1	2	2	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
Gamma-glutamyltranspeptidase	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	0	0
non-haem dioxygenase	1	1	1	1	2	1	1	2	1	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	0
Pyridine nucleotide-disulphide oxidoreductase	1	1	1	1	2	1	0	1	1	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	0
Zinc-binding dehydrogenase	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	0	0
N-acetyltransferase	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	1	1	1	1	0	0	1	2	2	0	0
Unknown	2	2	1	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2
PrsW protease	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0

PF06742: domain of unknown function

- Found in **pathogenic bacteria** (*e.g. Mycobacterium tuberculosis*, *Listeria monocytogenes* and *Salmonella enterica*) and fungi (e.g. *Aspergillus flavus*)
- Ortholog in *Chlamydomonas reinhardtii* is dramatically **up-regulated under sulfur depletion** conditions

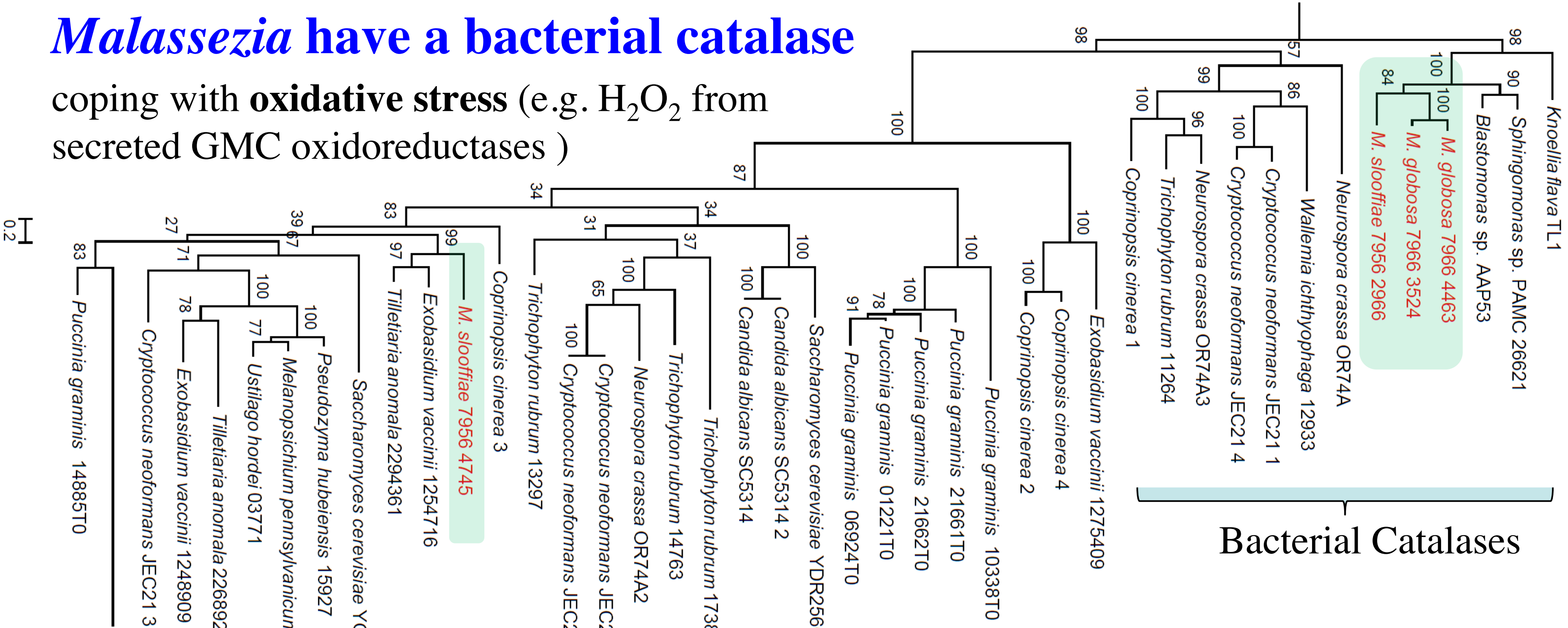




- **Secreted protein** (signal peptide match)
- Based on homology modelling, likely a **glycosyl hydrolase** (EC 3.2.1.x)
- Could modulate beta-glucan levels in the fungal cell wall (**cross-talk with immune system**)

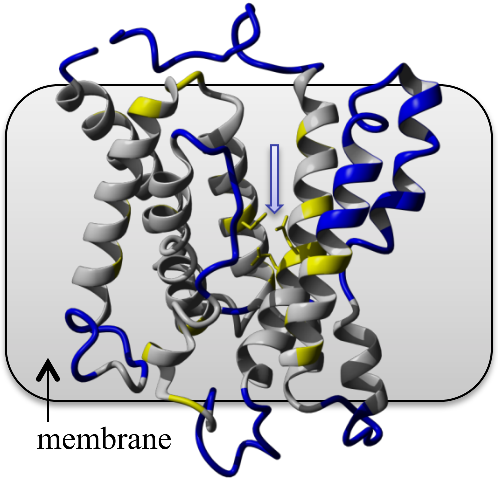
Malassezia have a bacterial catalase

coping with **oxidative stress** (e.g. H_2O_2 from secreted GMC oxidoreductases)



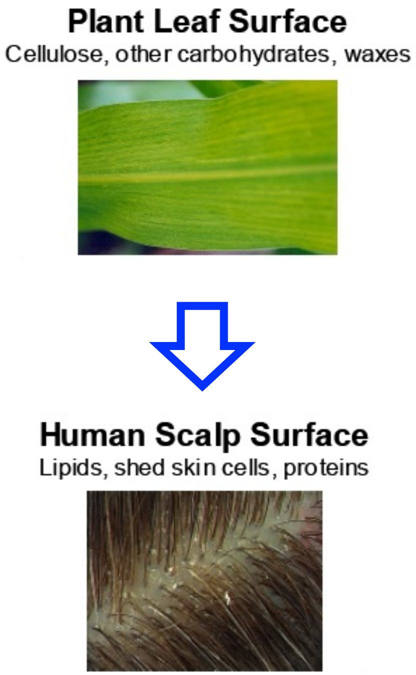
PF13367 found only in cluster B

and **in skin resident bacteria** (e.g. *Propionibacterium*, *Streptococcus* and *Staphylococcus*)



Peptidase - homology to proteins that cleave anti-microbial peptides

Gene loss in *Malassezia* (741)
- adapting to life on skin



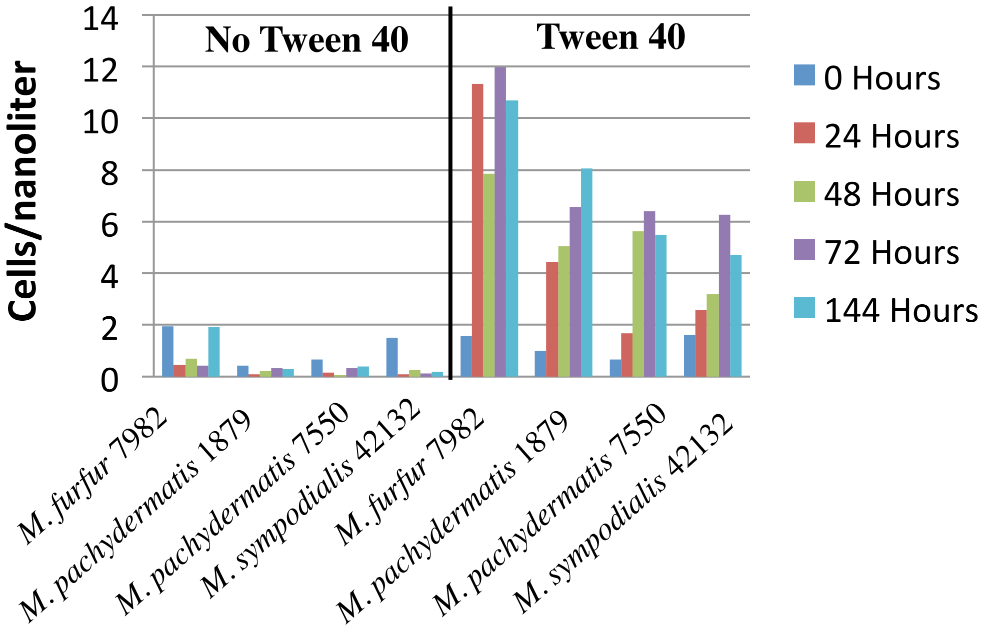
Plant Leaf Surface
Cellulose, other carbohydrates, waxes

Human Scalp Surface
Lipids, shed skin cells, proteins

Enrichment for enzymes involved in **carbohydrate metabolic process** (q-value < 4.5×10^{-4}) and in **hydrolysis activity** (hydrolyzing O-glycosyl compounds; q-value < 4.5×10^{-4})

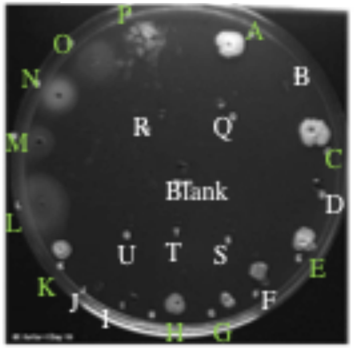
Fatty acid synthase (FAS) missing in all *Malassezia*

=> genus is lipid-dependent and not just lipophilic

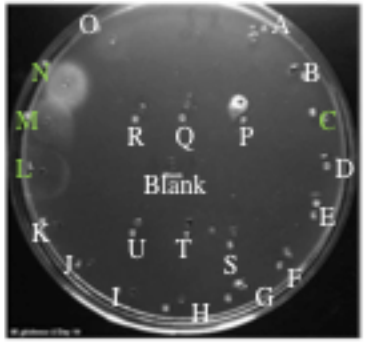


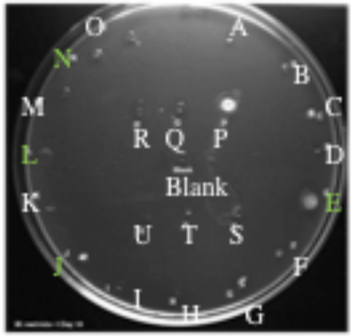
Malassezia have distinct lipid preferences

M. furfur

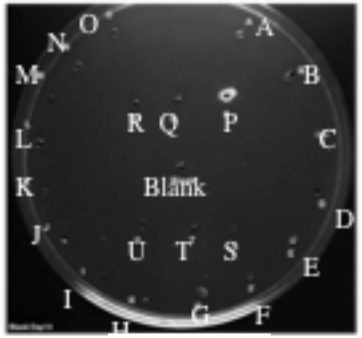


M. globosa





M. restricta

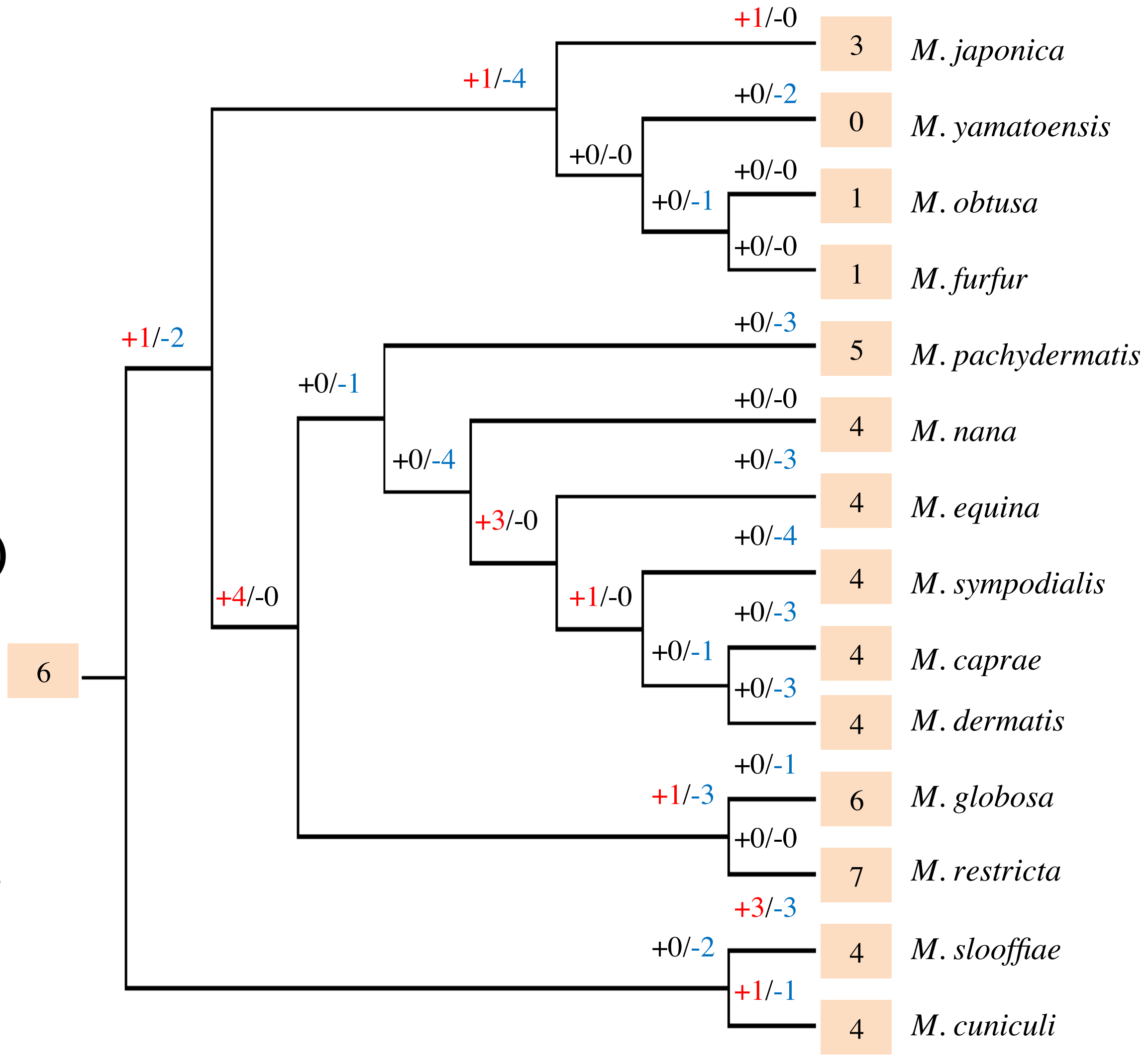


Blank

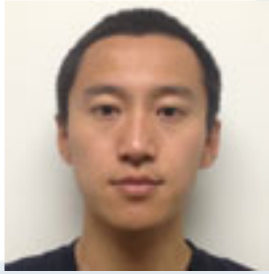
Compounds tested:	A. Olive oil	B. Propylene glycol	C. Artificial sebum	D. Triolein	E. Wheat germ oil	F. Squaline	G. Lard	H. Coconut oil	I. 0.1 % Rincinoleic	J. 0.1% Stearic	K. 0.1% Palmitic
	L. Tween 20	M. Tween 40	N. Tween 60	O. Tween 80	P. Tween 85	Q. 0.1% Oleic A (74%)	R. 0.1% Oleic A (Pure)	S. Linoleic	T. 0.1% Petroselinic	U. Palmitoleic	Blank. Cells with no lipid

Extensive turnover in Lipases

- **PF04185** (phospholipase)
- PF03583 (secretory lipase)
- PF01764 (other lipases)



"Genus-wide comparative genomics of *Malassezia* delineates its phylogeny, physiology, and niche adaptation on human skin", Guangxi Wu *et al.*, PLoS Genetics 2015



Wu Guangxi
Nagarajan Lab, GIS

Tom Dawson
Institute of Medical
Biology, A*STAR

Sebastian Maurer-Stroh
Bioinformatics Institute, A*STAR

P&G

Agency for
Science, Technology
and Research
SINGAPORE

http://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1005614


Genome