

Using and developing bioinformatics in deep level phylogenomic reconstructions and Development

Matthew W. Brown
Donald L. Hall Professor of Biology

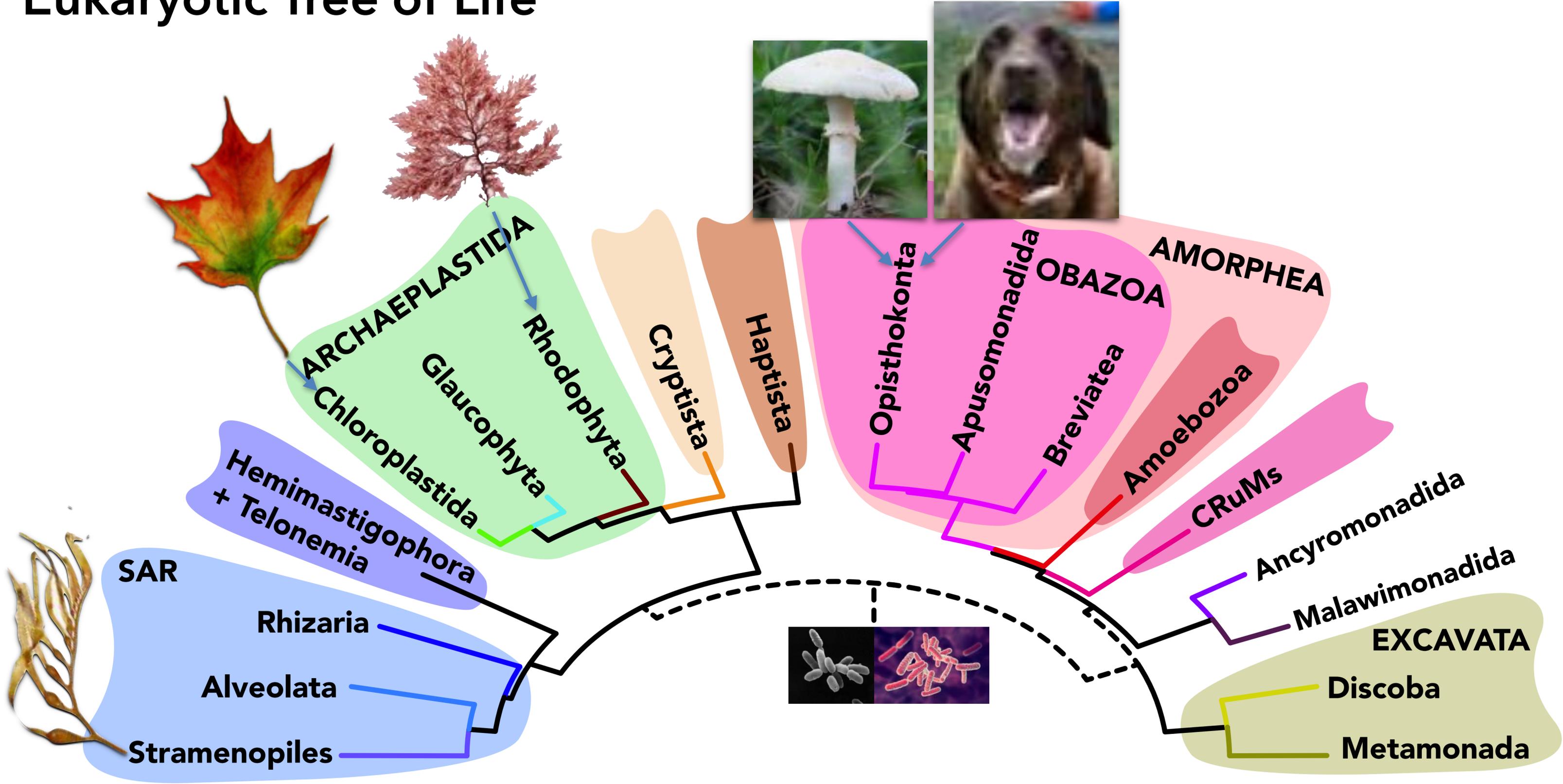
Fellow - High Performance Computing Collaboratory (HPC²)

Department of Biological Sciences, Mississippi State University

<http://amoeba.msstate.edu>

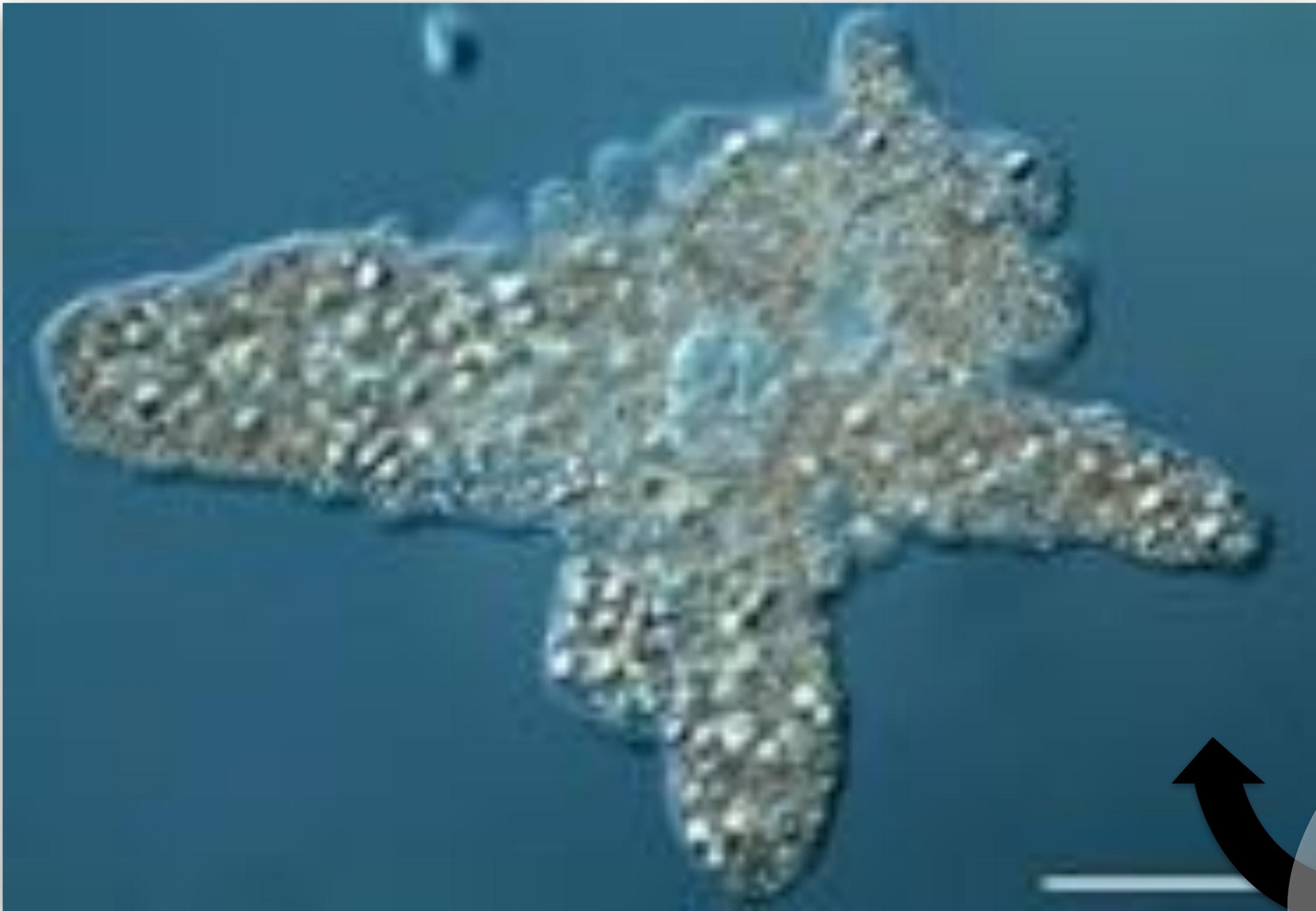


Eukaryotic Tree of Life

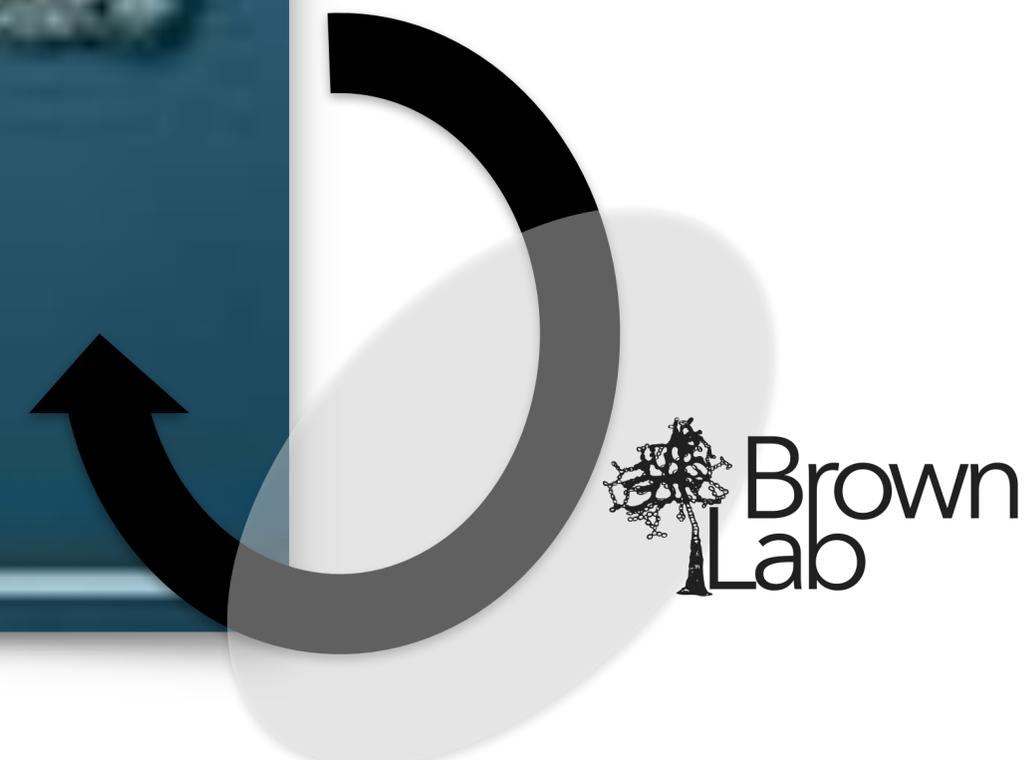


Burki et al. 2021 *TREE* | Tice et al. 2021 *PLoS Biology* | redrawn from Tice & Brown 2022. *Current Biology*

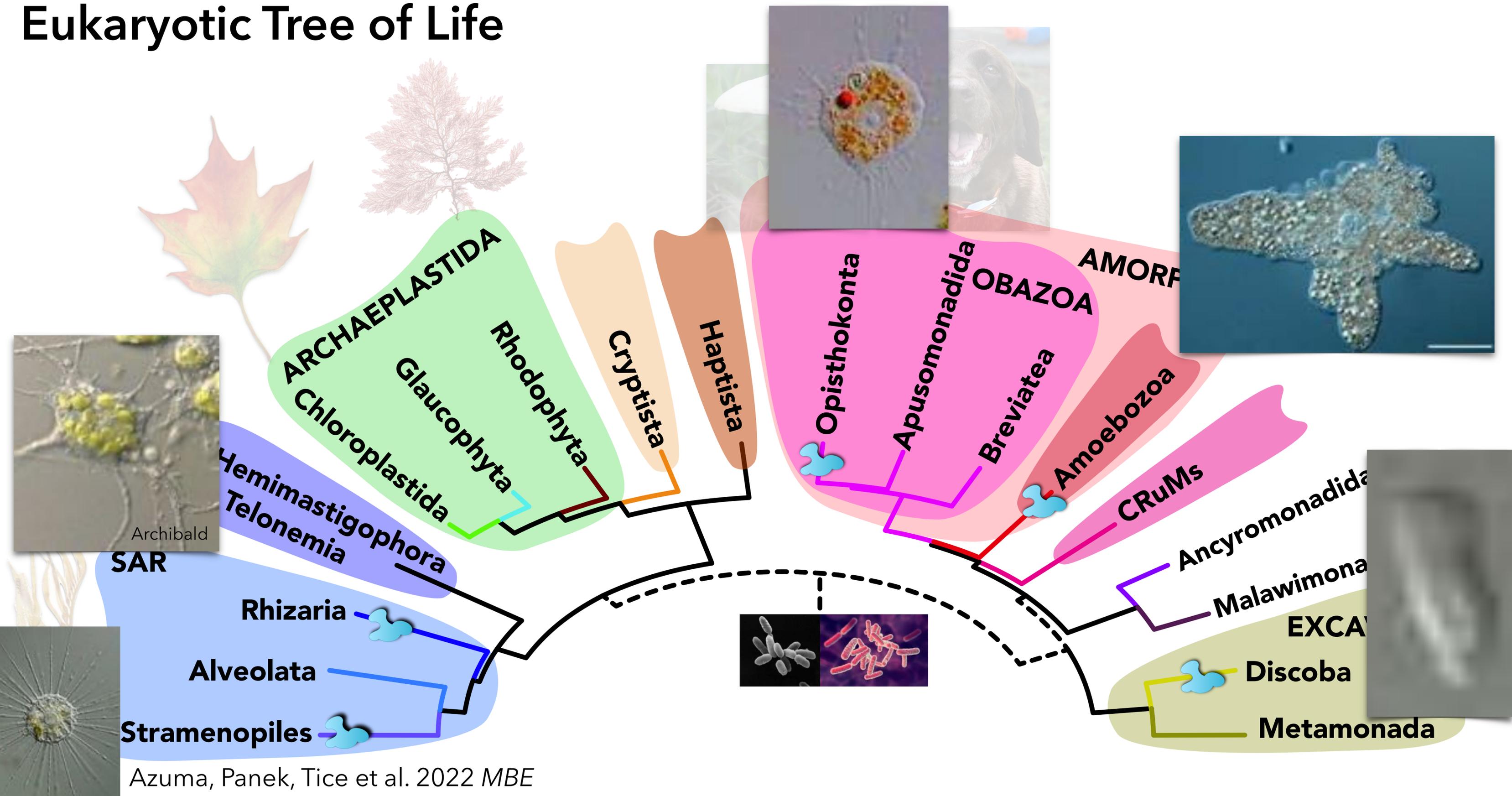
Amoeboid protists



Amoeba proteus



Eukaryotic Tree of Life

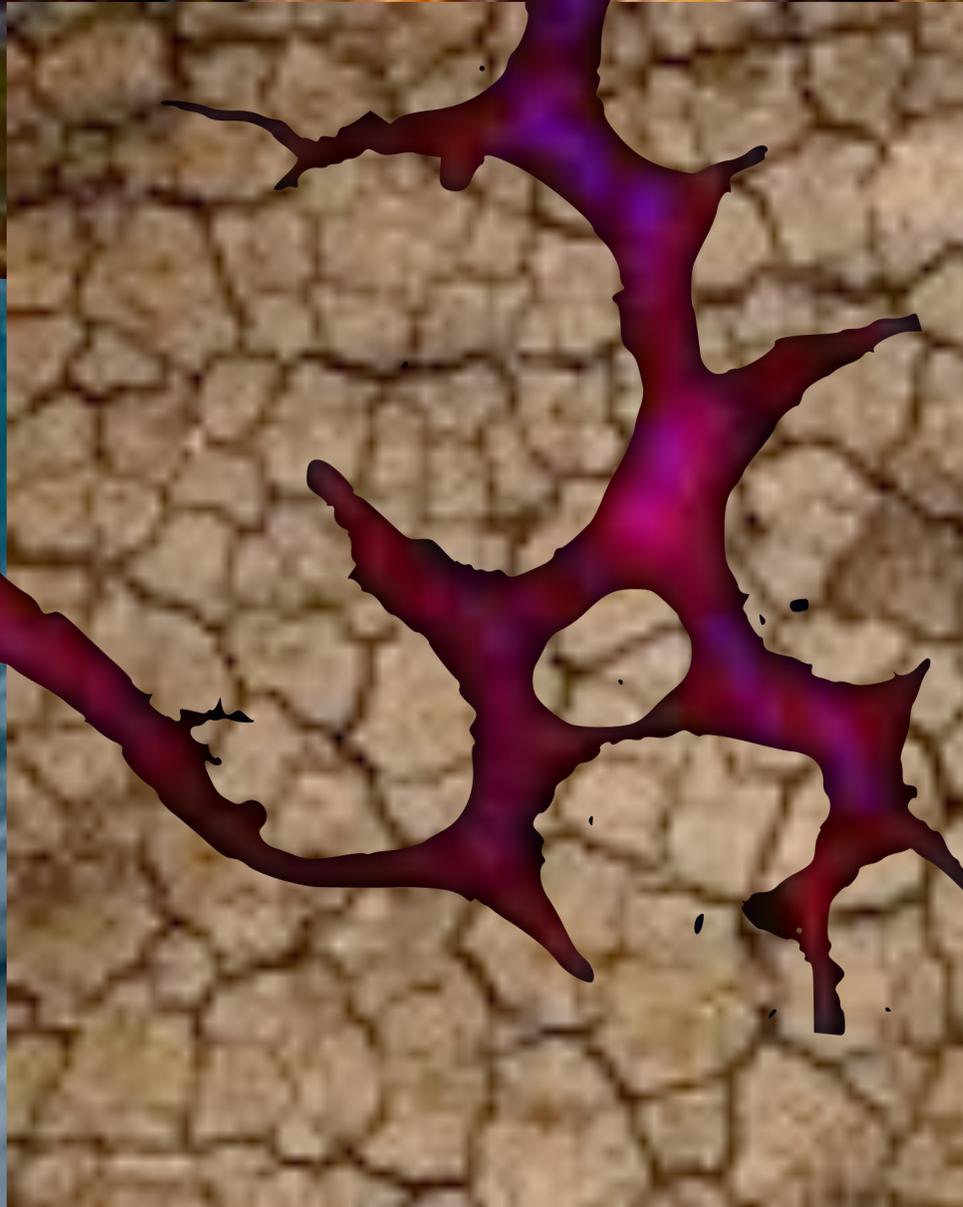
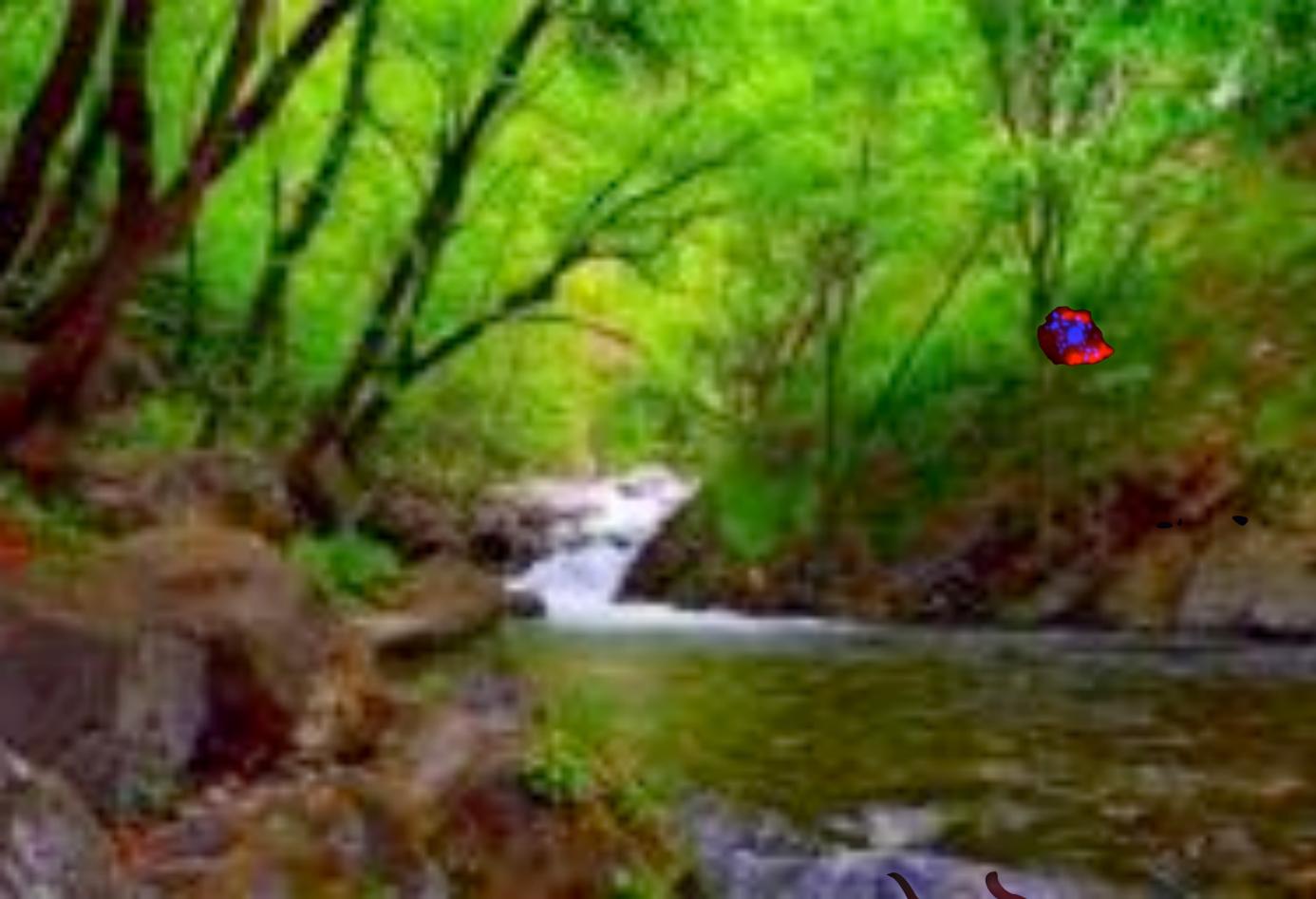


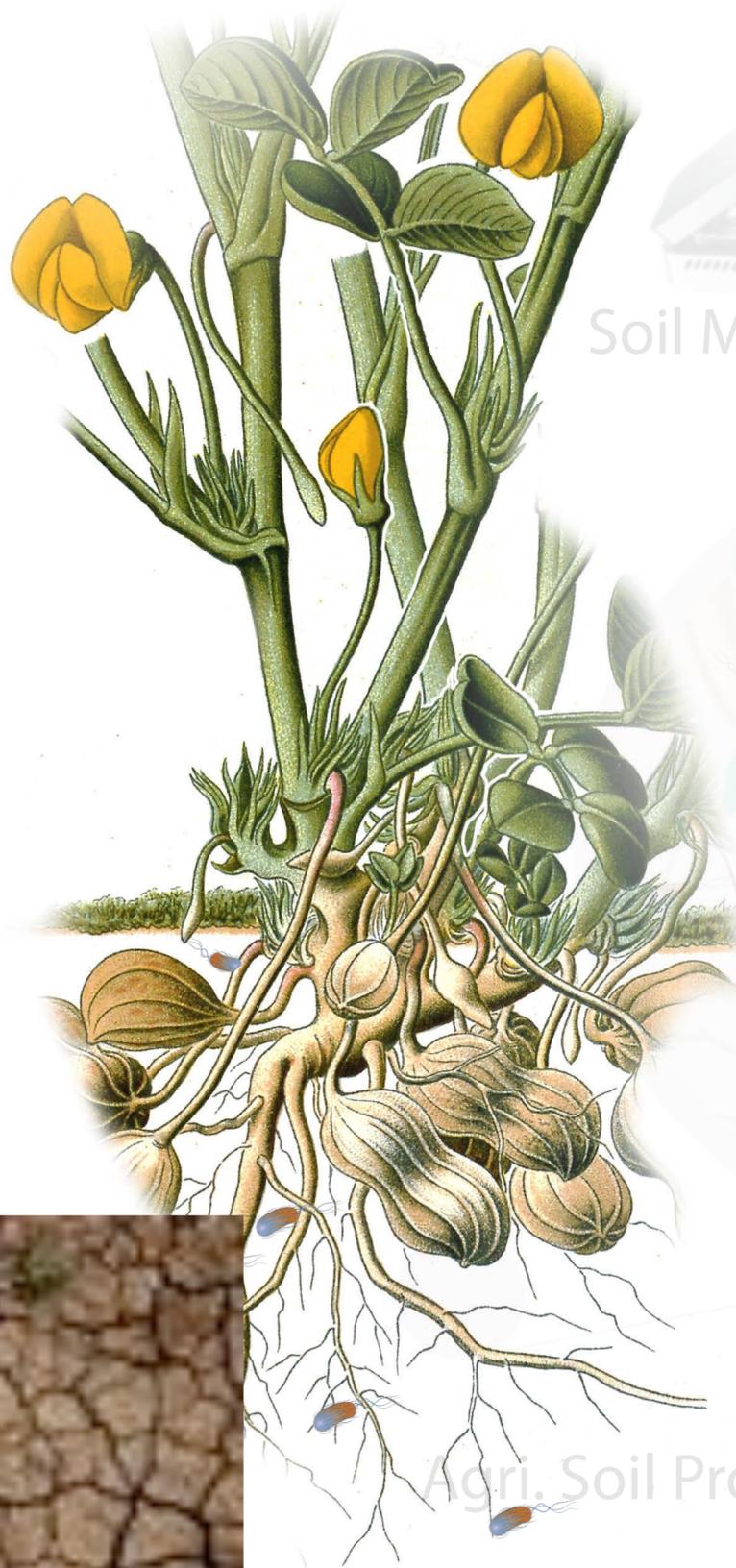
Azuma, Panek, Tice et al. 2022 *MBE*

Burki et al. 2021 *TREE* | Tice et al. 2021 *PLoS Biology* | redrawn from Tice & Brown 2022 *Current Biology*

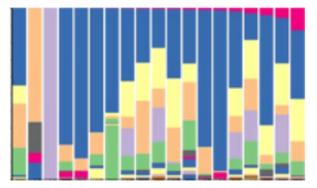
AMOEBOZOEA







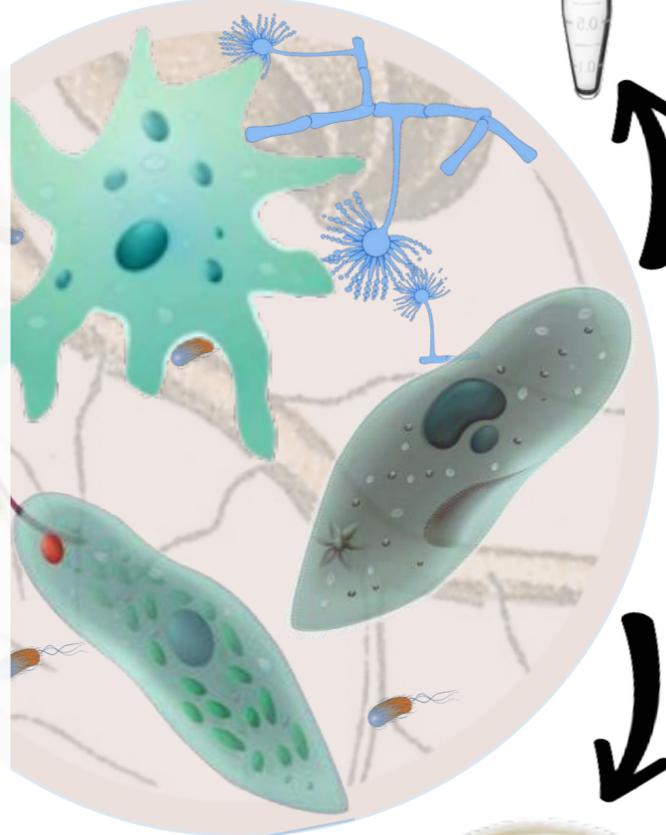
Community Composition



Soil Microbiome Seq.



Soil DNA



Zurweller, Ph.D.

MISSISSIPPI STATE UNIVERSITY
MS AGRICULTURAL AND FORESTRY EXPERIMENT STATION



Tice, Ph.D.

MISSISSIPPI STATE UNIVERSITY
COLLEGE OF ARTS & SCIENCES



Quentin Blandenier, Ph.D.



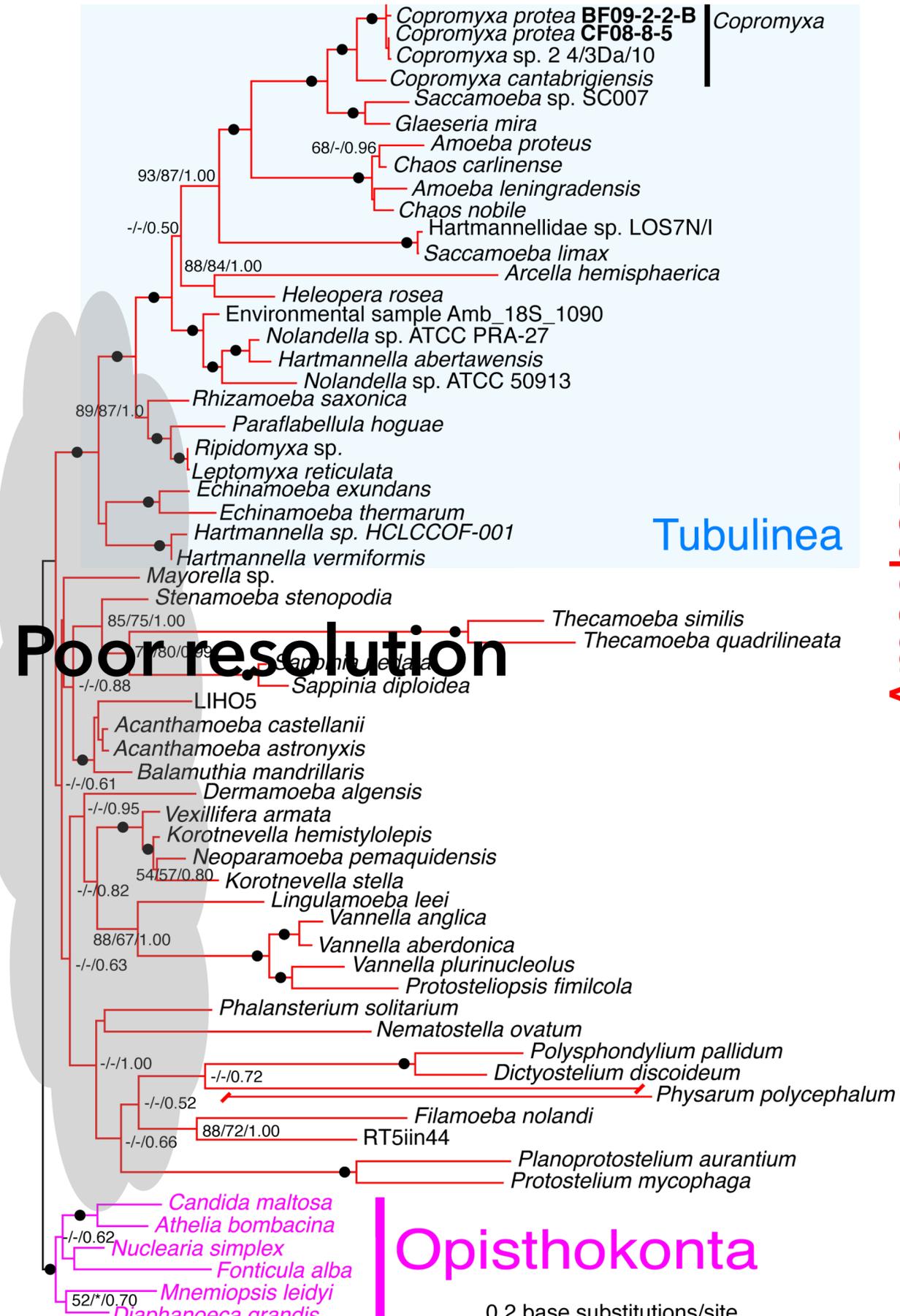
Nicholas Gibson

Agri. Soil Protist Culture

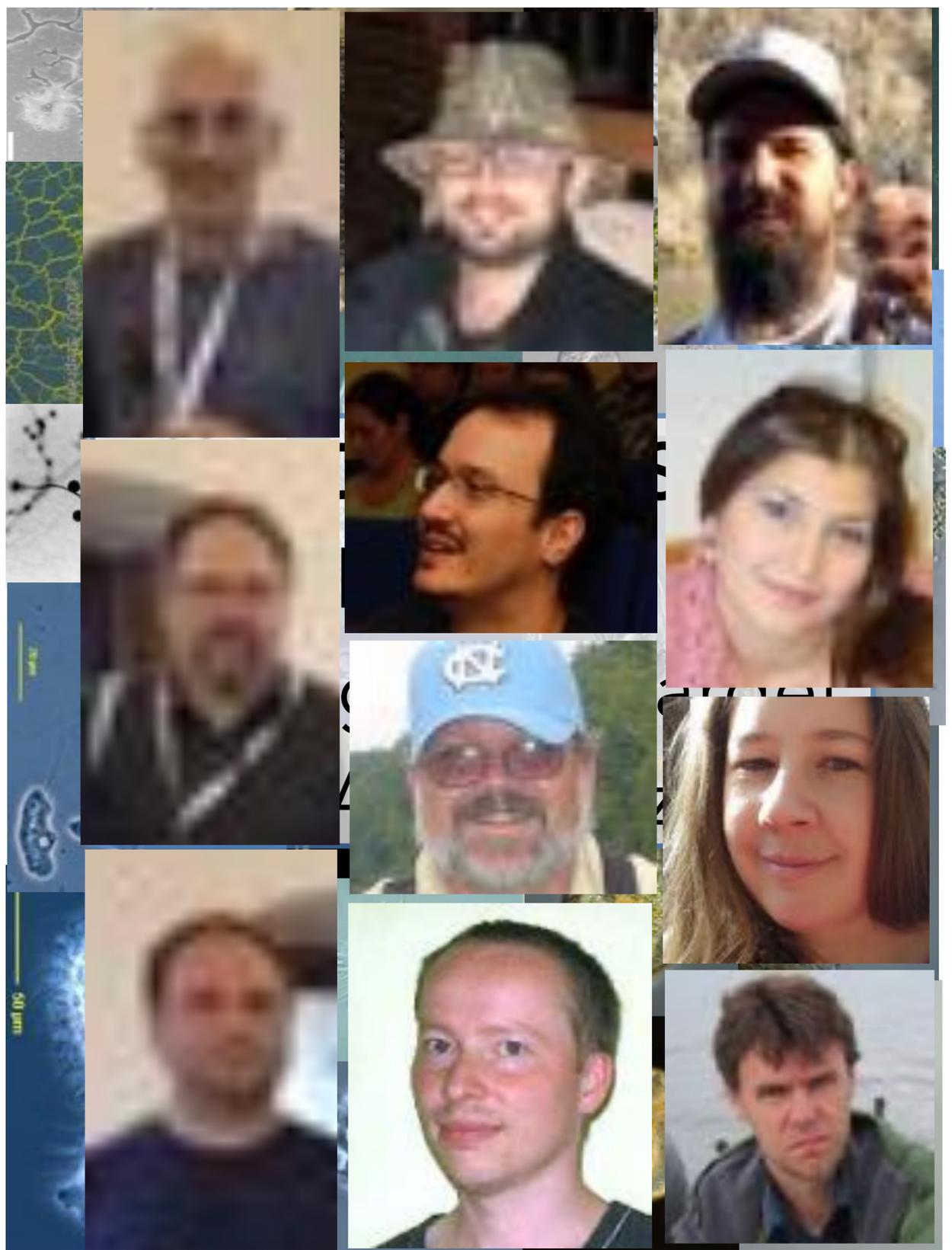




Seungho Kang



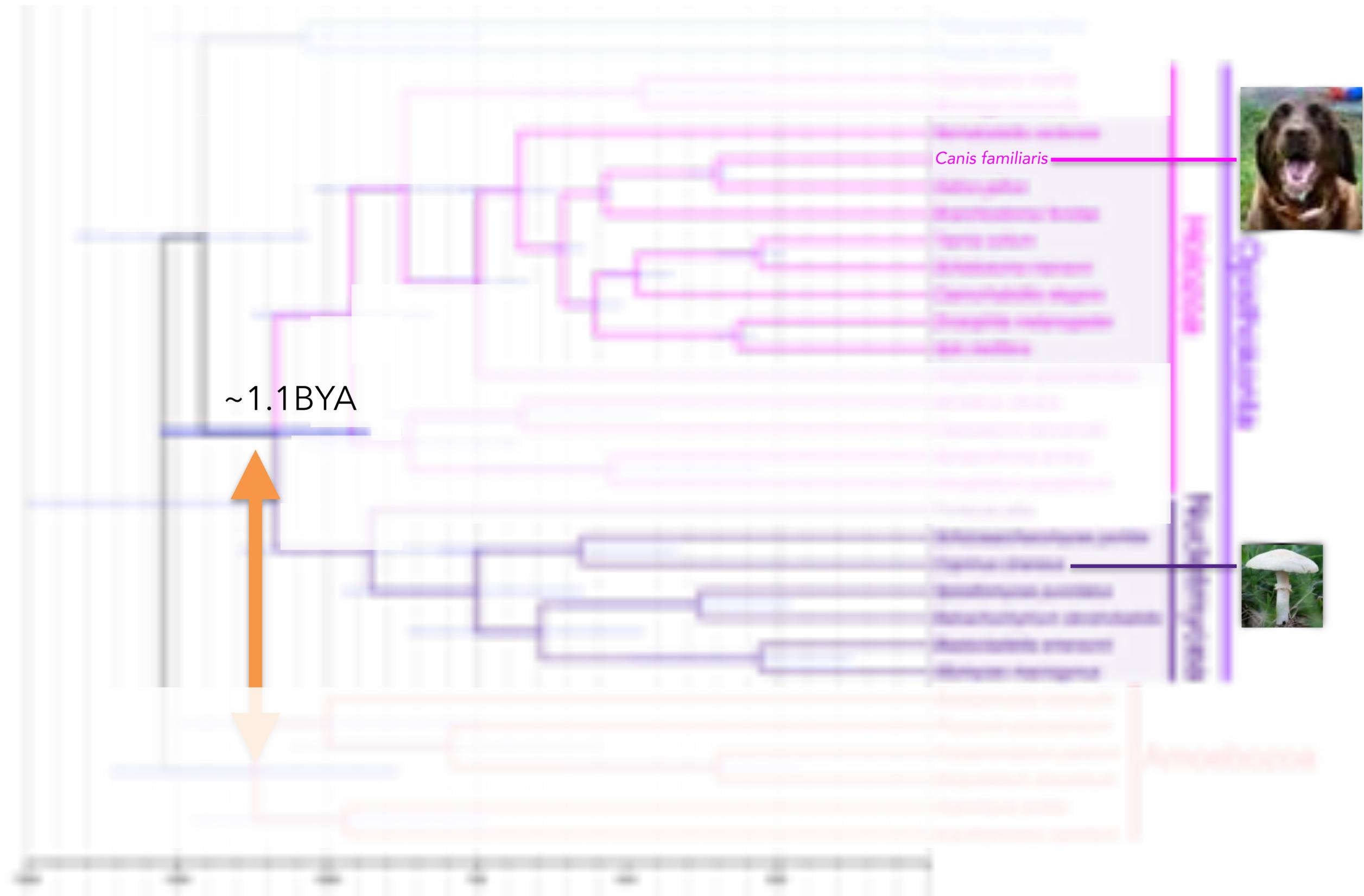
Amoebozoa



● = >95/>95/1.0

SSU rDNA: GTR + Γ + I: RaxML/GarliML/MrBayes
Brown et al. 2011. *Protist*

Amoebozoa evolving as long as divergence between **Metazoa** and **Fungi**



Age Estimates (Myr)

Eme et al. 2014. *CSHL Persp. Biol.*; Sharpe et al. 2015

Phylogenetics of Amoebozoa et al.

- Restricted to SSU rDNA (18S) phylogenetics
 - Poor resolution
- Very limited genomic-level data
- Major goal 1:
 - Phylogenomics of Amoebozoa
 - Robust phylogeny
- Major goal 2:
 - Phylogenies to inform evolutionary events
 - Mapping of morphological characters
 - Gene distributions and genomic innovations



Methods - Single Cell

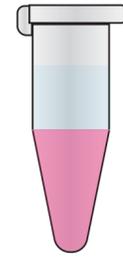
Uncultured or Eukaryote predators



Pelomyxa



Amoeba



TriZOL

Single cell isolation
(Photo and Video Voucher)



SMARTseq 2 (Picelli et al. 2014
Onsbring et al. 2021)

NexteraXT (Illumina)

QC (1.8x TBE), QPCR (KAPA), and library pooling

Ship



HiSeqX

Paired End (125bp)

Panek et al. 2016: *MPE*
Tice et al. 2016: *Biology Direct*
Kang et al. 2017: *MBE*
Onsbring, Tice et al. 2021: *BMC Genomics*

Single Cell

Cultured Cell



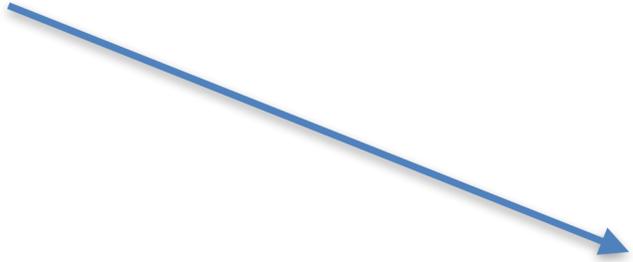
TransDecoder

de novo Assembly +
Trimmomatic quality trimming

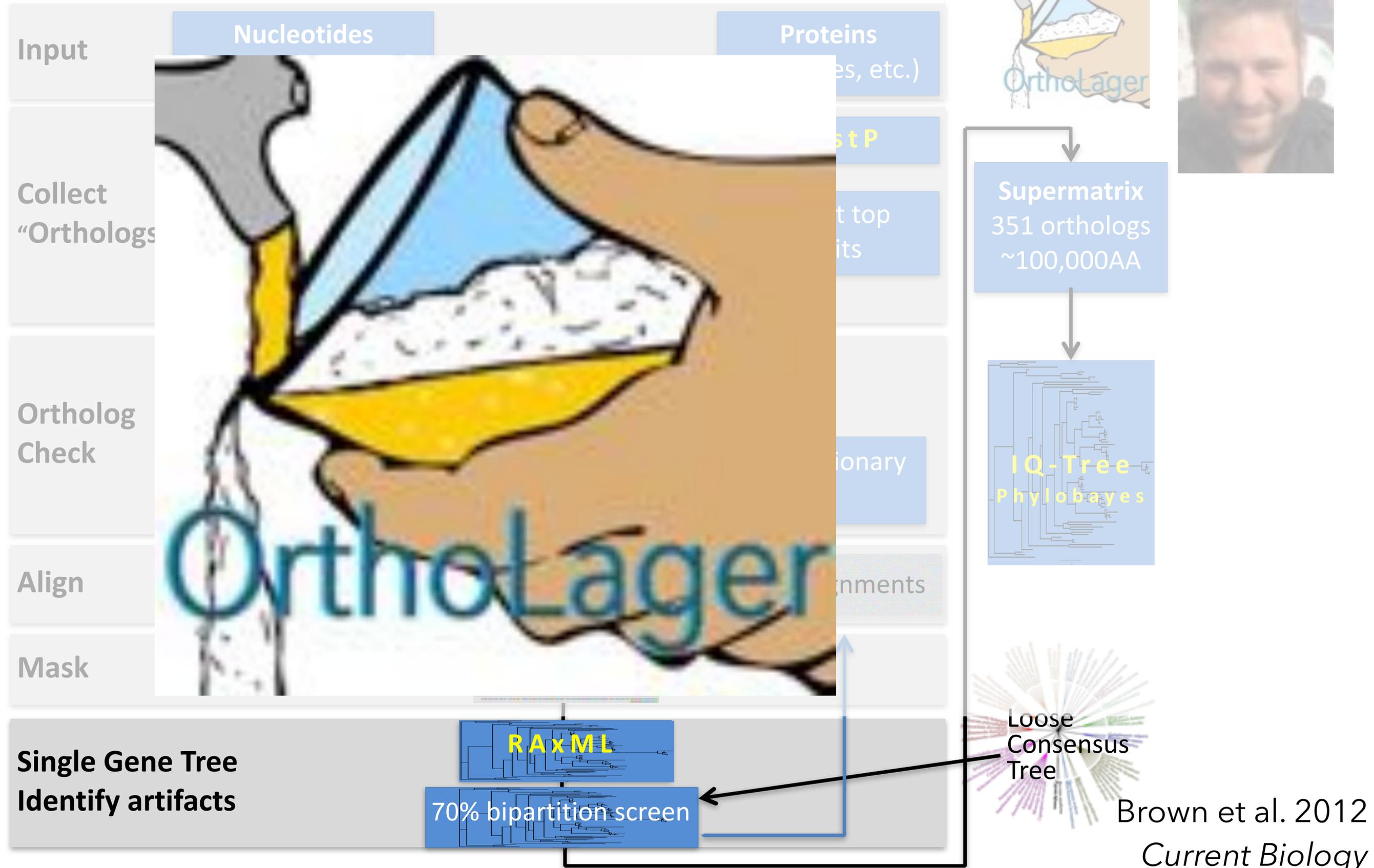
ORF Prediction/translation
PFAM domains



ORFS
translation (Amino Acids)

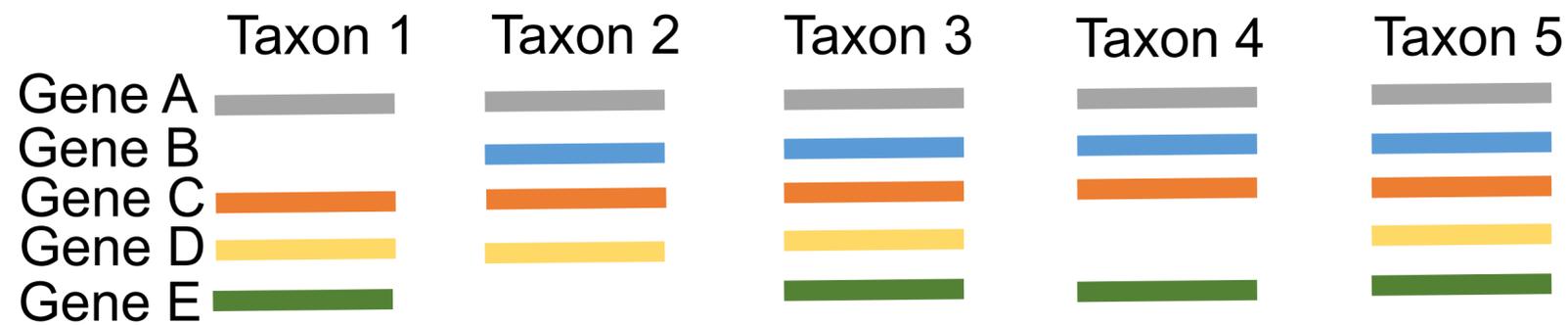


OrthoLager: A Phylogenomic Pipeline (Brown & Kolisko)

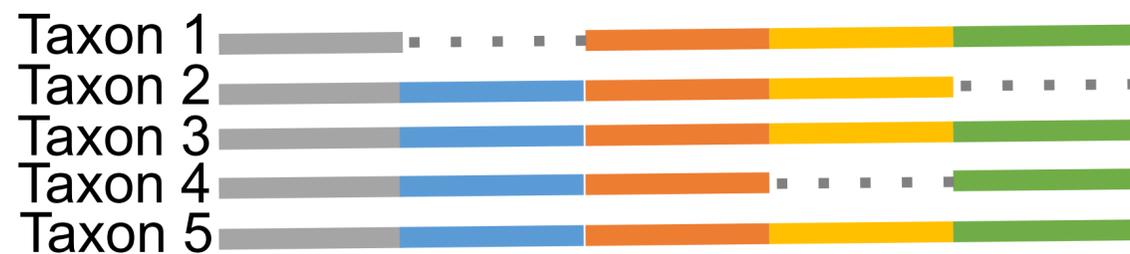
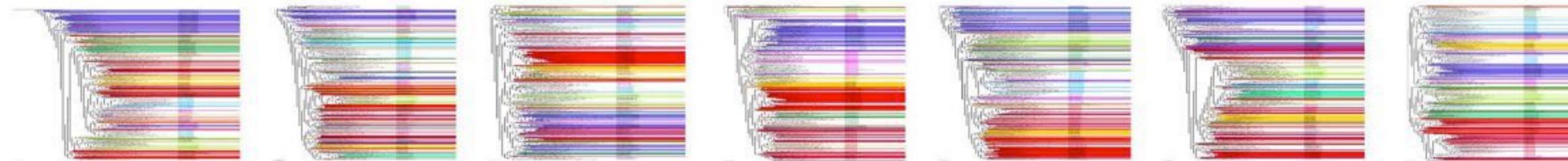


Multigene Phylogenetics - Phylogenomics

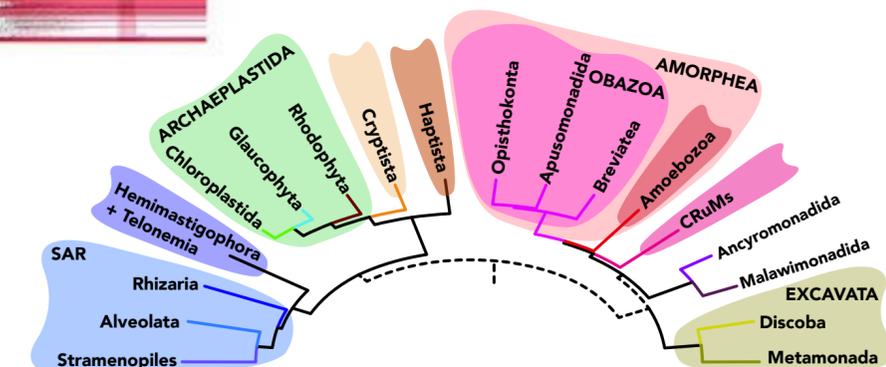
- Take multiple genes and infer a phylogeny
 - Genes A+B+C+D+E
- Offers more data
- Can handle missing data from taxa
- Concatenation
 - implies 1 taxon = 1 orthologous sequence



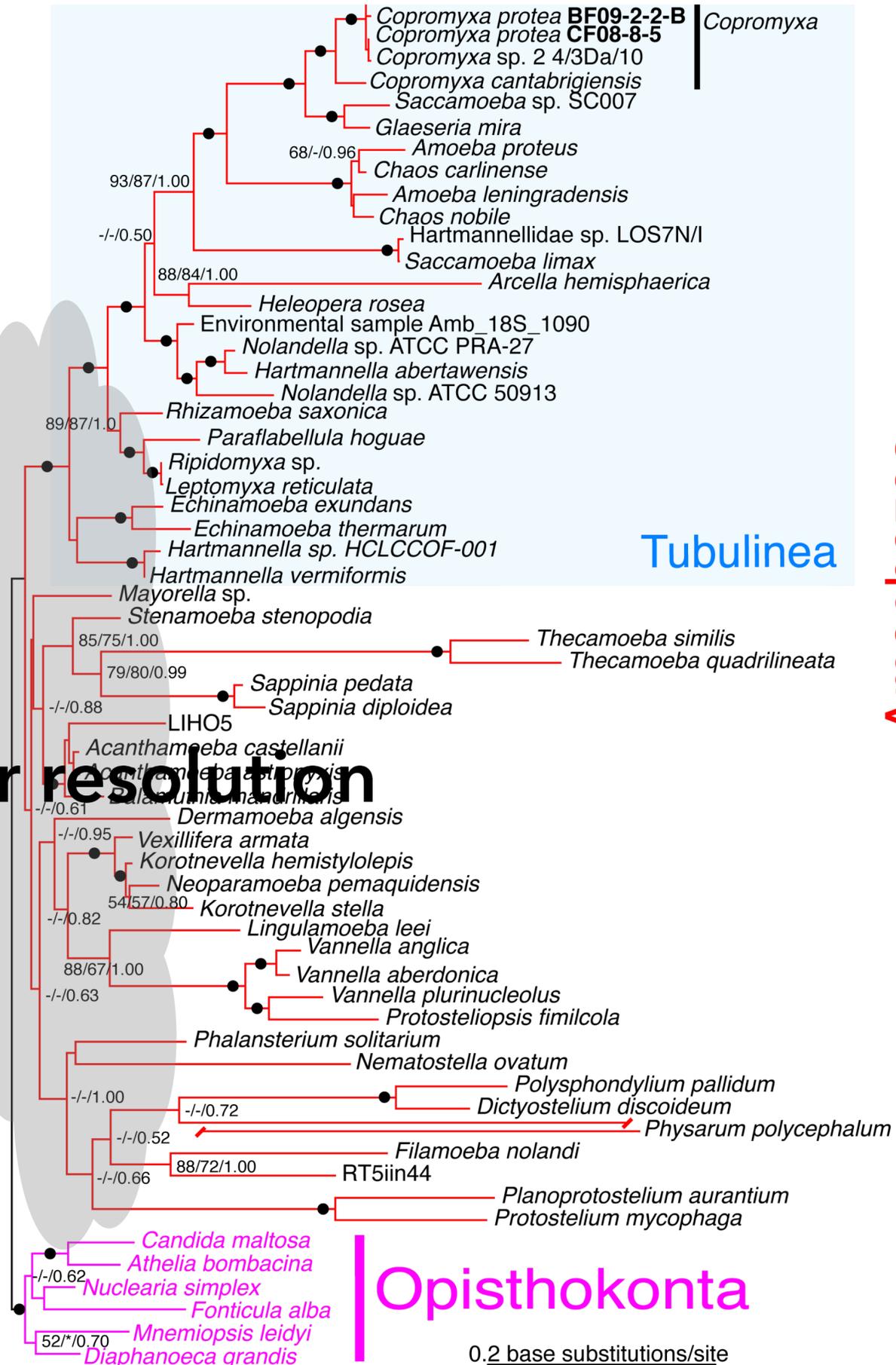
GENE SAMPLING



CONCATENATION
(Gene Matrix)



Poor resolution



Tubulinea

Amoebozoa

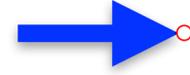
Opisthokonta

0.2 base substitutions/site

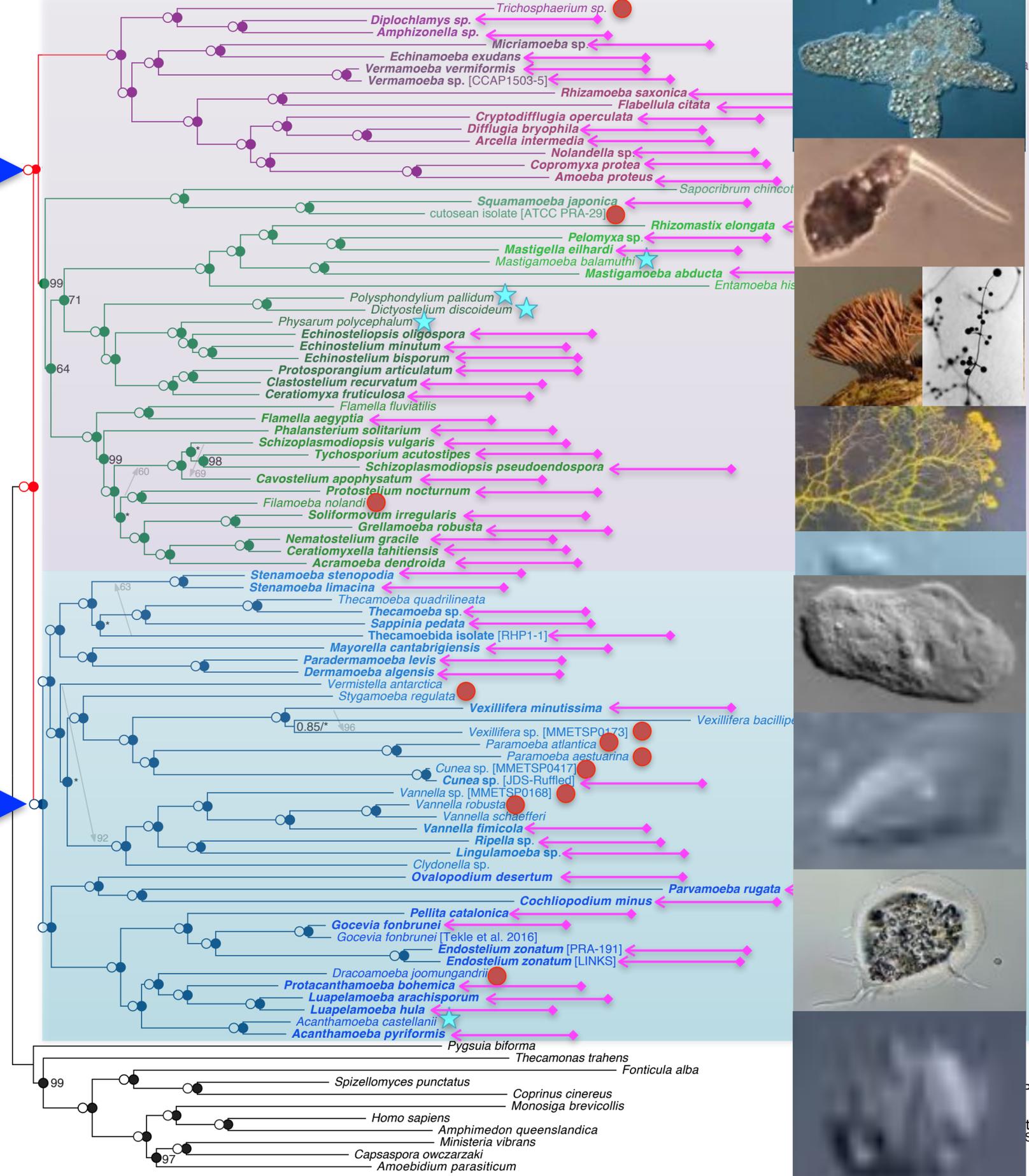
● = >95/>95/1.0

SSU rDNA: GTR + Γ + I: RaxML/GarliML/MrBayes
Brown et al. 2011. *Protist*

Tevosa



Discosea



Tubulina

Evosea

Discosea

Amoebozoa

Obazoa

★ Genome



↔ Our Data

325 Proteins
PhyloBayes CAT-GTR
IQ-Tree ML LG+G4+C60+PSMF

Kang et al. 2017, MBE

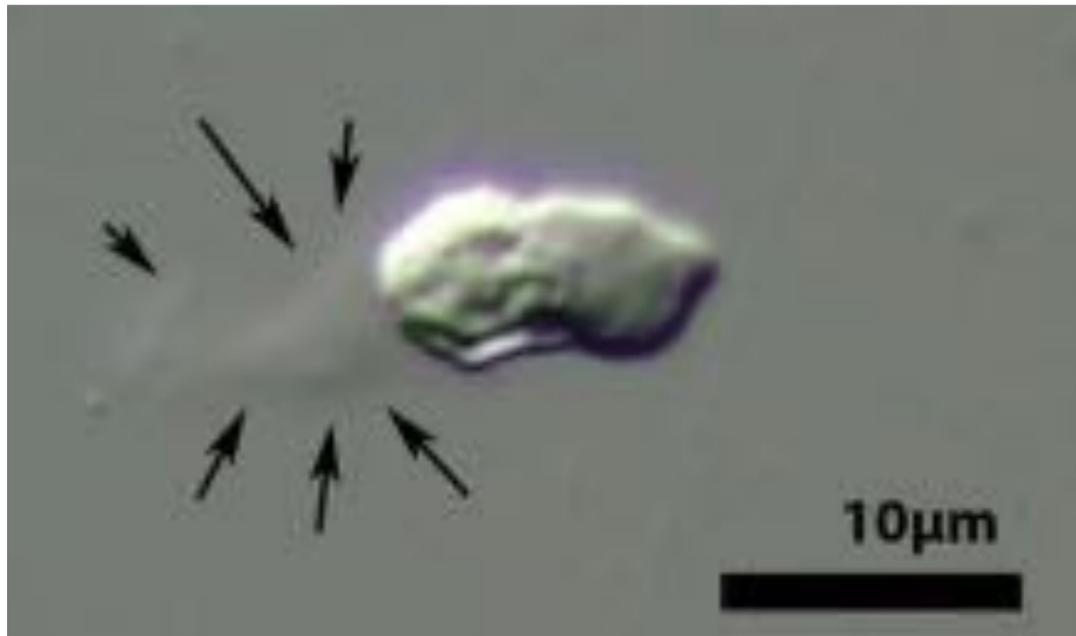
0% 50% 100%
Site Coverage

A fully resolved Tree of Amoebozoa

- 2 deepest clades Discosea and Tevosa
- Fully-supported clades in Maximum Likelihood (IQ-Tree) and Bayesian inferences (PhyloBayes)
- All known amoebozoan clades have been sampled
- Goal of RNAseq projects from **87** taxa
 - Sequenced 61 taxa (shown)
 - **Another 180+ sequenced**
- Kang et al. 2017 *MBE*



NEW ORDER

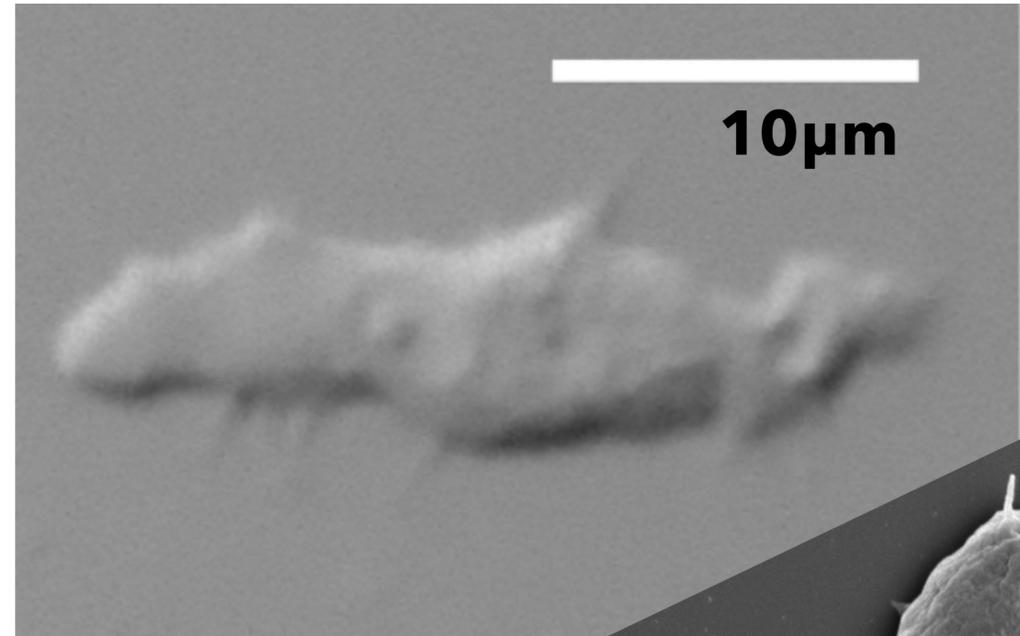


"Janelia velum"



Robert Jones

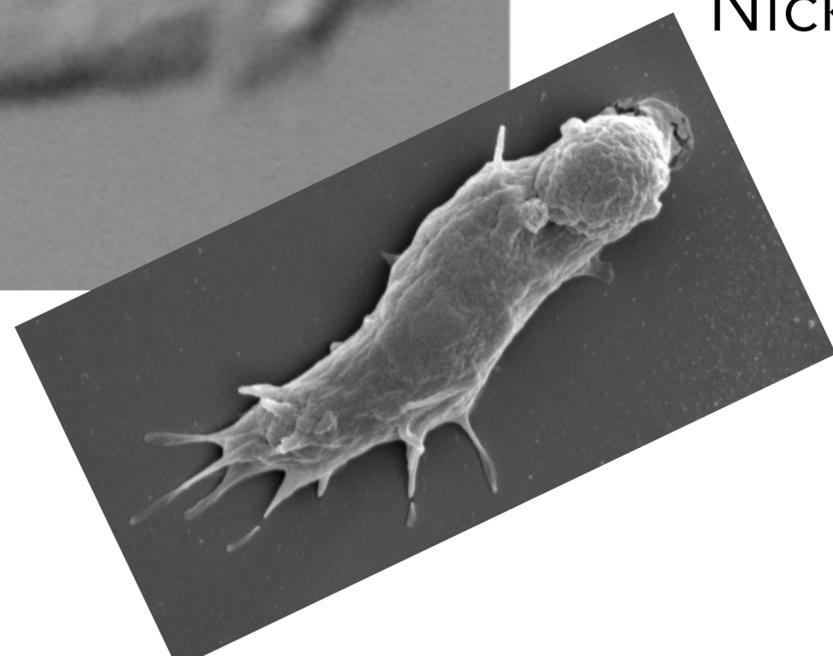
NEW FAMILY



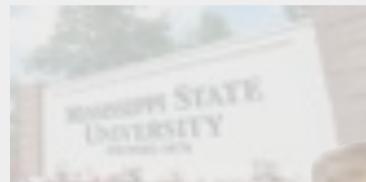
Nick Fry



Peter Kooienga



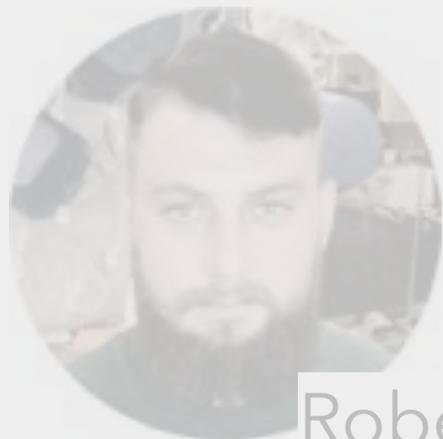
NEW ORDER



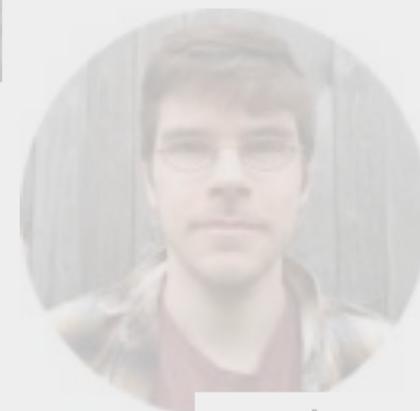
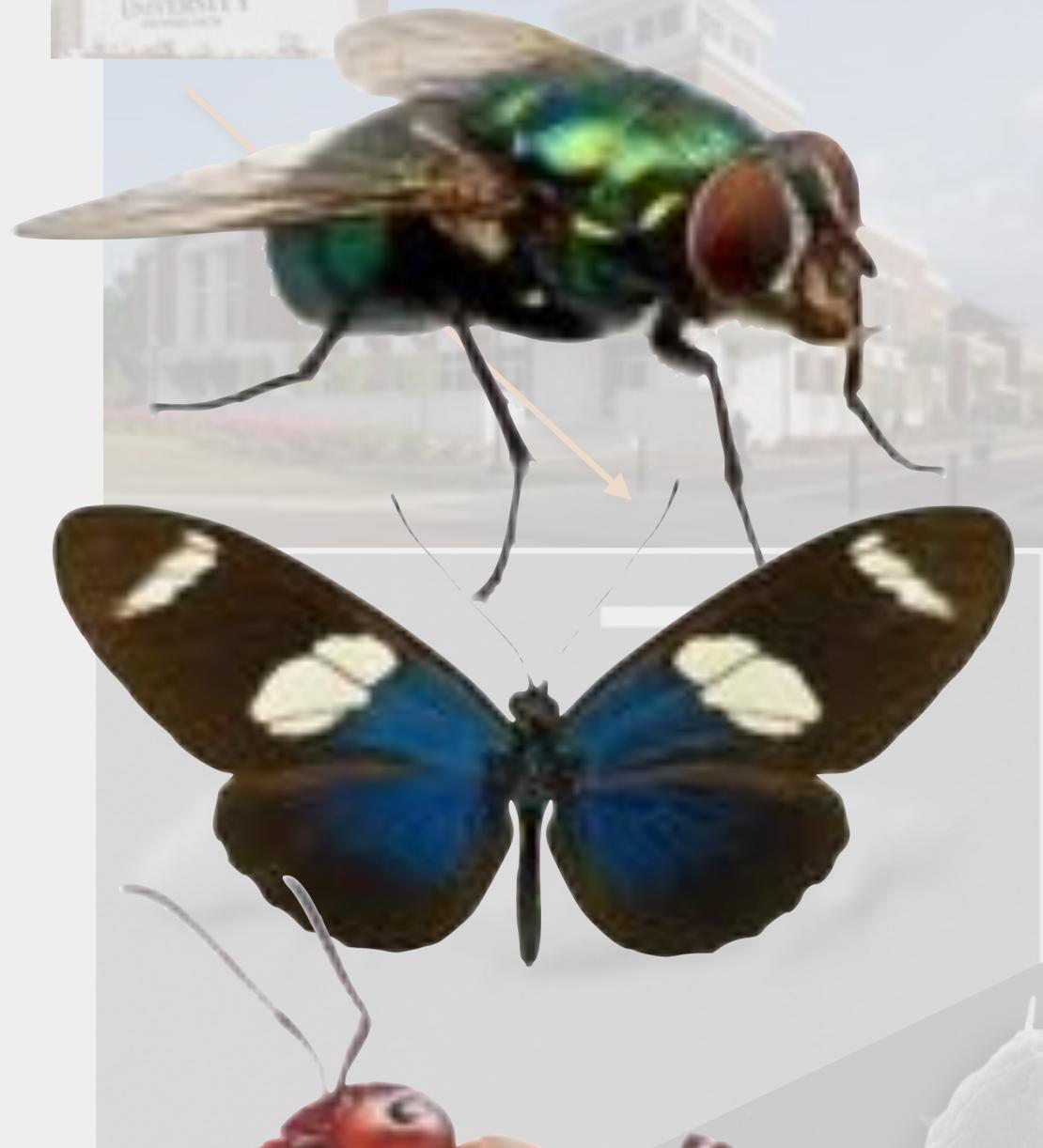
NEW FAMILY



"Janelia velum"



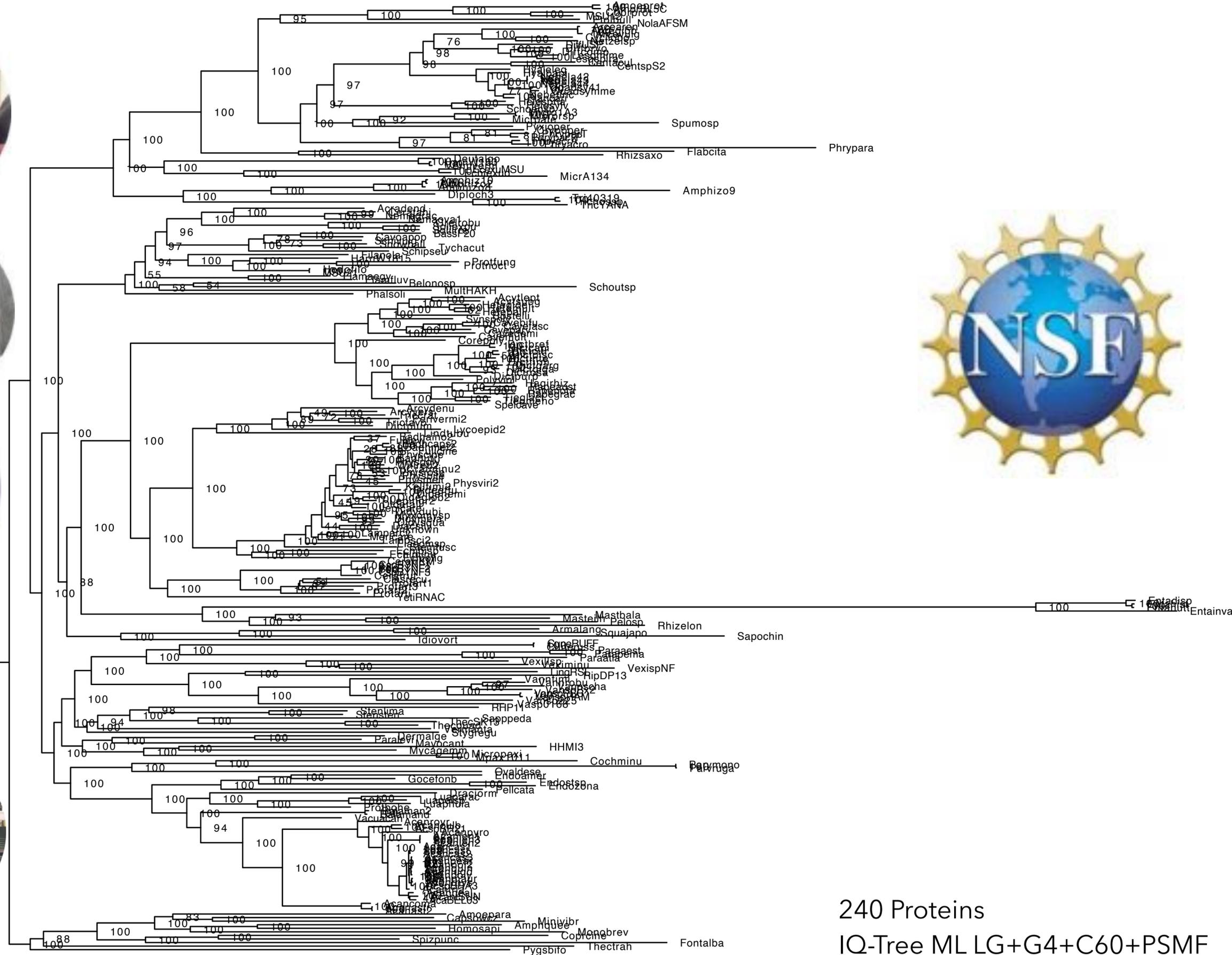
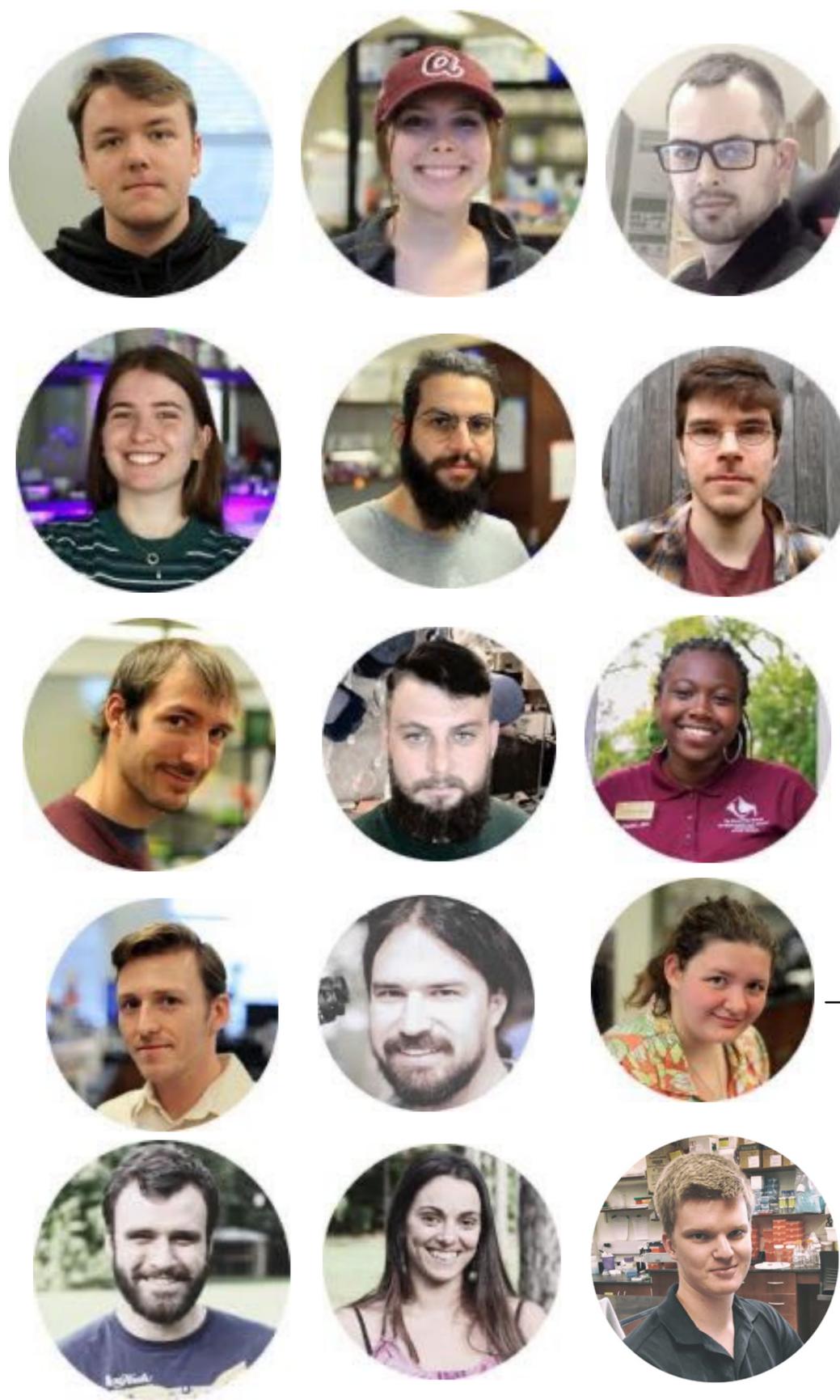
Robert Jones



Nick Fry

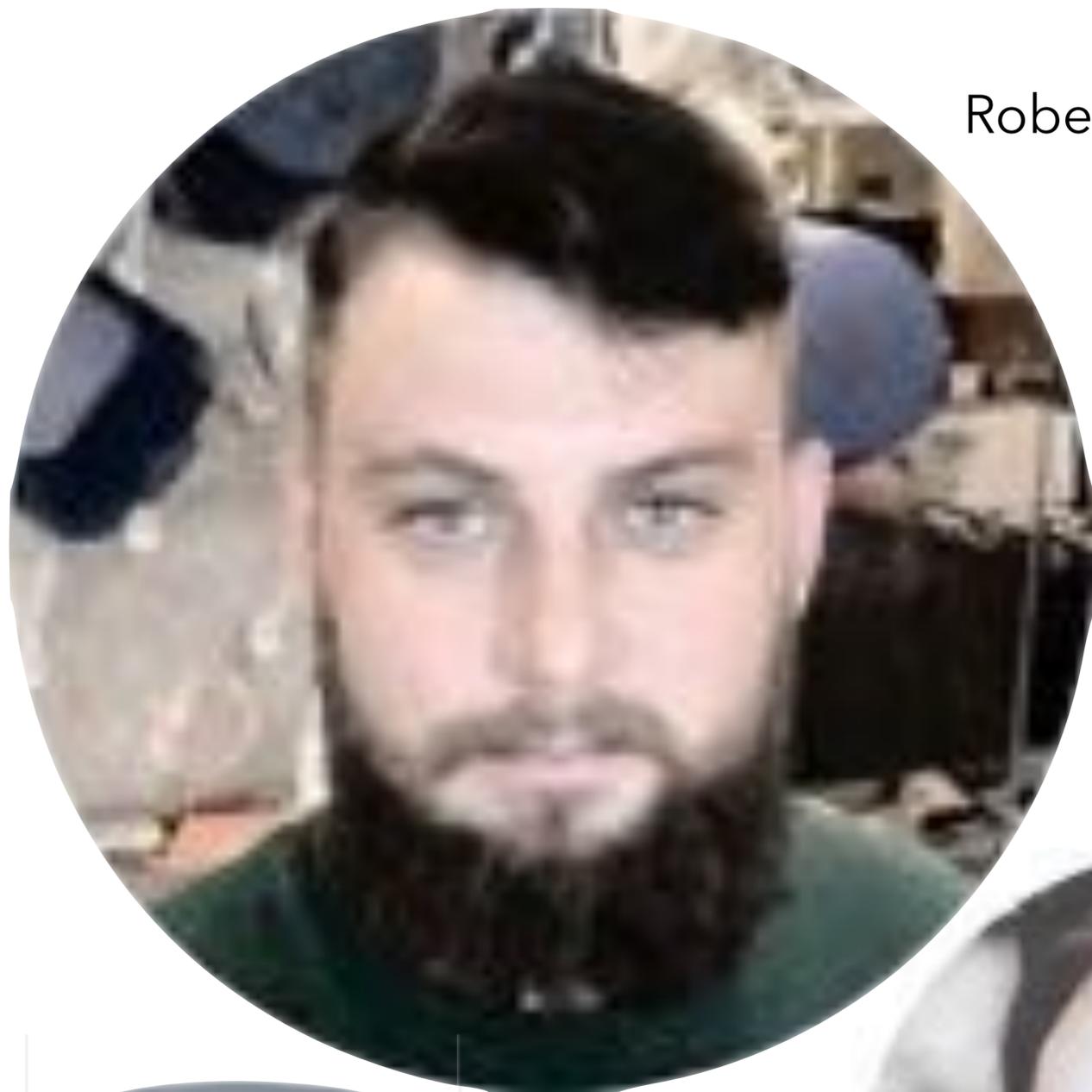


Peter Kooienga

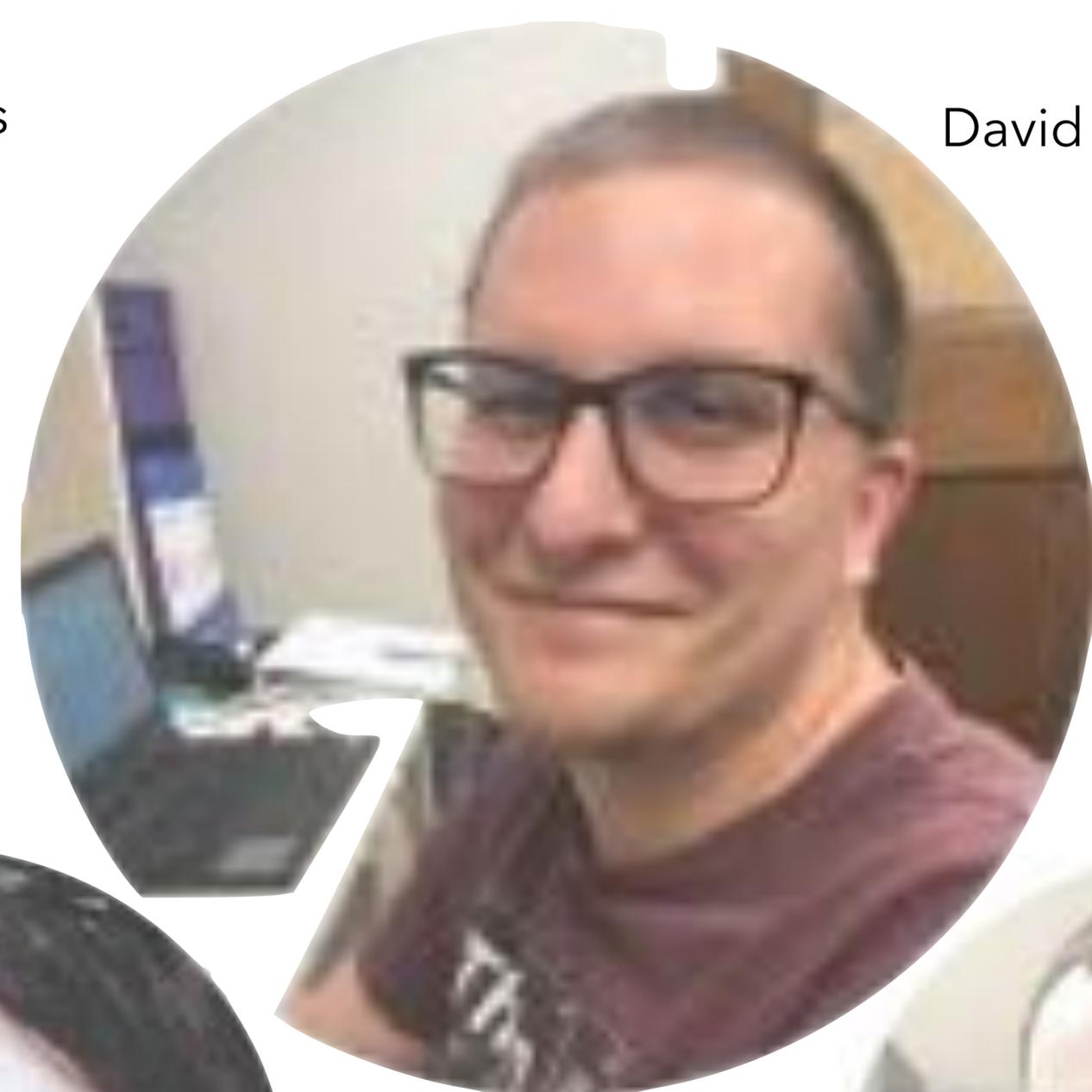


240 Proteins
 IQ-Tree ML LG+G4+C60+PSMF
 100 Real Bootstraps

0.3



Robert Jones



David Žihala



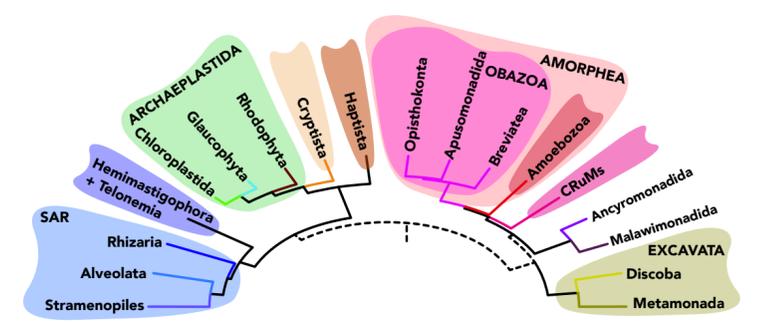
Alexander Tice, Ph.D.



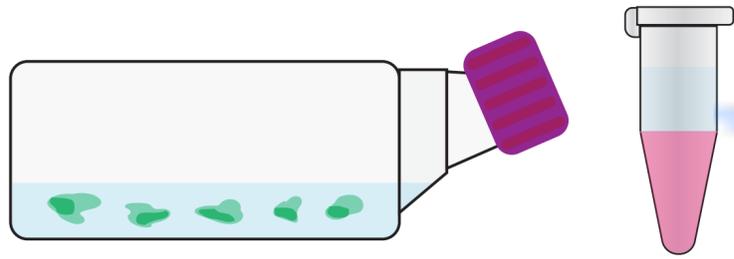
Tomáš Pánek, Ph.D.



PhyloFisher Plug

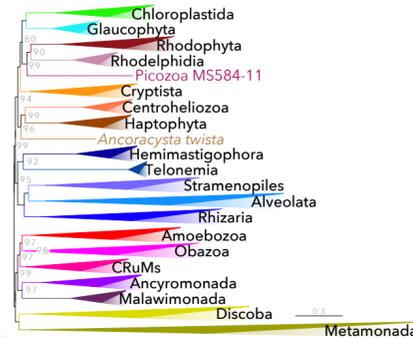


Tice, Panek, Žihala, Jones, et al. 2021 *PLoS Biology*



PHYLOFISHER

Phylogenomic dataset/tree



YouTube Seminars on PhyloFisher

<https://www.youtube.com/watch?v=TcKhLiUxBTU&t=11s>

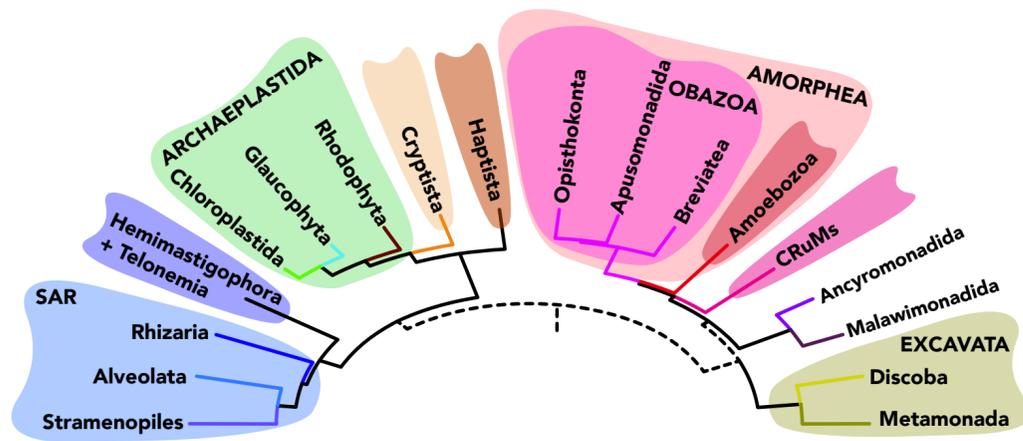
<https://www.youtube.com/watch?v=MJfb5CaIWPA>

<https://github.com/TheBrownLab/PhyloFisher>



PHYLOFISHER

- Easily installed and simple usage
- Ships with our dataset
- **Includes Paralogs for tree building, more accurate identification**
- Your own gene sets can be incorporated or used independently
- Tools for post-phylogenomic analyses

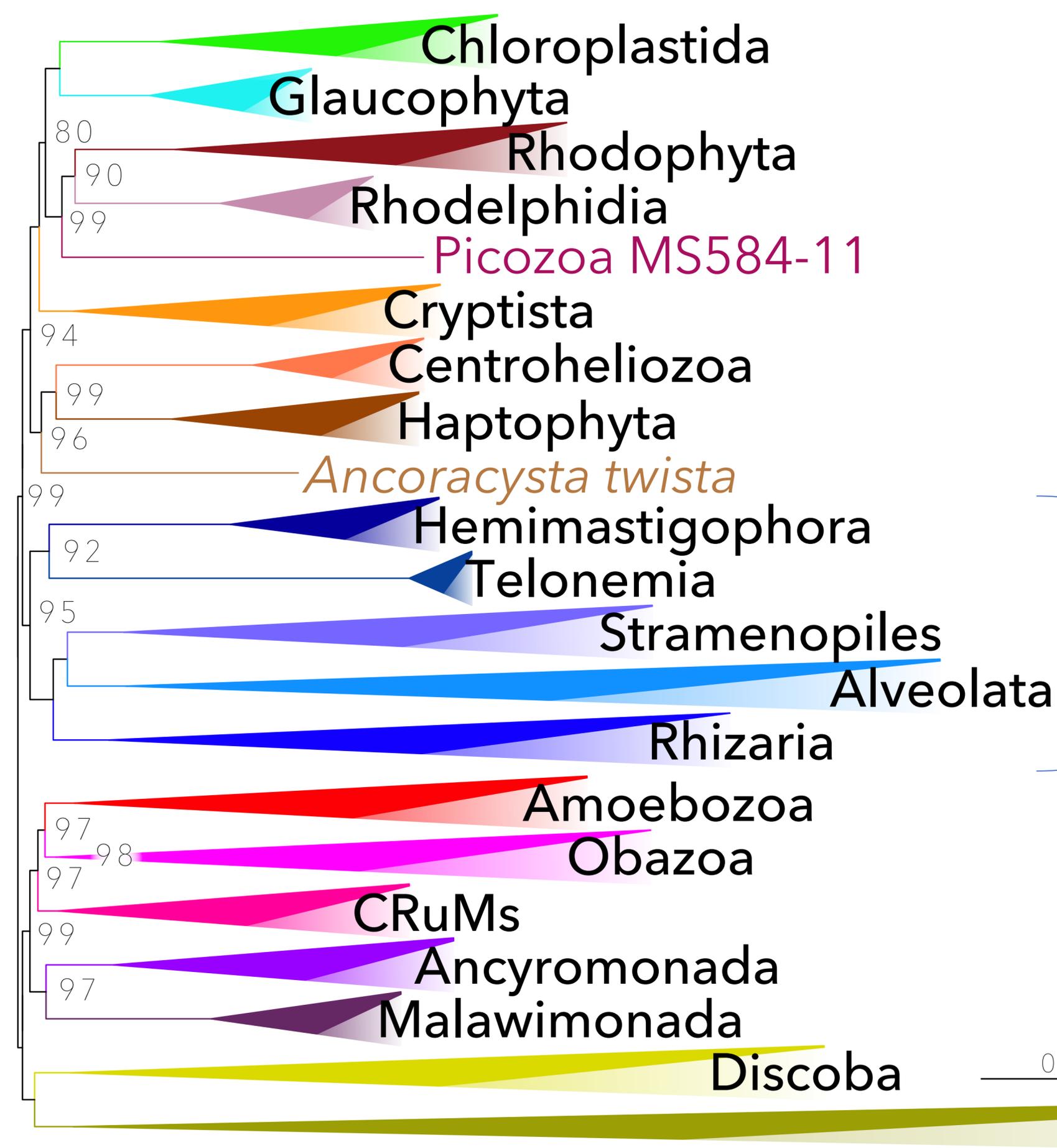


304 Taxa across all known diversity with their predicted proteomes



PHYLOFISHER

Phylogenomic Tree



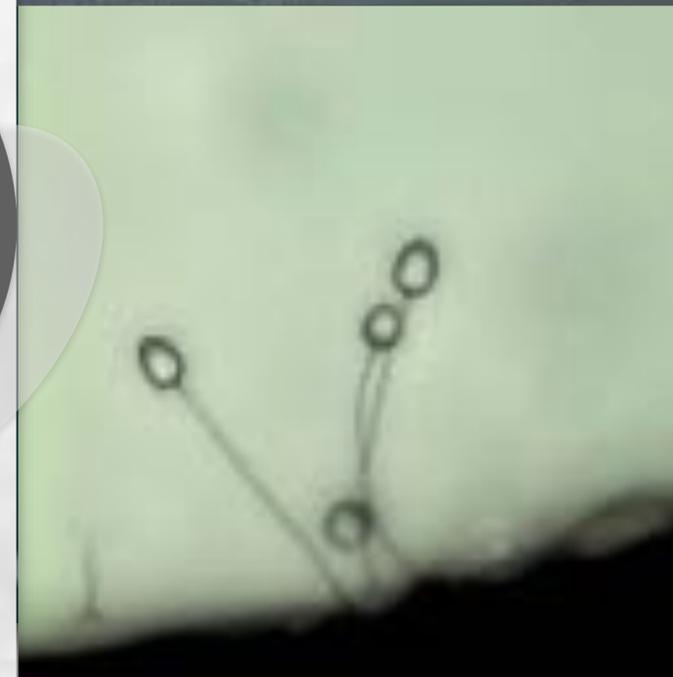
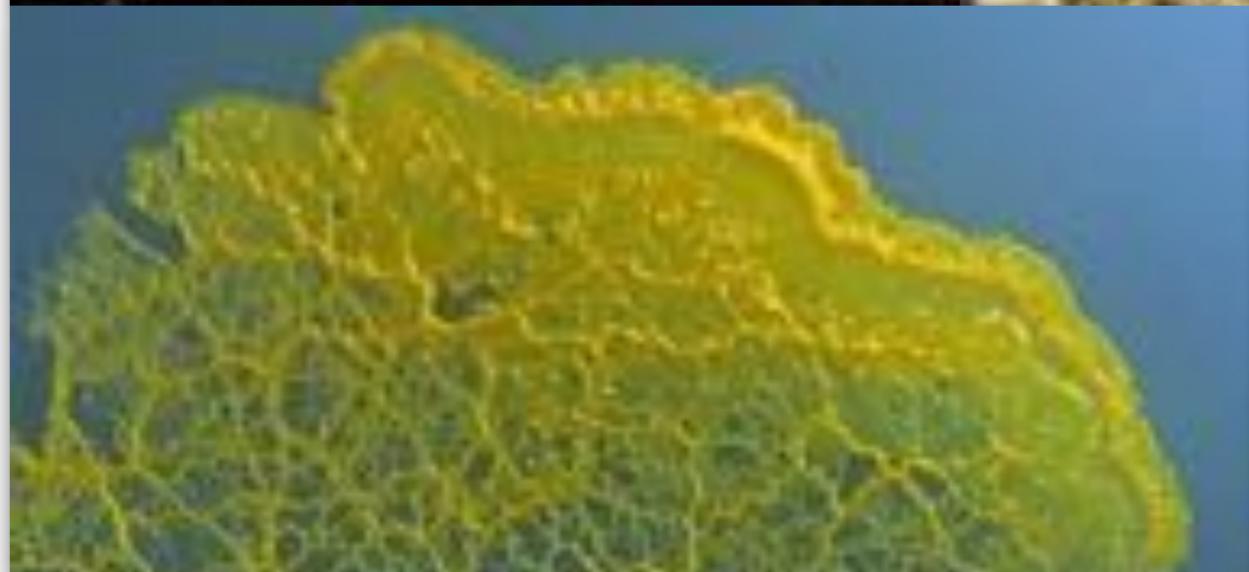
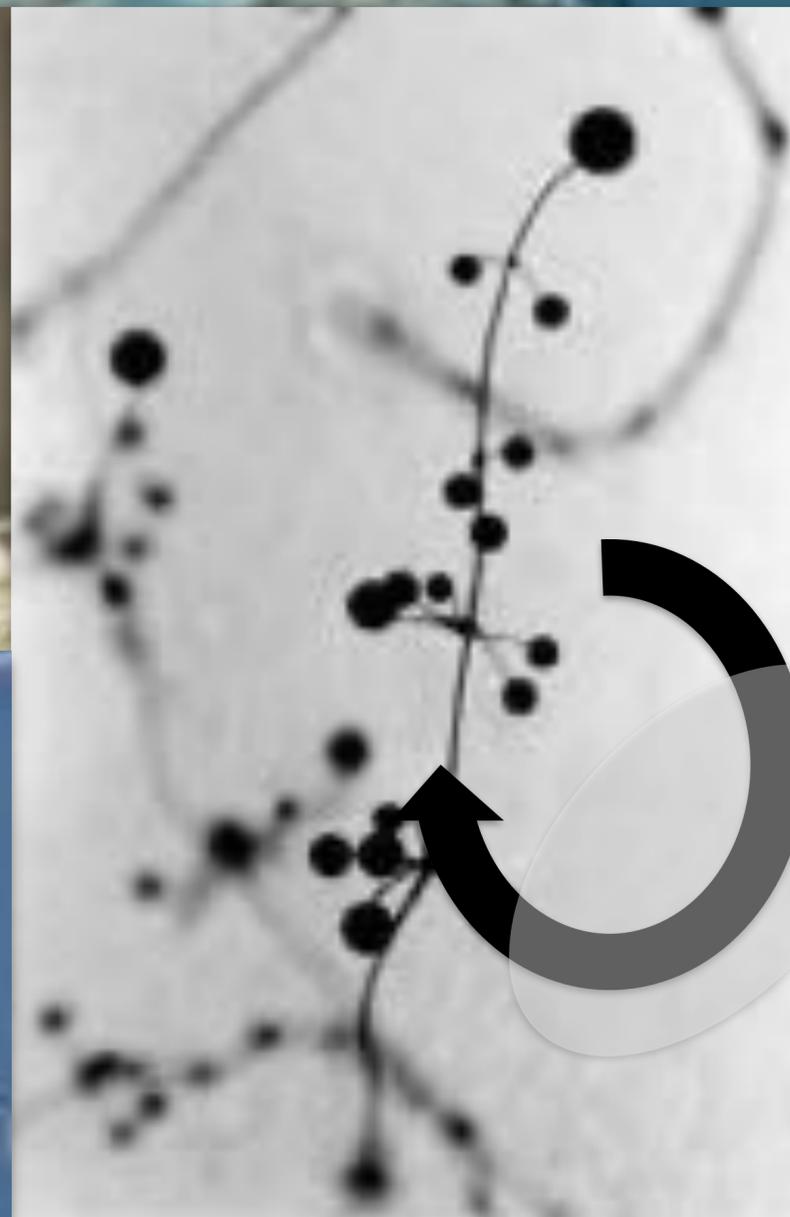
Hemimastigophora-TSAR

Tice, Panek, Žihala, Jones,
et al. 2021 *PLoS Biology*

0.3

240 orthologs: 305 taxa: 72,632AA sites
IQtree- LG+G+c20 => LG+G+c60(PMSF)
350 Bootstrap Reps

Amoebae are much more complicated

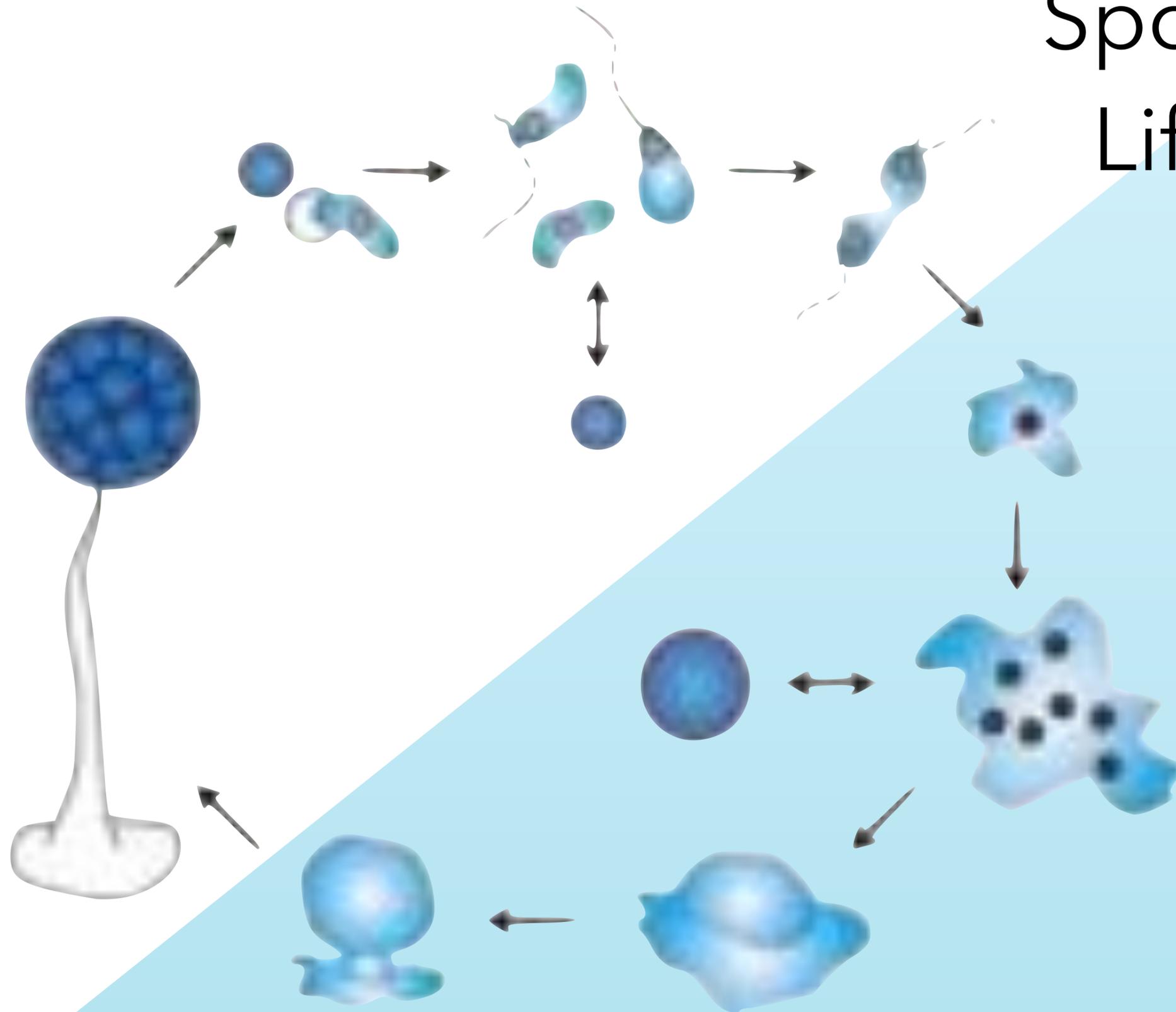




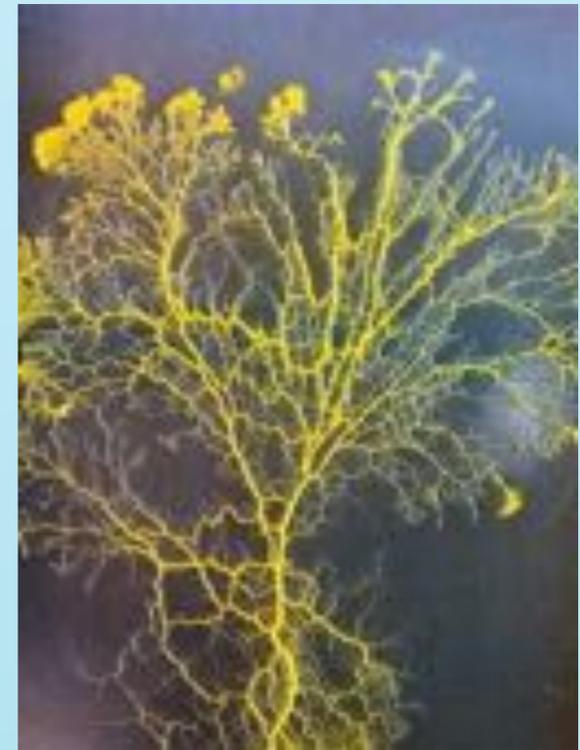
n

Sporocarpic Life Cycle

$2n$



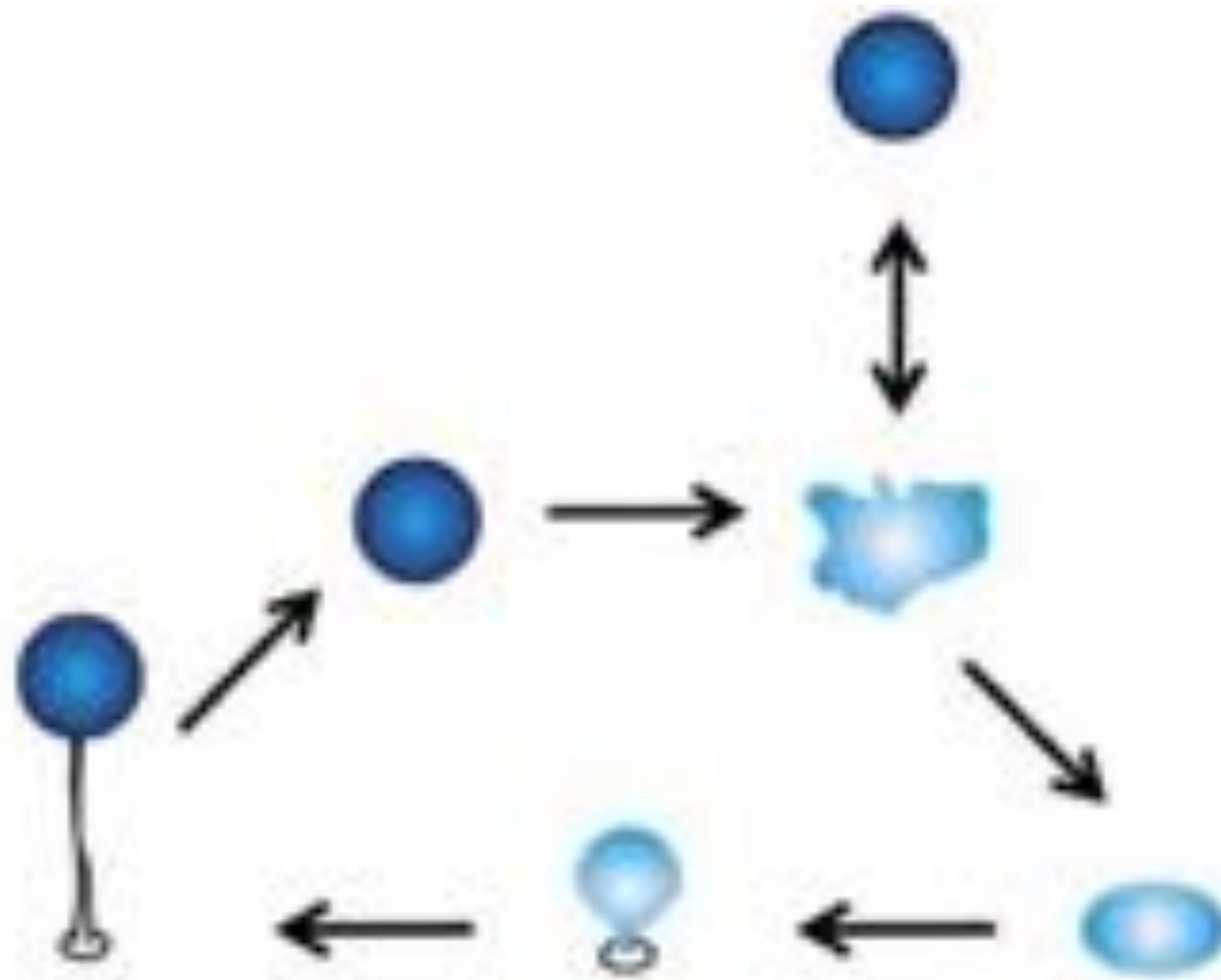
myxogastriid (myxomycete) amoebae
~1200 species



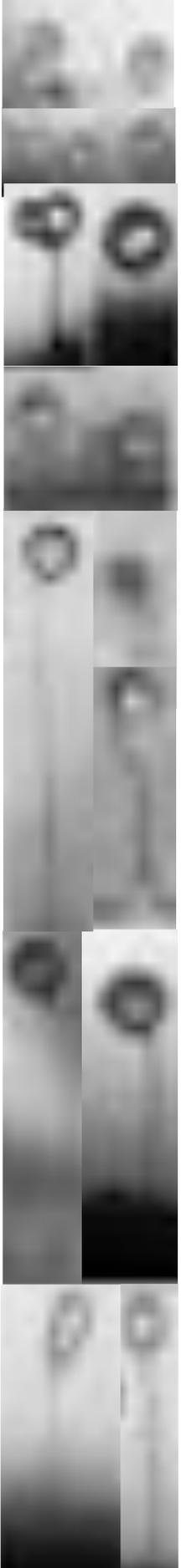


myxogastriid (myxomycete) amoebae
~1200 species

Sporocarpic Life Cycle



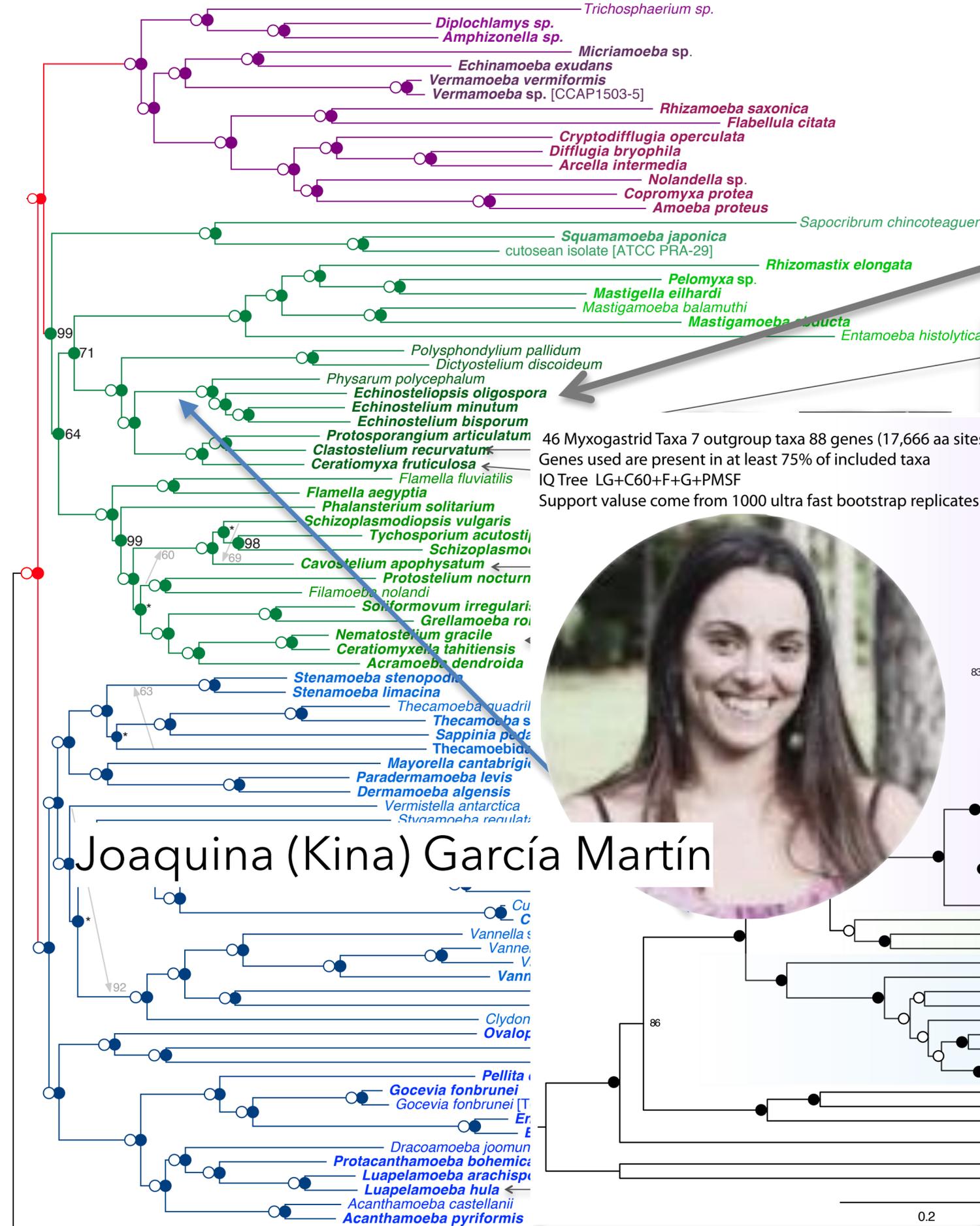
protosteloid amoebae
32 species (10+ to be described)



protosteloid amoebae
32 species (10+ to be described)

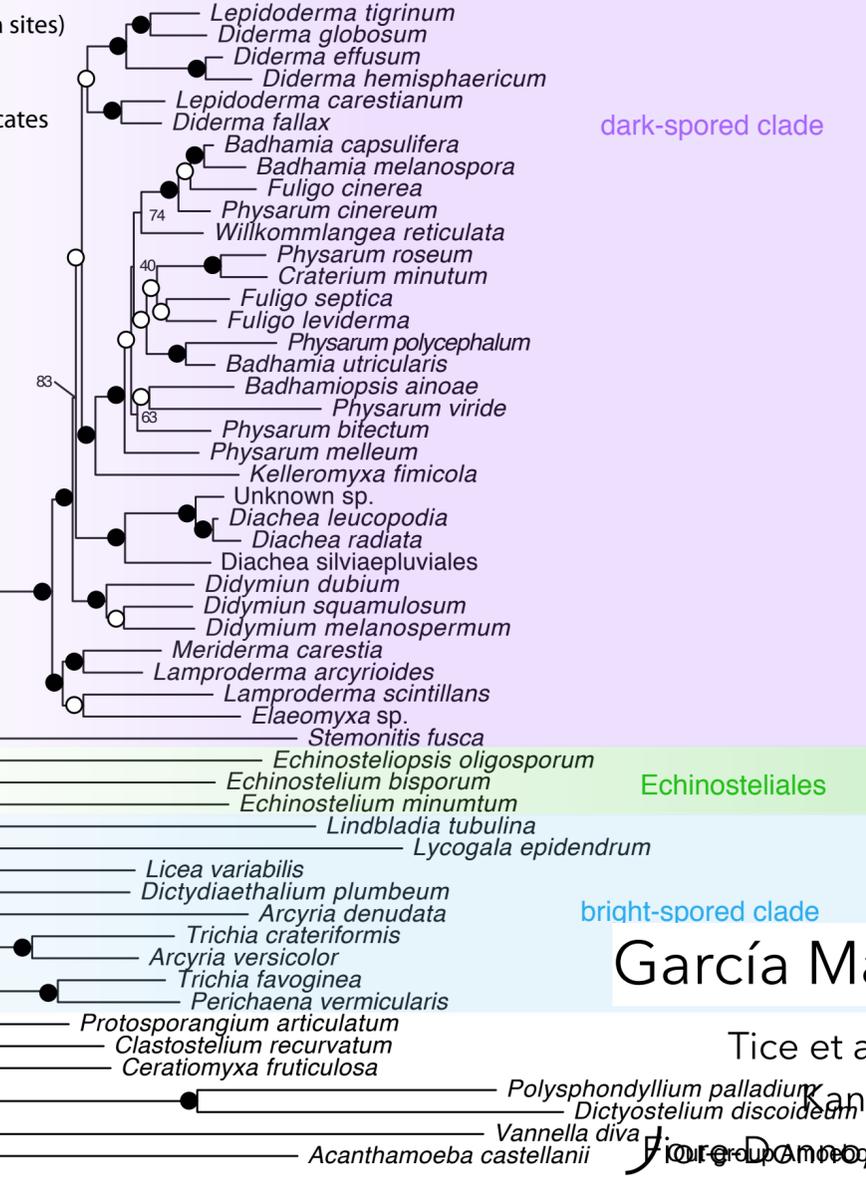
Sporocarps

Myxogastrids



Joaquina (Kina) García Martín

46 Myxogastrid Taxa 7 outgroup taxa 88 genes (17,666 aa sites)
 Genes used are present in at least 75% of included taxa
 IQ Tree LG+C60+F+G+PMSF
 Support values come from 1000 ultra fast bootstrap replicates



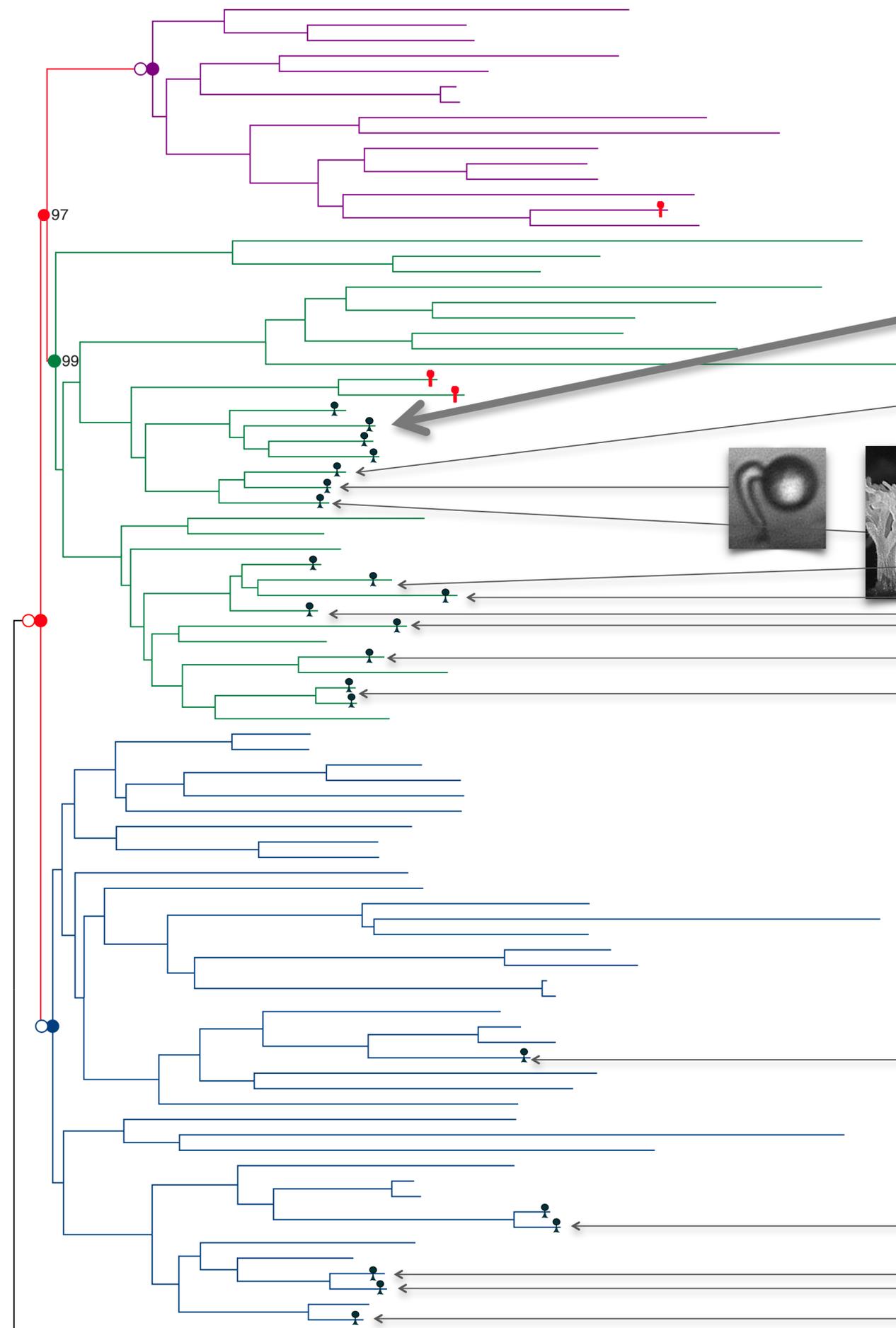
Eumycetozoa
Myxozoa

García Martín et al. In Prep

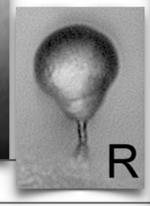
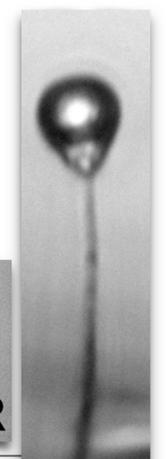
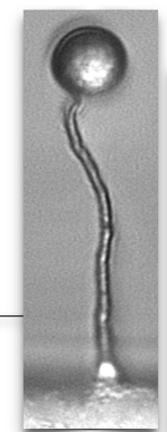
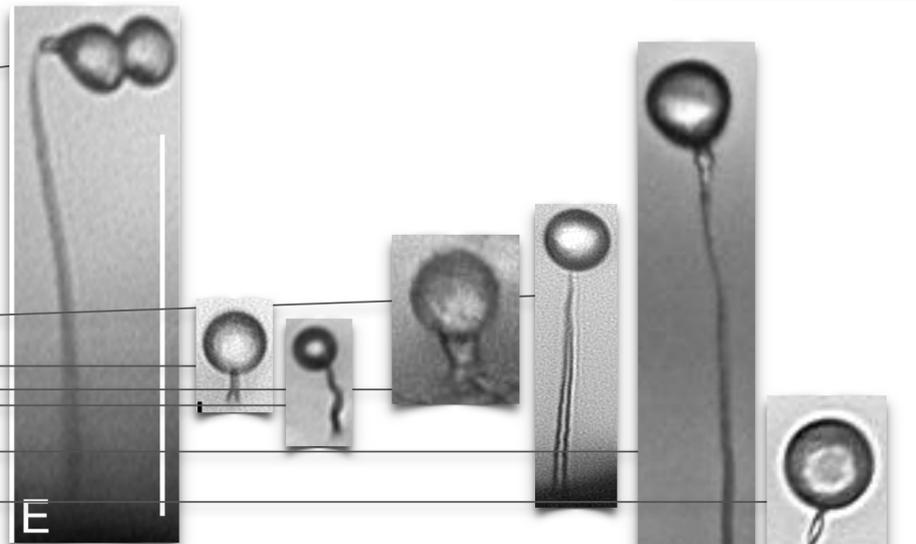
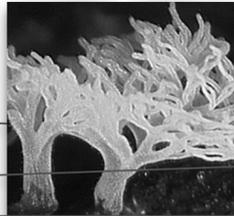
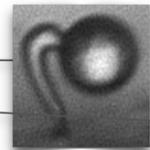
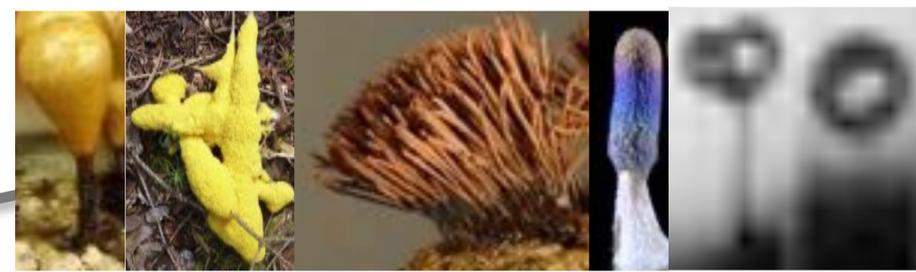
Tice et al. 2016: *Biology Direct*

Kang et al. 2017: *MBE*

Moreno-Dorado, Tice, & Brown. 2019. *JEMB*



Myxogastrids

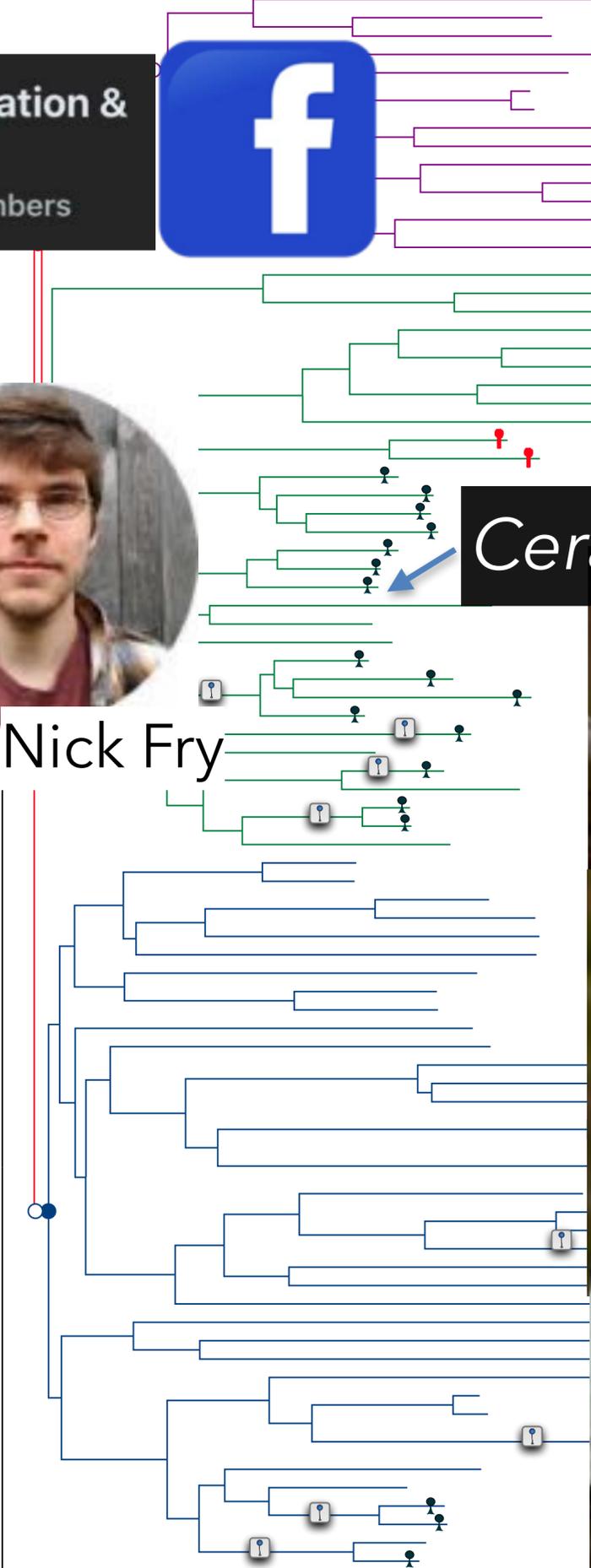


Slime Mold Identification & Appreciation

Public group · 37.1K members

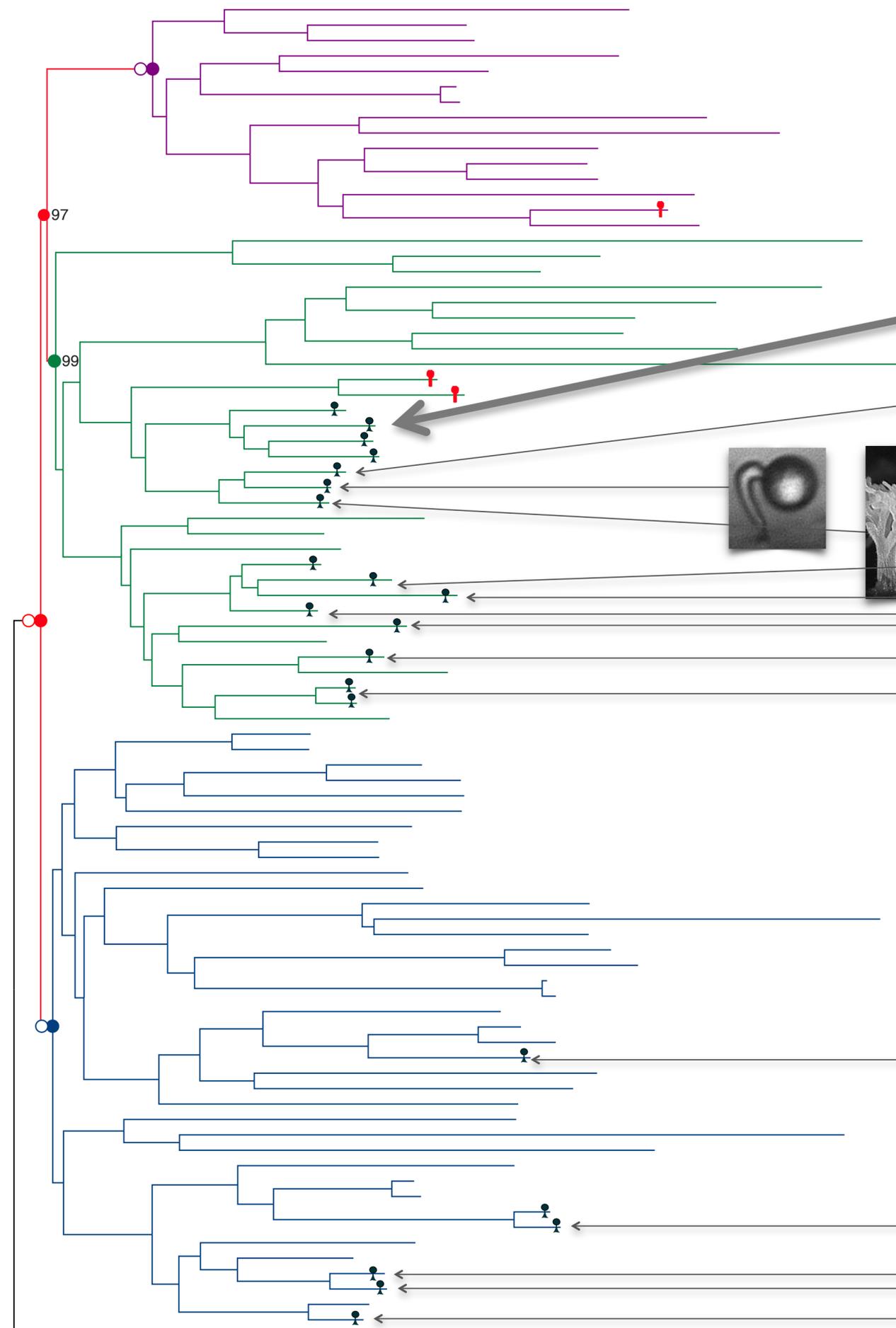


Nick Fry

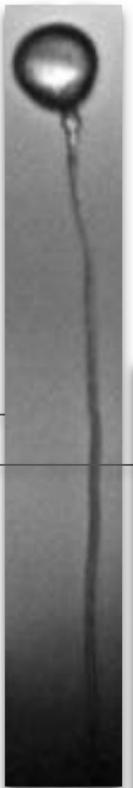
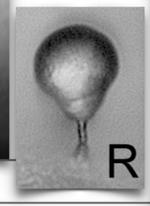
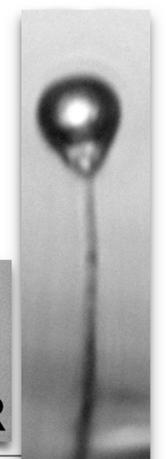
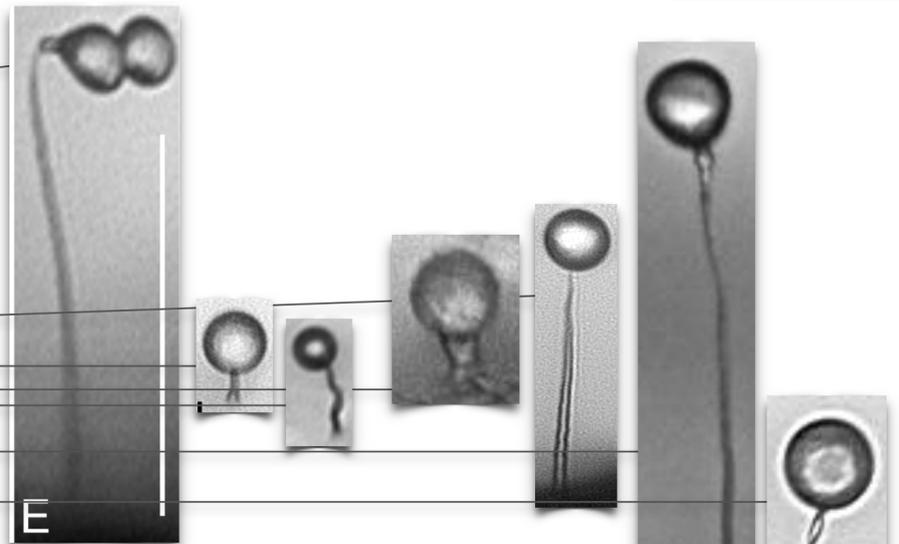
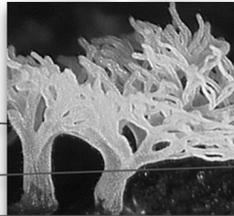
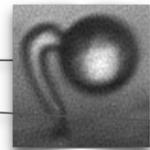
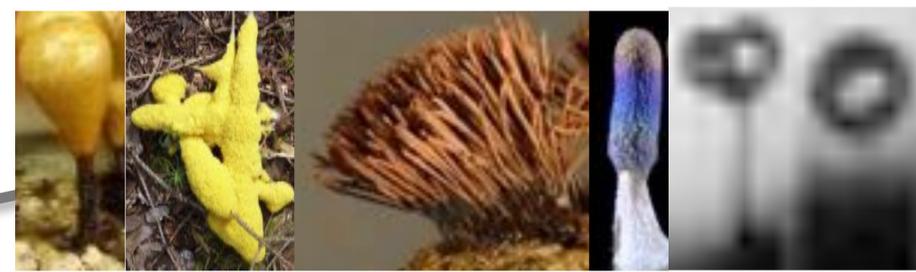


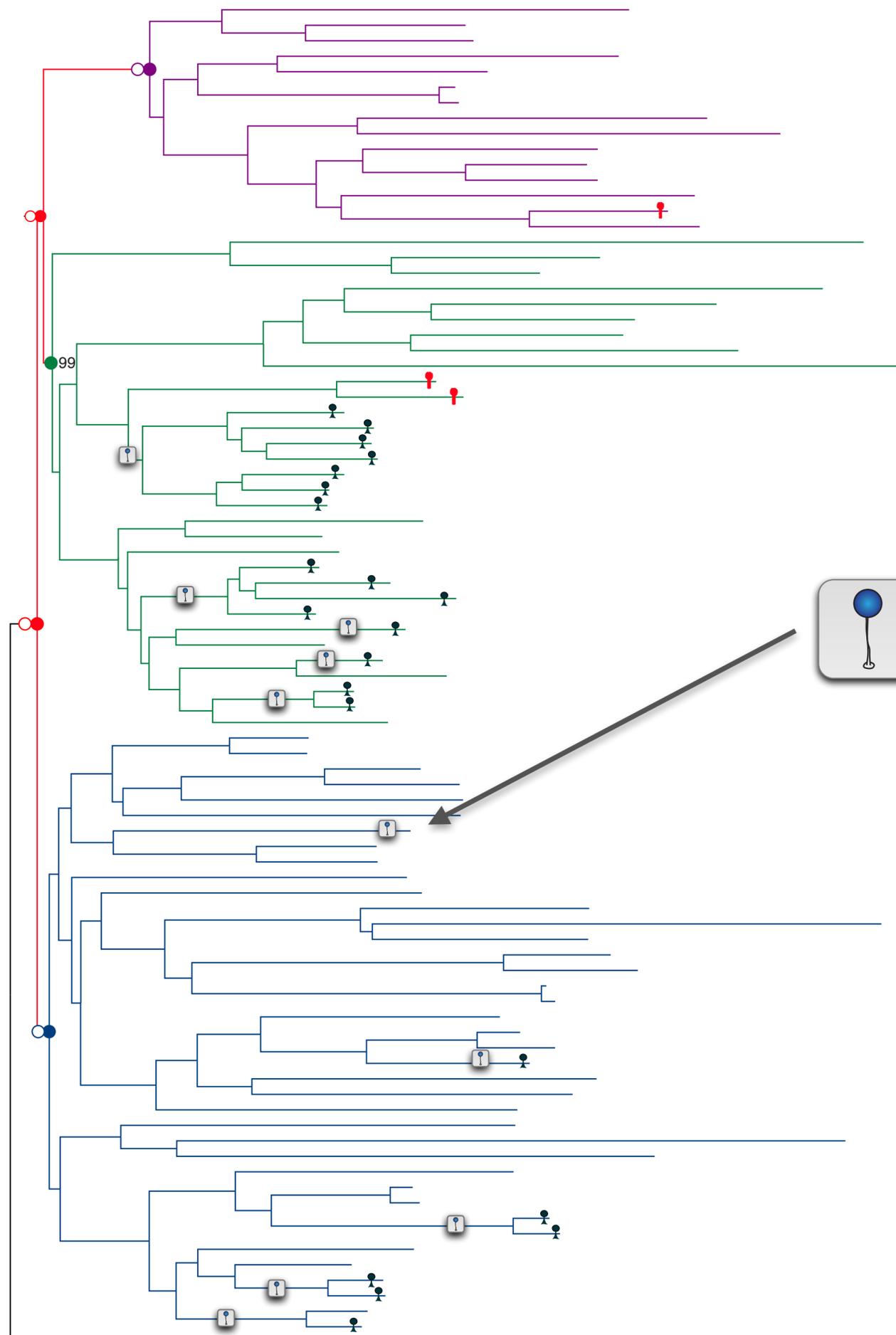
Ceratiomyxa fruticulosa



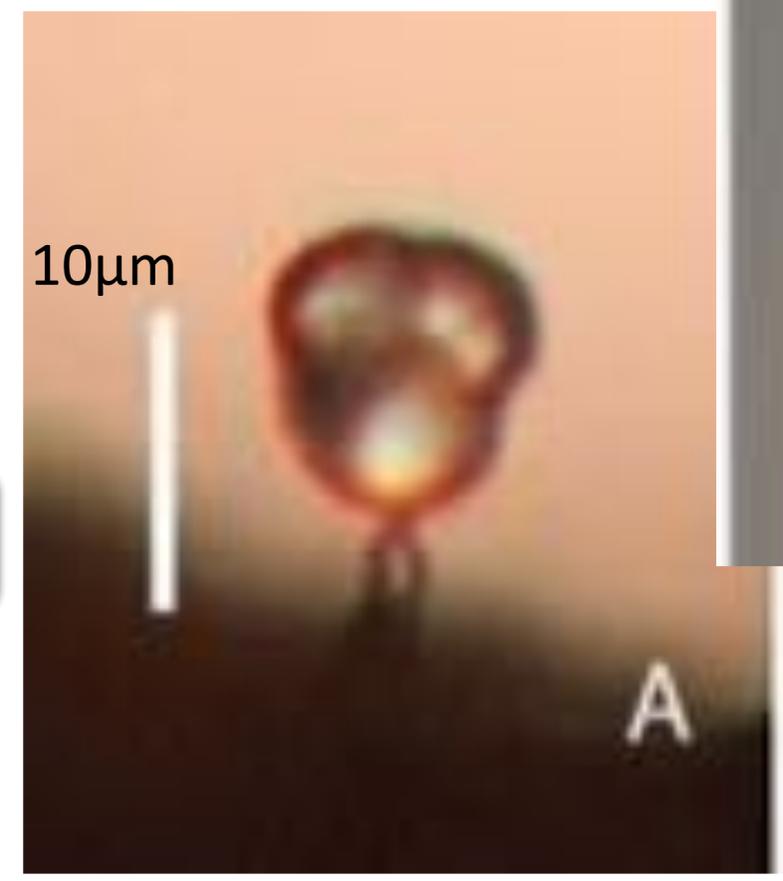


Myxogastrids

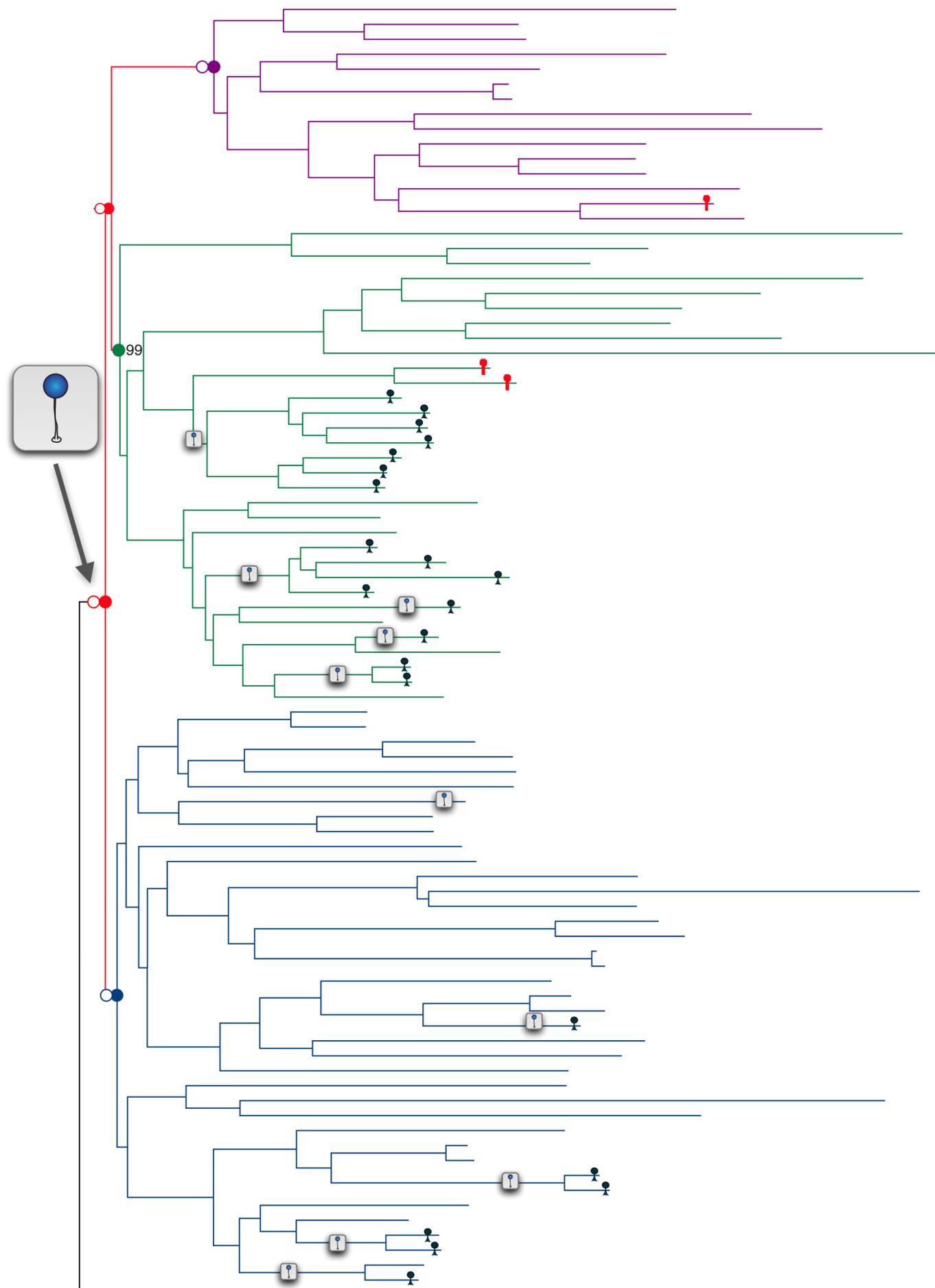




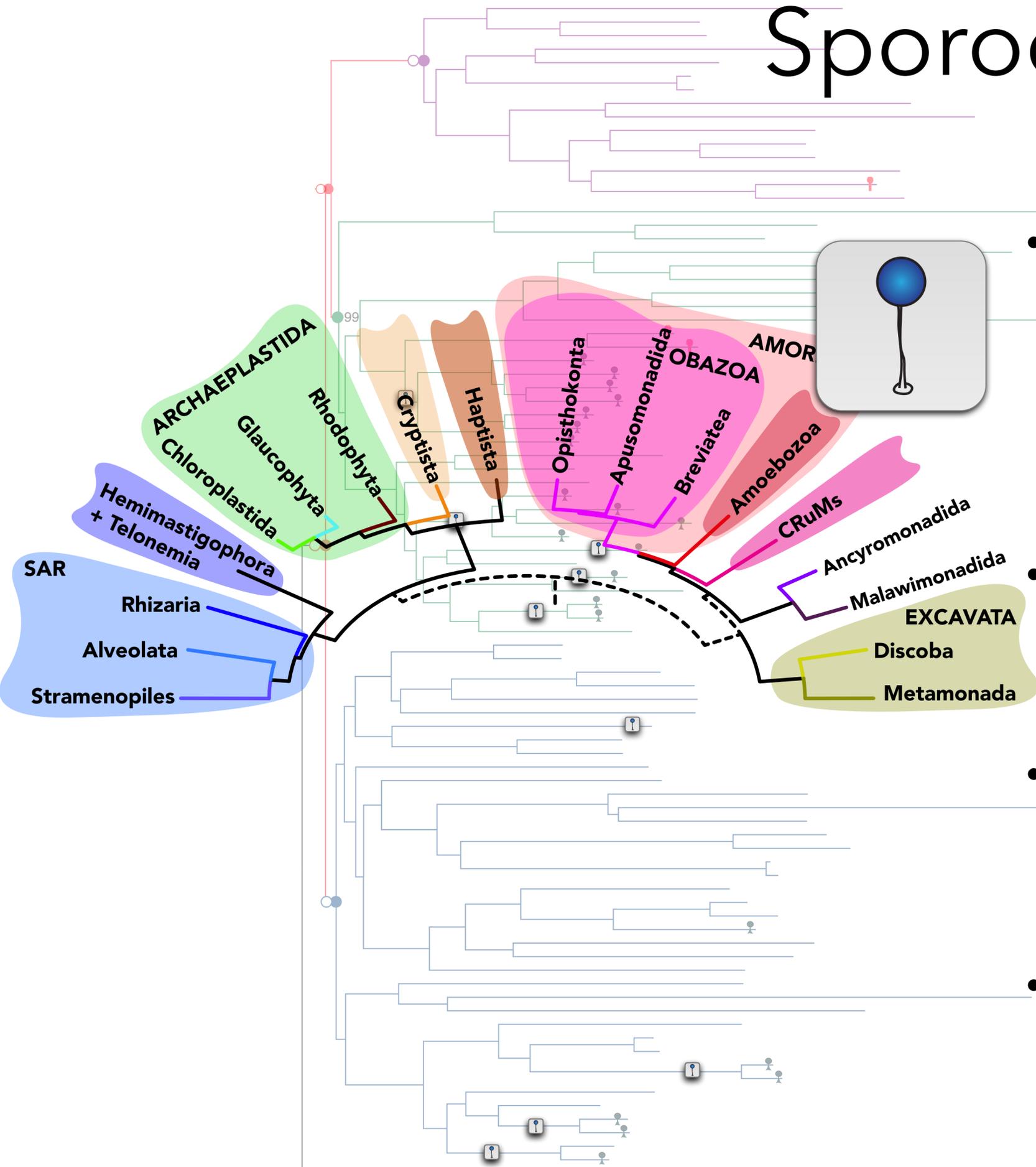
Microglomus paxillus



Tice & Brown. Submitted

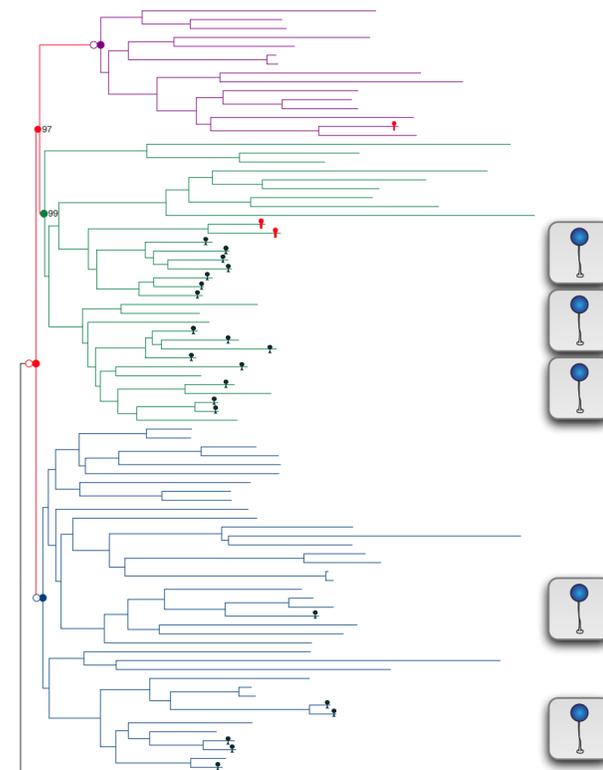
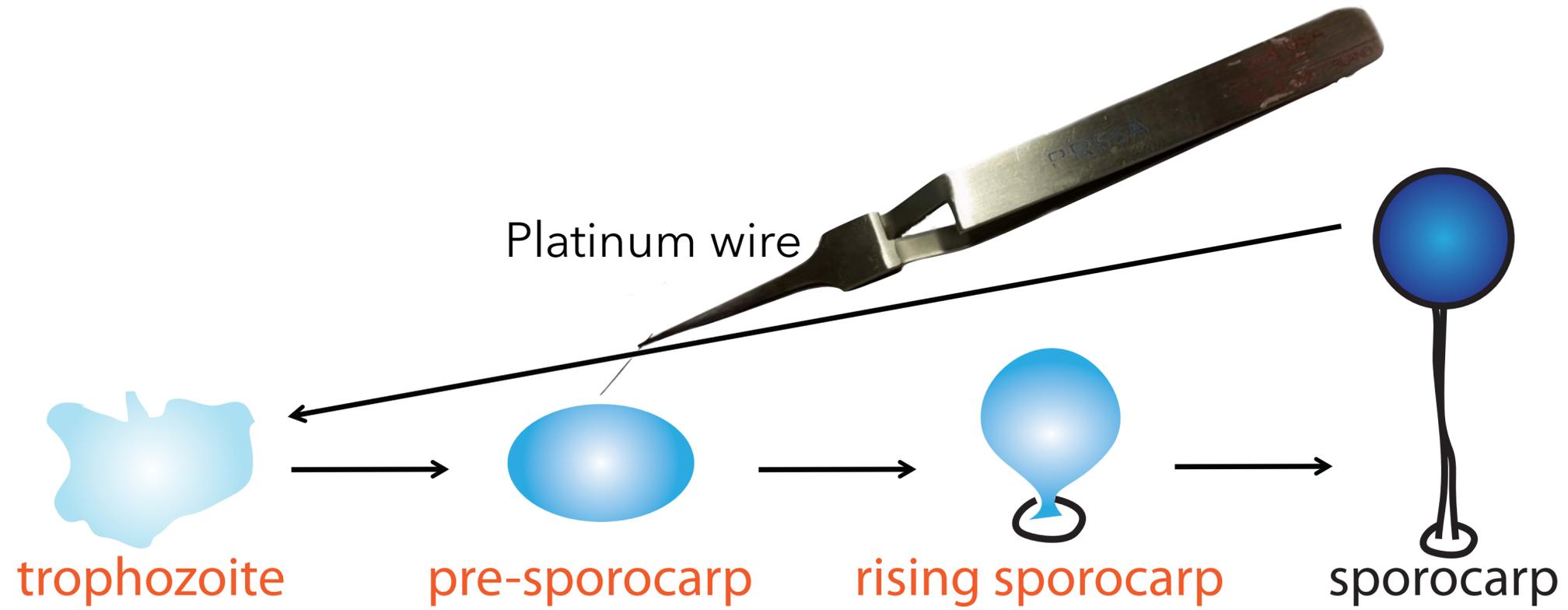


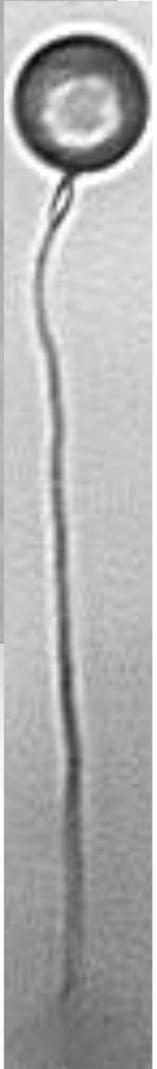
Sporocarpic Hypothesis



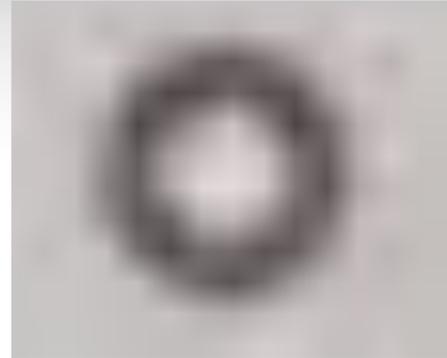
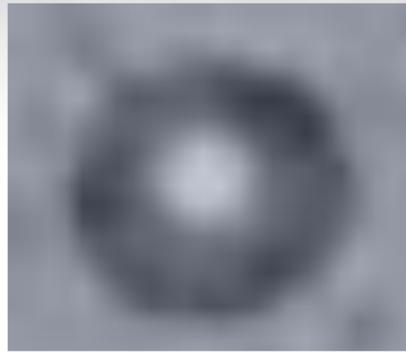
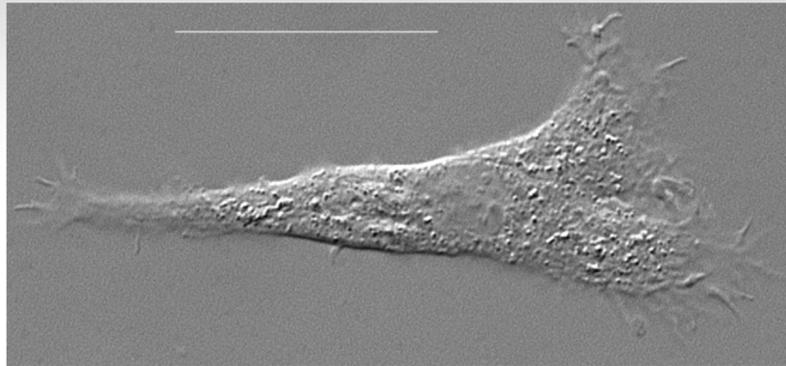
- Single origin of sporocarpy
 - Restricted to Amoebozoa
 - Life cycles are complex and varied
 - Stalk bearing structures are complex, but similar across taxa
- Hard to invent?
 - Unlike sorocarpy, sporocarpy is restricted to Amoebozoa
- What is more parsimonious?
 - single origin, many losses?
 - 9 independent origins?
- We can examine this using developmental transcriptomics!

SMART-Seq2 Picelli et al. 2014, 2015, Onsbring & Tice et al. 2021



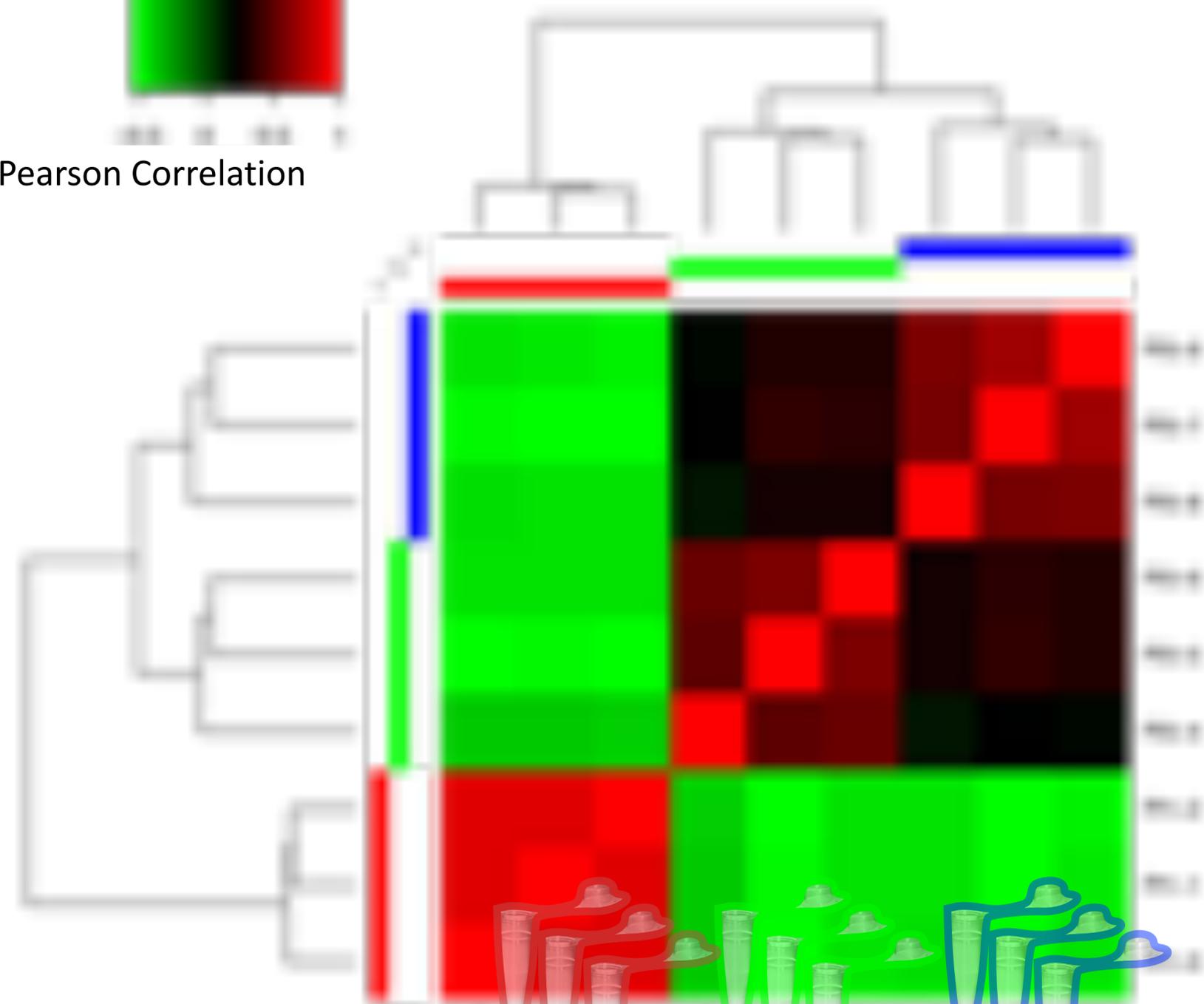
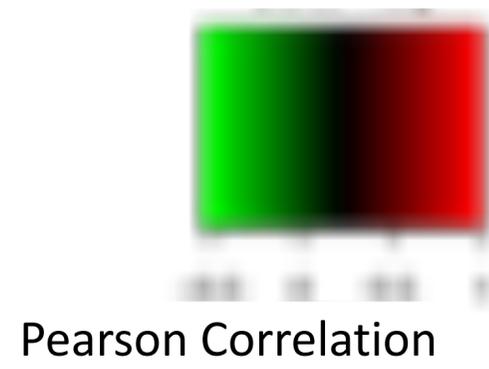


1:30h timelapse

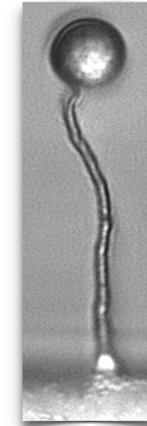


Soliformovum irregularis

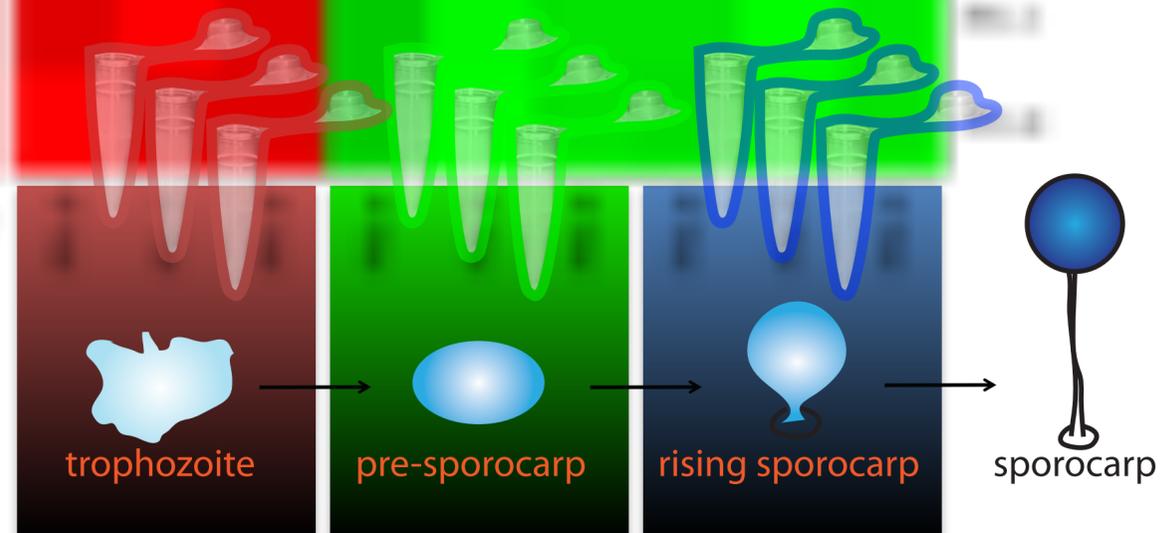
Sample Correlation Matrix



***Vannella
fimicola***

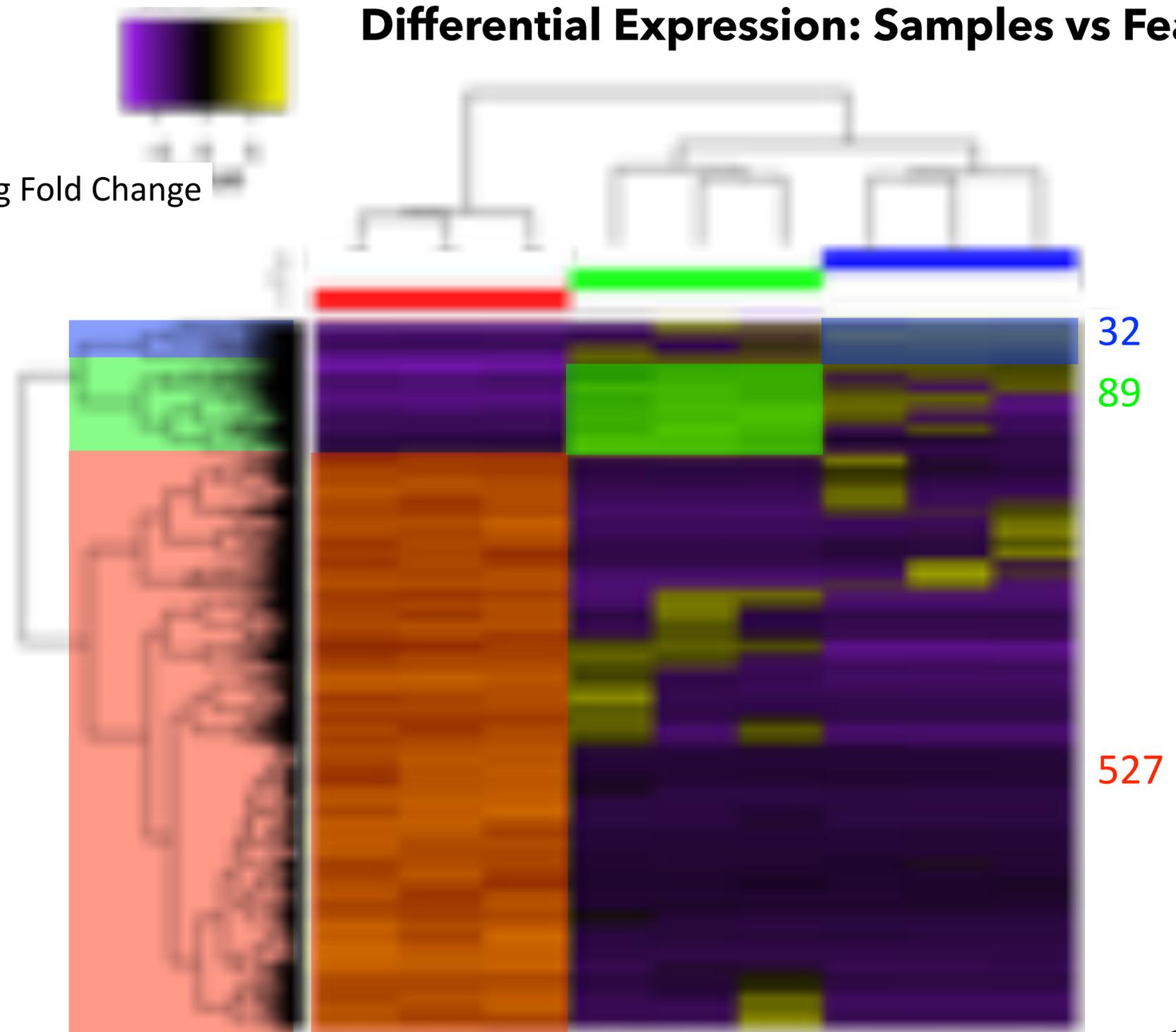


Pairwise FC = 4
p-value < 0.001

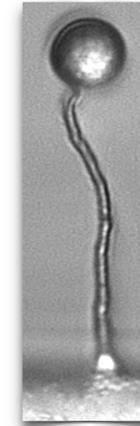


Differential Expression: Samples vs Features

Log Fold Change



*Vannella
fimicola*

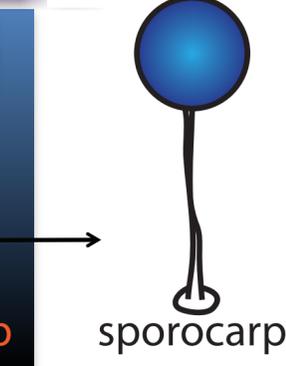
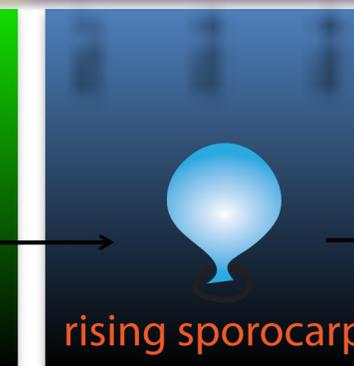
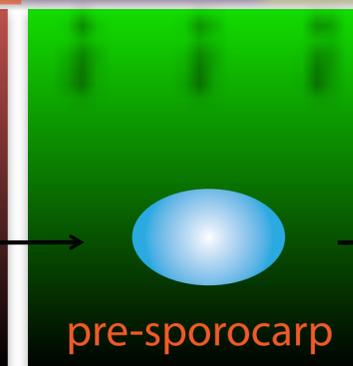
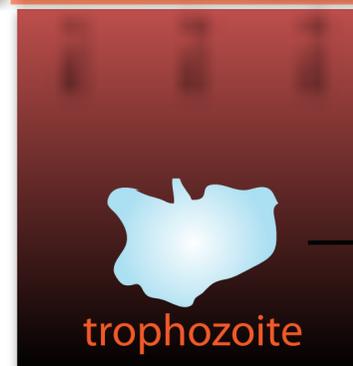


19916 Transcripts

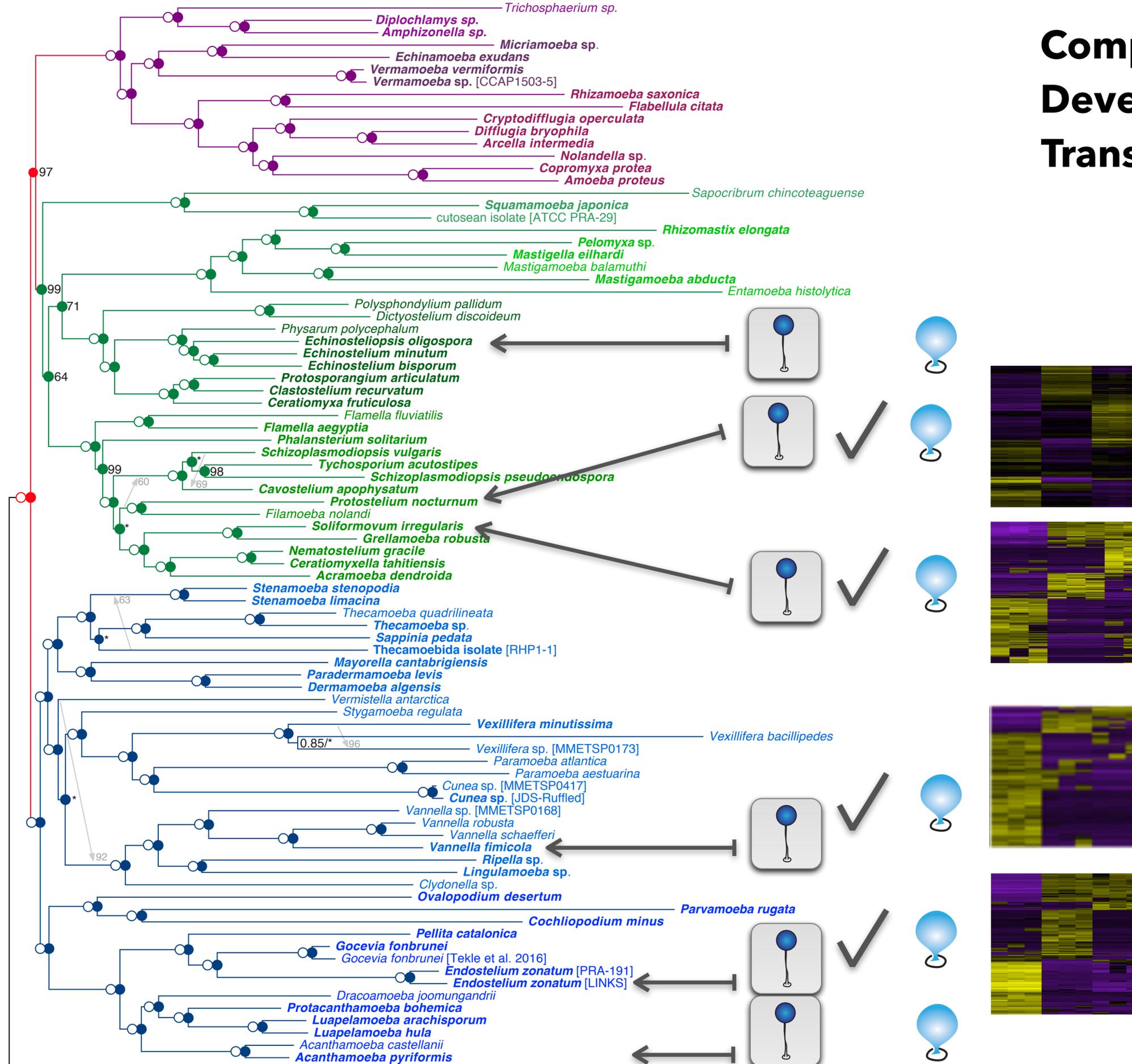
648 transcripts DE

Pairwise FC = 4

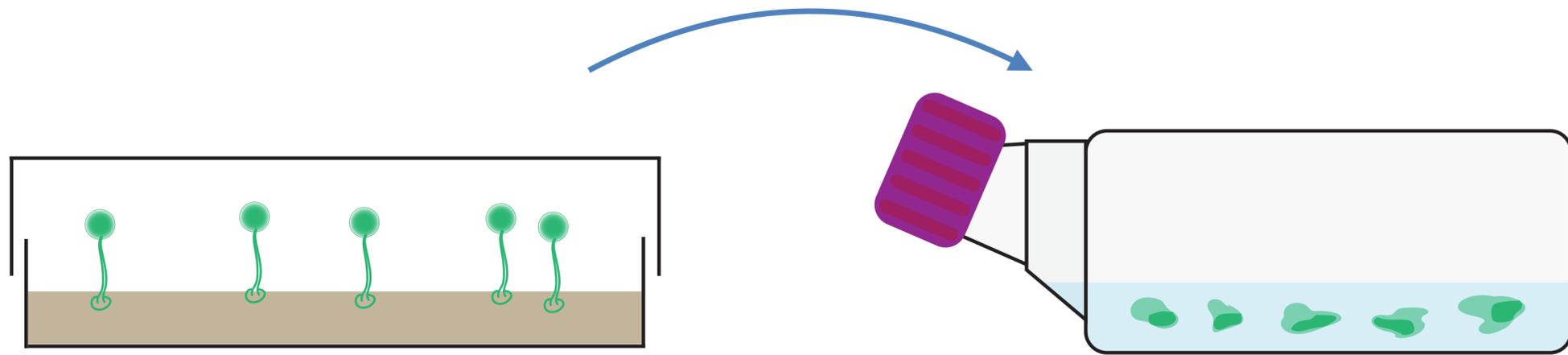
p-value < 0.001



Comparative Developmental Transcriptomics



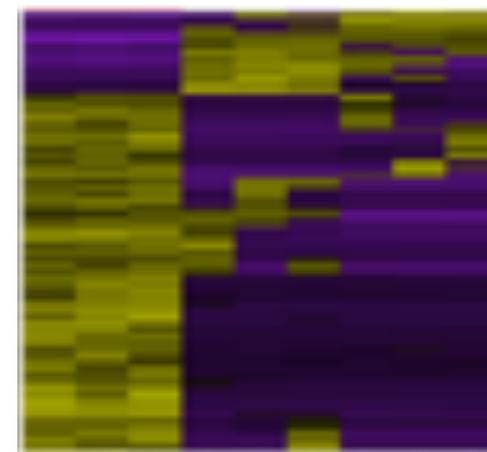
Synchronization of Sporocarpic Amoebae



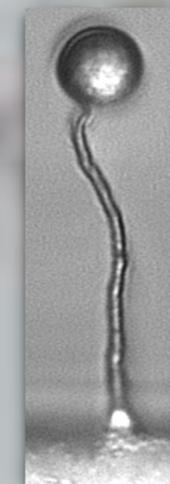
Vannella fimicola



Felicity Kleitz-Singleton

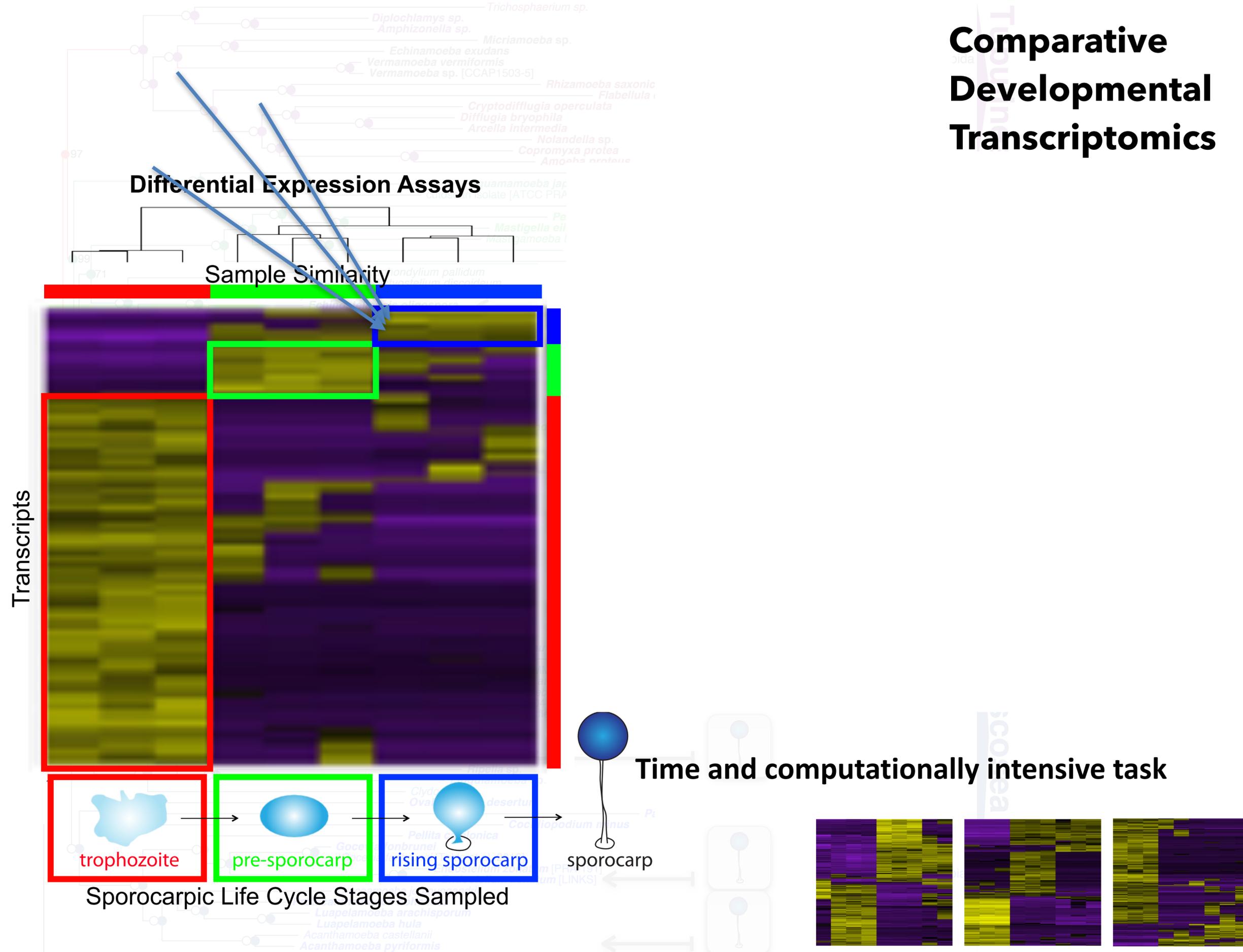


***Vannella
fimicola***

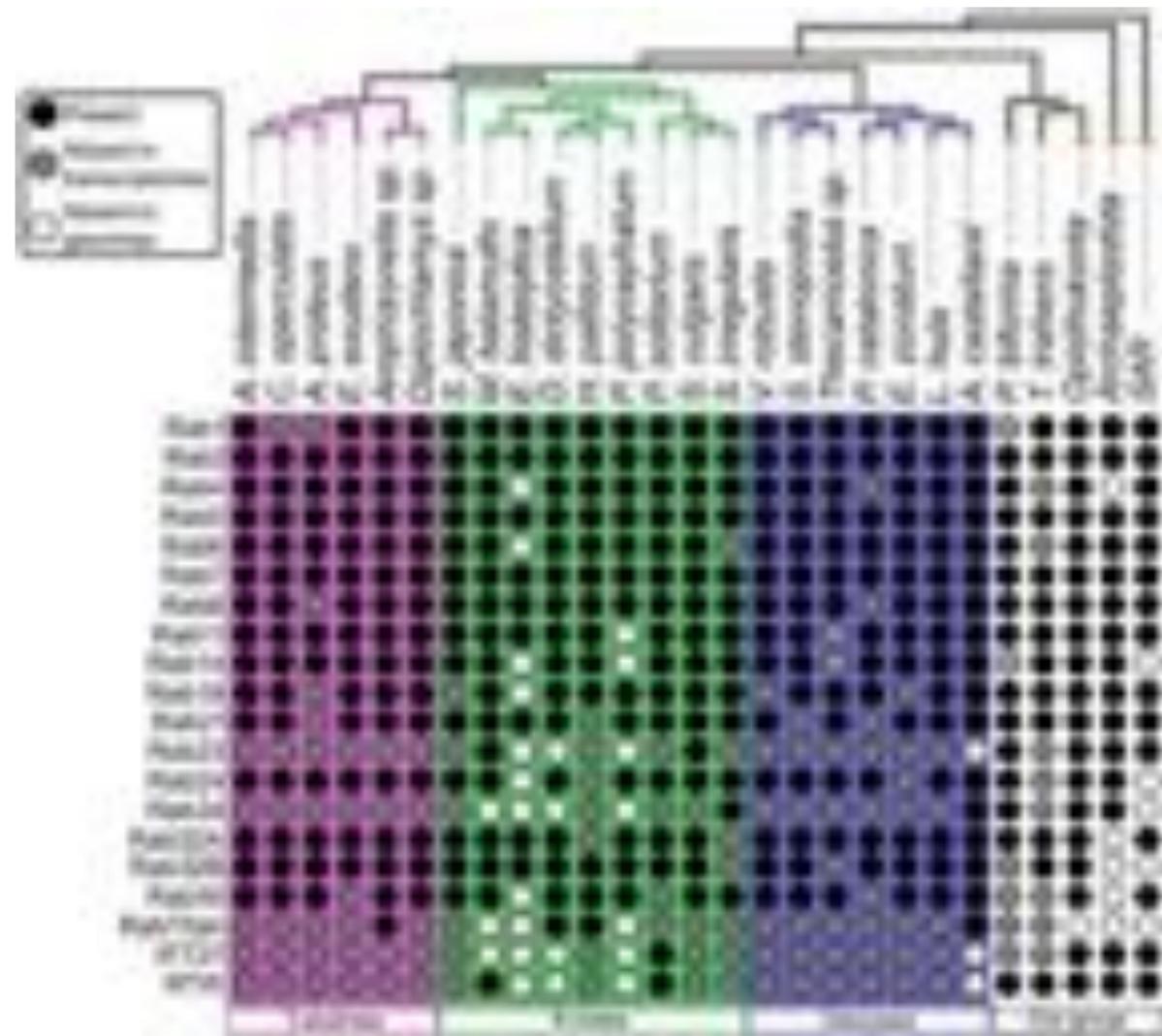
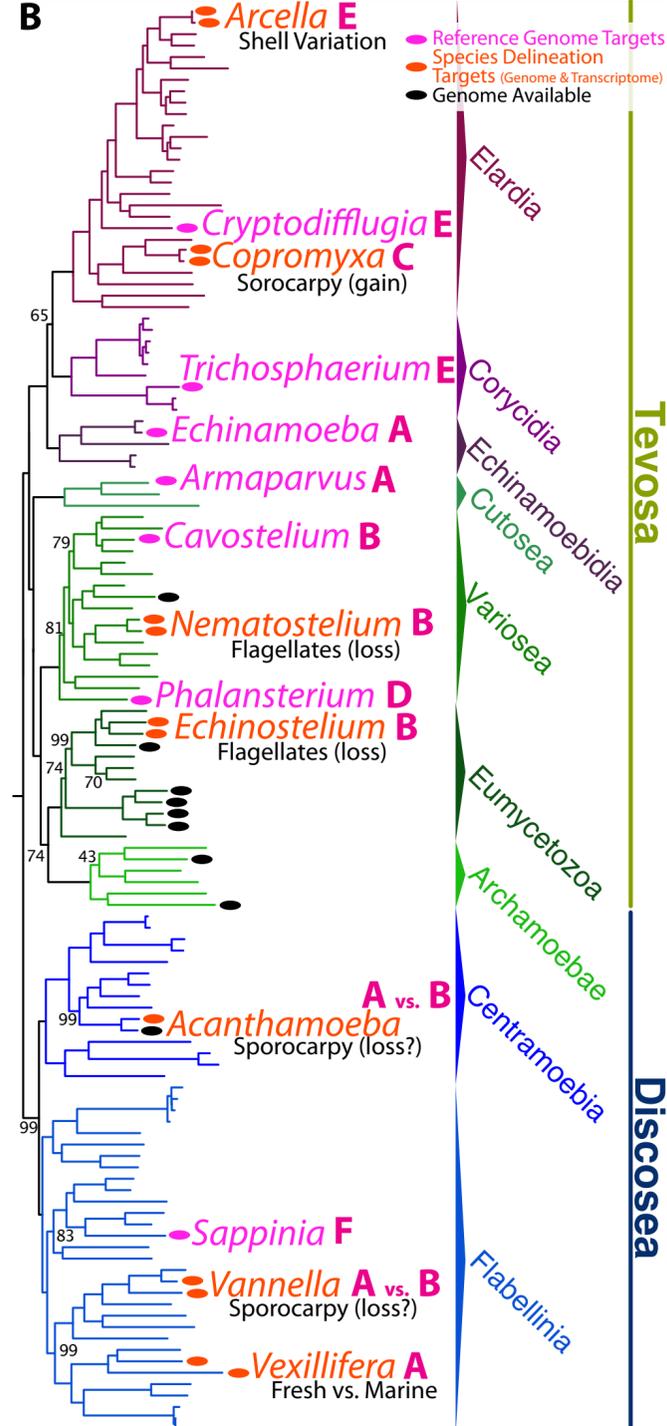
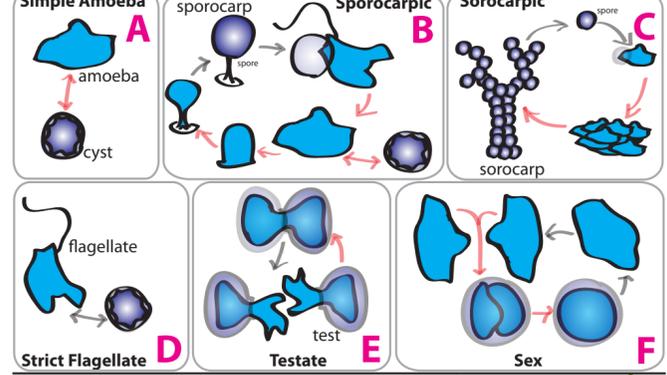


25.5hr time lapse
1 sec = 15min

Comparative Developmental Transcriptomics



A GENERALIZED LIFE CYCLE TYPES



Rab GTPases

Porfírio-Sousa et al. 2021.

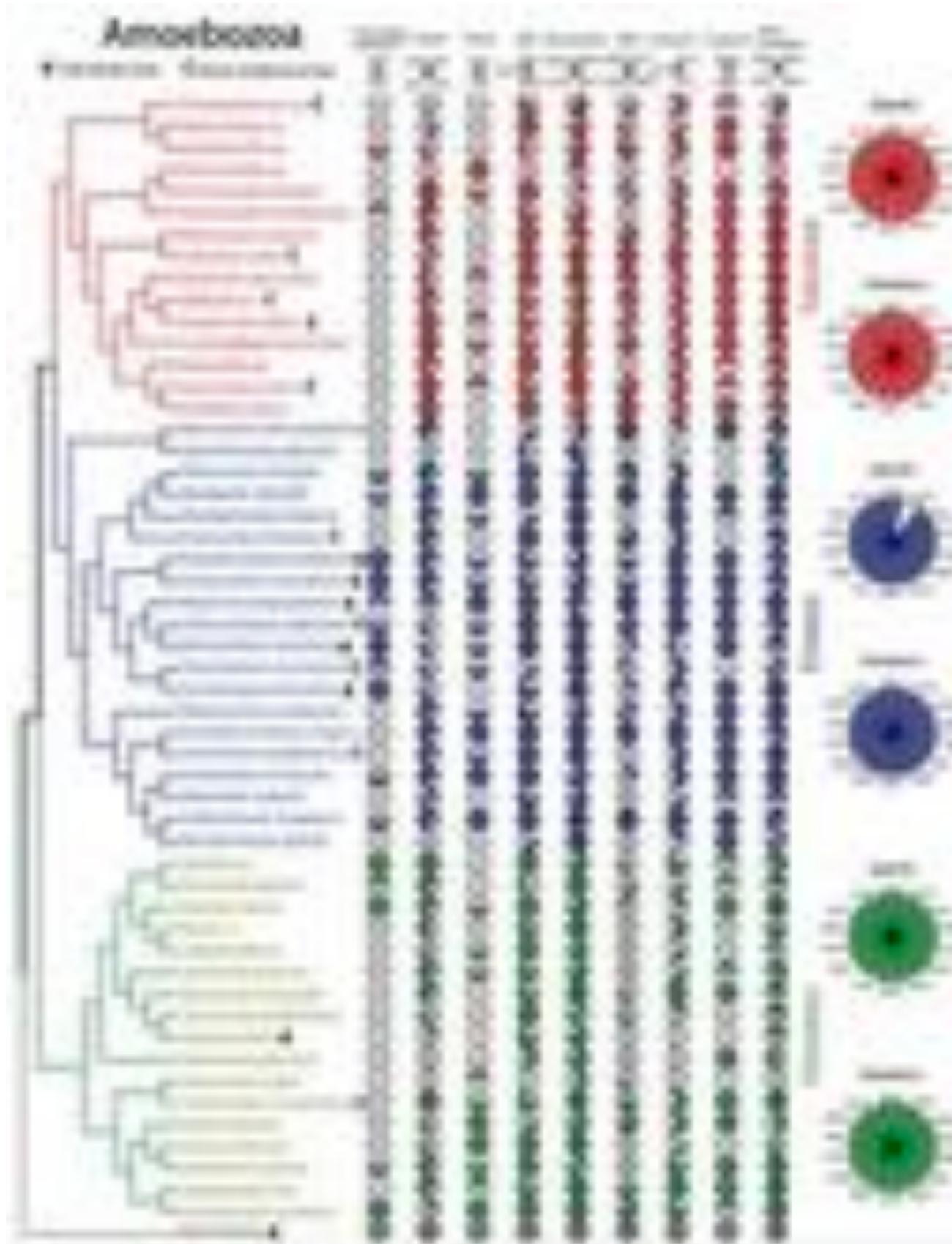
Small GTPases



GENOMES!

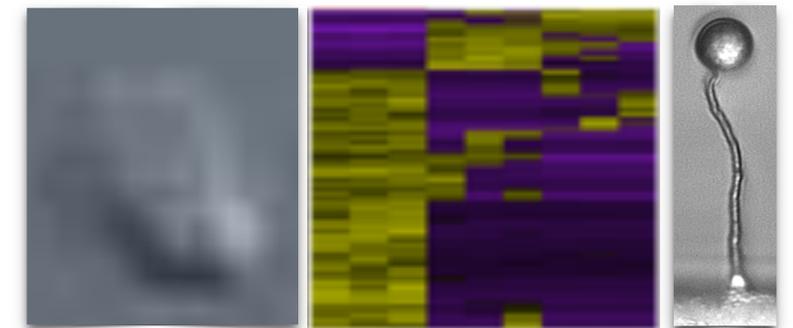
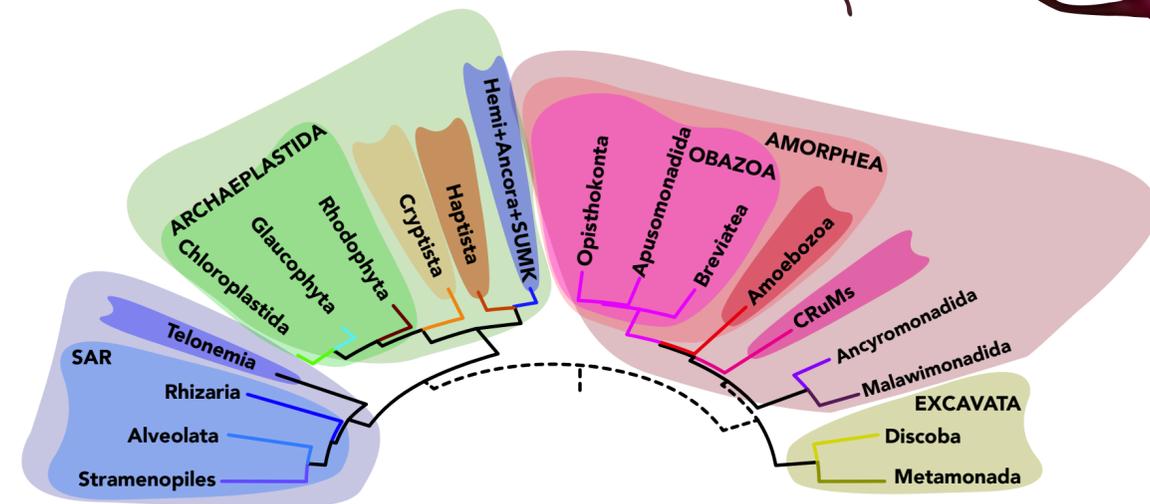
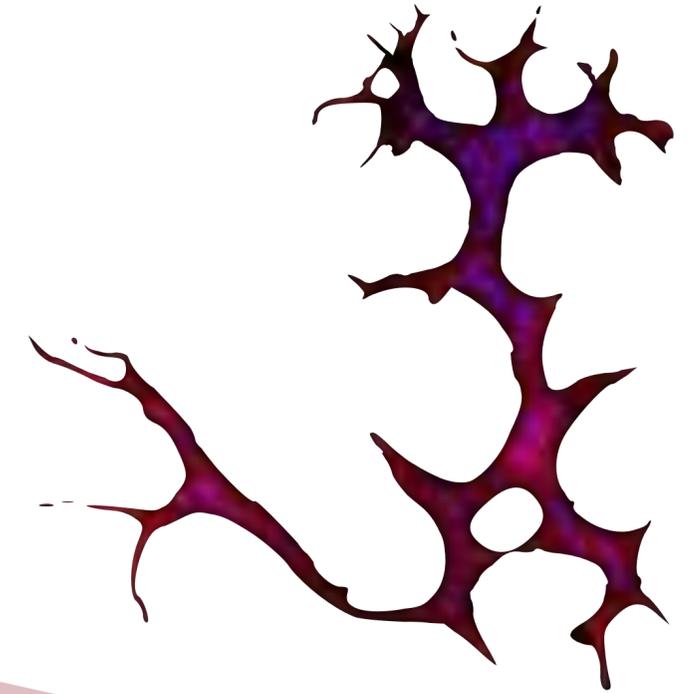
SEX - Meiosis

Hoffstater, Brown, Lahr. 2019. *GBE*



Conclusions

- Fully resolved phylogeny of a deep eukaryotic group
 - Amoebozoa
 - Complex last common ancestor
 - Genomically
 - Life History
 - Sampling diversity deeply is now possible
- PhyloFisher tool for phylogenomics
 - More diversity will change views
- Developmental Expression Profiling, a low input tool
 - Sporocarpic amoebae - Homologous?





Acknowledgments



- Alexander Tice, Ph.D.
- Quentin Blandenier, Ph.D.
- Seungho Kang
- Robert Jones
- Tomáš Pánek, Ph.D.
- Tristan Henderson
- Nicholas Fry
- Gabriel Schuler
- Kina Martin
- Alfredo de Sousa
- Chris Lebeau
- Peter Kooienga
- Emily Foreman
- Robert Berry
- Sammira Rais-Rohani
- Priya Patel
- Pamela Watson
- Stephanie Sorrell
- David Žihala
- Emily Dale
- Felicity Kleitz-Singleton
- Maggie Shepherd
- Violet Jira
- Nick Gibson



- **National Science Foundation** (DEB-SBS)
- Fulbright 
- College of Arts & Sciences – SRI
- Institute for Genomics, Biocomputing & Biotechnology
- Department of Biological Sciences
- Texas EcoLab  **Texas Ecological Laboratory**
- Frederick W. Spiegel, Ph.D.
- Daniel J. G. Lahr, Ph.D.
- Brendan Zurweller, Ph.D.



QUESTIONS?