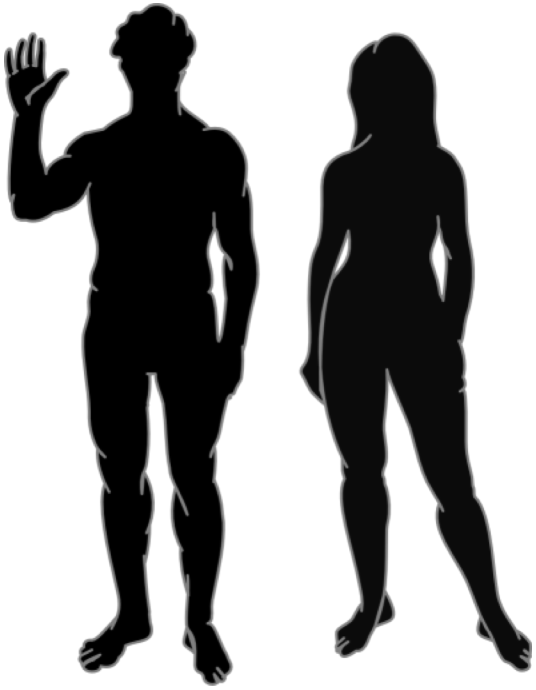
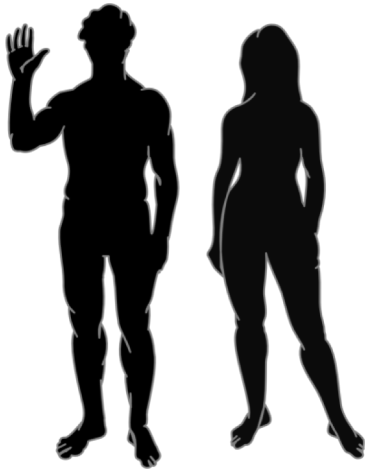
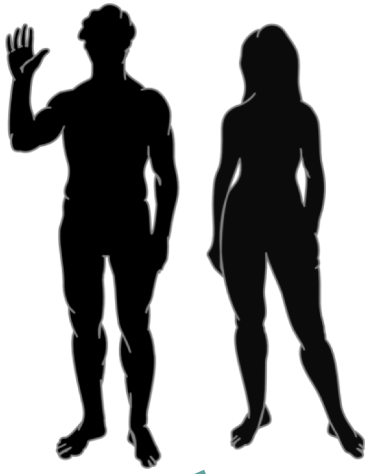
Conflict and concordance in phylogenomic datasets

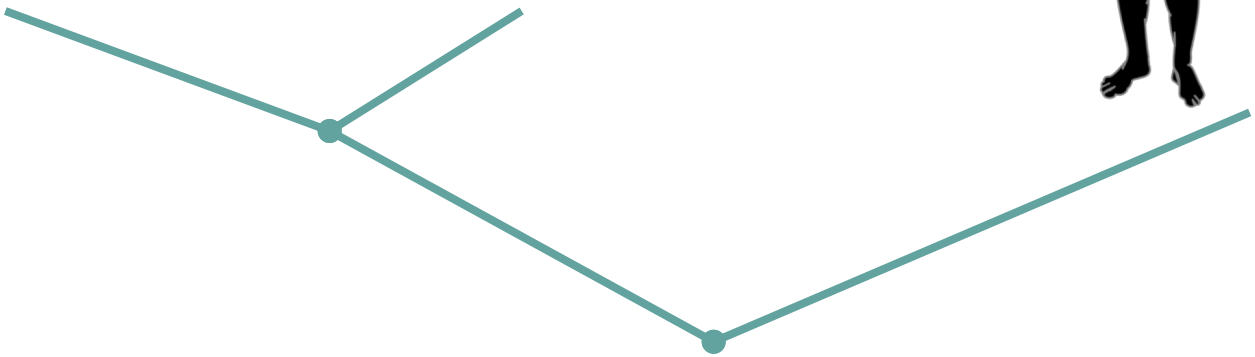
DR. STEPHEN A. SMITH

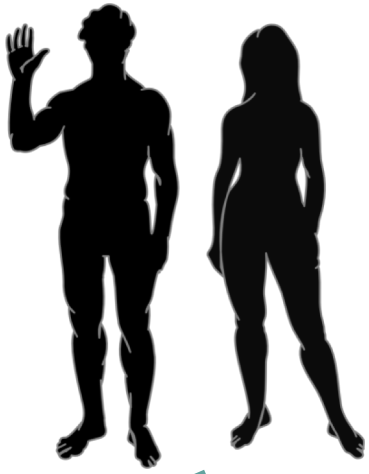
UNIVERSITY OF MICHIGAN











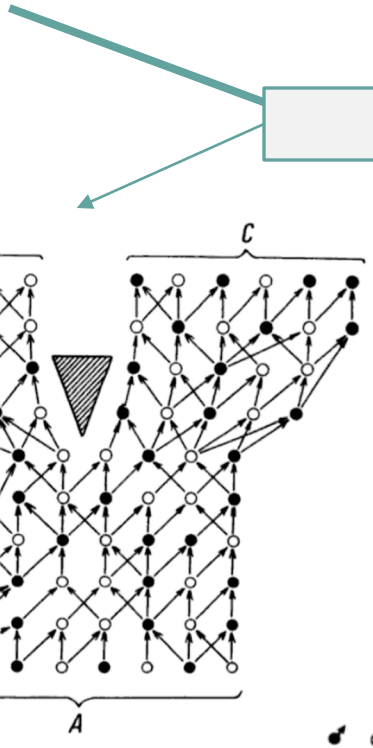
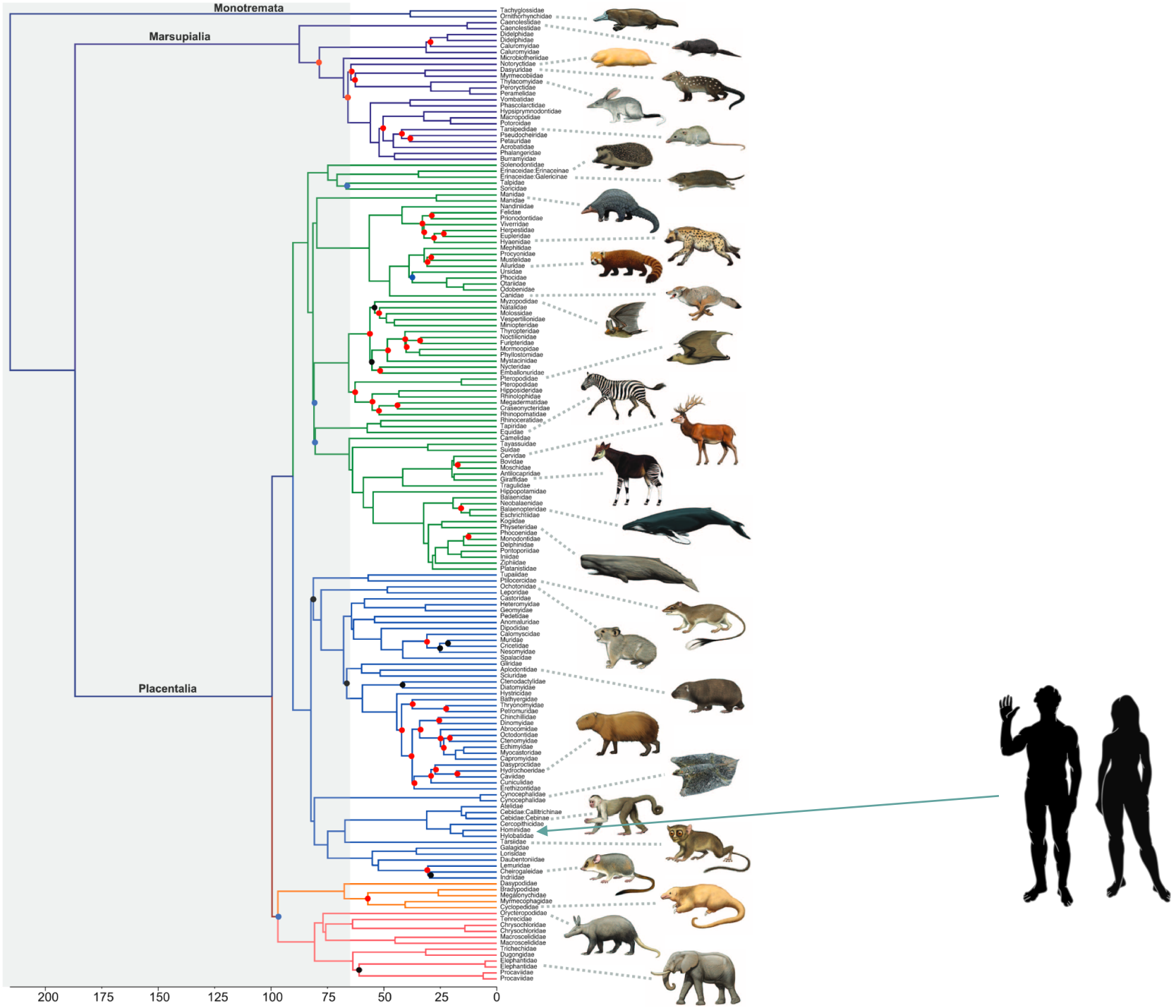
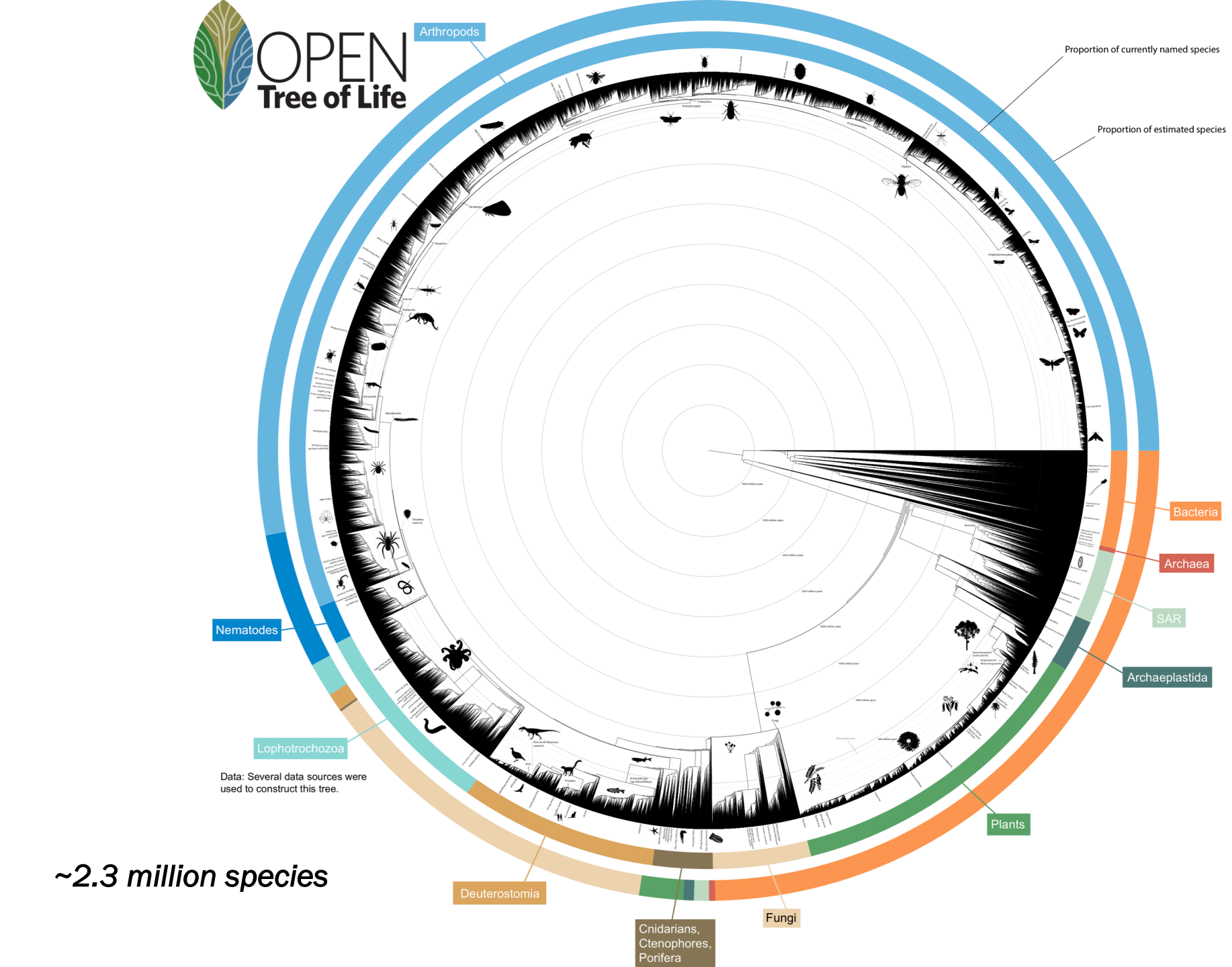


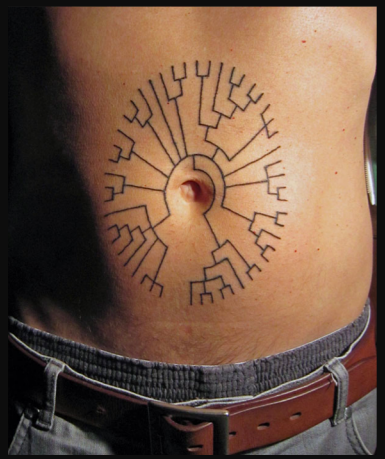
Figure 4. The process of species cleavage.


Monotremata
Marsupialia
Placentalia
Tachyglossidae
Ornithorhynchidae
Caenolestidae
Didelphidae
Didelphidae
Caluromyidae
Caluromyidae
Microbiotheriidae
Notoryctidae
Dasyuridae
Myrmecobiidae
Thylacomyidae
Peroryctidae
Peramelidae
Vombatidae
Phascolarctidae
Hypsiprymnodontidae
Macropodidae
Potoroidae
Tarsipedidae
Pseudocheiridae
Petauridae
Acrobatidae
Phalangeridae
Burramyidae
Solenodontidae
Erinaceidae:Erinaceinae
Erinaceidae:Galericinae
Talpidae
Soricidae
Manidae
Manidae
Nandiniidae
Felidae
Prionodontidae
Viverridae
Herpestidae
Eupleridae
Hyaenidae
Mephitidae
Procyonidae
Mustelidae
Ailuridae
Ursidae
Phocidae
Otariidae
Odobenidae
Canidae
Myzopodidae
Natalidae
Molossidae
Vespertilionidae
Miniopteridae
Thyropteridae
Noctilionidae
Furipteridae
Mormoopidae
Phyllostomidae
Mystacinidae
Nycteridae
Emballonuridae
Pteropodidae
Pteropodidae
Hippocrepidae
Rhinolophidae
Megadermatidae
Craseonycteridae
Rhinopomatidae
Rhinoceratidae
Tapiridae
Equidae
Camelidae
Tayassuidae
Suidae
Cervidae
Bovidae
Moschidae
Antilocapridae
Giraffidae
Tragulidae
Hippopotamidae
Balaenidae
Neobalaenidae
Balaenopteridae
Eschrichtiidae
Kogiidae
Physeteridae
Phocoenidae
Monodontidae
Delphinidae
Peritoporiidae
Iniidae
Ziphiidae
Platanistidae
Tupaiidae
Ptilocercidae
Ochotonidae
Leporidae
Castoridae
Heteromyidae
Geomyidae
Pedetidae
Anomaluridae
Dipodidae
Calomyscidae
Muridae
Cricetidae
Nesomyidae
Spalacidae
Gliridae
Aplodontidae
Sciuridae
Ctenodactylidae
Diatomyidae
Hystricidae
Bathyergidae
Thryonomyidae
Petromuridae
Chinchillidae
Dinomyidae
Abrocomidae
Octodontidae
Ctenomyidae
Echimyidae
Myocastoridae
Capromyidae
Dasyproctidae
Hydrochoeridae
Caviidae
Cuniculidae
Erethizontidae
Cynocephalidae
Cynocephalidae
Atelidae
Cebidae:Callitrichinae
Cebidae:Cebinae
Cercopithicidae
Hominidae
Hylobatidae
Tarsiidae
Galagidae
Lorisidae
Daubentoniidae
Lemuridae
Cheirogaleidae
Indriidae
Dasypodidae
Bradypodidae
Megalonychidae
Myrmecophagidae
Cyclopedidae
Orycteropodidae
Tenrecidae
Chrysochloridae
Chrysochloridae
Macroscelididae
Macroscelididae
Trichechidae
Dugongidae
Elephantidae
Elephantidae
Procaviidae
Procaviidae
200
175
150
125
100
75
50
25
0

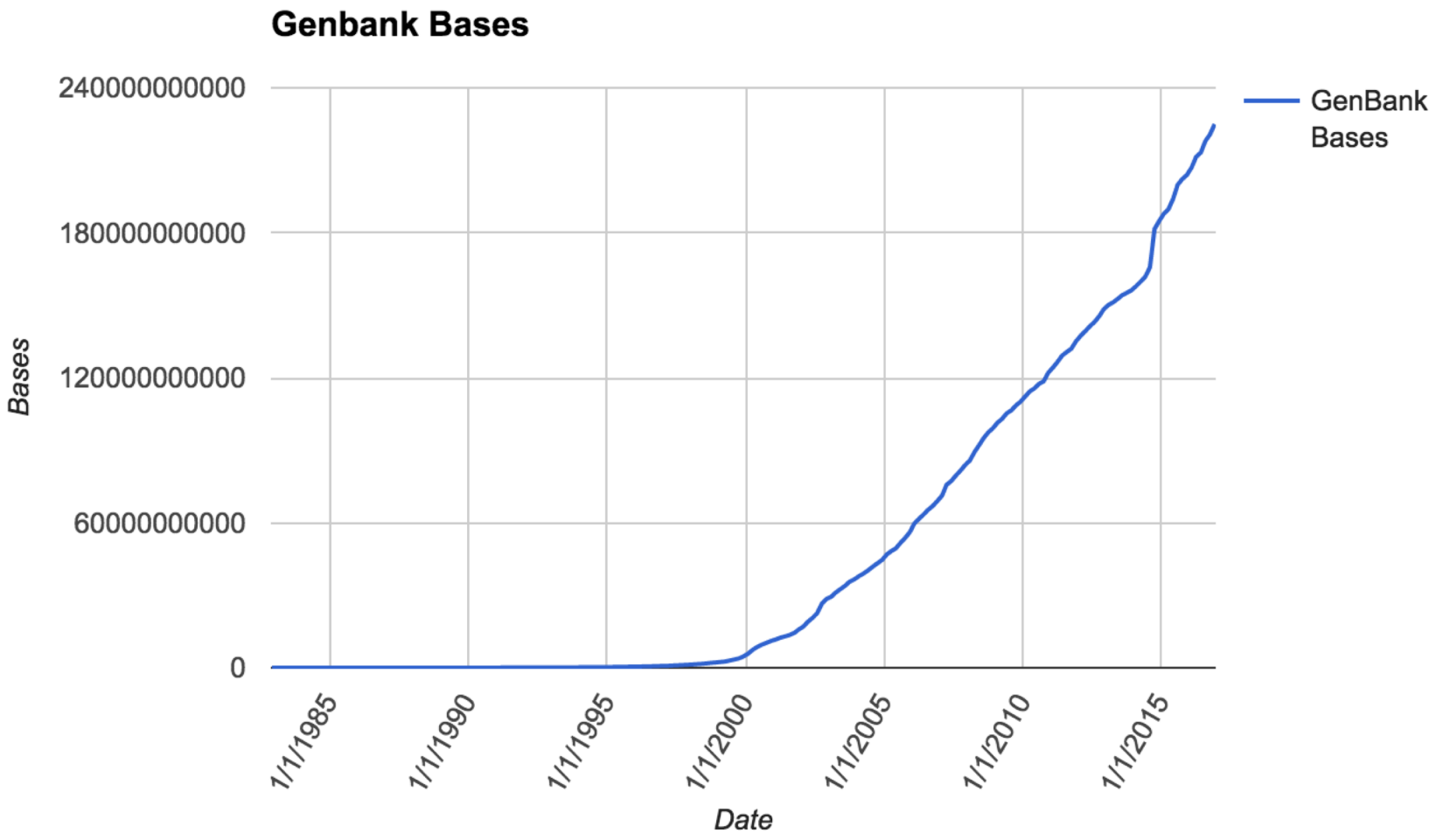


~2.3 million species

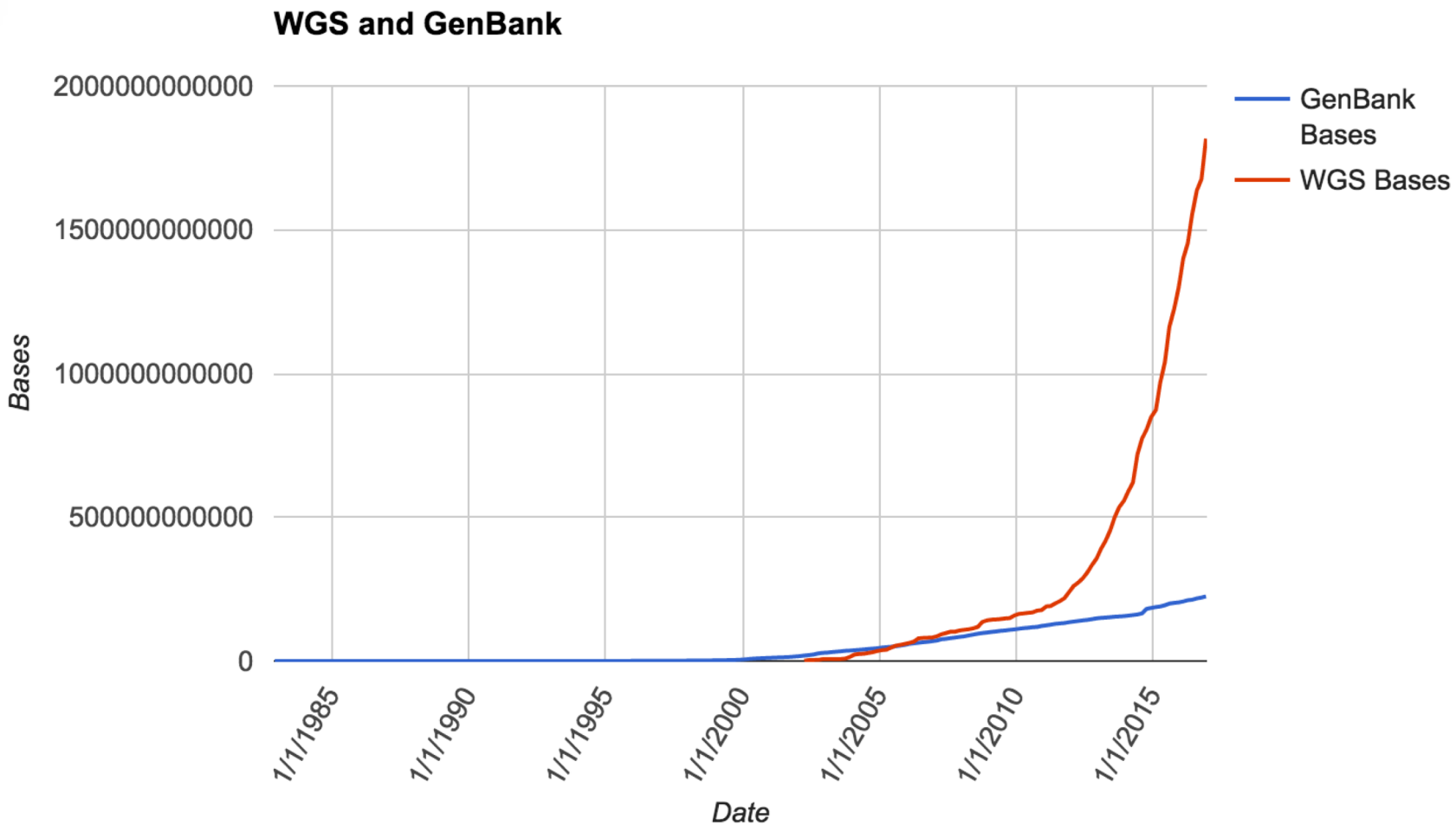

WE ARE HERE



Molecular data availability



Molecular data availability (with whole genomes)



(GIGA)n SCIENCE

Articles Submit ▾ Alerts About ▾ All GigaSci

Article Contents

ACCEPTED MANUSCRIPT

10KP: A Phylodiverse Genome Sequencing Plan

Shifeng Cheng ✉, Michael Melkonian, Stephen A Smith, Samuel Brockington, John M Archibald, Pierre-Marc Delaux, Fay-wei Li, Barbara Melkonian, Evgeny V Mavrodiev, Wenjing Sun ... Show more

Author Notes

GigaScience, giy013, https://doi.org/1

Published: 20 February 2018

Freshwater alga in the genus Zygnema would be one target of sequencing project. NORBERT HÜLSMANN/FLICKR (CC BY-NC-SA 2.0)

Plant scientists plan massive effort to sequence 10,000 genomes

By Dennis Normile | Jul. 27, 2017 , 8:00 AM

Large datasets with many genes

Typical phylogenetic analyses

- 1-10 genes
- 17 genes. Plants (Soltis et al. 2011)
- 19 genes. Birds (Hackett et al. 2007)

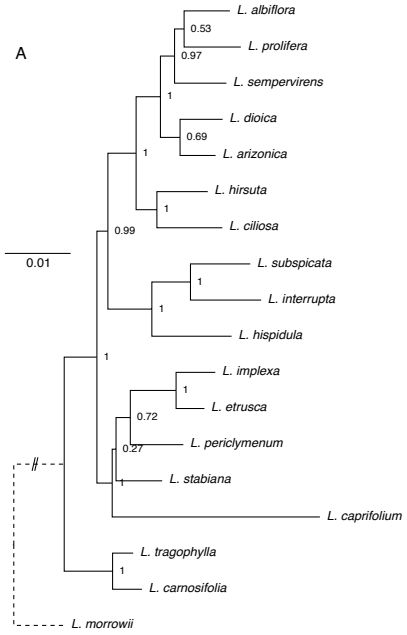
Transcriptomic and genomic phylogenetic analyses

- 140 genes. Metazoa (Dunn et al. 2008)
- 242 genes. Metazoa (Ryan et al. 2013)
- 248 genes. Turtles (Chiari et al., 2012)
- 1185 genes. Molluscs (Smith et al. 2011)
- 1720 genes. Rice (Cranston et al. 2007)
- 2970 genes. Seed plants (Lee et al. 2011)
- >8000 genes. Birds (Jarvis et al. 2014)
- 259 genes. Birds (Prum et al. 2015)
- 859 genes. Seed plants (Wickett et al. 2014)

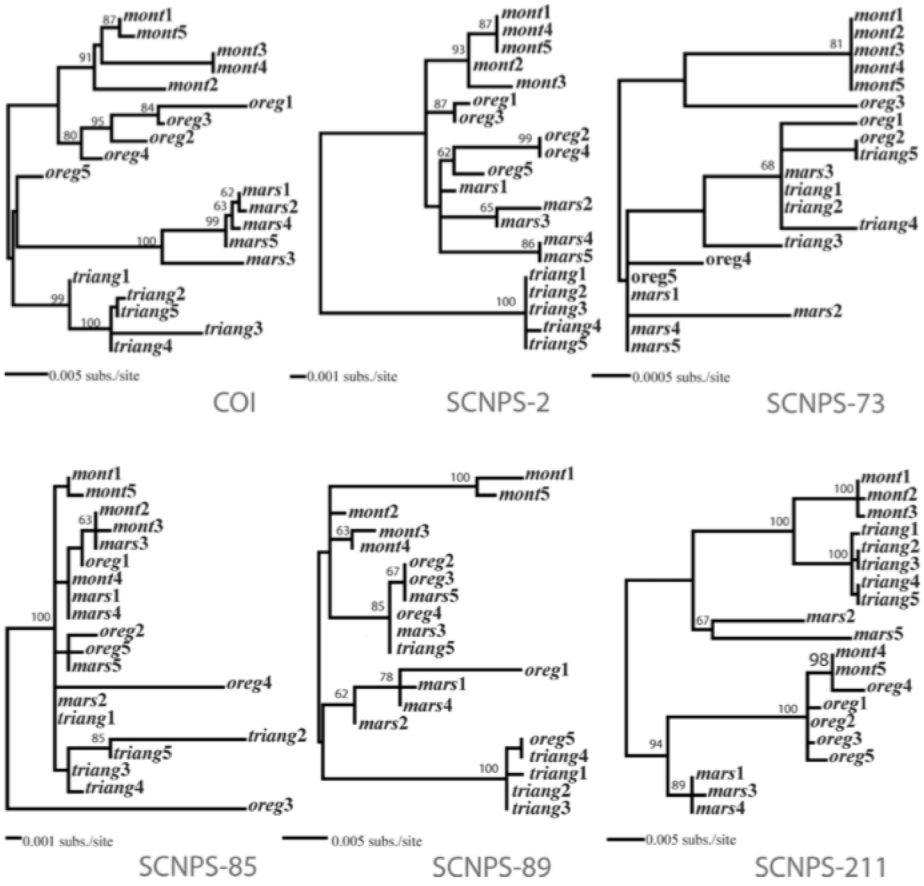
Concatenate genes to get more information

Gene 1 Gene 2 Gene 3 Gene 4 Gene 5

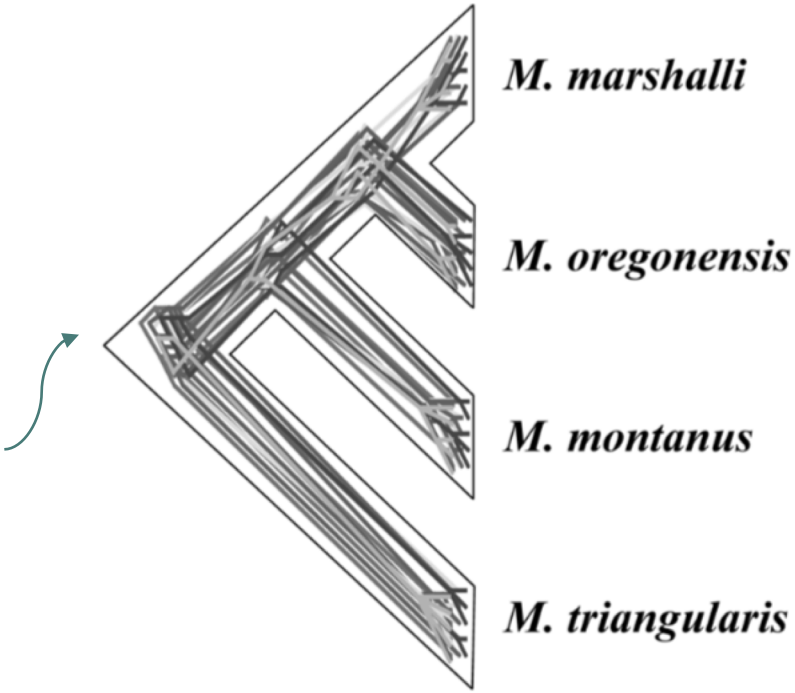
More information
(all the genes together)



Combining gene trees (using conflict as information)



Each gene contributes individually



Assume one underlying tree

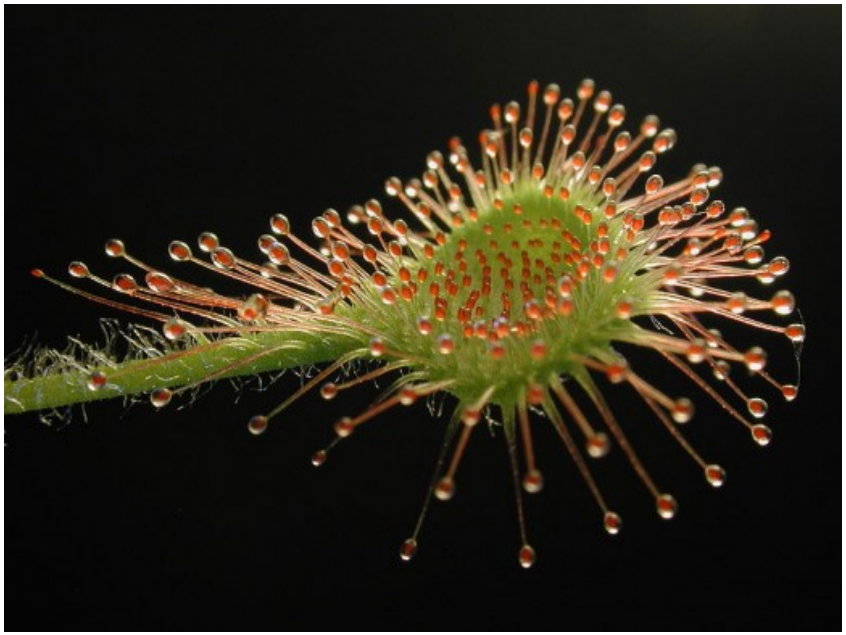
Triumphant phylogenomics

EXPLORING UNDERLYING MOLECULAR PATTERNS AND PROCESSES

Stephen A. Smith

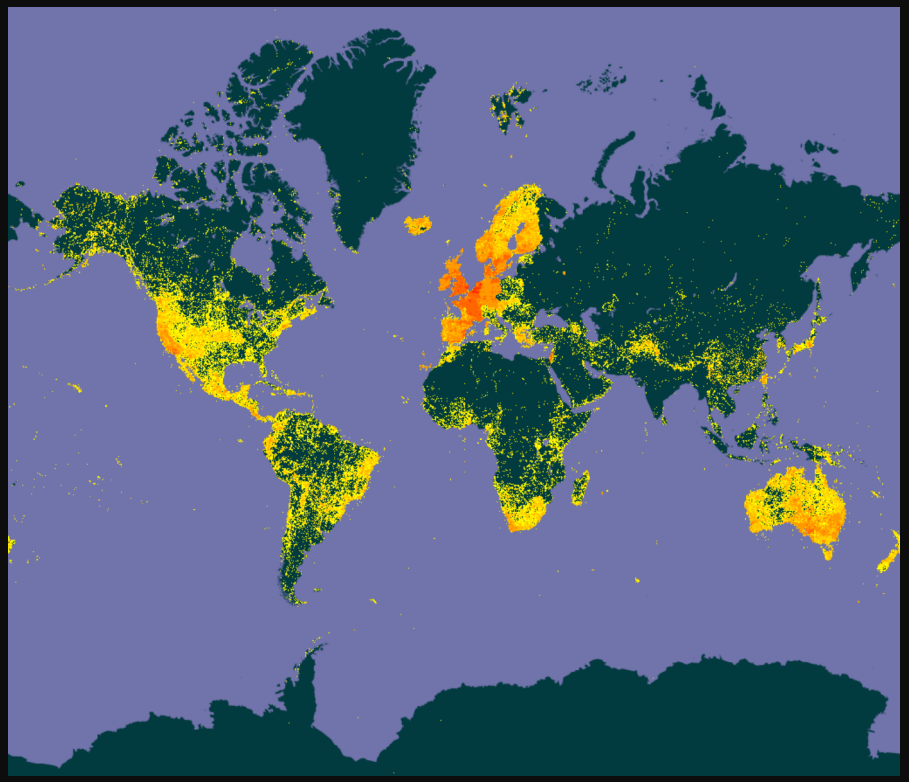
Caryophyllales

- >12,500 species in 39 families
- extreme disparity in life history and ecology



Carnivory

Cold environments

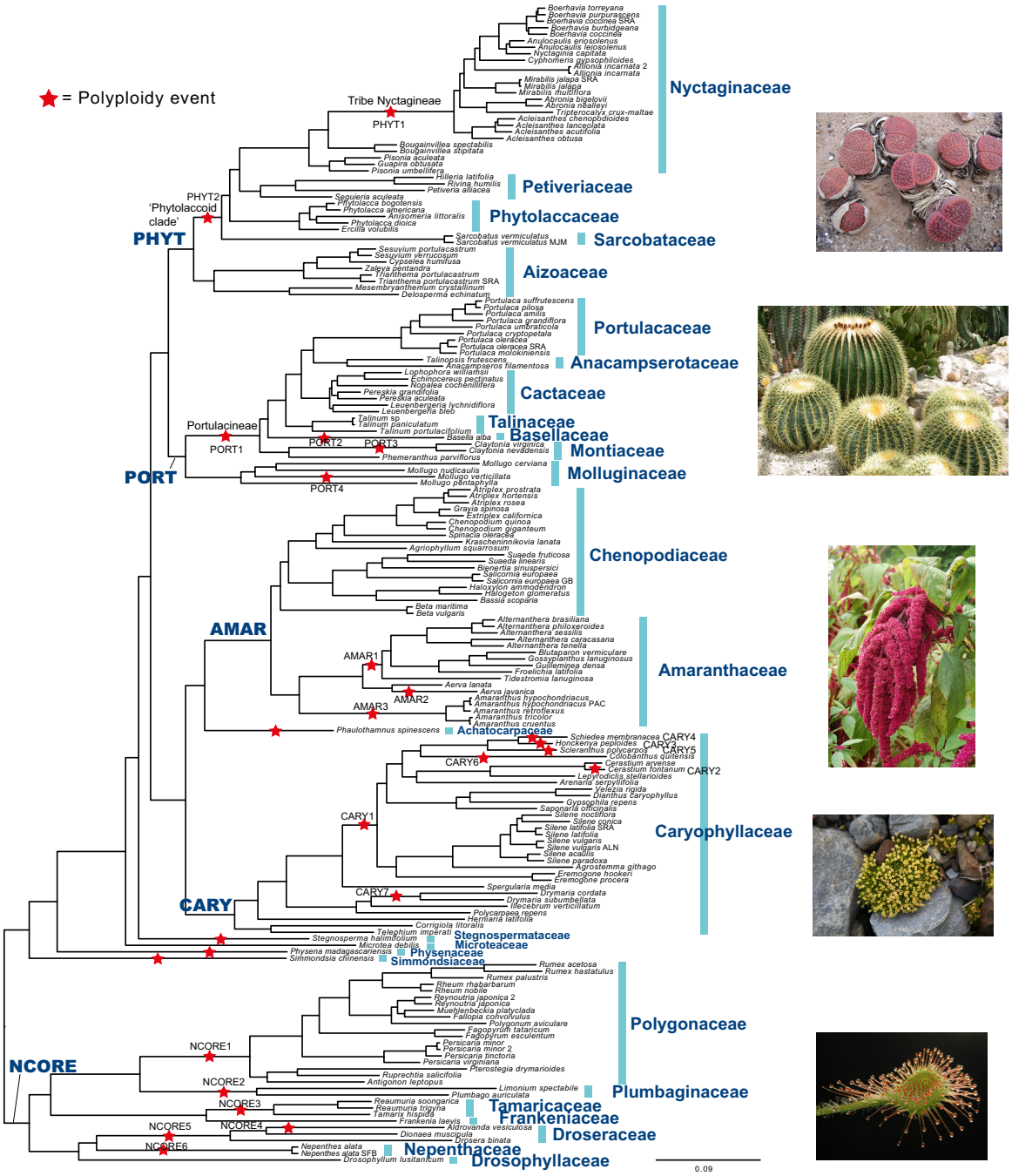


Photosynthetic modifications

Morphological modifications

Genome duplications in Caryophyllales

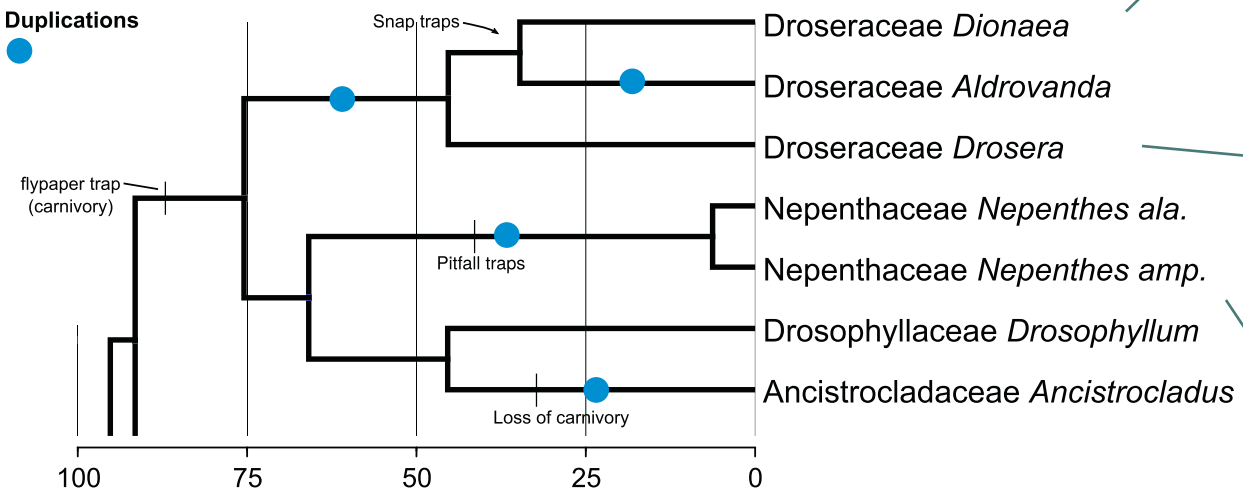
- 168 species
- 736 gene regions used for inference (from thousands identified)
- 26 duplications
- Yang et al. 2017



Duplications in carnivores

Walker et al. 2017

4 duplications associated with the evolution of carnivory

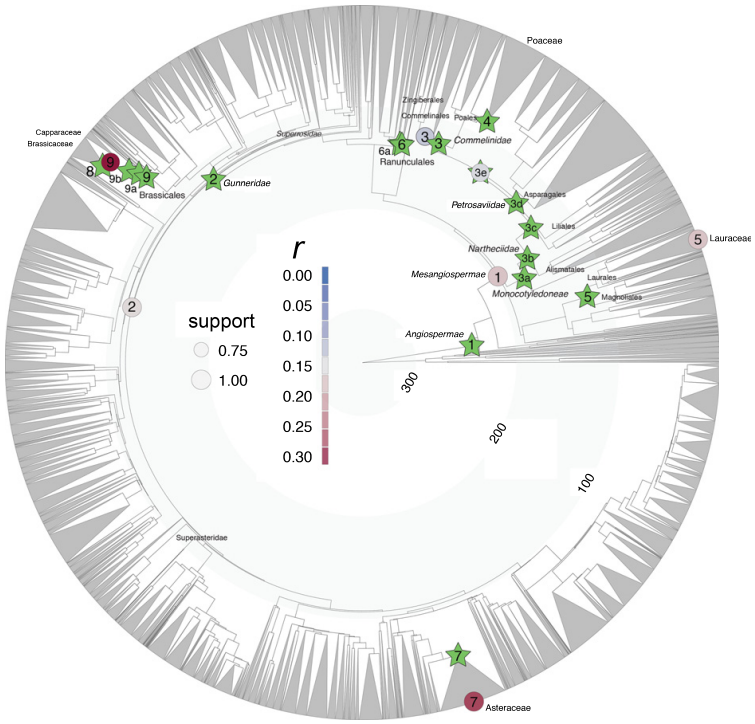


Duplications and diversification

- It has been suggested that **diversification** is associated with **genome duplications**

Are there diversification shifts associated with genome duplication?

Are there climatic shifts associated with genome duplication?



Caryophyllales: annual mean temperature

Nyctaginaceae,
Phytolaccaceae, etc.

Caryophyllaceae

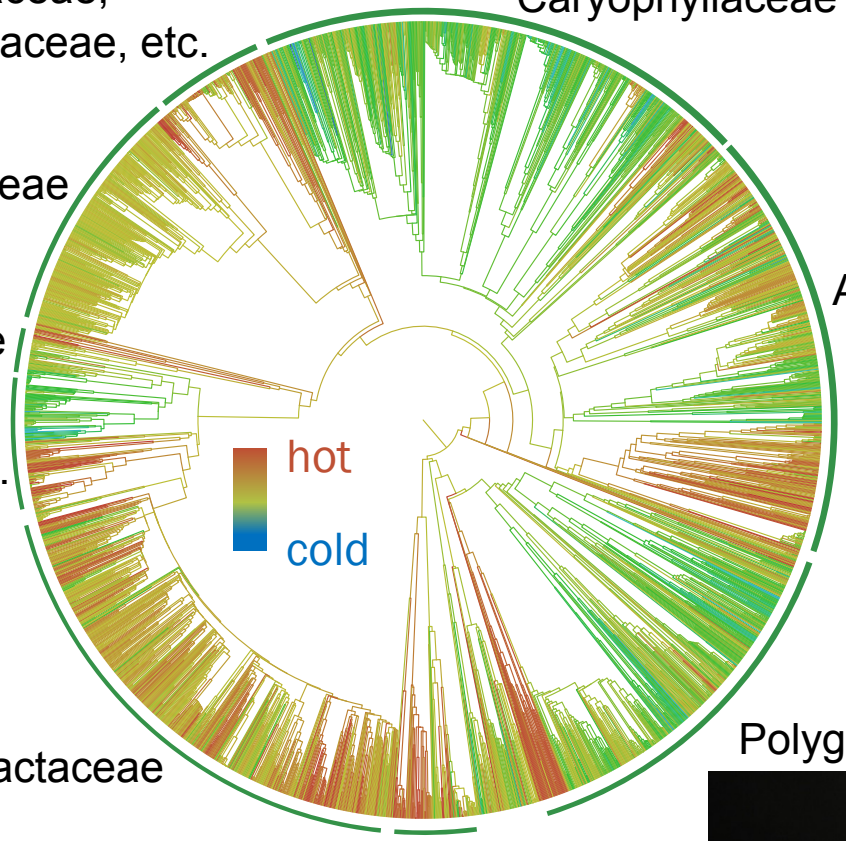
Aizoaceae

5036 taxa

Amaranthaceae

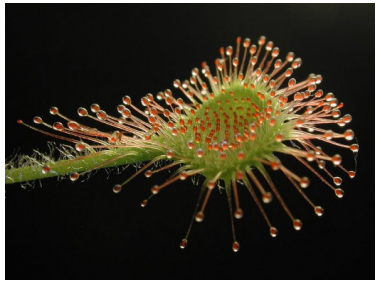
Molluginaceae

Portulacaceae,
Montiaceae, etc.

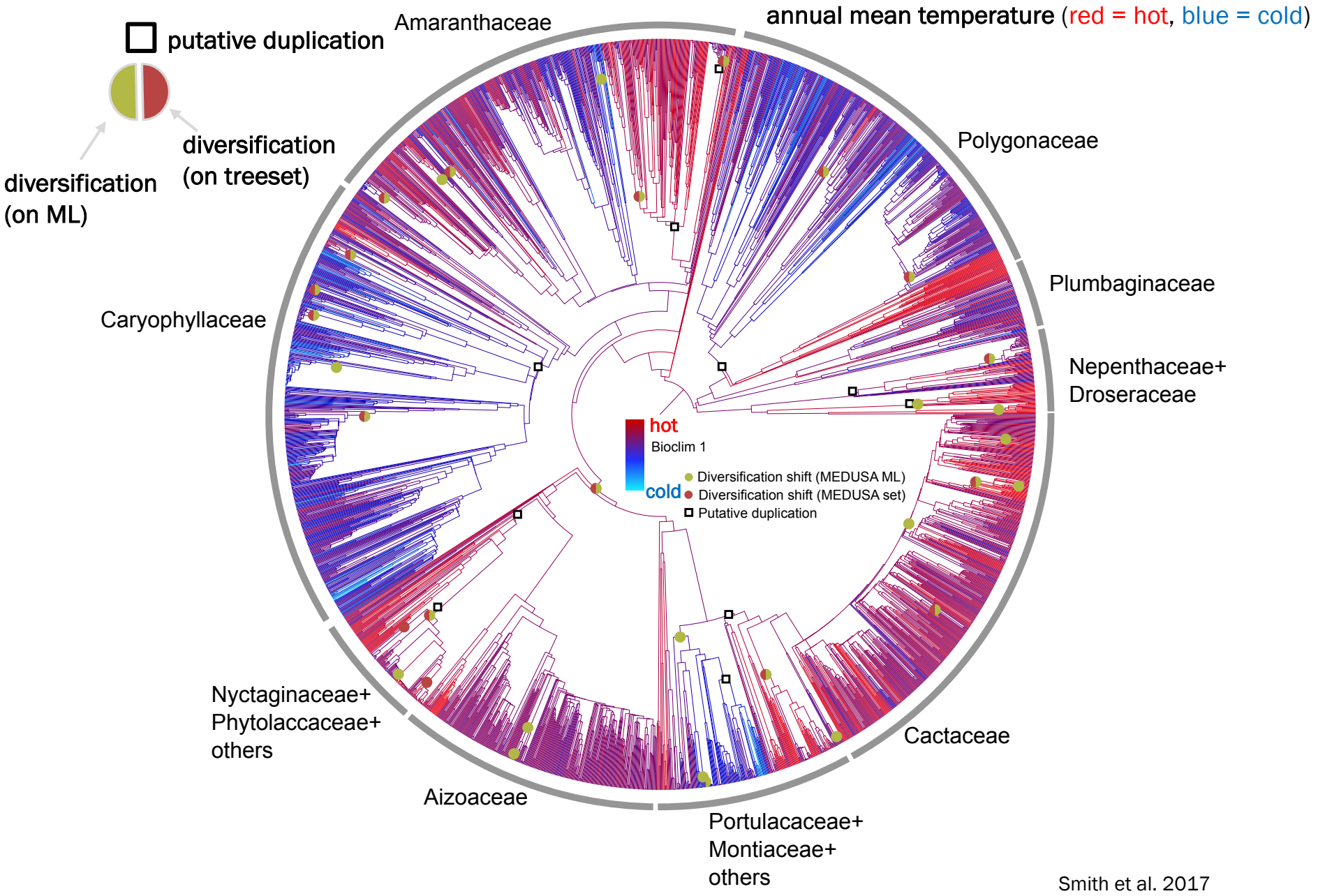


Cactaceae

Polygonaceae

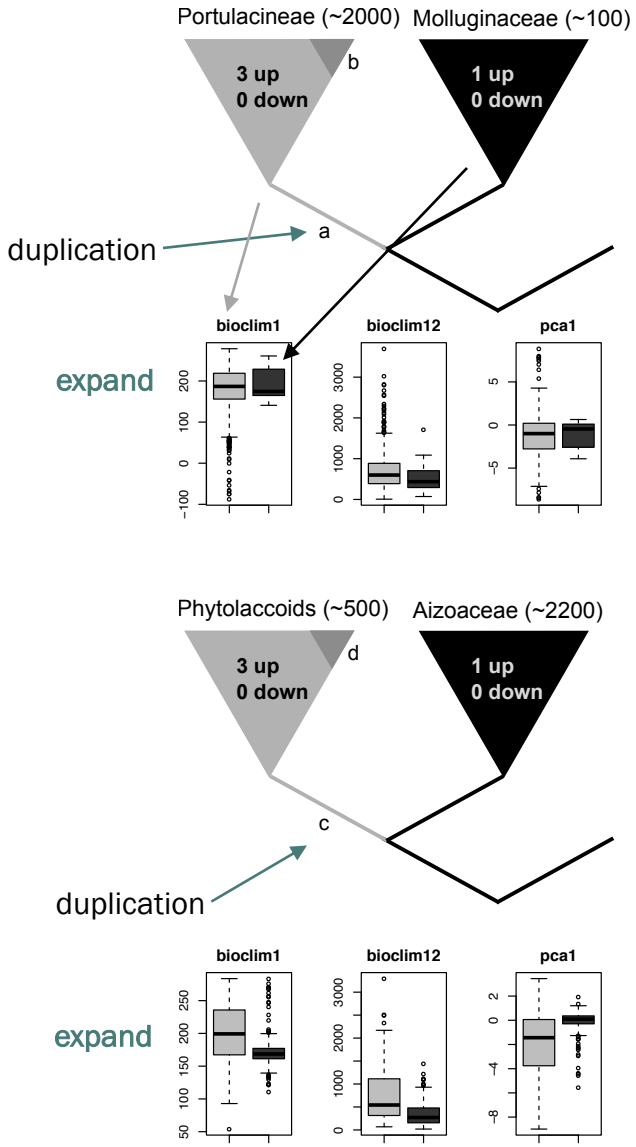


Carnivorous


putative duplication
diversification (on ML)
diversification (on treeset)
annual mean temperature (red = hot, blue = cold)
Amaranthaceae
Polygonaceae
Plumbaginaceae
Nepenthaceae+ Droseraceae
Caryophyllaceae
Nyctaginaceae+ Phytolaccaceae+ others
Aizoaceae
Portulacaceae+ Montiaceae+ others
Cactaceae
hot
Bioclim 1
cold
Diversification shift (MEDUSA ML)
Diversification shift (MEDUSA set)
Putative duplication
Smith et al. 2017

Summary of biological results

- Duplications **are not** associated with faster speciation
- Duplications occur **before** increases in speciation
- ***Many duplications associated with expanded climate ranges***

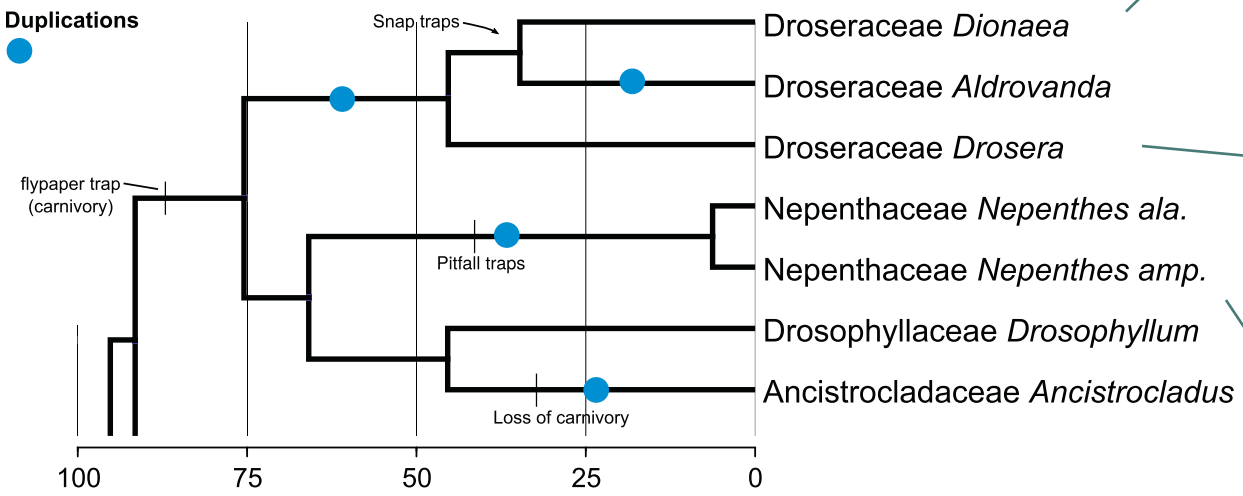


There are cracks on the horizon...

Duplications in carnivores

Walker et al. 2017

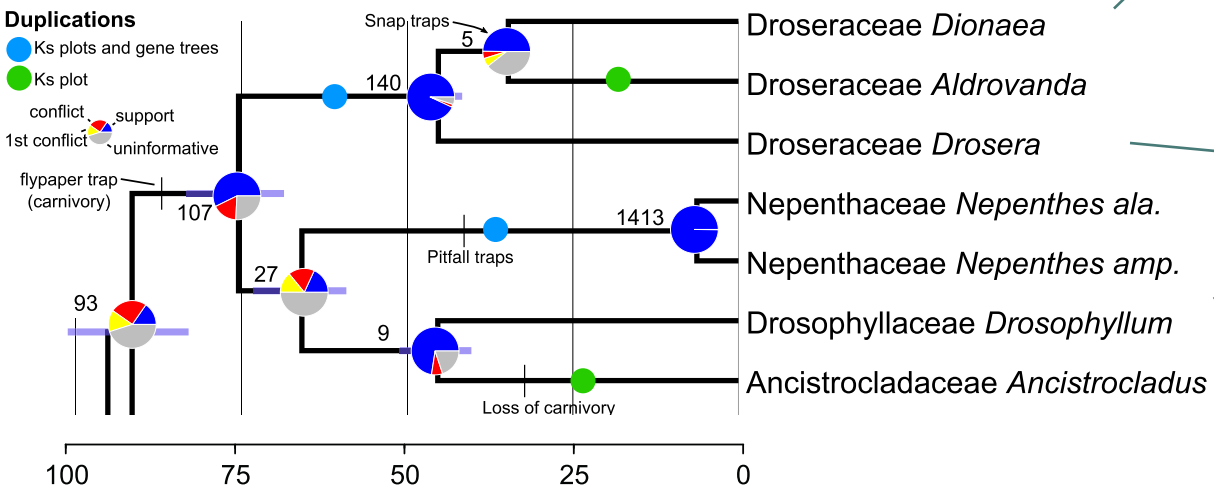
7 duplications associated with the evolution of carnivory

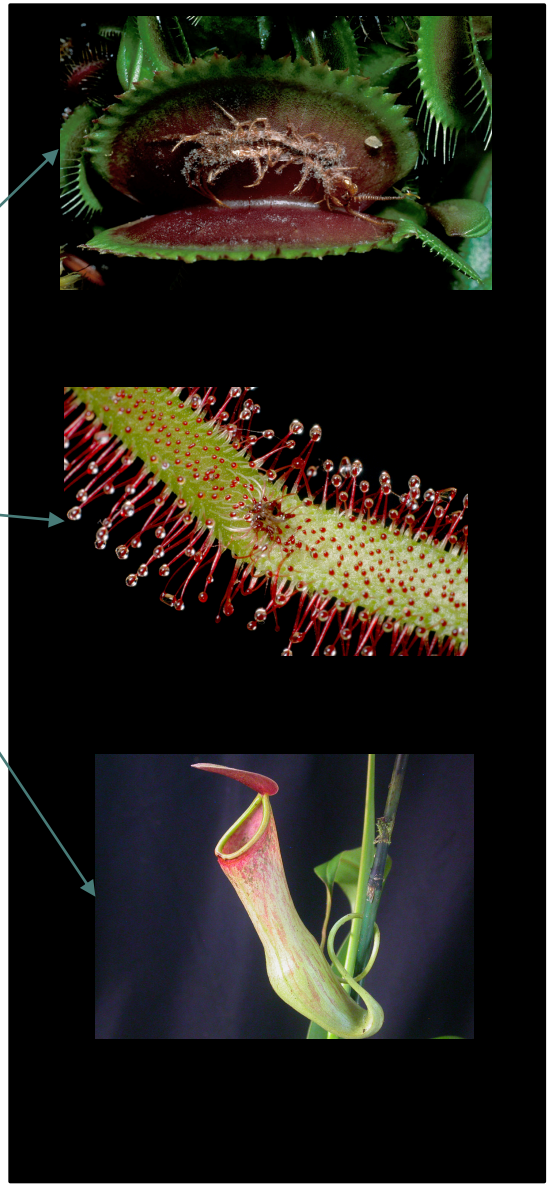


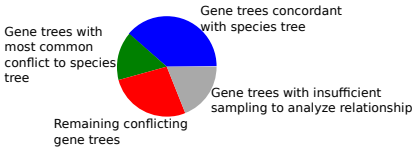
Duplications in carnivores

Walker et al. 2017

7 duplications associated with the evolution of carnivory







Most gene trees conflict with this relationship

Nyctaginaceae

Boerhavia torreyana
Boerhavia purpurascens
Boerhavia coccinea
Boerhavia burbidgeana
Okenia hypogea
Boerhavia ciliata
Anulocaulis annulatus
Anulocaulis leiosolenus
Anulocaulis eriosolenus
Nyctaginia capitata
Cyphomeris gypsophiloides
Allionia incarnata
Commicarpus scandens
Mirabilis jalapa
Mirabilis pringlei
Mirabilis multiflora
Abronia glabrifolia
Abronia bigelovii
Abronia carletonii
Abronia fragrans
Abronia latifolia
Abronia maritima
Abronia umbellata
Tripterocalyx carneus
Tripterocalyx crux-maltae
Acleisanthes lanceolata
Acleisanthes purpusiana
Acleisanthes chenopodioides
Acleisanthes obtusa
Bougainvillea stipitata
Bougainvillea spectabilis
Guapira obtusata
Neea psychotrioides
Pisonia aculeata
Pisonia umbellifera
Colignonia ovalifolia
Salpianthus purpurascens
Petiveriaceae *Monococcus echinophorus*
Petiveriaceae *Petiveria alliacea*
Petiveriaceae *Hilleria latifolia*
Petiveriaceae *Trichostigma octandrum*

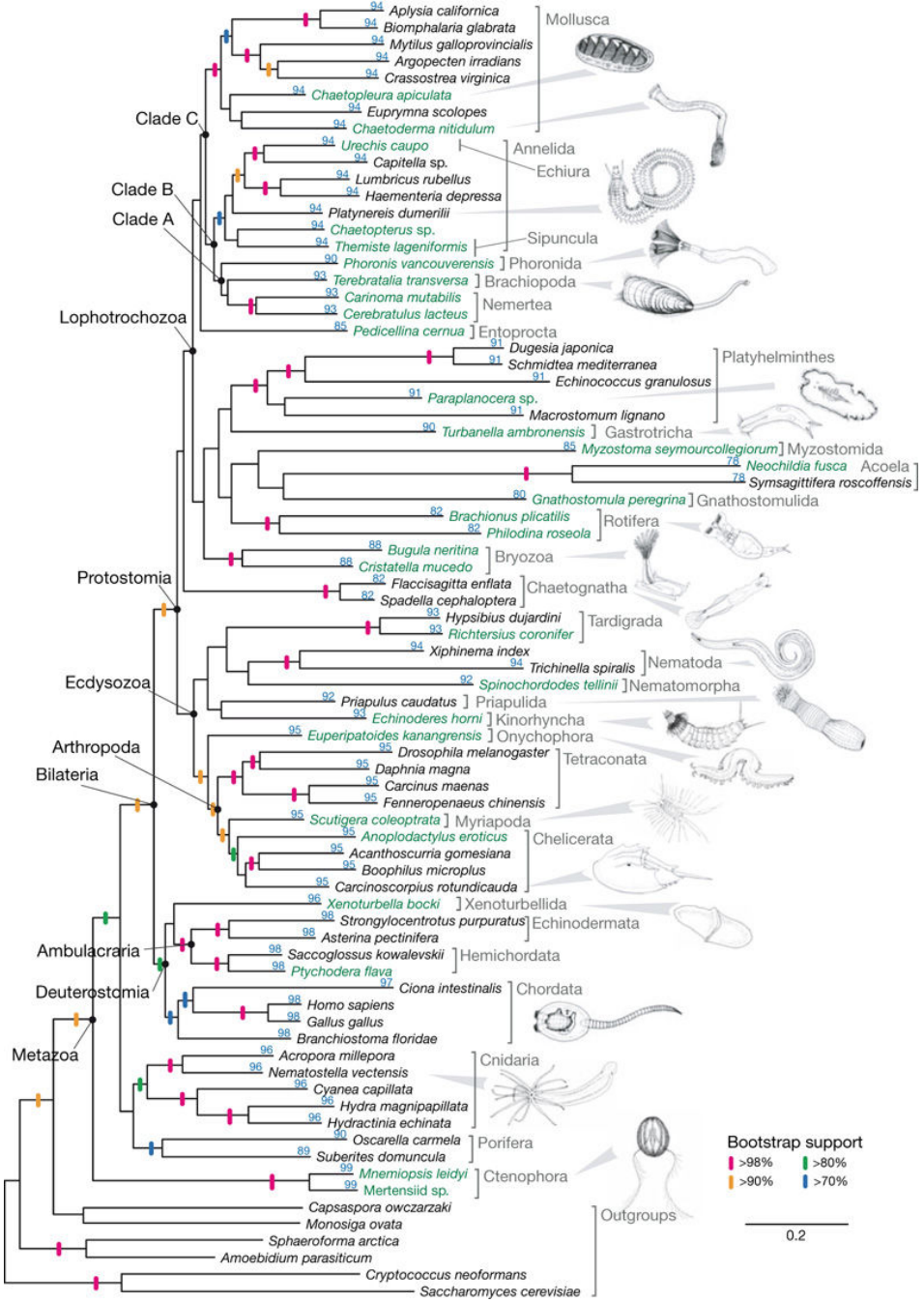
0.05

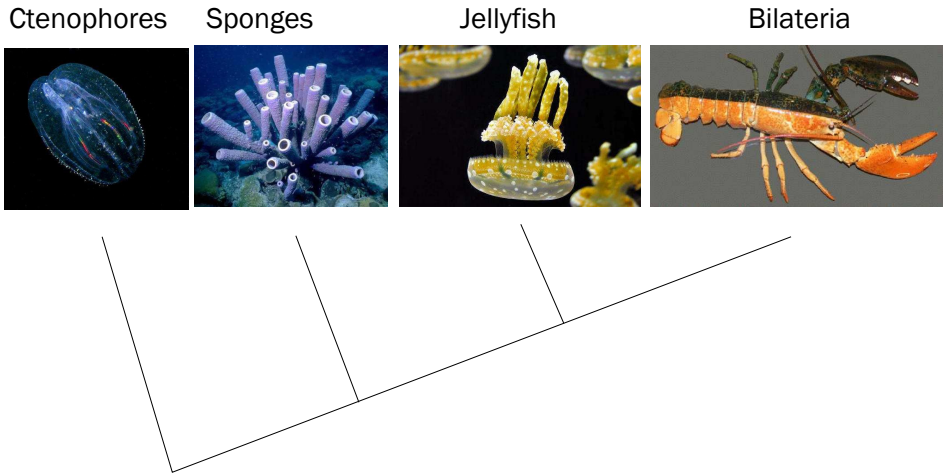

Nyctaginia capitata


Mirabilis multiflora


Abronia umbellata


Pisonia umbellifera





Dunn et al. 2008

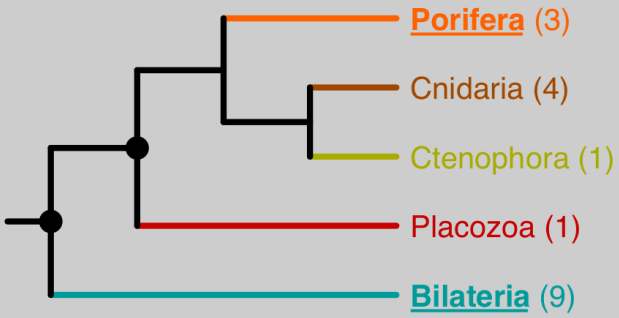
Alternative views

Just showed you this one

Schierwater et al. (2009)

15 mitochondrial and 34 nuclear genes

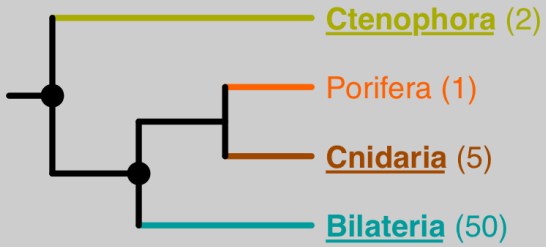
GTR and WAG models



Dunn et al. (2008)

6 mitochondrial and 144 nuclear genes

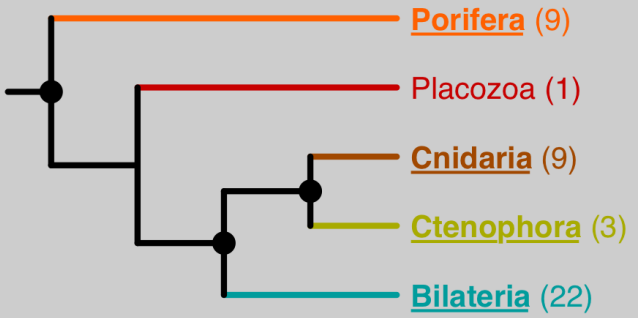
WAG and CAT models



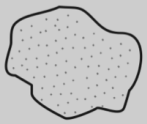
Philippe et al. (2009)

128 nuclear genes

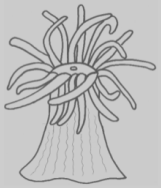
CAT model



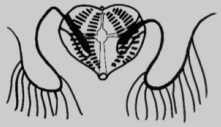
Porifera



Placozoa



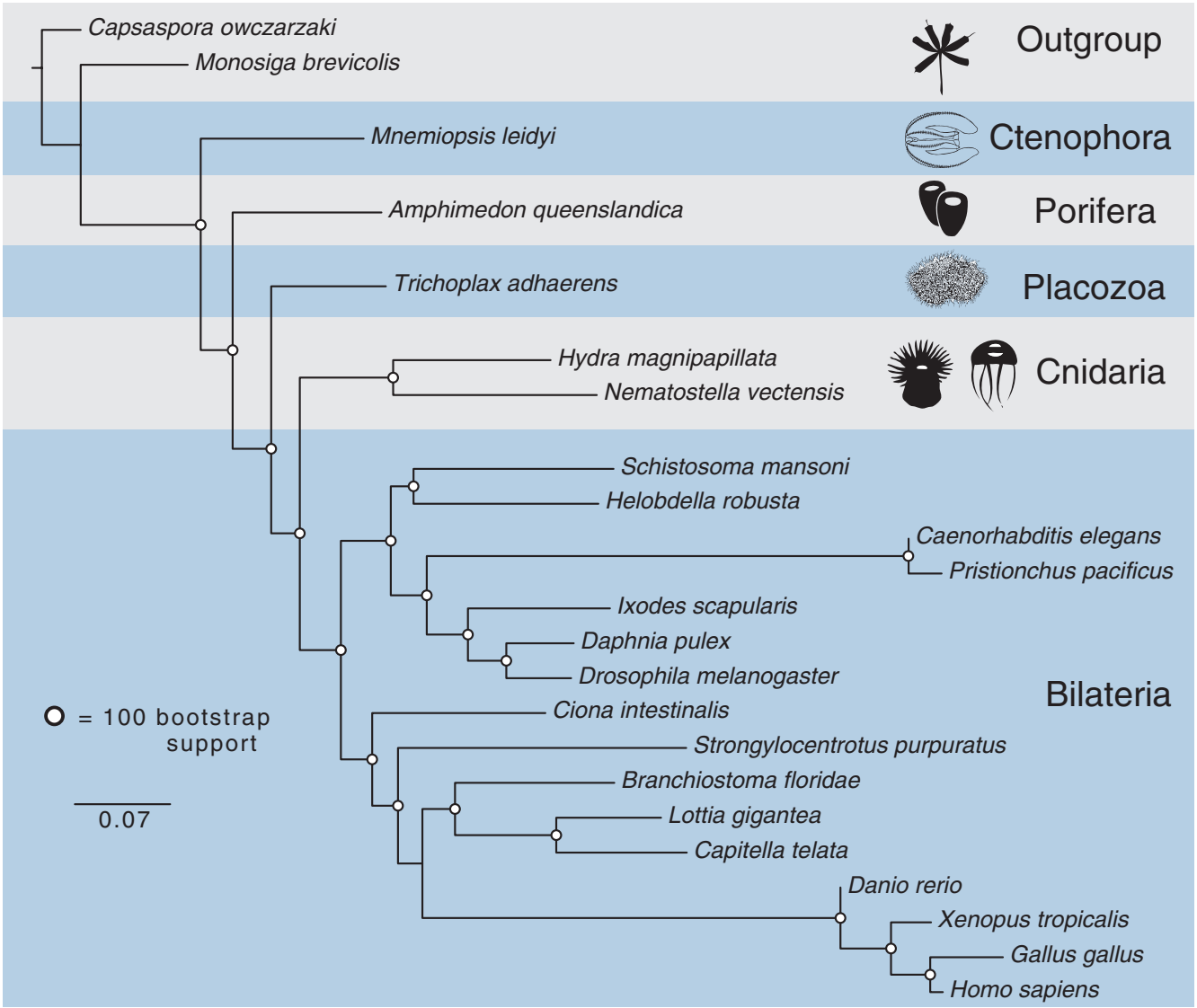
Cnidaria



Ctenophora

Bilateria

Metazoan phylogeny using 242 genes



Ryan et al. 2013

← Sequenced the whole genome

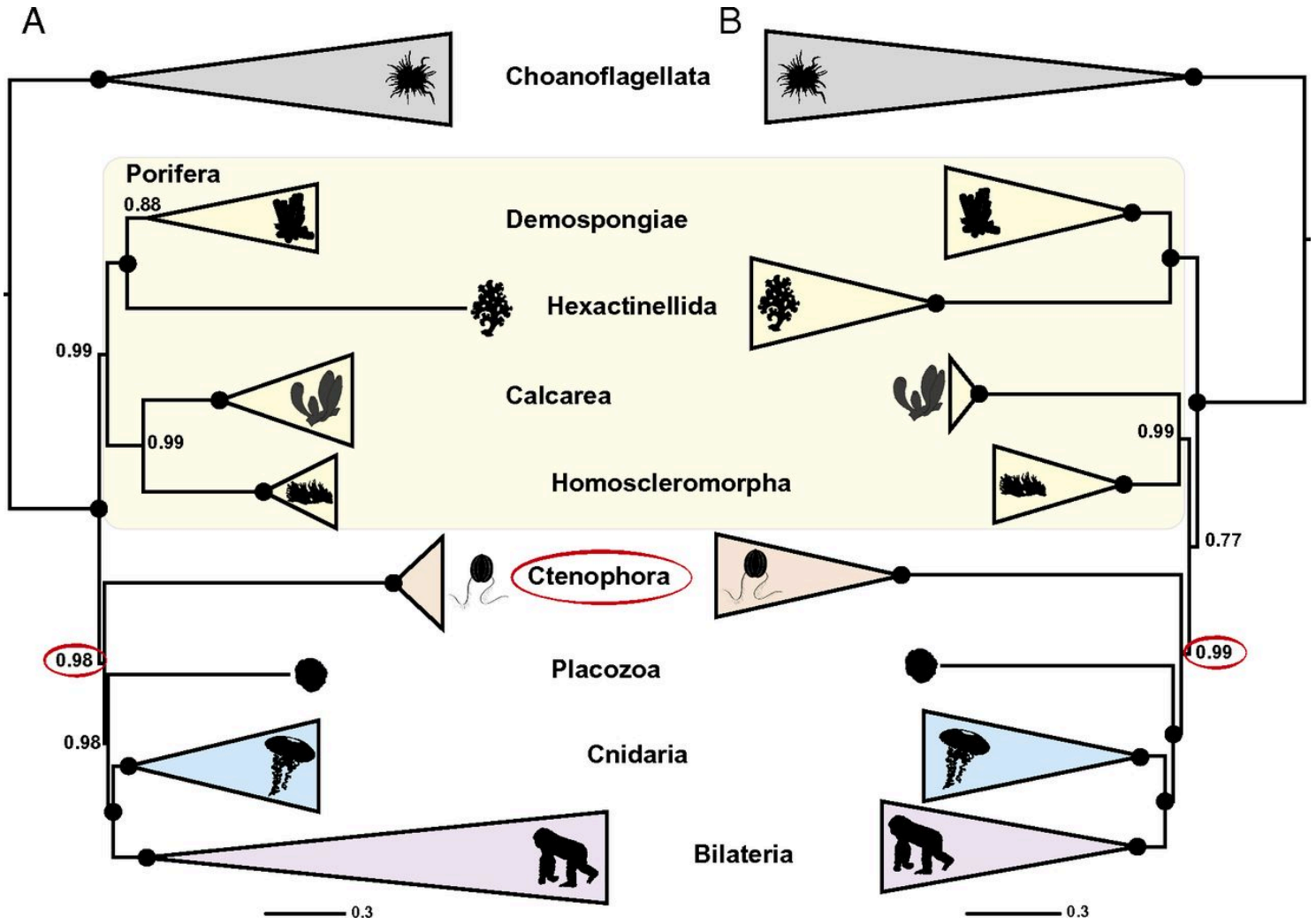
Is data the problem?

Genomic data do not support comb jellies as the sister group to all other animals

Davide Pisani, Walker Pett, Martin Dohrmann, Roberto Feuda, Omar Rota-Stabelli, Hervé Philippe, Nicolas Lartillot, and Gert Wörheide

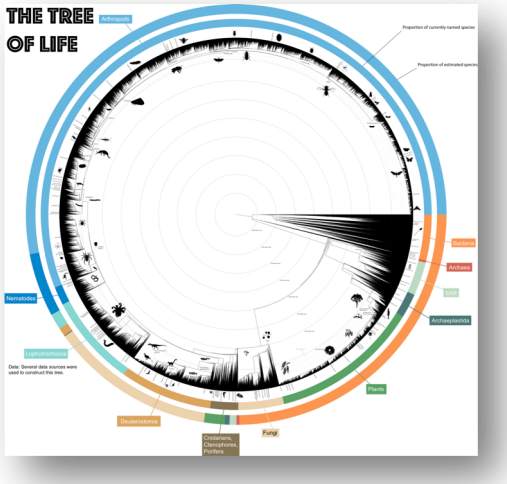
PNAS 2015 December, 112 (50) 15402-15407. https://doi.org/10.1073/pnas.1518127112

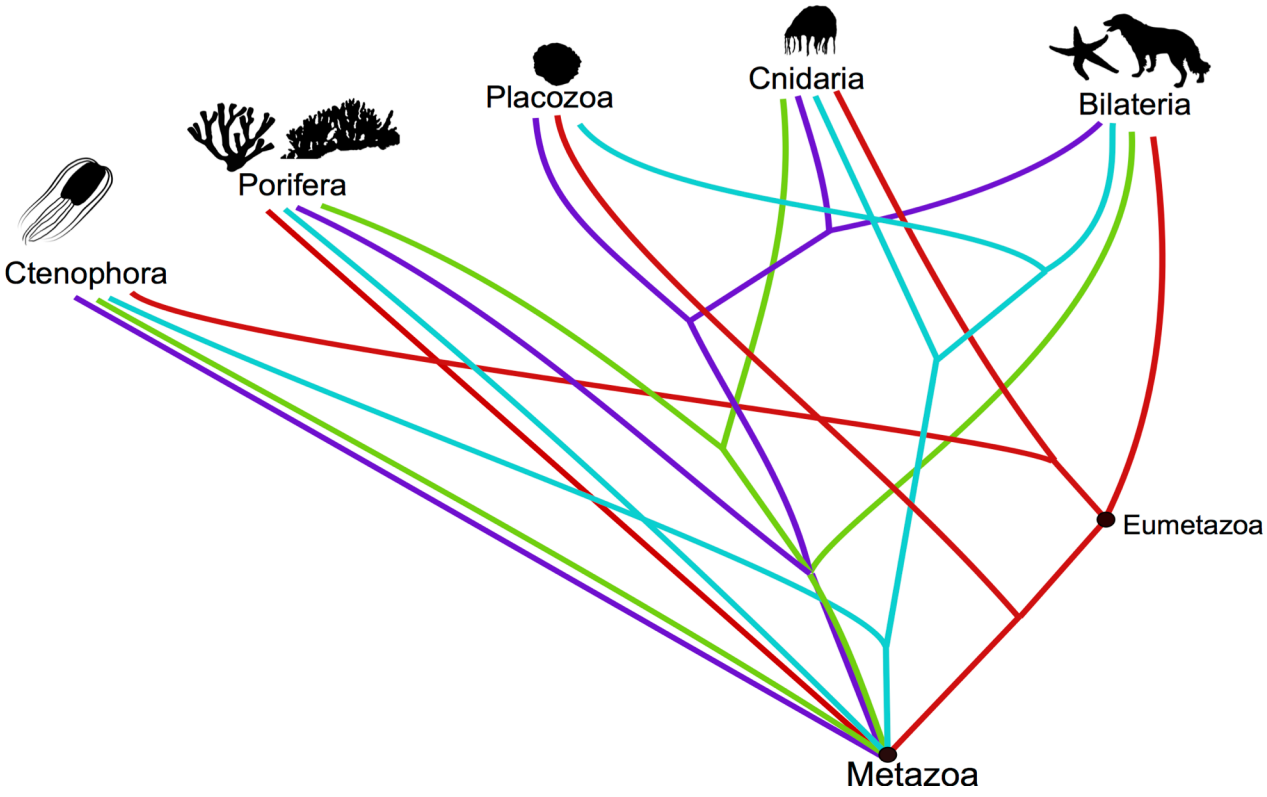
Are molecular models the problem?



Building a more complete view of the Tree of Life

Open Tree of Life tried to accommodate this but is there a way to resolve any of this *confidently?*





Dunn et al., 2008

Pick et al. 2010

Ryan et al. 2013

Phillipe et al. 2009

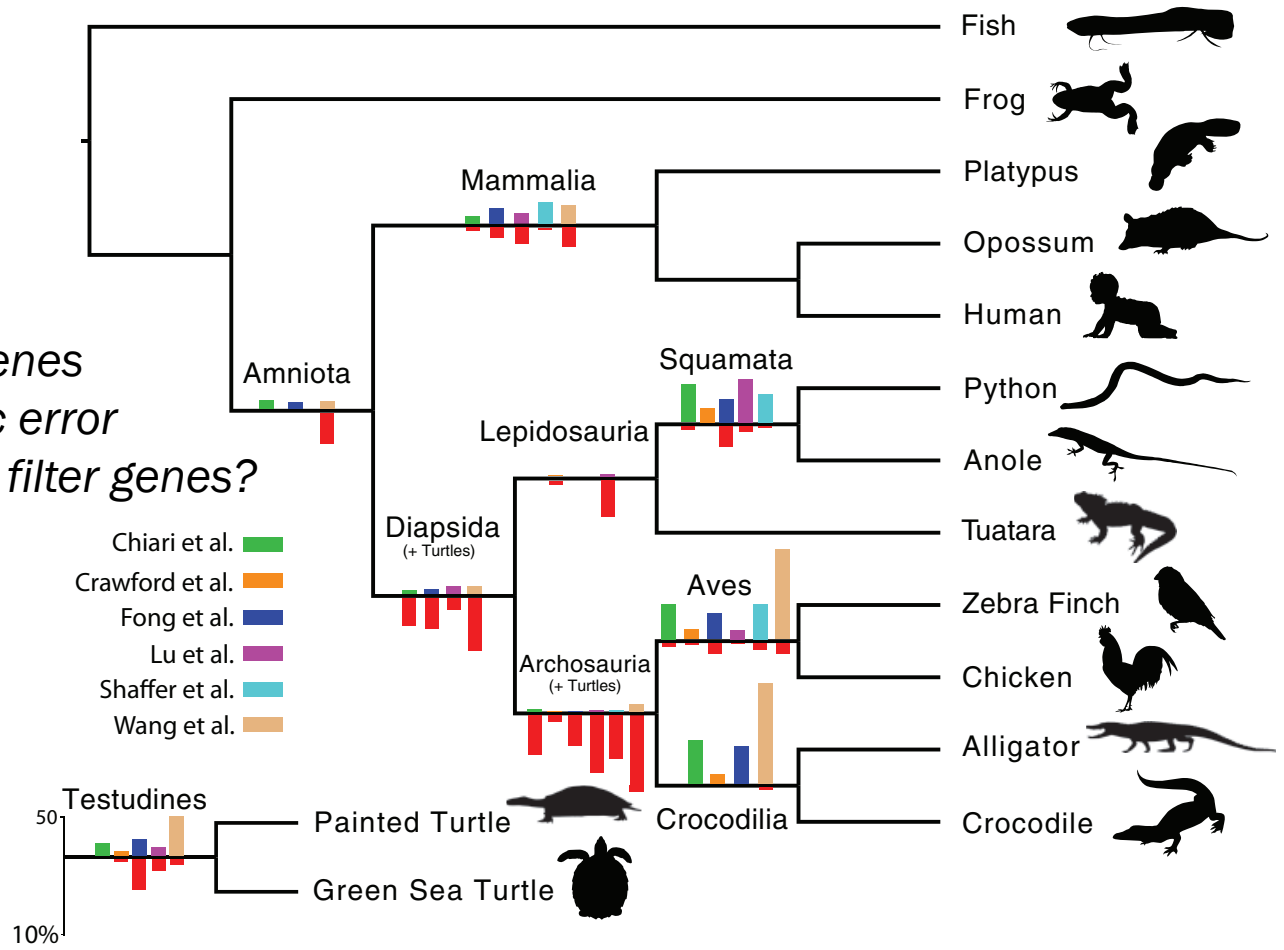
Syst. Biol. 66(4):517–530, 2017

DOI:10.1093/sysbio/syw101
Advance Access publication November 16, 2016

Bayes Factors Unmask Highly Variable Information Content, Bias, and Extreme Influence in Phylogenomic Analyses

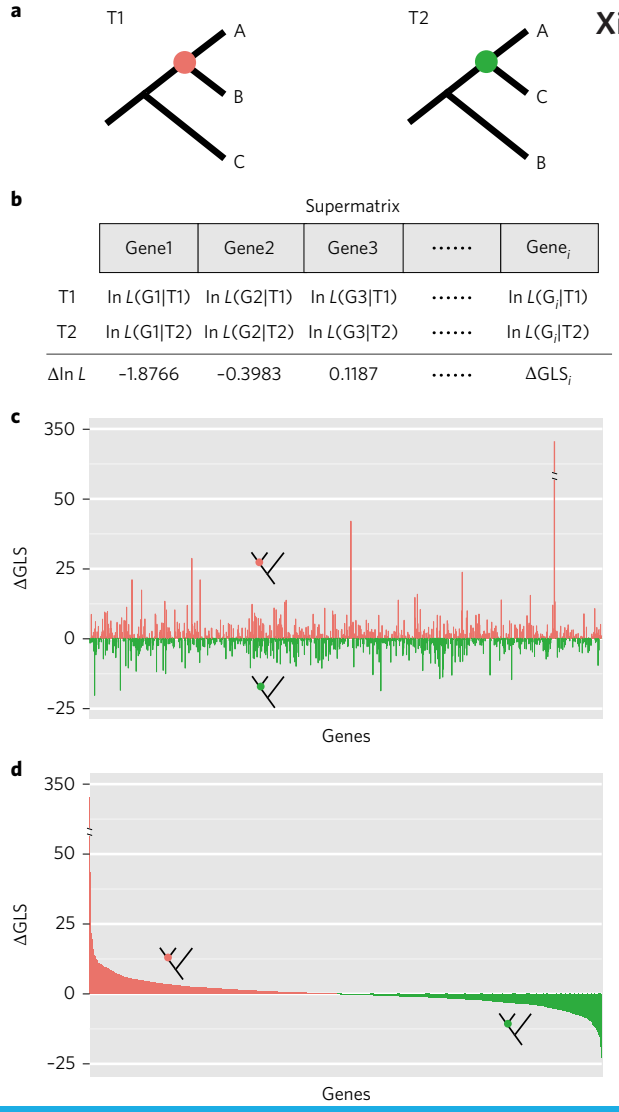
JEREMY M. BROWN[1,*] AND ROBERT C. THOMSON[2]

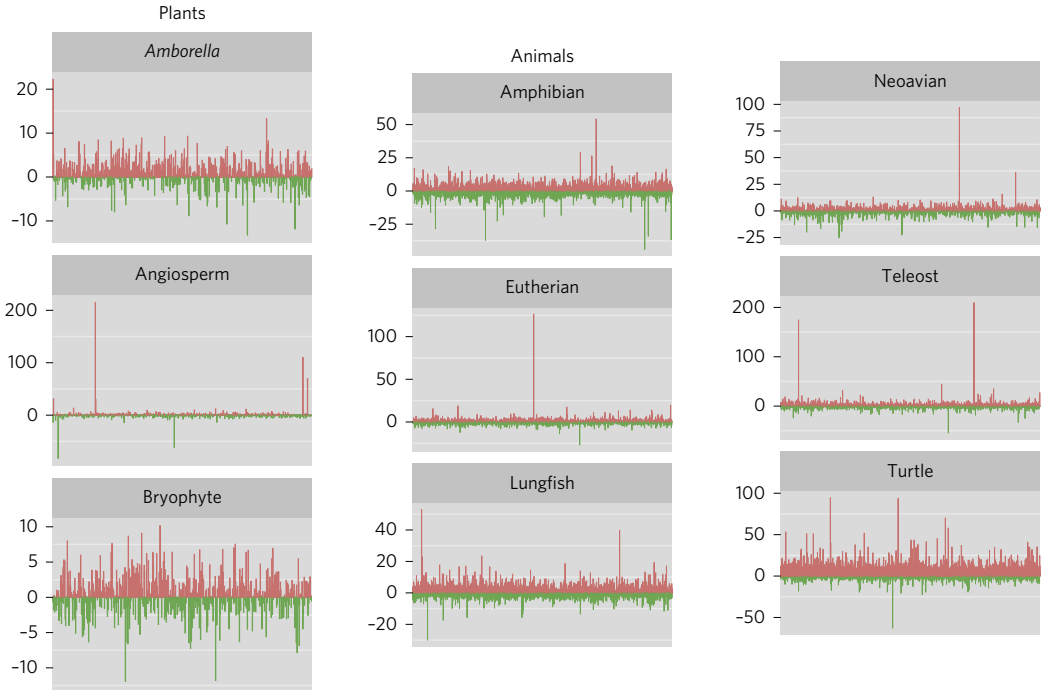
- *Outlying genes*
- *Systematic error*
- *Should we filter genes?*



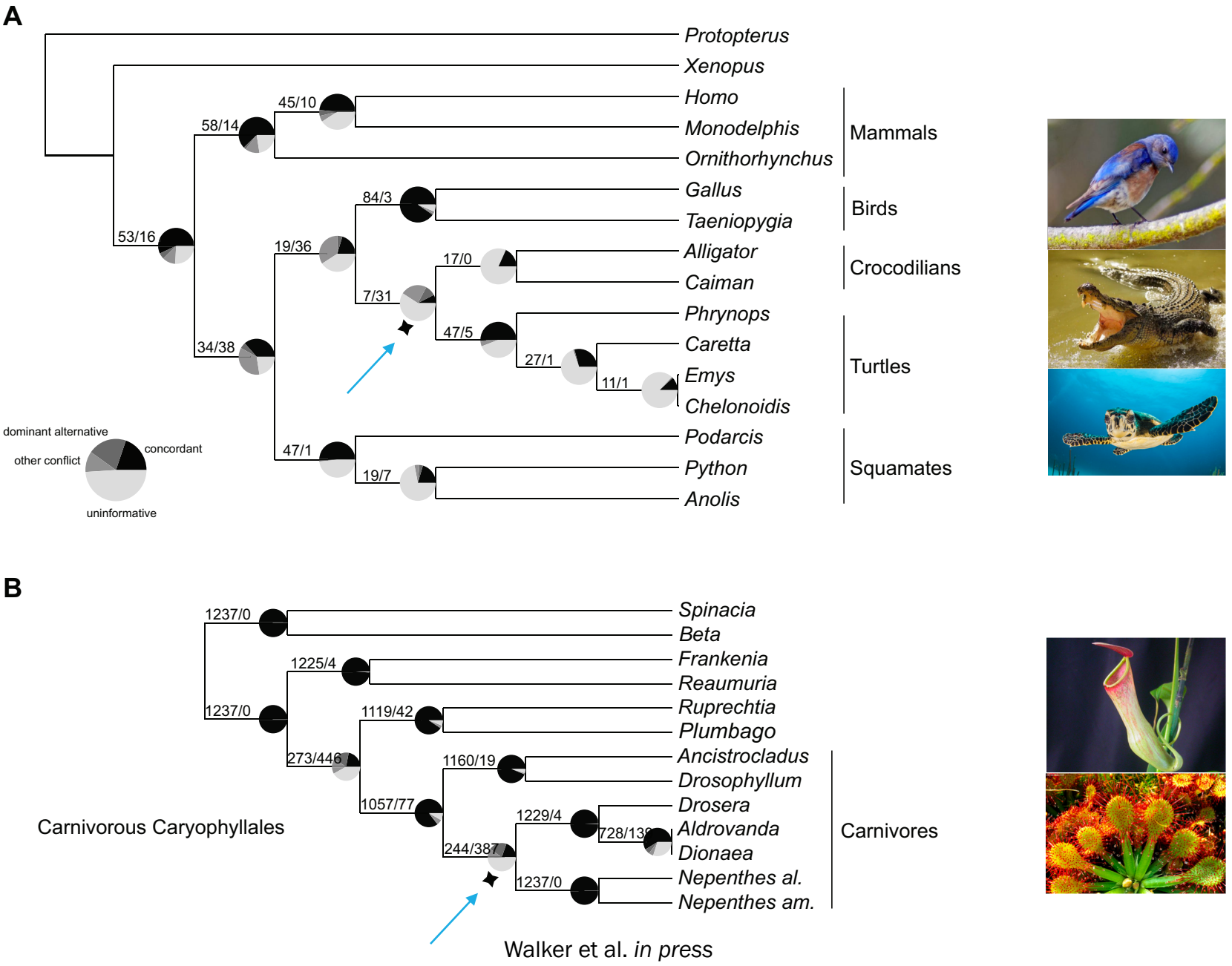
Contentious relationships in phylogenomic studies can be driven by a handful of genes

Xing-Xing Shen[1], Chris Todd Hittinger[2] and Antonis Rokas[1]*





- *Outlying genes*
- *Should we do two (or few) topology tests?*

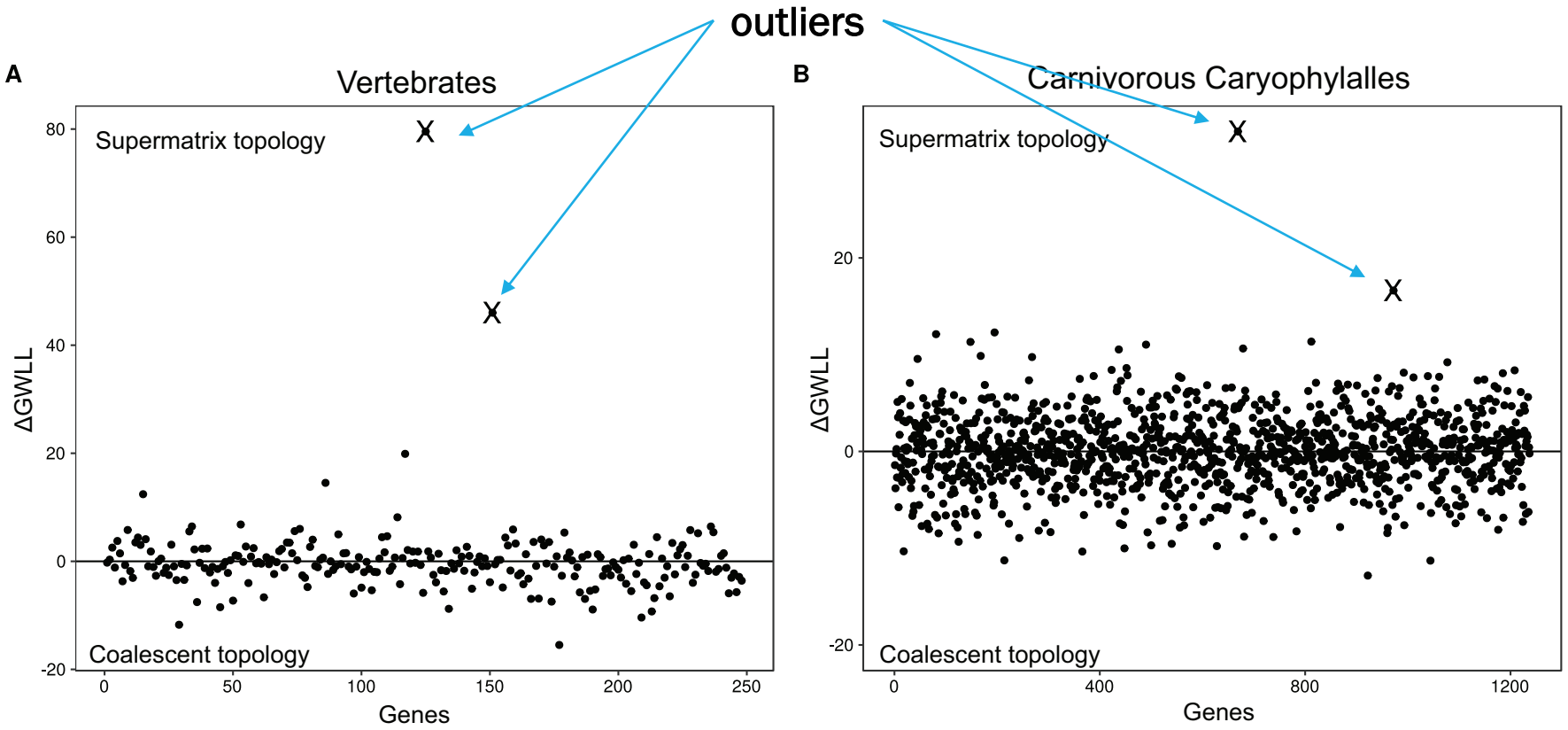

A
Protopterus
Xenopus
Homo
Monodelphis
Ornithorhynchus
Mammals
Gallus
Taeniopygia
Birds
Alligator
Caiman
Crocodilians
Phrynops
Caretta
Emys
Chelonoidis
Turtles
Podarcis
Python
Anolis
Squamates
45/10
58/14
53/16
84/3
19/36
17/0
7/31
47/5
27/1
11/1
34/38
47/1
19/7
dominant alternative
other conflict
concordant
uninformative
B
Spinacia
Beta
Frankenia
Reaumuria
Ruprechtia
Plumbago
Ancistrocladus
Drosophyllum
Drosera
Aldrovanda
Dionaea
Nepenthes al.
Nepenthes am.
Carnivores
1237/0
1225/4
1237/0
1119/42
273/446
1160/19
1057/77
1229/4
728/13
244/387
1237/0
Carnivorous Caryophyllales
Walker et al. in press

Coalescent topologies were favored for biological reasons

Supermatrix topologies were assumed to reflect error

Two topology comparisons exposed “outlier” genes

When removed, supermatrix topologies match coalescent topologies



Walker et al. *in press*

Important lessons

Just concatenation and gene tree / species tree (ASTRAL) are probably not going to cut it

Systematic error

- Researchers need to be significantly more careful about the underlying data
- Check for errors
  - in homology/orthology
  - Alignment
  - Heterogeneity in molecular evolution

Biological sources of error

- Gene duplication and loss
- ILS
- Hybridization

Potential “outlier” genes

- One or a few genes can drive phylogenetic inference
- Especially problematic with concatenation

Can we take a different approach to phylogenomic data analysis?

Edge-based look at plants

Wickett et al. 2014 initial 1KP paper

Conducted ASTRAL and supermatrix analyses on many different datasets

ASTRAL and supermatrix trees are largely congruent but disagree

This didn't settle all the arguments

Phylotranscriptomic analysis of the origin and early diversification of land plants

Norman J. Wickett[a,b,1,2], Siavash Mirarab[c,1], Nam Nguyen[c], Tandy Warnow[c], Eric Carpenter[d], Naim Matasci[e,f], Saravanaraj Ayyampalayam[g], Michael S. Barker[f], J. Gordon Burleigh[h], Matthew A. Gitzendanner[h,i], Brad R. Ruhfel[h,j,k], Eric Wafula[l], Joshua P. Der[l], Sean W. Graham[m], Sarah Mathews[n], Michael Melkonian[o], Douglas E. Soltis[h,i,k], Pamela S. Soltis[h,i,k], Nicholas W. Miles[k], Carl J. Rothfels[p,q], Lisa Pokorny[p,r], A. Jonathan Shaw[p], Lisa DeGironimo[s], Dennis W. Stevenson[s], Barbara Surek[o], Juan Carlos Villarreal[t], Béatrice Roure[u], Hervé Philippe[u,v], Claude W. dePamphilis[l], Tao Chen[w], Michael K. Deyholos[d], Regina S. Baucom[x], Toni M. Kutchan[y], Megan M. Augustin[y], Jun Wang[z], Yong Zhang[v], Zhijian Tian[z], Zhixiang Yan[z], Xiaolei Wu[z], Xiao Sun[z], Gane Ka-Shu Wong[d,z,aa,2], and James Leebens-Mack[g,2]

Sister to land plants

Zygn-sister
(Timme et al 2012)

Char-sister
(Karol et al 2001)

Col-sister
(Finet et al 2010)

ANA-grade angiosperms

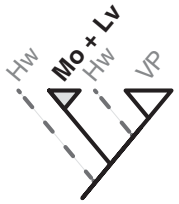
Amborella + *Nuphar*
(Qiu et al 2010)

Amb-sister
(Soltis et al 2011)

Bryophytes

Lv-basal
(Qiu et al 1998)

Hw-sister
(Chang & Graham 2011)



Moss + Lv
(Renzaglia et al 2000)

Bryo monophy.
(Nishiyama et al 2004)

Hw-basal
(Nickrent et al 2000)

Gymnosperms

Gnepine
(Bowe et al 2000)

Conifers monophy.
(Lee et al 2011)

Gnetifer
(Chaw et al 1997)

Gnetales-sister
(Lee et al 2011)

Angiosperms

Eudi + Mag
(Lee et al 2011)

Eudi + Mag/Chl
(Burleigh et al 2009)

Mono + Eudi
(Soltis et al 2011)

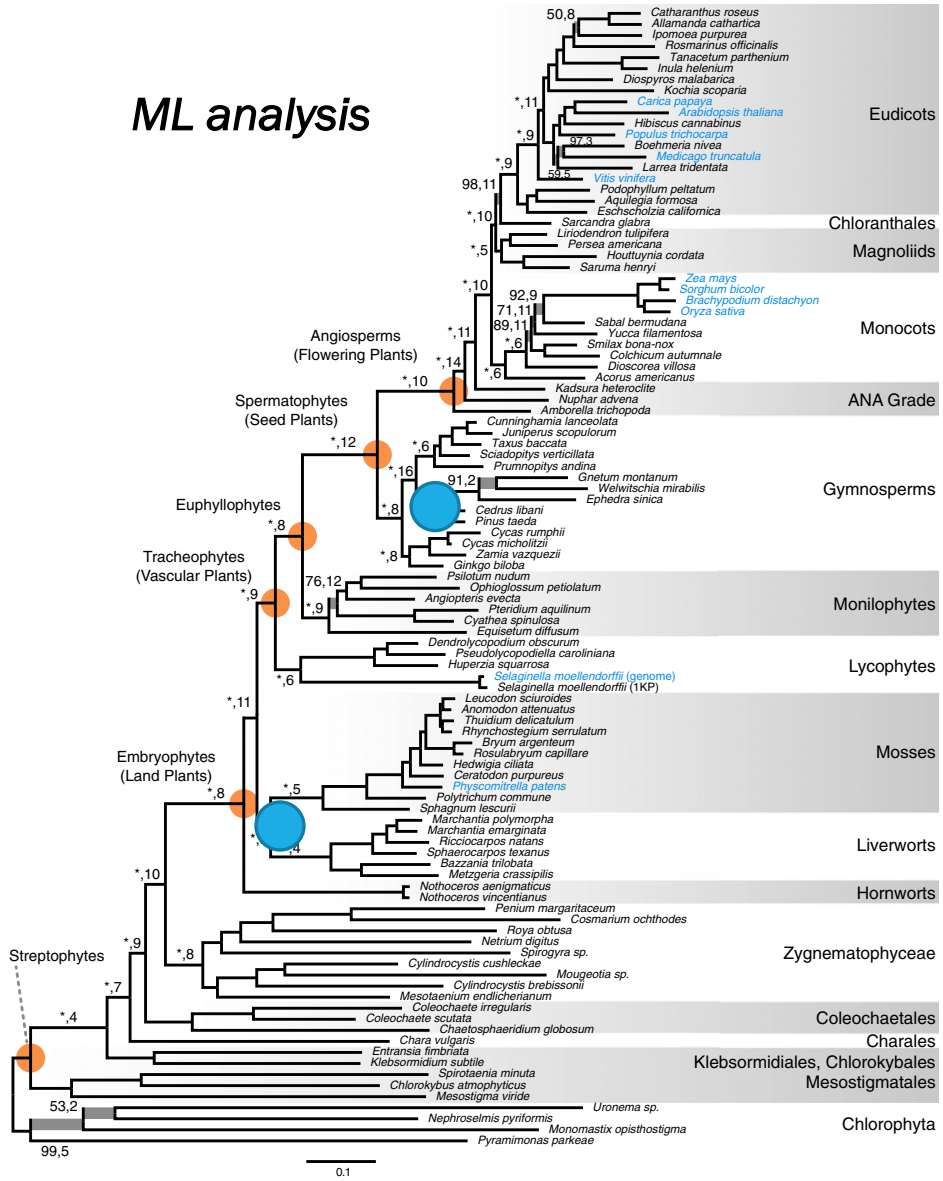
Mag + Chl, Mono
(Soltis et al 2000)

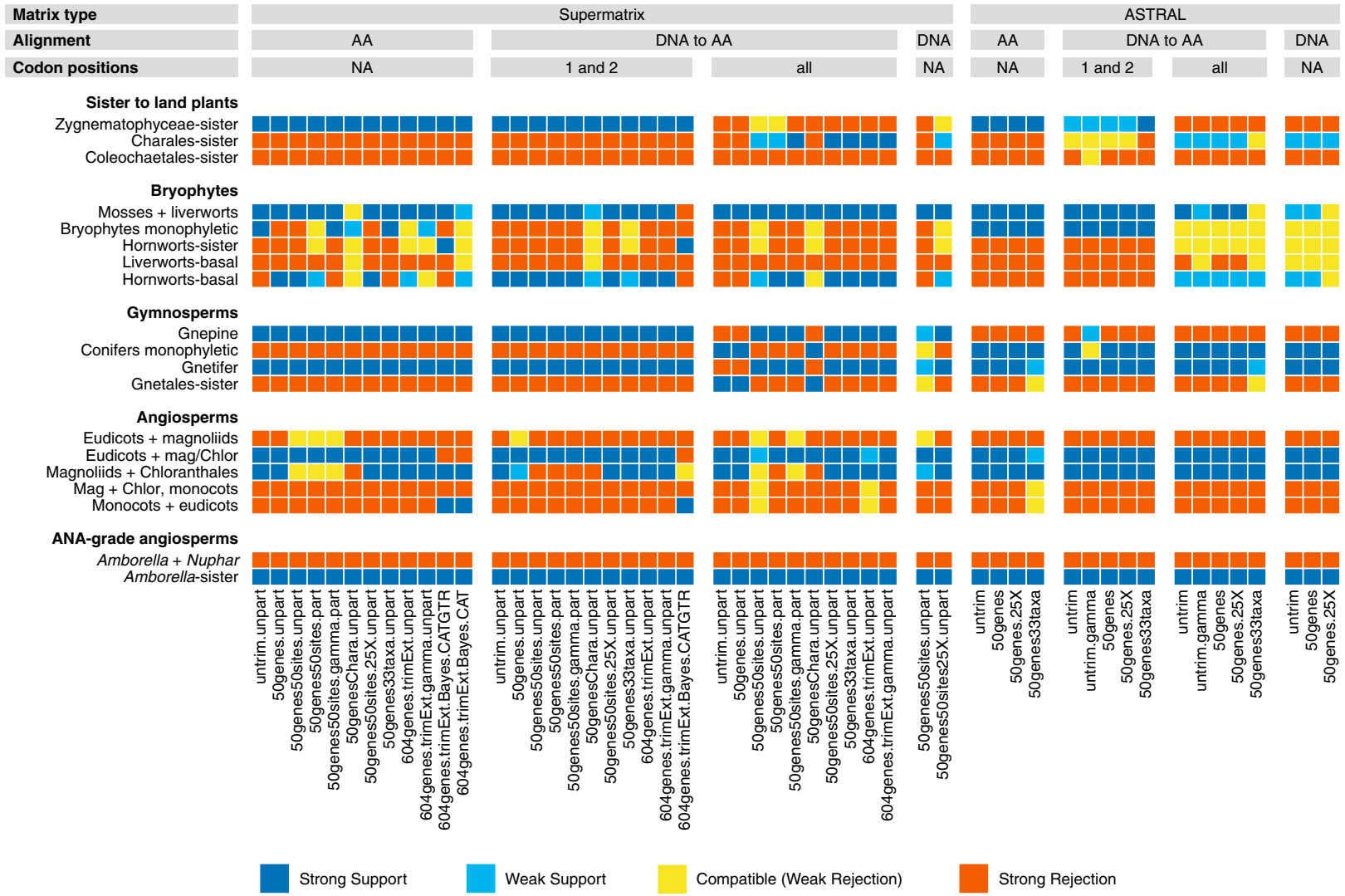
Mag + Chl
(Soltis et al 2011)

ANA-grade angiosperms

Amborella + *Nuphar*
(Qiu et al 2010)

Amb-sister
(Soltis et al 2011)


ML analysis
Eudicots
Chloranthales
Magnoliids
Monocots
ANA Grade
Gymnosperms
Monilophytes
Lycophytes
Mosses
Liverworts
Hornworts
Zygnematophyceae
Coleochaetales
Charales
Klebsormidiales, Chlorokybales
Mesostigmatales
Chlorophyta
Angiosperms (Flowering Plants)
Spermatophytes (Seed Plants)
Euphyllophytes
Tracheophytes (Vascular Plants)
Embryophytes (Land Plants)
Streptophytes
Catharanthus roseus
Allamanda cathartica
Ipomoea purpurea
Rosmarinus officinalis
Tanacetum parthenium
Inula helenium
Diospyros malabarica
Kochia scoparia
Carica papaya
Arabidopsis thaliana
Hibiscus cannabinus
Populus trichocarpa
Boehmeria nivea
Medicago truncatula
Larrea tridentata
Vitis vinifera
Podophyllum peltatum
Aquilegia formosa
Eschscholzia californica
Sarcandra glabra
Liriodendron tulipifera
Persea americana
Houttuynia cordata
Saruma henryi
Zea mays
Sorghum bicolor
Brachypodium distachyon
Oryza sativa
Sabal bermudana
Yucca filamentosa
Smilax bona-nox
Colchicum autumnale
Dioscorea villosa
Acorus americanus
Kadsura heteroclite
Nuphar advena
Amborella trichopoda
Cunninghamia lanceolata
Juniperus scopulorum
Taxus baccata
Sciadopitys verticillata
Prumnopitys andina
Gnetum montanum
Welwitschia mirabilis
Ephedra sinica
Cedrus libani
Pinus taeda
Cycas rumphii
Cycas micholitzii
Zamia vazquezii
Ginkgo biloba
Psilotum nudum
Ophioglossum petiolatum
Angiopteris evecta
Pteridium aquilinum
Cyathea spinulosa
Equisetum diffusum
Dendrolycopodium obscurum
Pseudolycopodiella caroliniana
Huperzia squarrosa
Selaginella moellendorffii (genome)
Selaginella moellendorffii (1KP)
Leucodon sciuroides
Anomodon attenuatus
Thuidium delicatulum
Rhynchostegium serrulatum
Bryum argenteum
Rosulabryum capillare
Hedwigia ciliata
Ceratodon purpureus
Physcomitrella patens
Polytrichum commune
Sphagnum lescurii
Marchantia polymorpha
Marchantia emarginata
Ricciocarpos natans
Sphaerocarpos texanus
Bazzania trilobata
Metzgeria crassipilis
Nothoceros aenigmaticus
Nothoceros vincentianus
Penium margaritaceum
Cosmarium ochthodes
Roya obtusa
Netrium digitus
Spirogyra sp.
Mougeotia sp.
Cylindrocystis cushleckae
Cylindrocystis brebissonii
Mesotaenium endlicherianum
Coleochaete irregularis
Coleochaete scutata
Chaetosphaeridium globosum
Chara vulgaris
Entransia fimbriata
Klebsormidium subtile
Spirotaenia minuta
Chlorokybus atmophyticus
Mesostigma viride
Uronema sp.
Nephroselmis pyriformis
Monomastix opisthostigma
Pyramimonas parkeae
0.1
ASTRAL
Eudicots
Chloranthales
Magnoliids
Monocots
ANA Grade
Gymnosperms
Monilophytes
Lycophytes
Mosses
Liverworts
Hornworts
Zygnematophyceae
Coleochaetales
Charales
Klebsormidiales, Chlorokybales
Mesostigmatales
Chlorophyta
Angiosperms (Flowering Plants)
Spermatophytes (Seed Plants)
Euphyllophytes
Tracheophytes (Vascular Plants)
Embryophytes (Land Plants)
Streptophytes


Matrix type
Supermatrix
ASTRAL
Alignment
AA
DNA to AA
DNA
AA
DNA to AA
DNA
Codon positions
NA
1 and 2
all
NA
NA
1 and 2
all
NA
Sister to land plants
Zygnematophyceae-sister
Charales-sister
Coleochaetales-sister
Bryophytes
Mosses + liverworts
Bryophytes monophyletic
Hornworts-sister
Liverworts-basal
Hornworts-basal
Gymnosperms
Gnepine
Conifers monophyletic
Gnetifer
Gnetales-sister
Angiosperms
Eudicots + magnoliids
Eudicots + mag/Chlor
Magnoliids + Chloranthales
Mag + Chlor, monocots
Monocots + eudicots
ANA-grade angiosperms
Amborella + Nuphar
Amborella-sister
untrim.unpart
50genes.unpart
50genes50sites.unpart
50genes50sites.part
50genes50sites.gamma.part
50genesChara.unpart
50genes50sites.25X.unpart
50genes33taxa.unpart
604genes.trimExt.unpart
604genes.trimExt.gamma.unpart
604genes.trimExt.Bayes.CATGTR
604genes.trimExt.Bayes.CAT
untrim.unpart
50genes.unpart
50genes50sites.unpart
50genes50sites.part
50genes50sites.gamma.part
50genesChara.unpart
50genes50sites.25X.unpart
50genes33taxa.unpart
604genes.trimExt.unpart
604genes.trimExt.gamma.unpart
604genes.trimExt.Bayes.CATGTR
untrim.unpart
50genes.unpart
50genes50sites.unpart
50genes50sites.part
50genes50sites.gamma.part
50genesChara.unpart
50genes50sites.25X.unpart
50genes33taxa.unpart
604genes.trimExt.unpart
604genes.trimExt.gamma.unpart
50genes50sites.unpart
50genes50sites25X.unpart
untrim
50genes
50genes.25X
50genes33taxa
untrim
untrim.gamma
50genes
50genes.25X
50genes33taxa
untrim
untrim.gamma
50genes
50genes.25X
50genes33taxa
untrim
50genes
50genes.25X
Strong Support
Weak Support
Compatible (Weak Rejection)
Strong Rejection

Analyses on these data

Look at the nucleotide data

852 gene regions

3rd positions removed because of molecular evolution issues with *Selaginella*

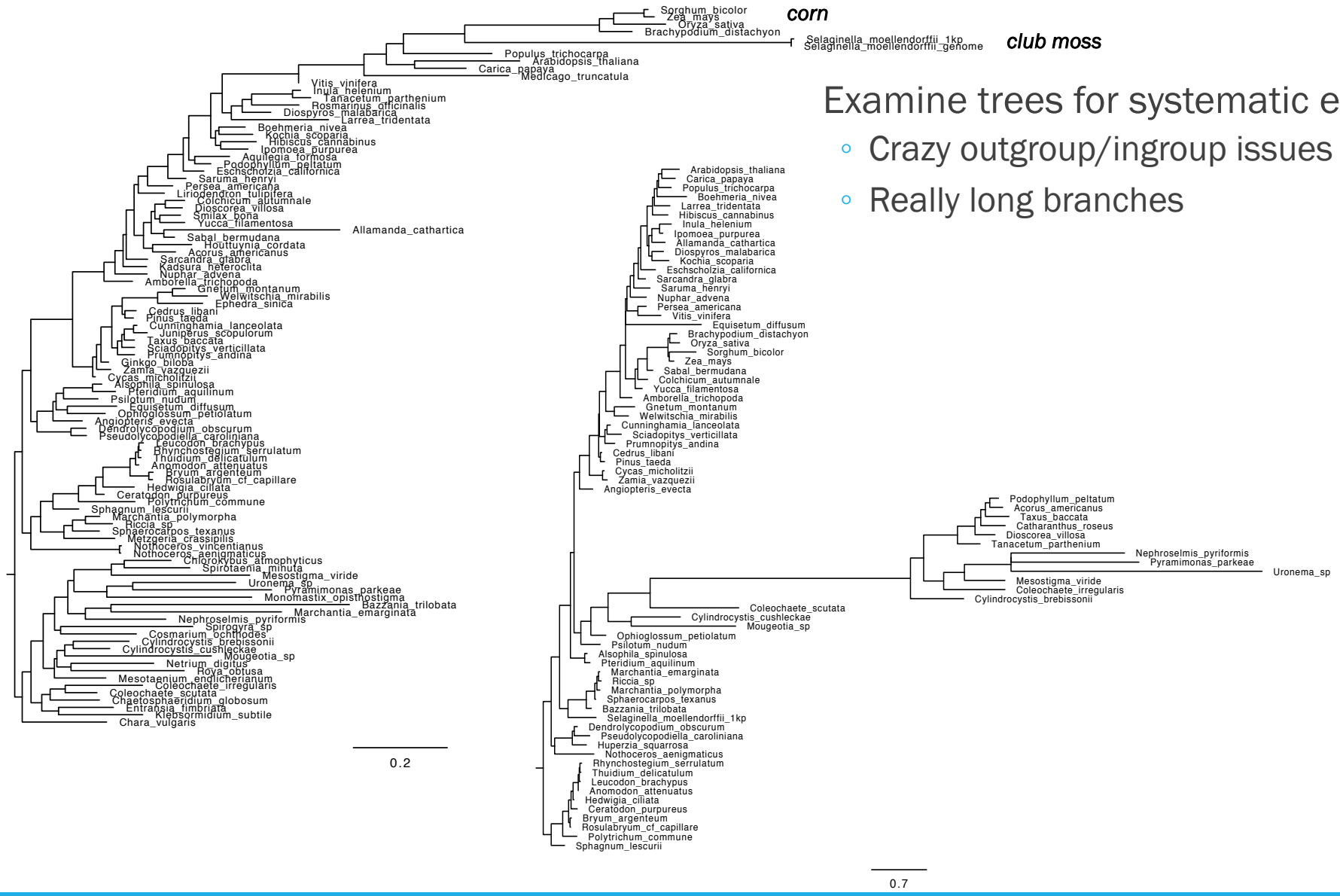
Take an edge-based view of the reconstruction of the relationships without requiring the genes to share topology beyond the edge of interest

Each gene is partially overlapping in taxonomic coverage

Answering the question: *What does the information in **this dataset** suggest for the resolution of several clades?*

Work with: Joseph Walker, Joseph Brown, Nat Hale (and me)

Gene tree examination exposes errors



Examine trees for systematic error

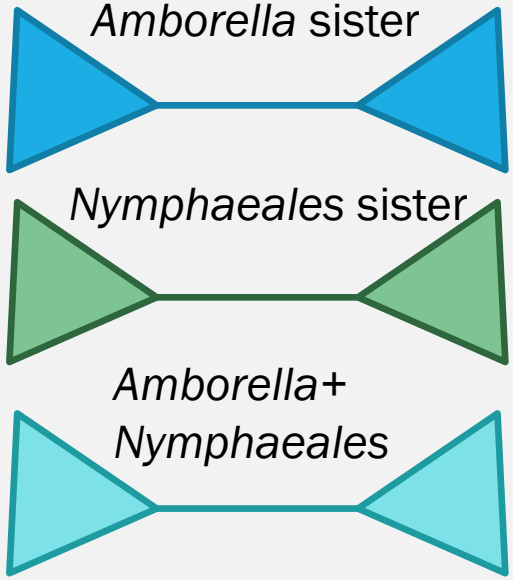
- Crazy outgroup/ingroup issues
- Really long branches

Analyses

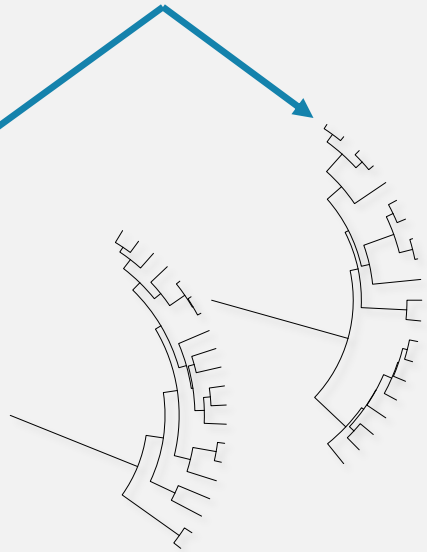
Instead of examining one or a few topologies

- Check concordance and conflict with ML gene trees (can filter on support)
- Constrain topologies based on the edges we are interested in and their alternatives

Constraints



Use constraint for ML gene runs



$\#gt * \#constraints$

Which constraint had the highest likelihood (and $> 2Lnl$)?

Check every gene

Count genes

Sum ΔLnl

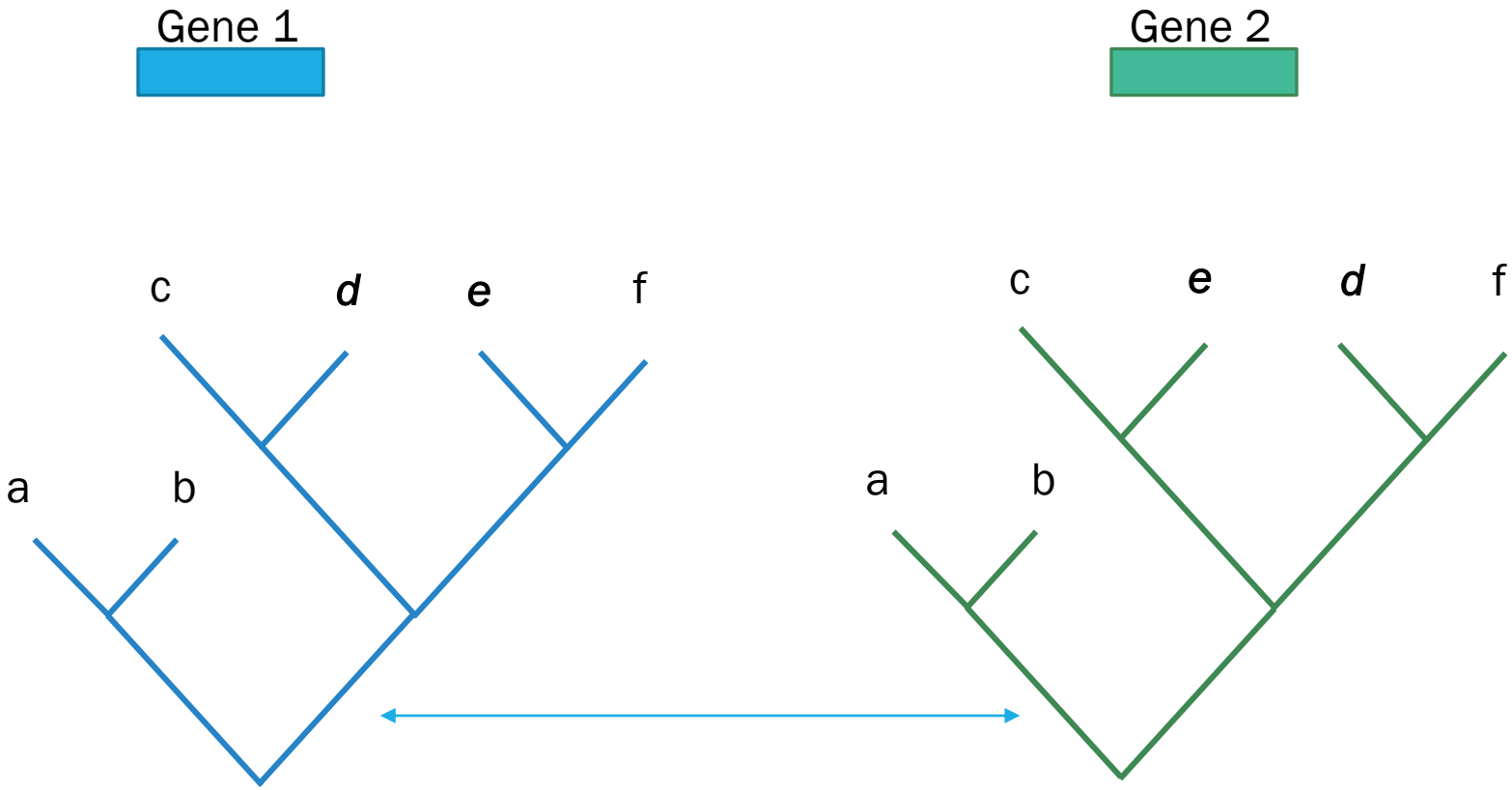
Must have necessary taxa
Can't have crazy outgroup/ingroup

Smith et al. in prep

Why not fixed topologies?

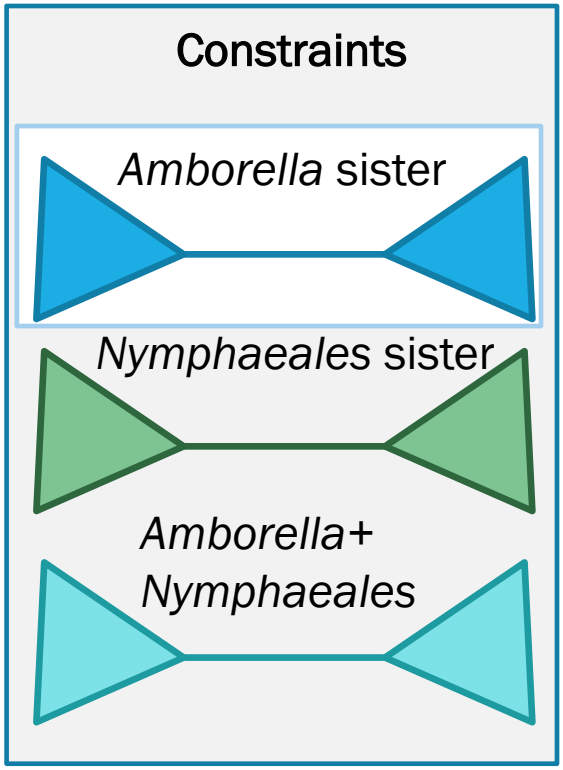
Several processes underlie these trees

As we incorporate more taxa or go deeper in the tree, we are likely to include even more complexity

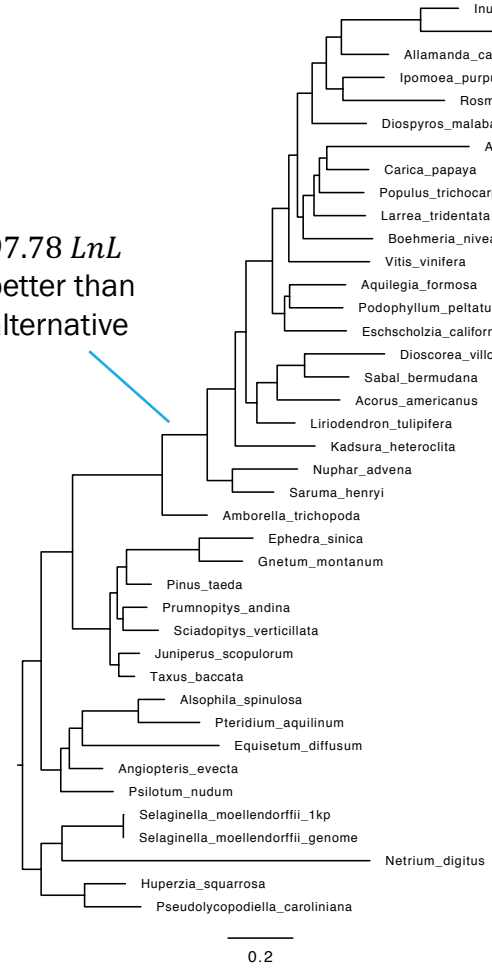


Analyses again...

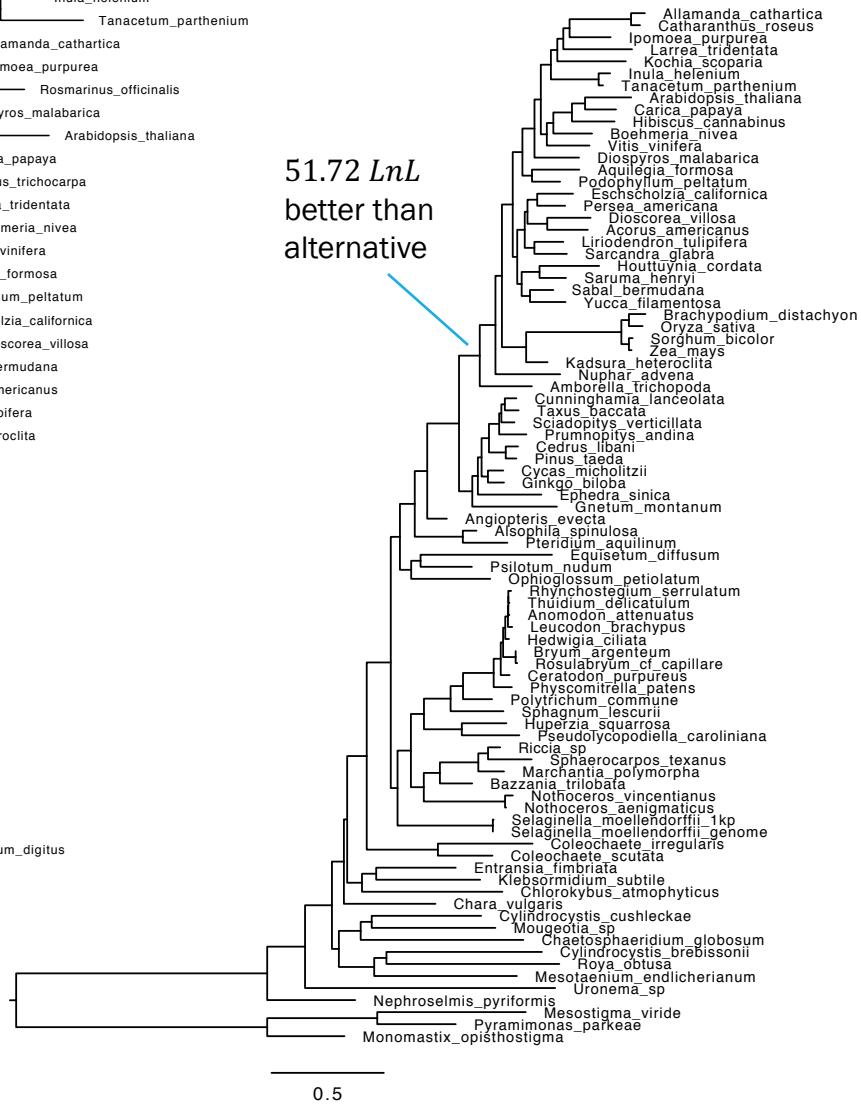
These are the better vs the alternatives and are concordant with the ML tree (so there isn't some even better alternative)



97.78 LnL better than alternative

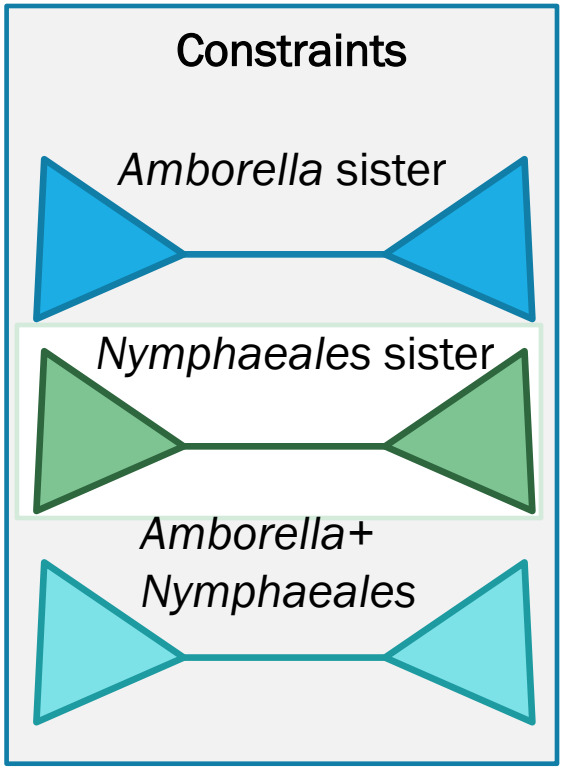


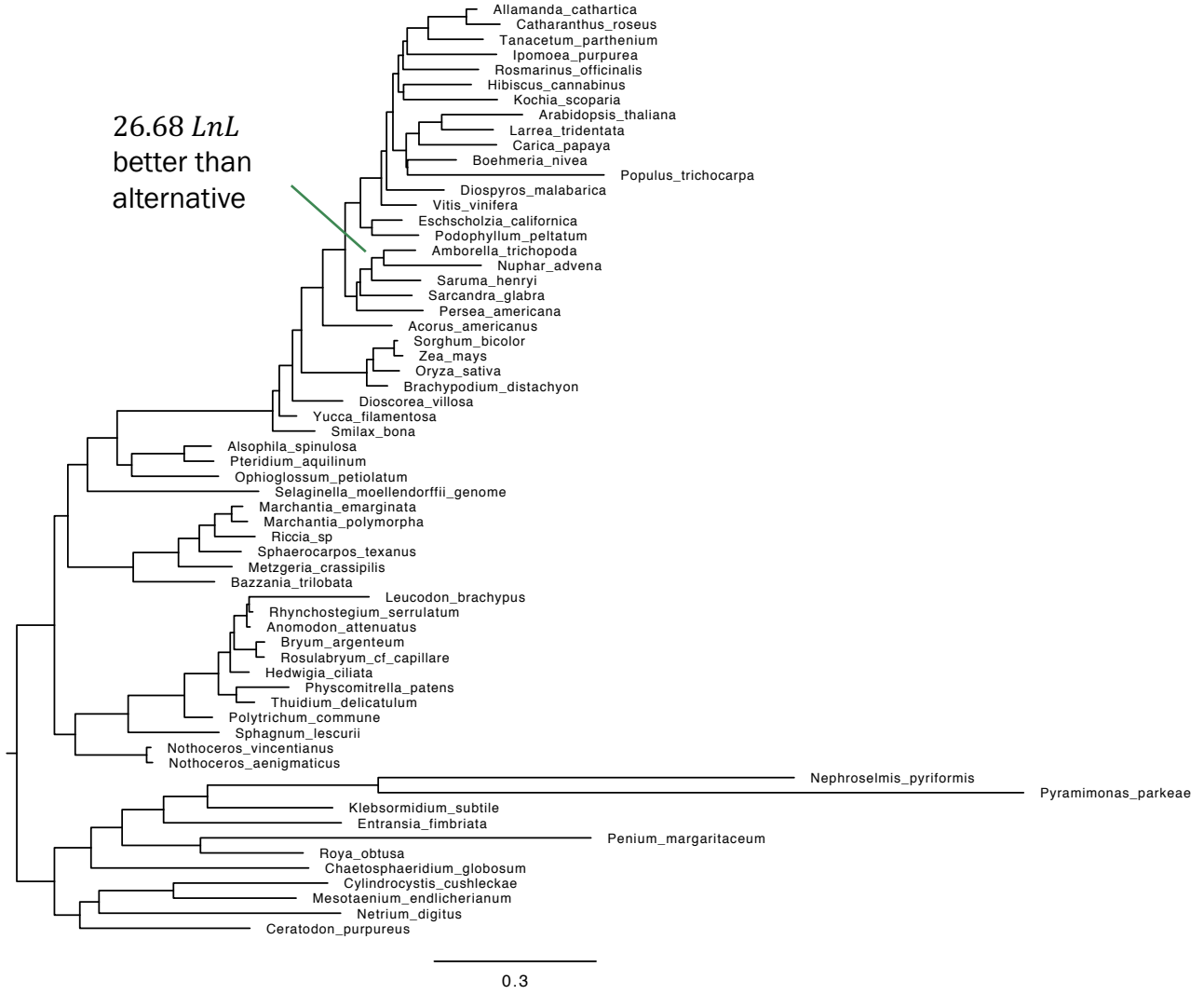
51.72 LnL better than alternative



Analyses again…

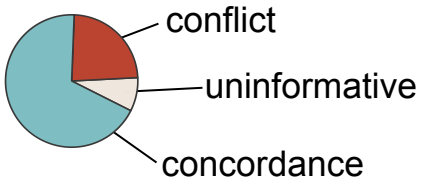
These are the better vs the alternatives and are concordant with the ML tree (so there isn't some even better alternative)

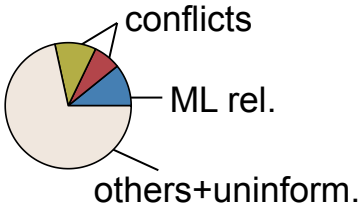




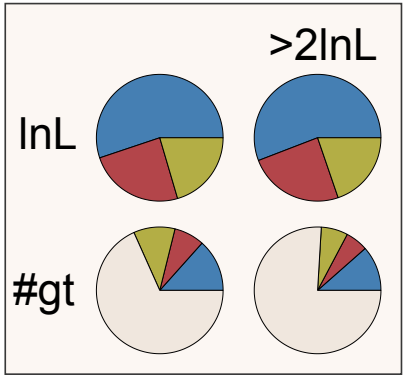
Results

Major clade	Resolutions	Genes	Genes (> 2lnL)	DlnL	DlnL > 2
Horworts	Hornworts sister*	110	83	677.6	654.1
	Liverworts sister	56	41	294.1	280.8
	Mosses+liverworts	81	40	228.9	190.2
	All monophyly	81	37	185.3	148.5
Gymnosperms	monophyly*	288	264	7259.0	7233.8
	Gnetum sister	45	31	229.8	216.0
	Cycas sister	39	18	120.3	105.2
Gymno relat.	Gnepine*	107	85	1017.2	994.4
	conifers	93	79	800.0	787.2
	Gnetifers	134	55	288.1	217.8
	Gnetales sister	76	40	211.2	176.3
Amborella	*Amborella* sister*	184	152	1501.1	1470.0
	Amborella+*Nuphar*	118	75	564.2	526.3
	Nuphar sister	111	62	392.2	345.2
Eudicots	Magnoliids+eudicots*	114	98	1223.4	1204.3
	Monocots+eudicots	66	49	541.5	526.5
	Monocots+magnoliids	90	58	453.3	425.5

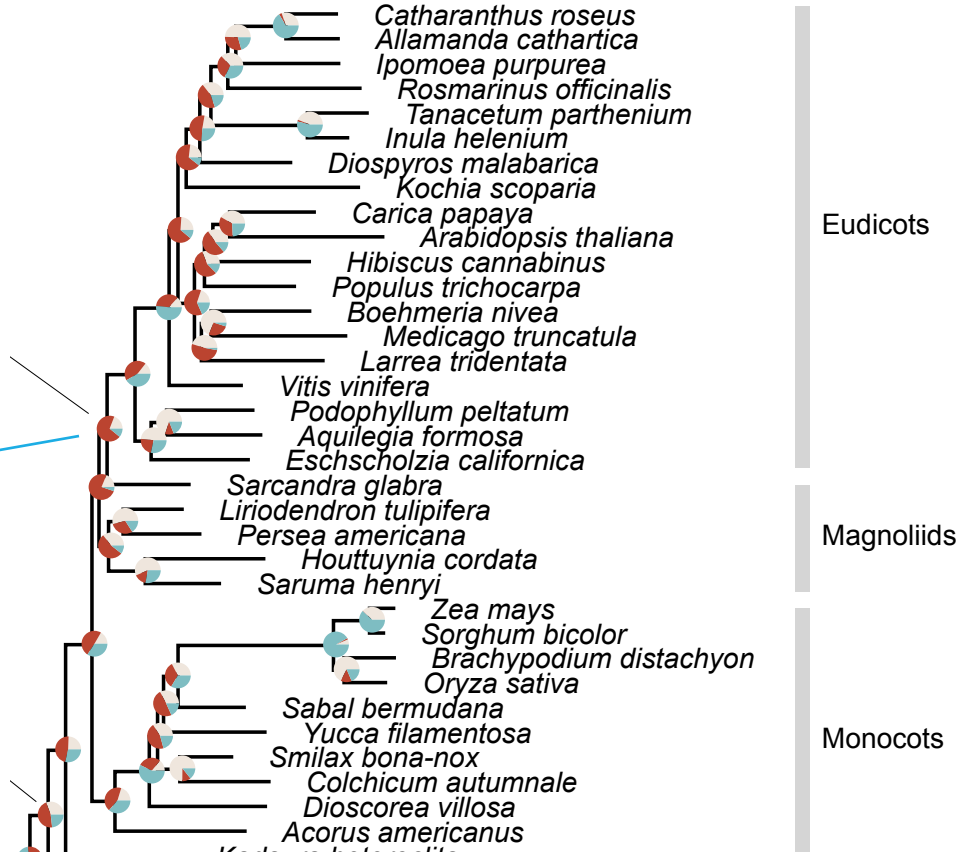




Major clade	Resolutions	Genes	Genes (> 2lnL)	DlnL	DlnL > 2
Eudicots	Magnoliids+eudicots*	114	98	1223.4	1204.3
	Monocots+eudicots	66	49	541.5	526.5
	Monocots+magnoliids	90	58	453.3	425.5



Many that were in conflict are uninformative (weakly in conflict, weird relationships, etc) or represent very minority relationships



In this dataset...

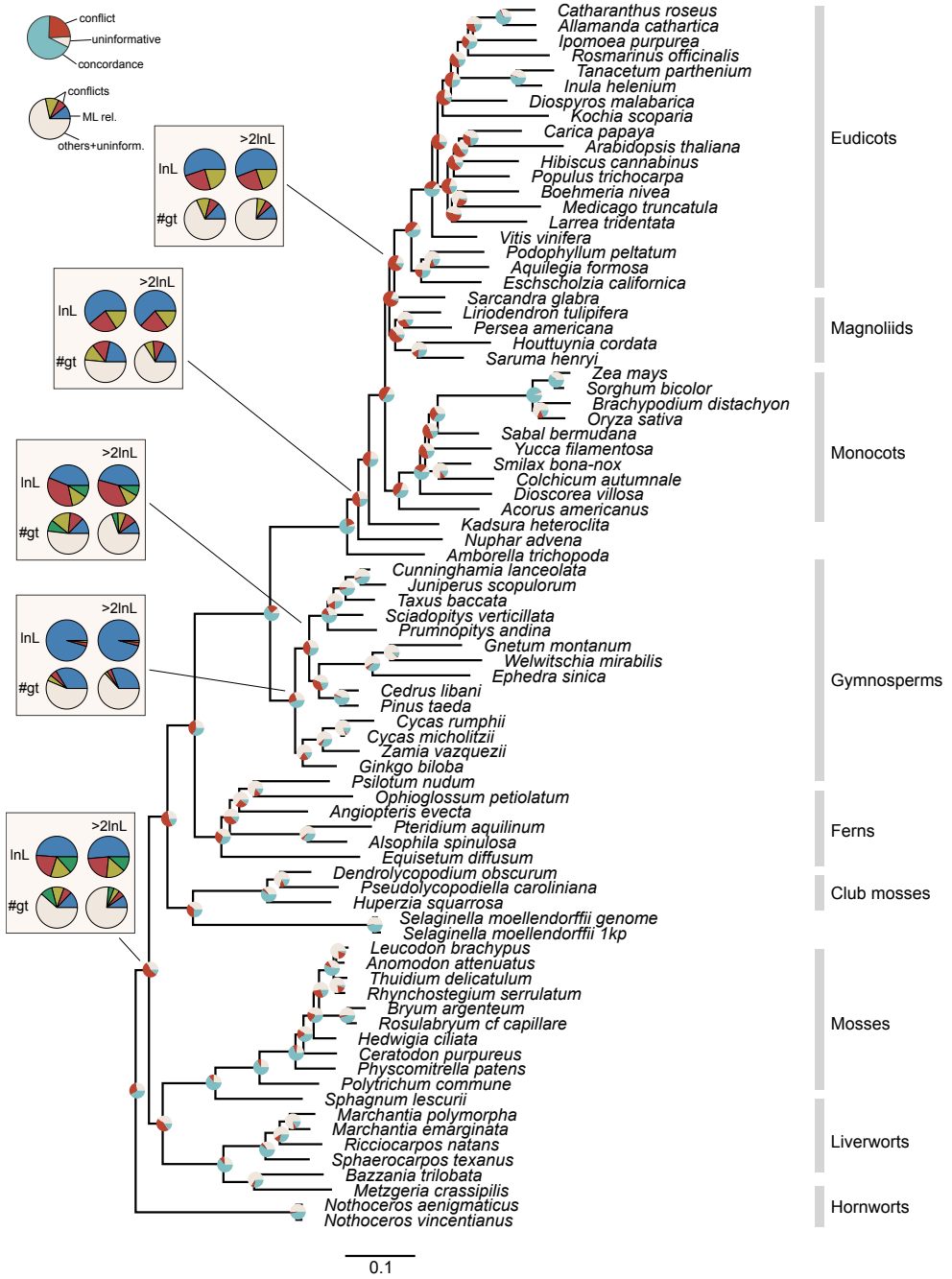
Magnoliids are sister to eudicots

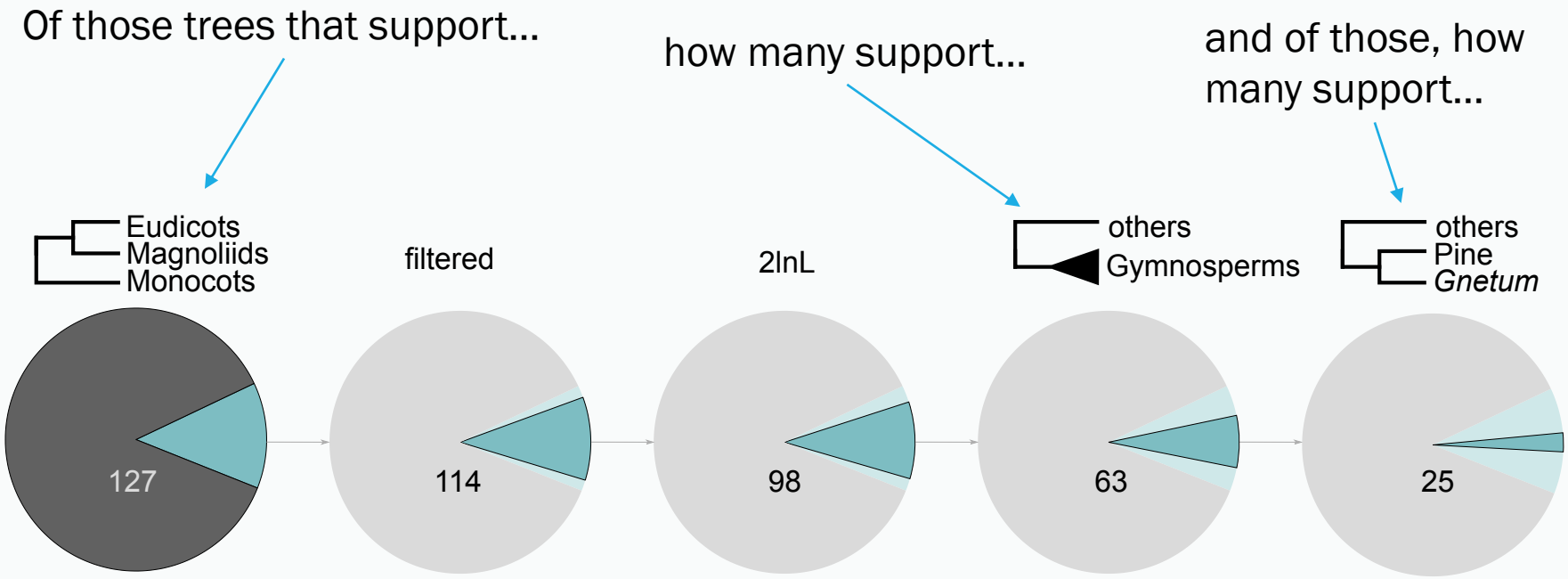
Amborella is sister to Angiosperms

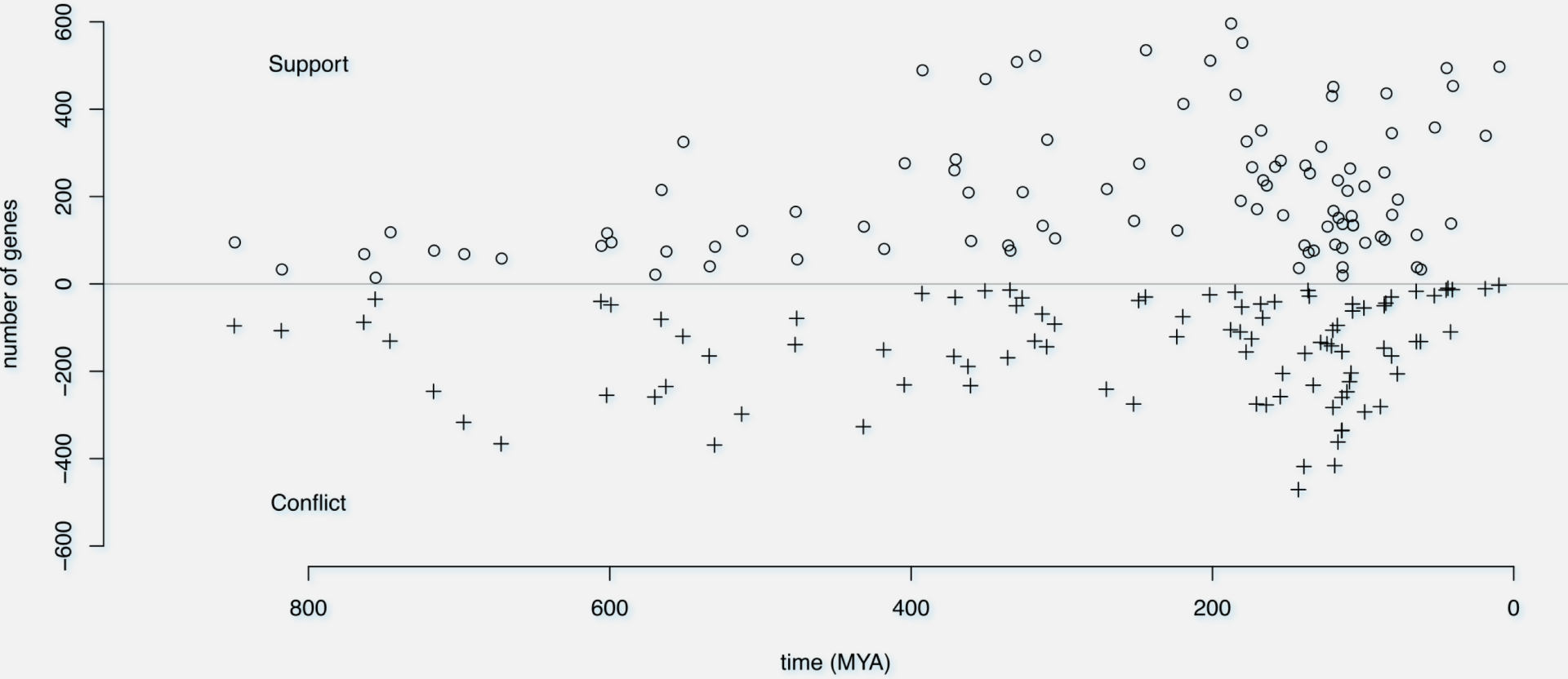
Relationships within Gymnosperms are not clear

Gymnosperms are monophyletic

Bryophytes are *not* monophyletic




Of those trees that support...
how many support...
and of those, how many support...
Eudicots
Magnoliids
Monocots
filtered
2lnL
others
Gymnosperms
others
Pine
Gnetum
127
114
98
63
25


Support
Conflict
number of genes
600
400
200
0
−200
−400
−600
800
600
400
200
0
time (MYA)

Conclusions

By examining individual relationships we may be able to more confidently make conclusions regarding ***datasets*** and ***relationships***

It is probably unreasonable to assume that all genes will speak to all the edges of a tree

- Rate of evolution
- Gene specific evolutionary processes

These approaches explored here are just the first steps, but

- They are tractable
- They are easily extensible
- They support a great deal of complexity

Questions remain

- Are these consistent with the coalescent?

Look for more coming soon!

Do you want to do anything that is here?

Check out the github and bitbucket under the

- User: blackrim
- Organization: FePhyFoFum

phyparts

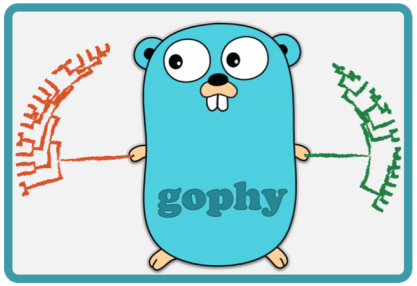
- General tool for conflict and concordance
- If you have duplications
- https://bitbucket.org/blackrim/phyparts

phyx

- Many tools but pxbp can be useful if you have mostly (all) overlapping taxa
- https://github.com/FePhyFoFum/phyx

gophy

- New tool for intended to be faster and more efficient
- https://github.com/FePhyFoFum/gophy



Look for more things coming out of the lab over the next couple months

Acknowledgements

- Funding sources
  - National Science Foundation
    - DEB
    - ABI
  - University of Michigan

- Collaborators
  - Smith lab
    - Grad students
      - Joseph Walker
      - Drew Larson
      - Lijun Zhao
    - Postdoc
      - Greg Stull
      - Ning Wang
      - Oscar Vargas
      - Diego Serrano
    - Undergrads
      - Sonia Ahluwalia
      - Julia Olivieri
      - EJ Huang
  - Former postdocs
    - Ya Yang
    - James Pease
    - Joseph Brown
    - Cody Hinchliff
  - Michael Moore
  - Sam Brockington
  - Douglas Soltis
  - Pam Soltis
  - All the Open Tree of Life folks

https://github.com/blackrim/ blackrim.org