Phylogenomic estimation of the Hemiptera phylogeny and the co-diversification with plants

Christopher L. Owen
Postdoctoral Researcher
George Washington University
Computational Biology Institute

OPEN Tree of Life

Outline

Introduction

- Review previous Hemiptera phylogeny estimates

Methods

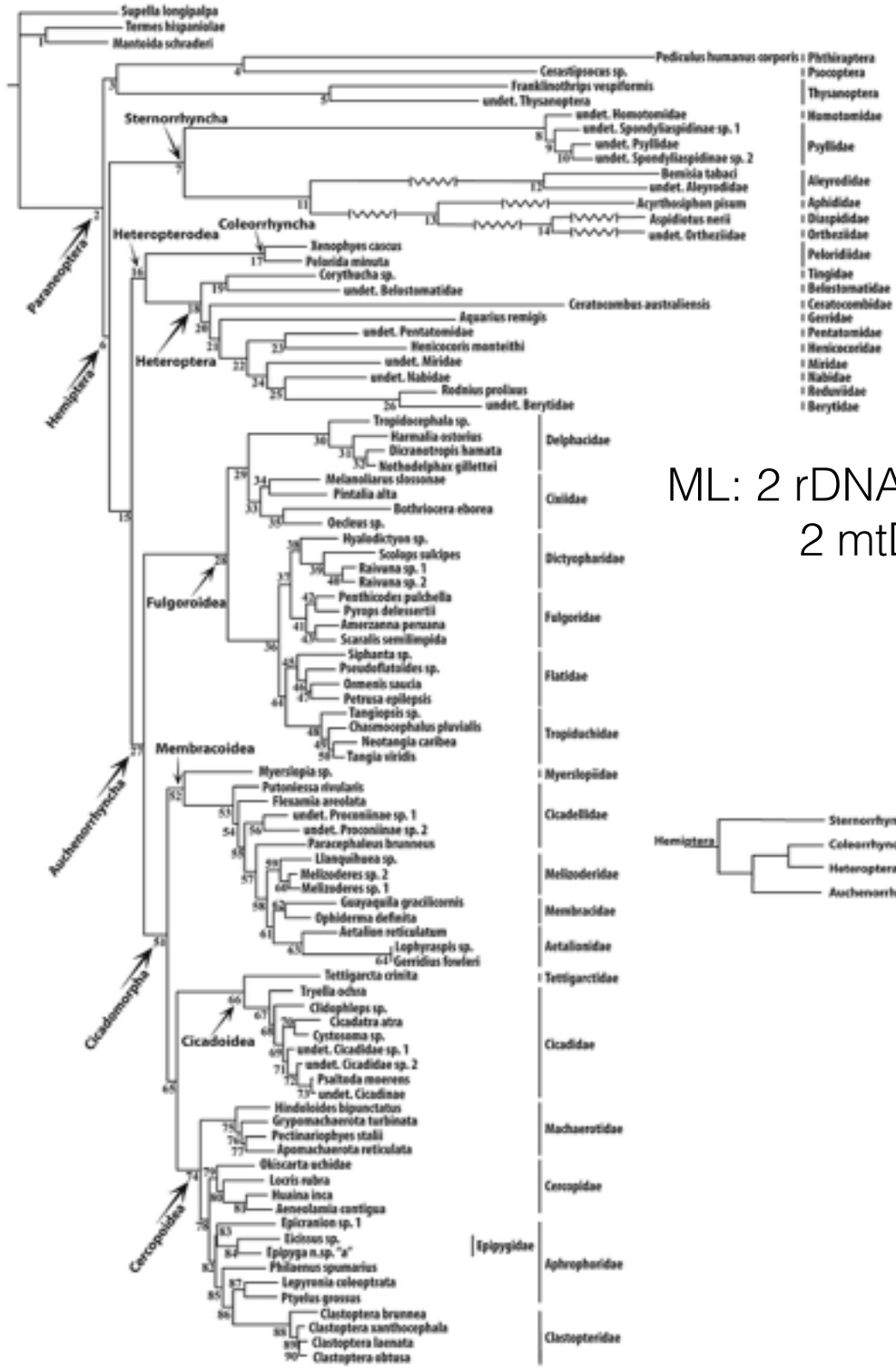
- Dataset assembly and curation
- Preliminary dating and diversification analyses

Results

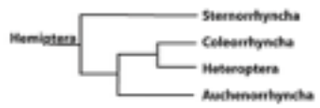
- Phylogeny of Hemiptera
- Gene tree/species tree discordance/concordance
- Preliminary chronogram
- Preliminary diversification results

Recent Hemiptera Phylogenies: Cryan and Urban 2012 J. Syst. Ent.

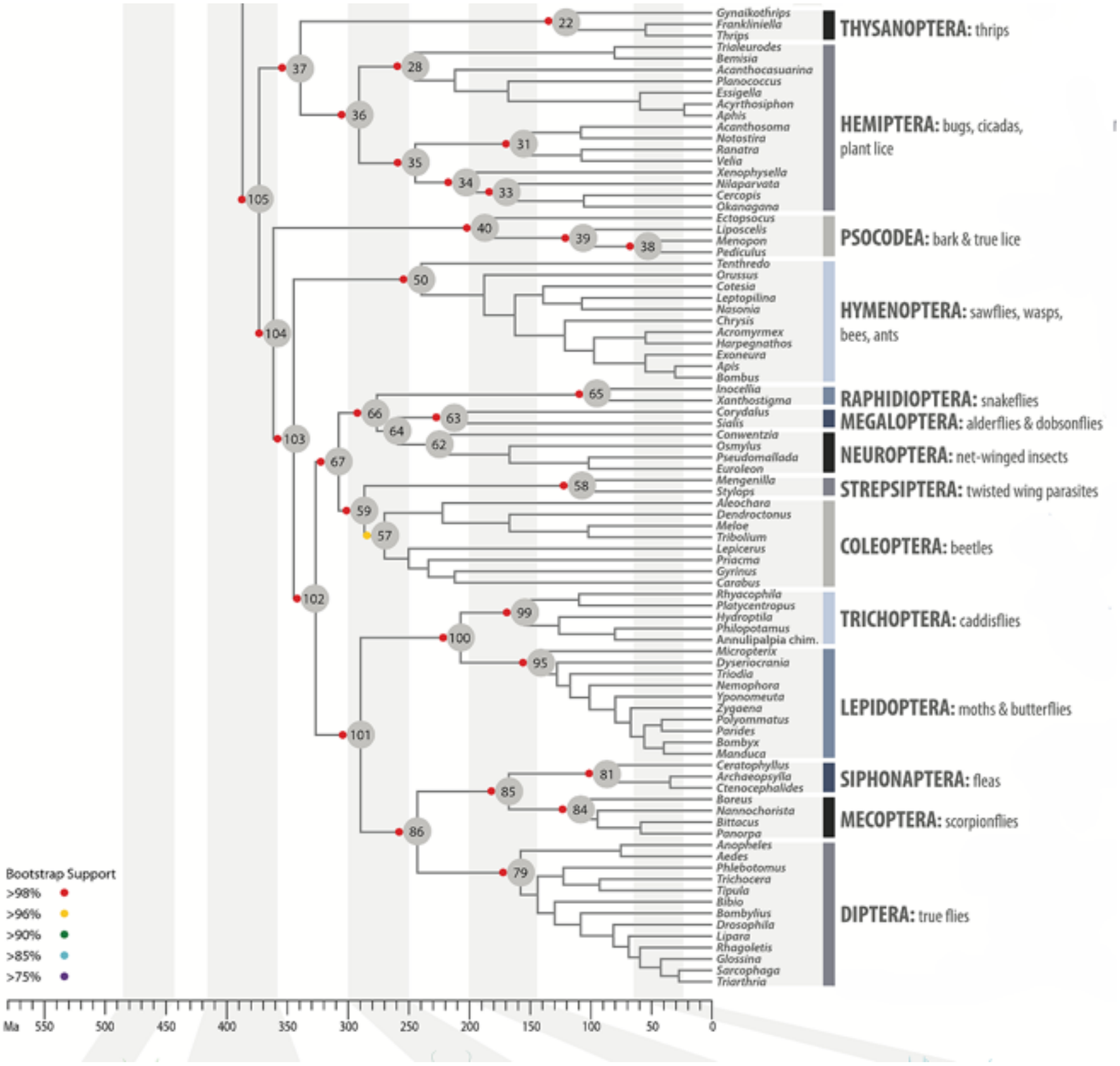
- Well-supported backbone for most higher taxa
- Lacking support for Hemiptera
- Lacking support for Paraneoptera
- Molecular evidence for long branches in Sternorrhyncha



ML: 2 rDNA, 3 nDNA, 2 mtDNA



Recent Hemiptera Phylogenies: Misof et al. 2014 Science



- 1,478 orthologs
- Lots of curation and analyses!
- 100 BS for Psocodea sister to Holometabola; poor support quartet mapping

Methods

Transcriptome assembly, ortholog identification, and phylogenetic analyses

<u>Transcriptome assembly</u>

- 57 Paraneoptera transcriptomes and NCBI ESTs
- Transcriptomes assembled from raw NGS reads
- Quality control->Error correction->Trinity

<u>Ortholog Identification</u>

- OrthoDB: *R. prolixus*, *P. humanus*, *A. pisum*
- HaMStR pipeline

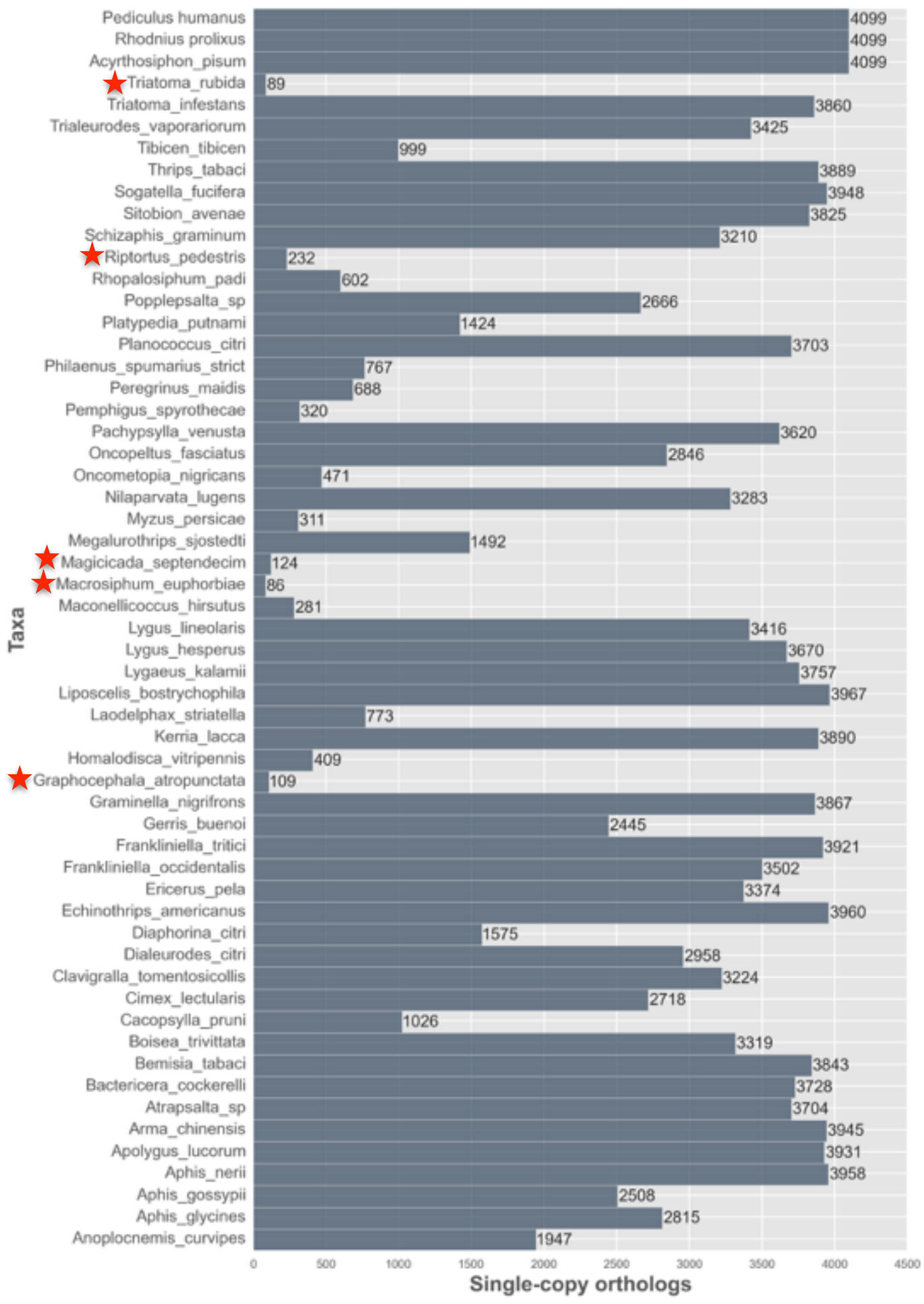
Alignments, trimming, phylogenetic analyses

- Mafft->GBlocks->Prottest3->PhyloBayes & RAxML->Decisiveness

Divergence time estimation and diversification rates

- treePL -> Medusa

Results



Taxon & Ortholog Sampling

<u>Outgroups</u>

- Thysanoptera: 5 species
- Phthiraptera: 1 species
- Psocoptera: 1 species

<u>Auchenorrhyncha</u>

- Cicadoidea: 5 species
- Cercopoidea: 1 species
- Membracoidea: 4 species
- Fulgoroidea: 4 species

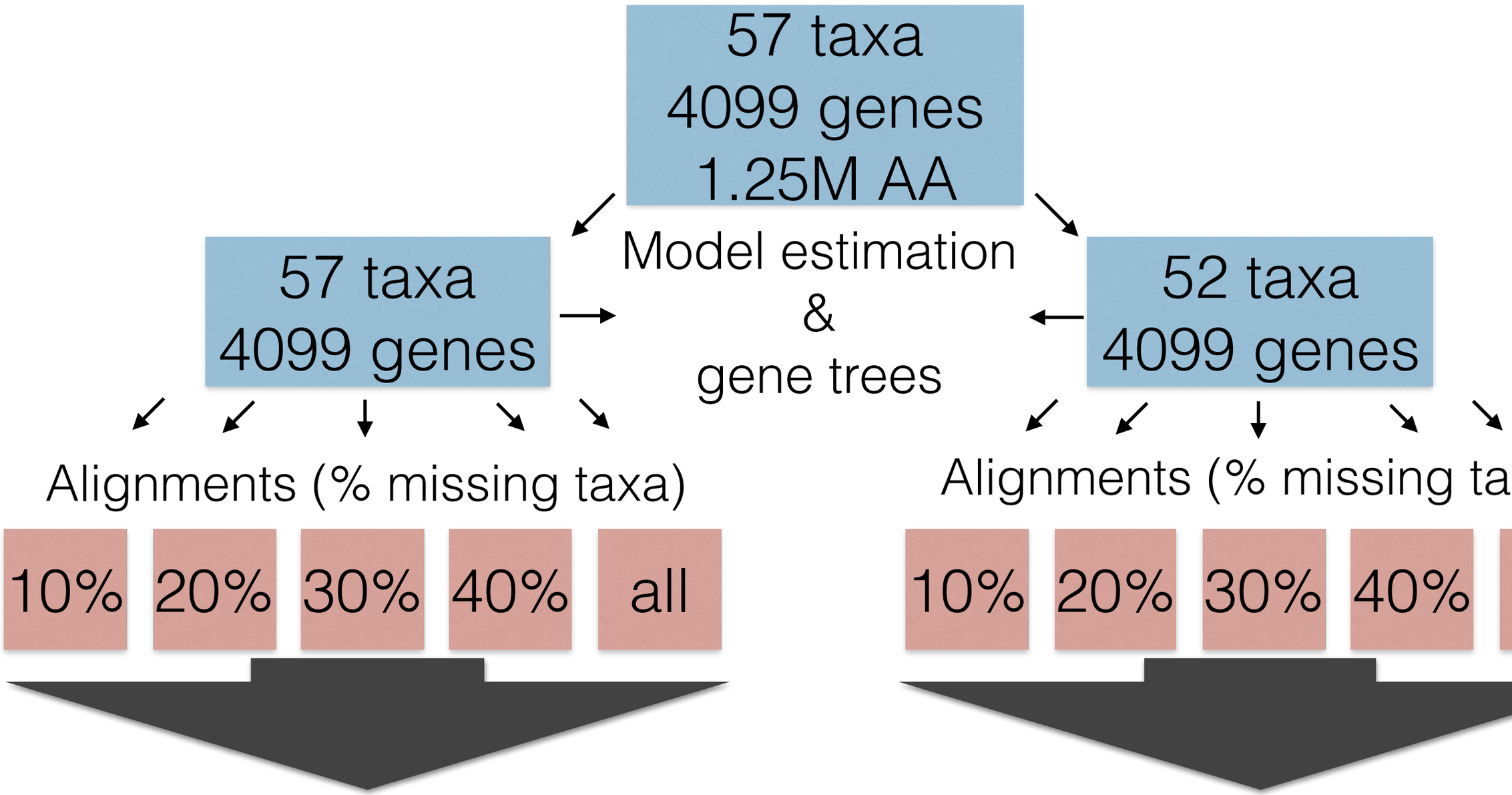
<u>Heteroptera</u>

- Pentatomoidea: 1 species
- Coreoidea: 4 species
- Lygaeoidea: 2 species
- Cimicoidea: 1 species
- Miroidea: 3 species
- Reduvioidea: 3 species
- Gerroidea: 1 species

<u>Sternorrhyncha</u>

- Coccoidea: 4 species
- Aphidoidea: 10 species
- Psylloidea: 4 species
- Aleyrodidae: 3 species

Datasets and gene removal:”Choosing the best genes for the job” (Collins 2005)



Choosing the best genes:

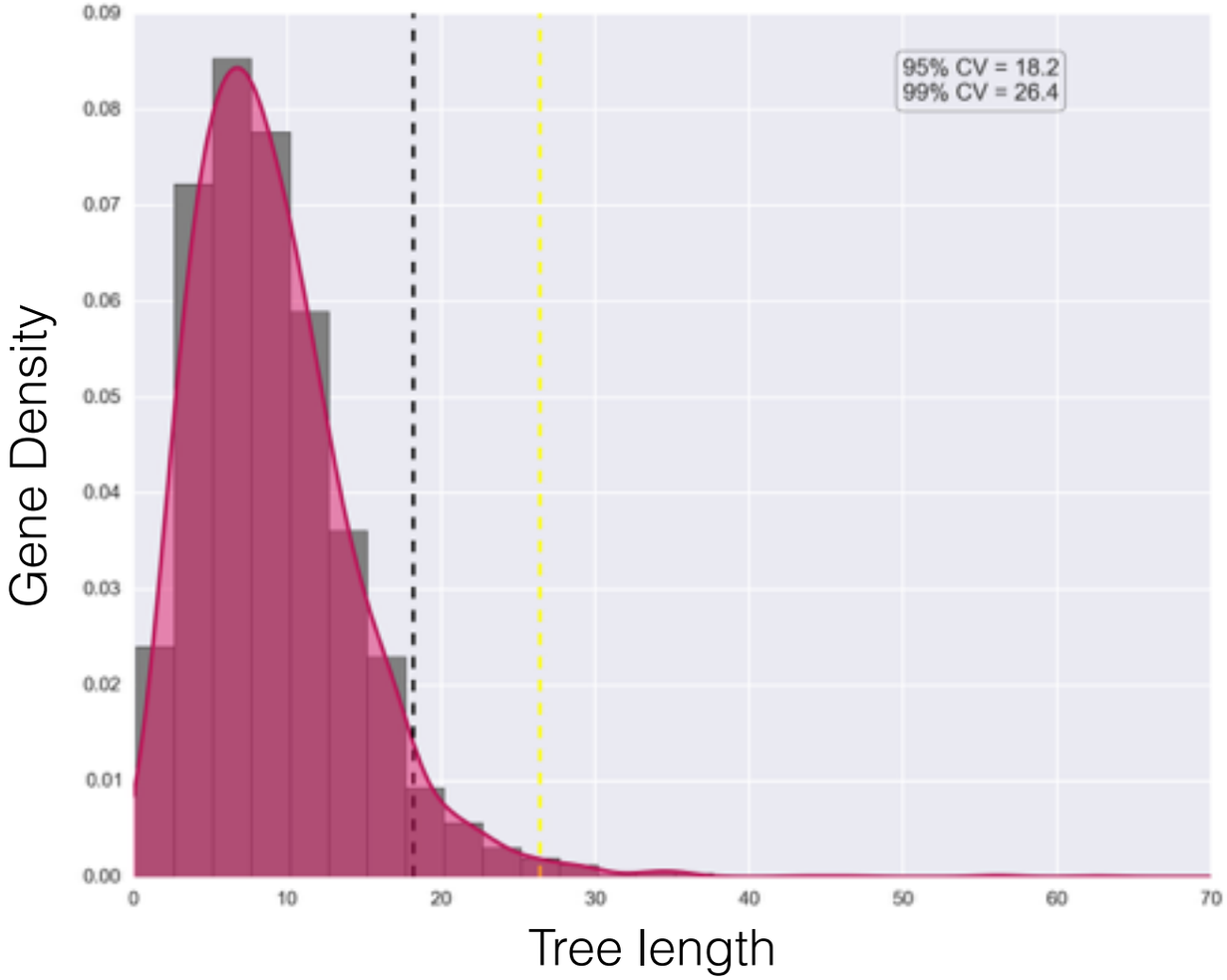
1. Rate of evolution
2. Among site rate variation
3. Compositional bias
4. Long branches

Choosing the best genes:

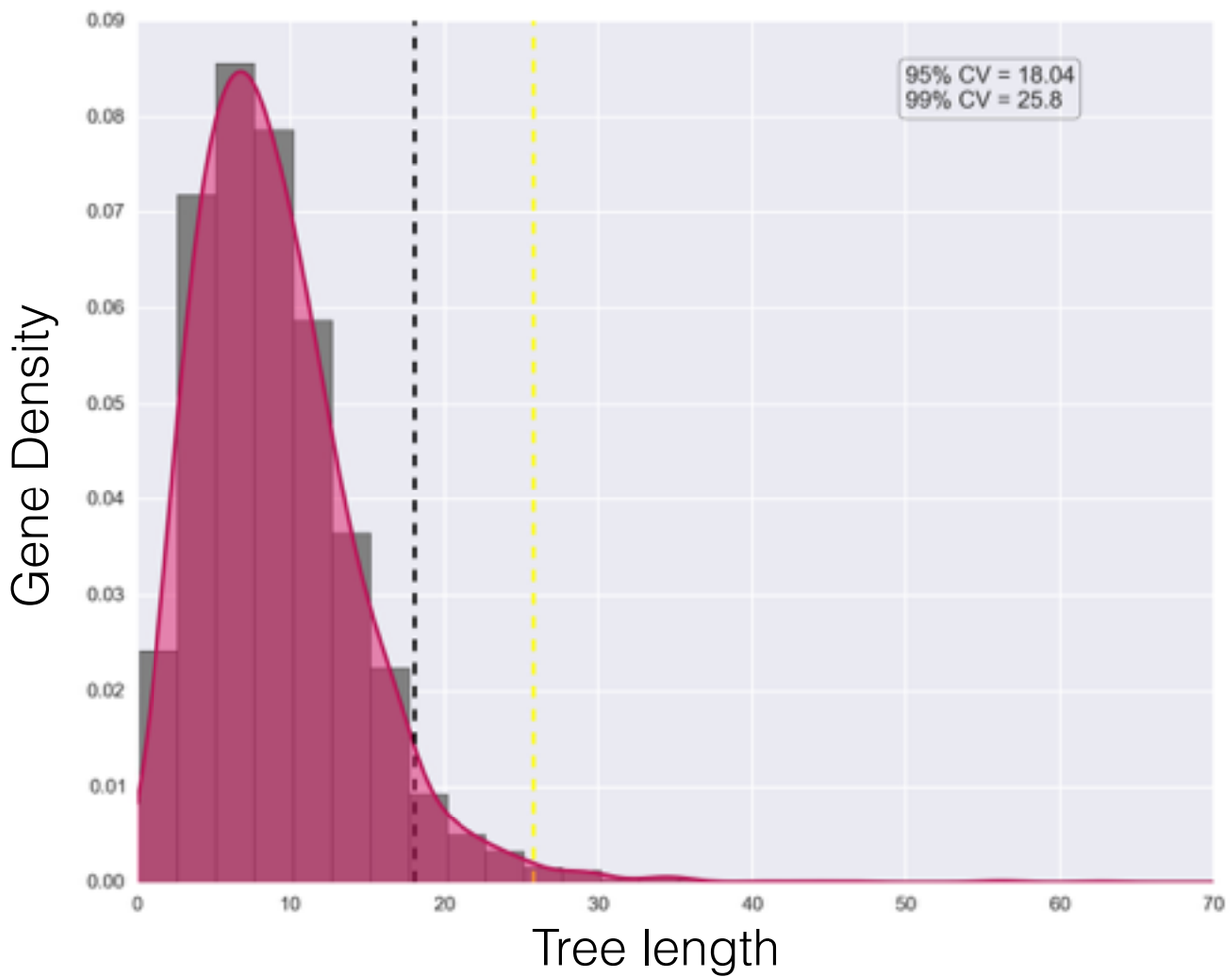
1. Rate of evolution
2. Among site rate variation
3. Compositional bias
4. Long branches

Choosing the best genes: rate of evolution

No taxa removed



5 taxa removed

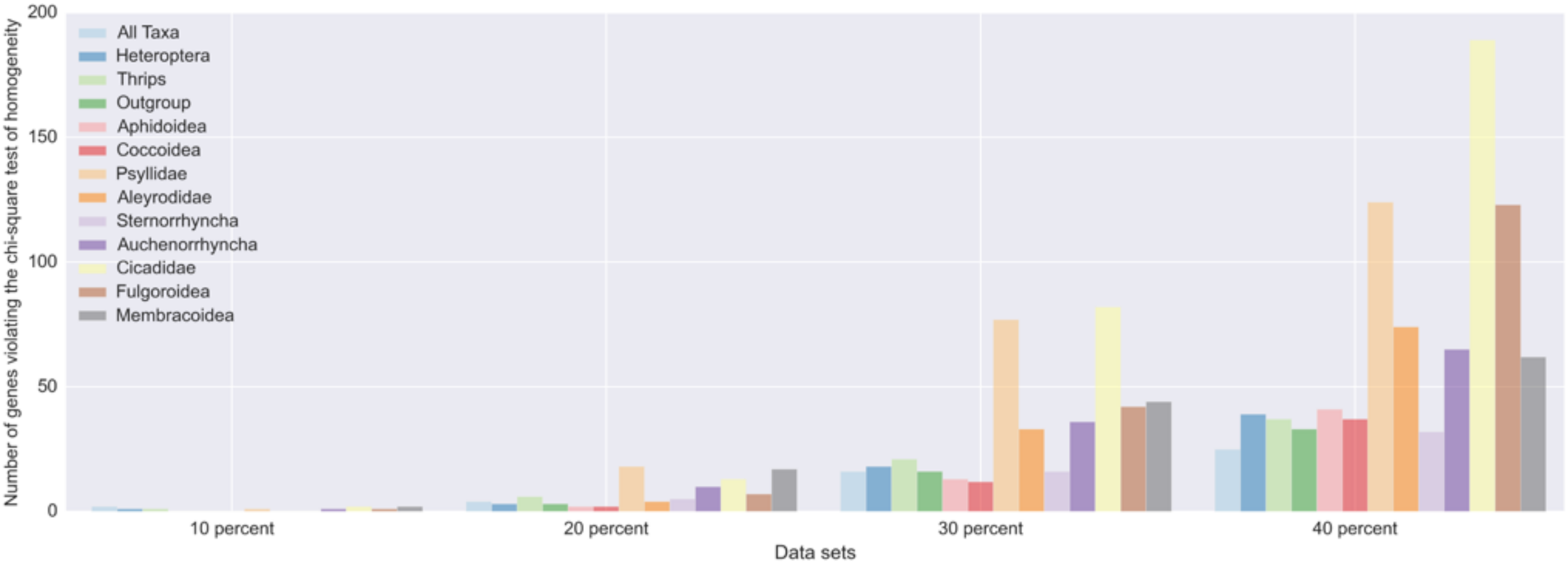


- Excluded genes with high rates of evolution (Oakley et al. 2013 MBE; Philippe et al. 2005)
- Arbitrary cutoff of TL 10

Choosing the best genes: among site rate variation

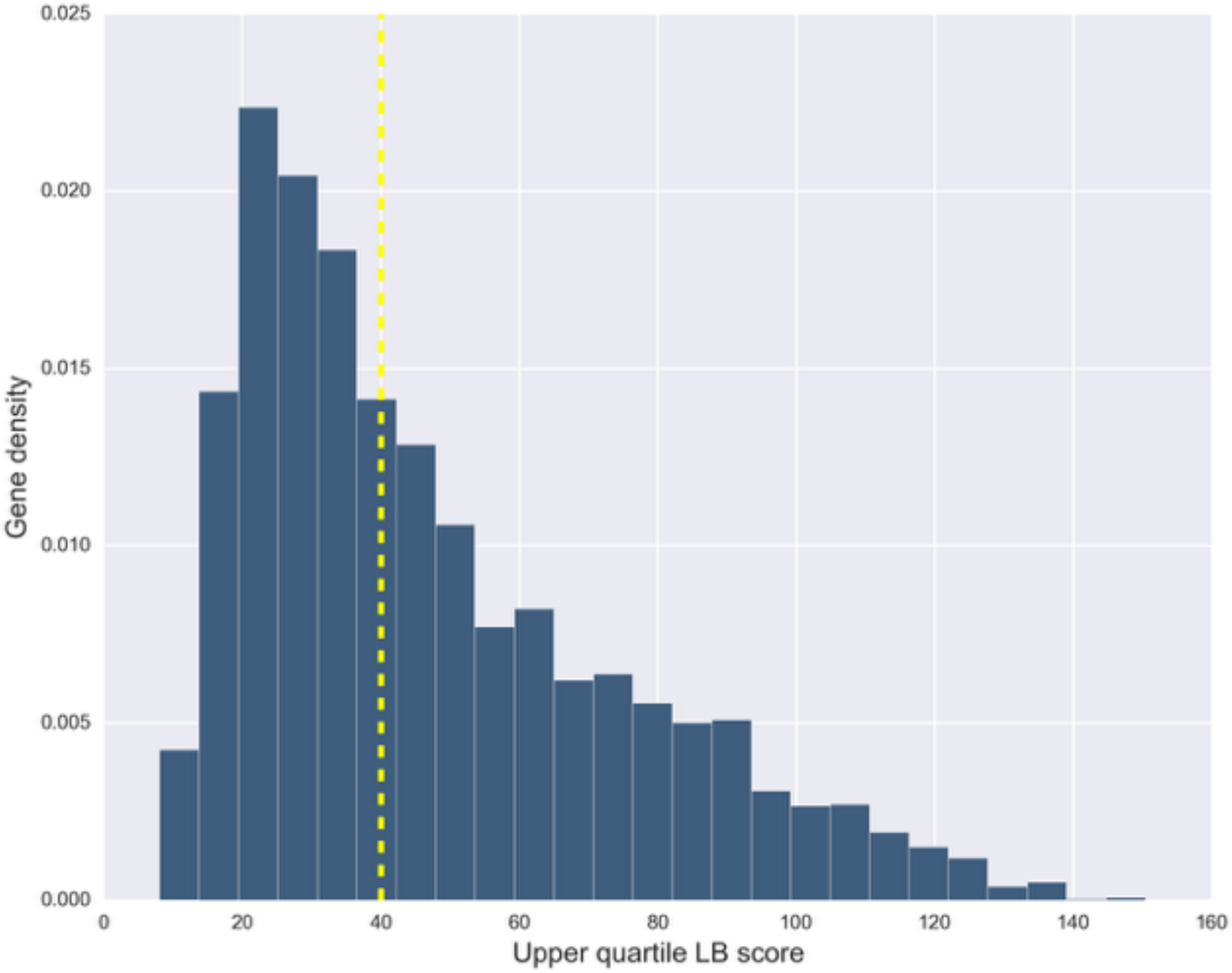
- removed genes alpha value < 1 (Yang 1994)

Choosing the best genes: compositional bias



- Chi-square test among taxa and higher taxa ($p<0.5$)
- suffers from type 2 error (Foster 2004)
- Foster 2004; Galtier & Gouy 1995; Lockhart et al. 1992

Choosing the best genes: long branches



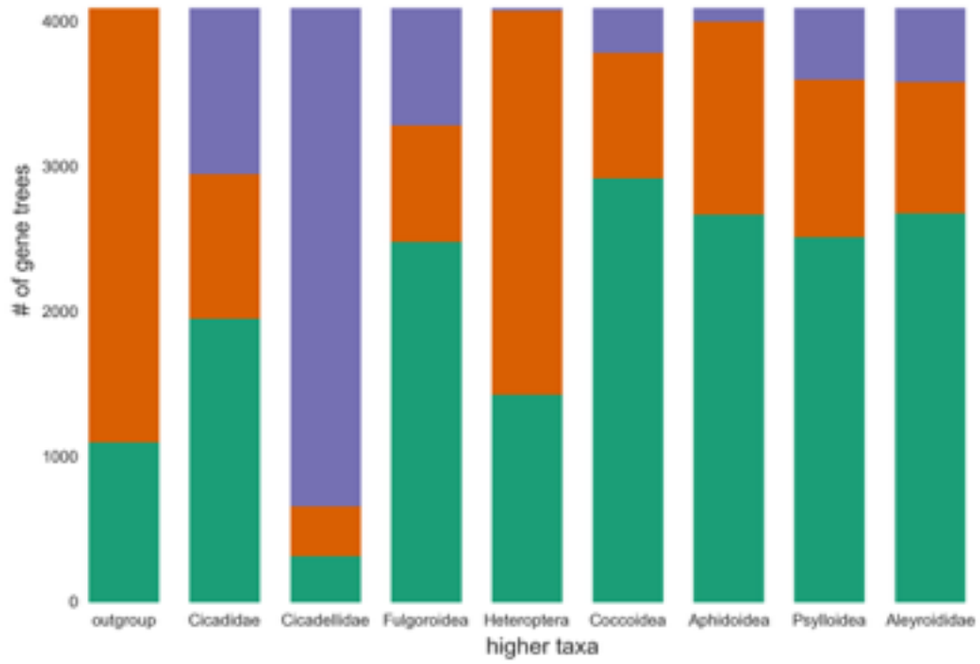
$$LB_i = \left(\frac{\overline{PD_i}}{\overline{PD_a}} - 1 \right) * 100$$

- LB score upper quartile of each partition
- unrooted gene trees
- Brinkman & Philippe 2008; Bergsten 2005

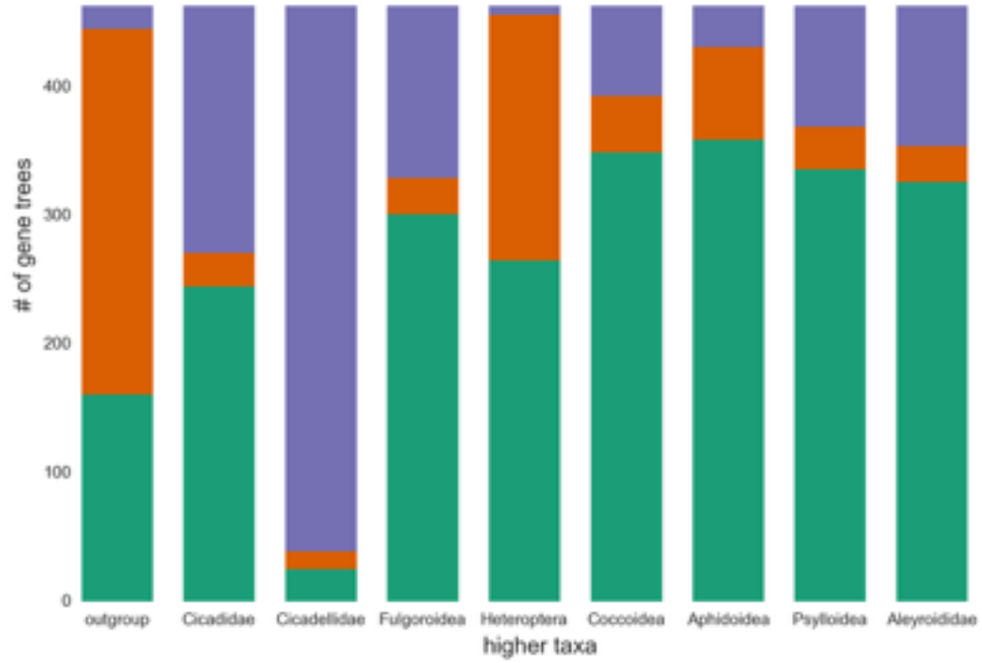
Comparison of monophyly among datasets

5 taxa rm

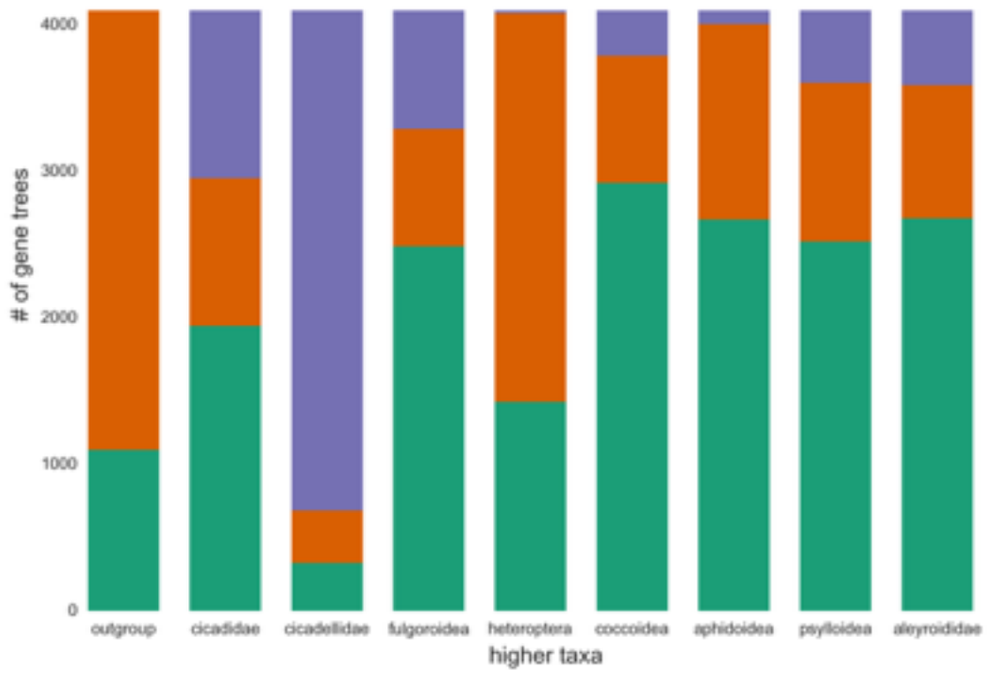
All genes

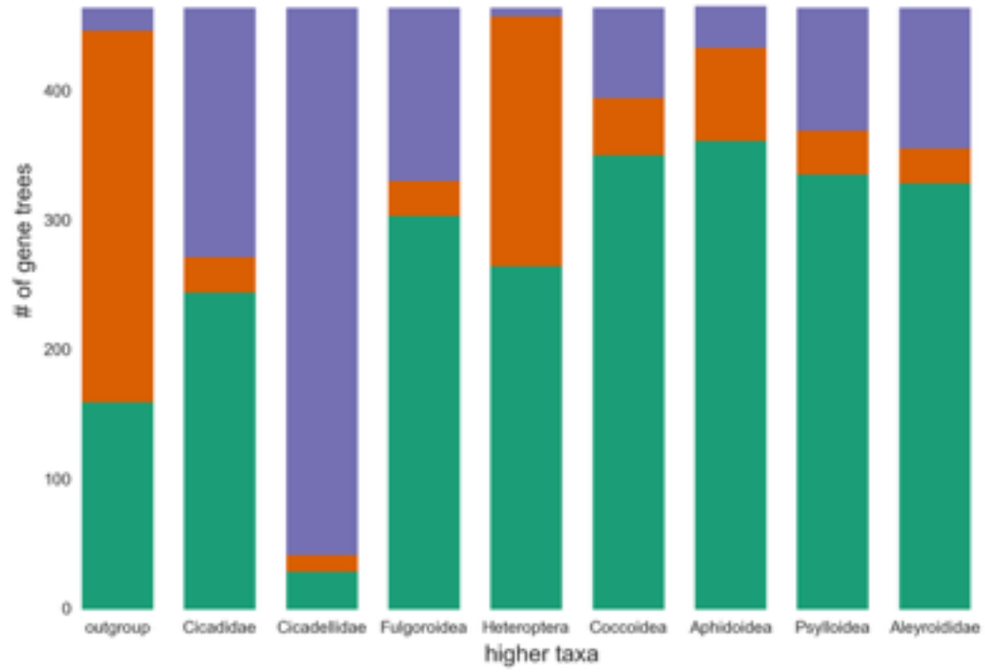


Curated dataset



All taxa

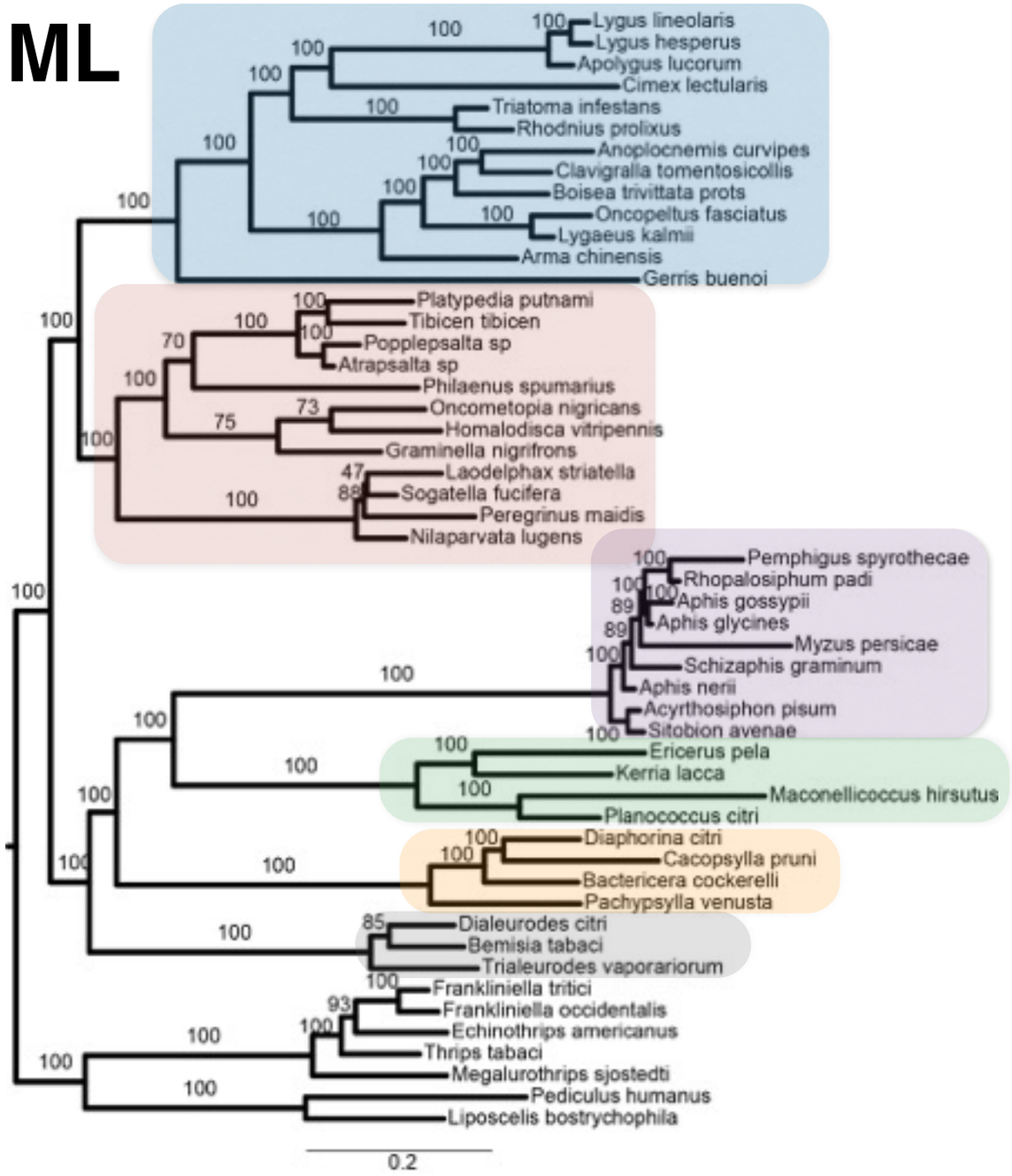




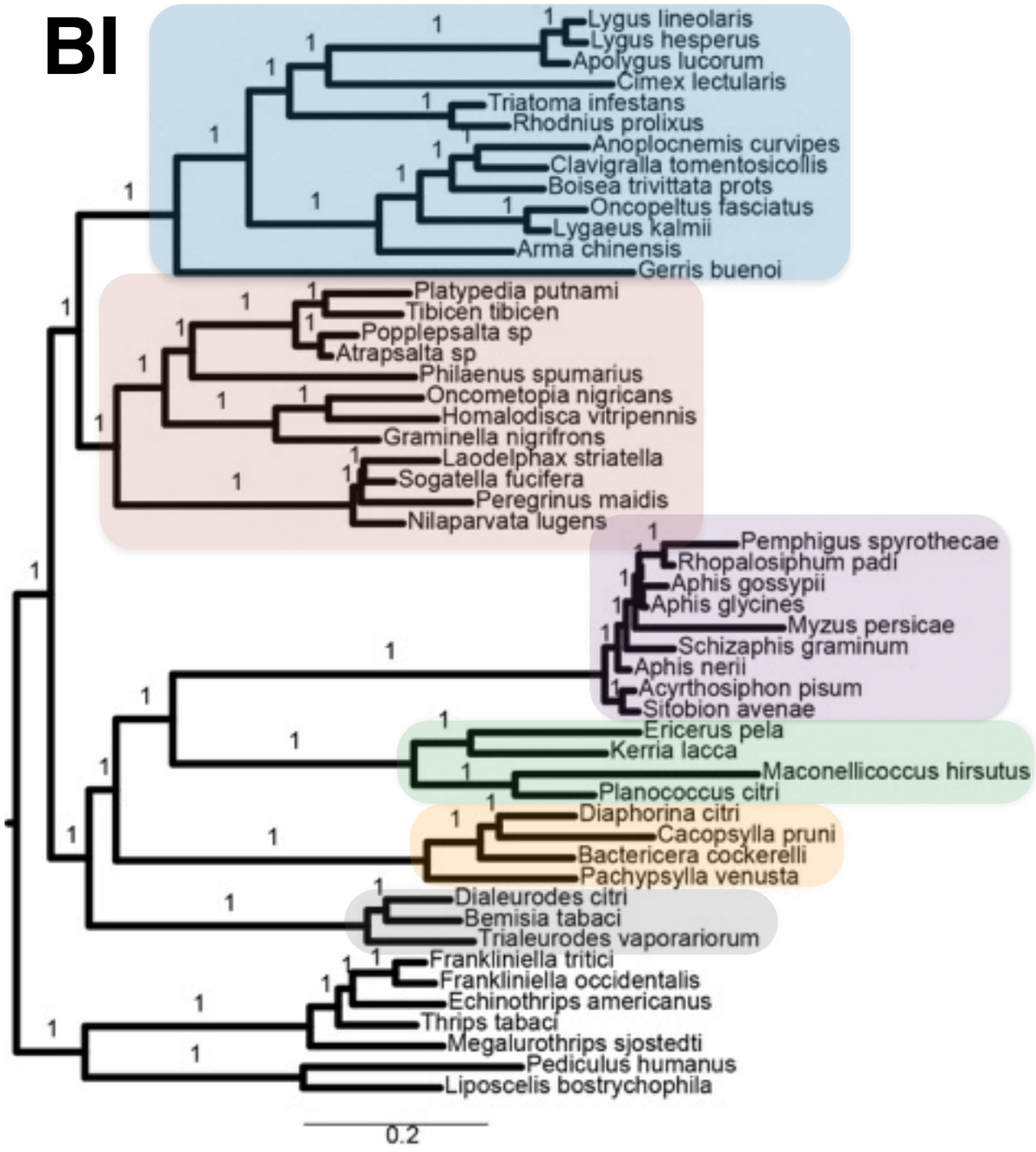
Hemiptera phylogeny: most curated dataset

Alignment: 463 genes; 106,740 AA; 52 taxa

ML



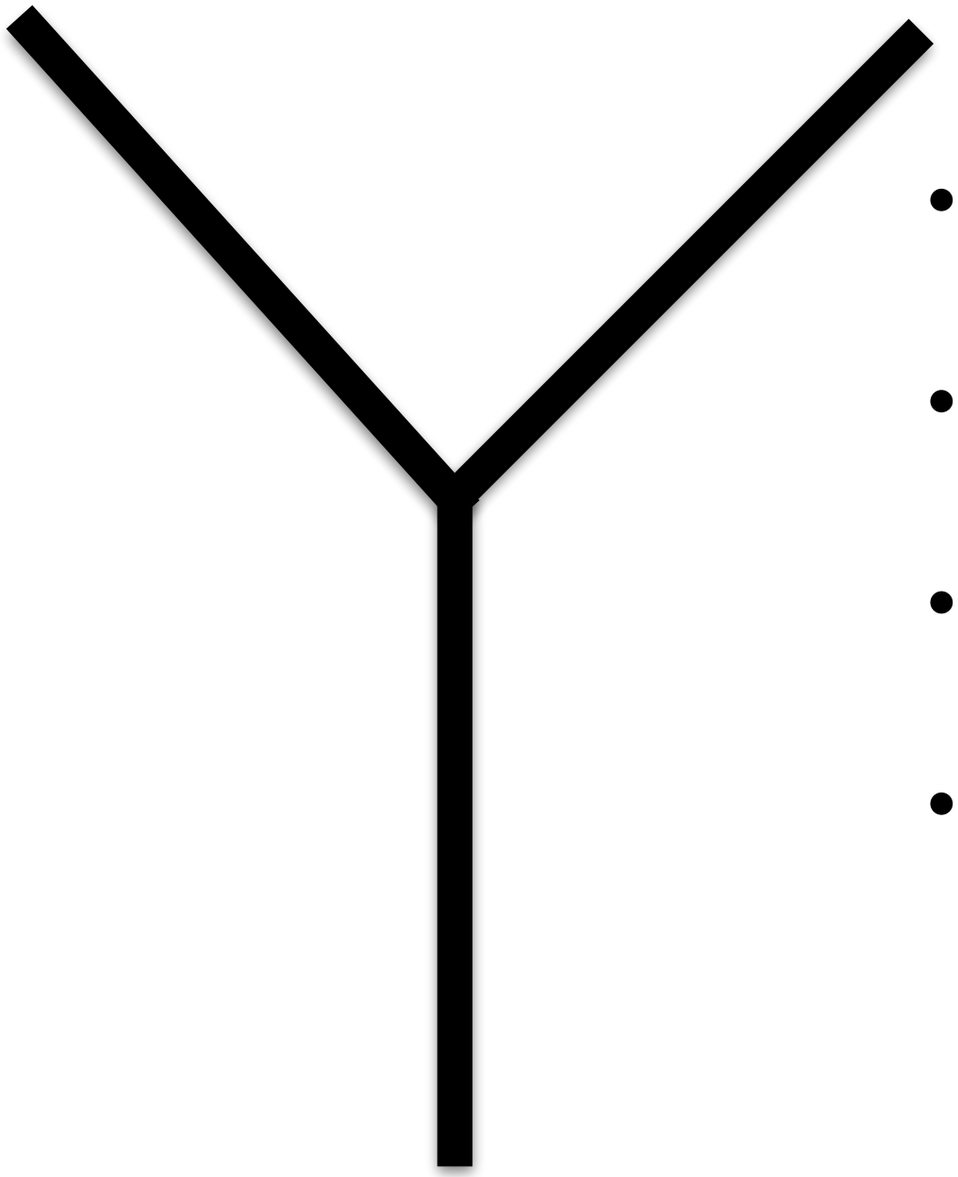
BI



Monophyly of Paraneoptera

Thysanoptera

Hemiptera

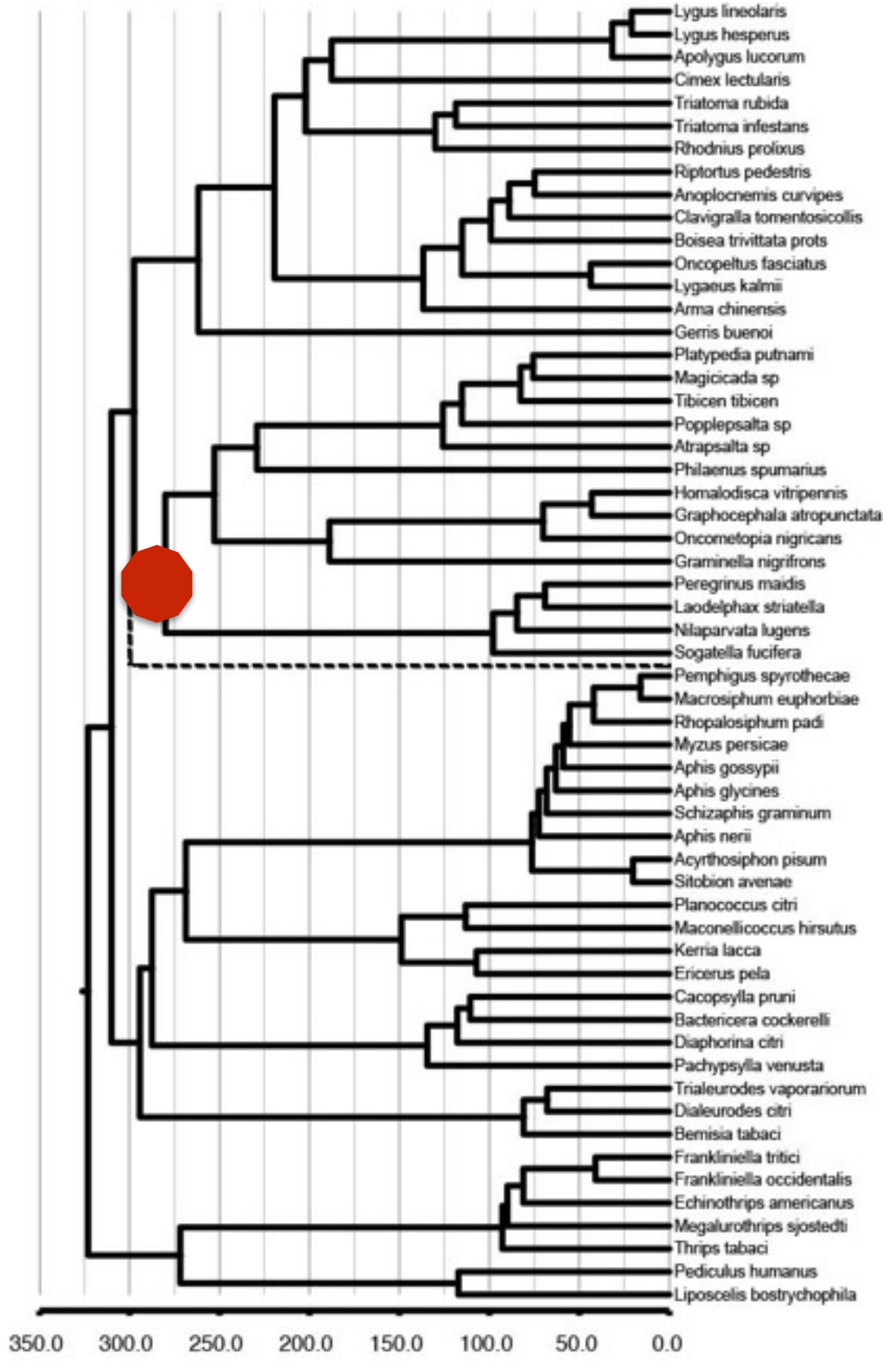


Psocodea

Dataset: 52 taxa; 463 genes

- 149 gene trees all three lineages are monophyletic
- 3/149 Psocodea more closely related to Thysanoptera
- 146/149 Hemiptera more closely related to Thysanoptera
- Distance from Psocodea to Hemiptera ~2x greater than the distance from Thysanoptera to Hemiptera

Divergence times and diversification



Divergence times: treePL

- calibrations: Li et al. 2012; Misof et al. 2014
- a little older than Misof et al. 2014
- need to validate fossil congruence

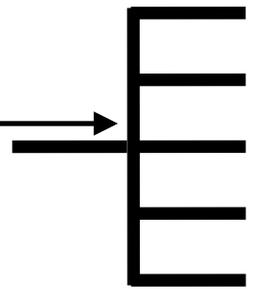
Diversification: Medusa

- Fit BD models to nodes
- compare models with AIC
- identified one change in diversification rate leading to Coleorrhyncha

Hemiptera phylogeny and diversificaton: future direction

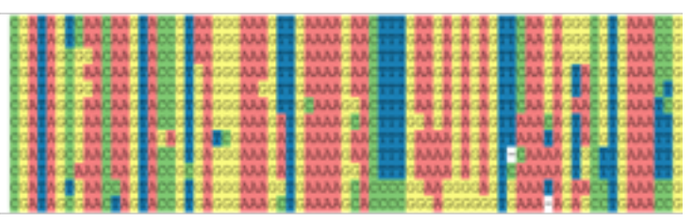
Hemiptera backbone phylogeny

Taxonomy



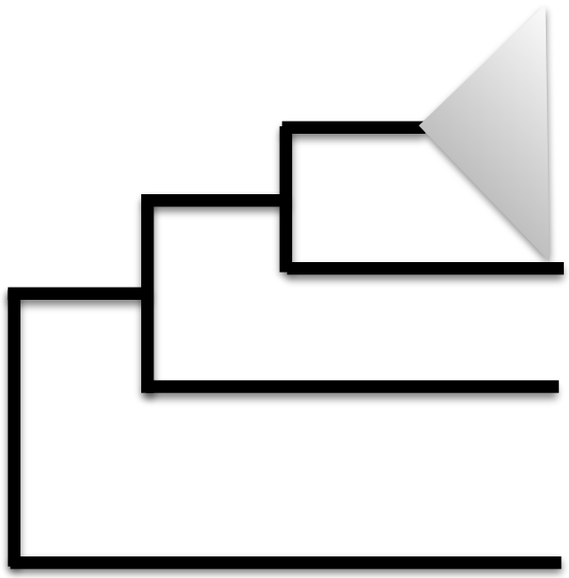
Taxon list from curated taxonomy

Query alignment database for taxa



Build gene trees and species trees

Hemiptera phylogeny



Questions?