



Phylogénie et délimitation d'espèces du complexe d'escargots terrestres *Pyramidula* par l'analyse de données RADseq

Sonet G¹, Razkin O², Breugelmans K¹, Madeira M J², Gómez-Moliner B J² & Backeljau T^{1,3}

¹Royal Belgian Institute of Natural Sciences, Brussels, Belgium.

²University of the Basque Country (UPV/EHU) & Biodiversity Research Group CIEA Lucio Lascaray, Alava, Spain.

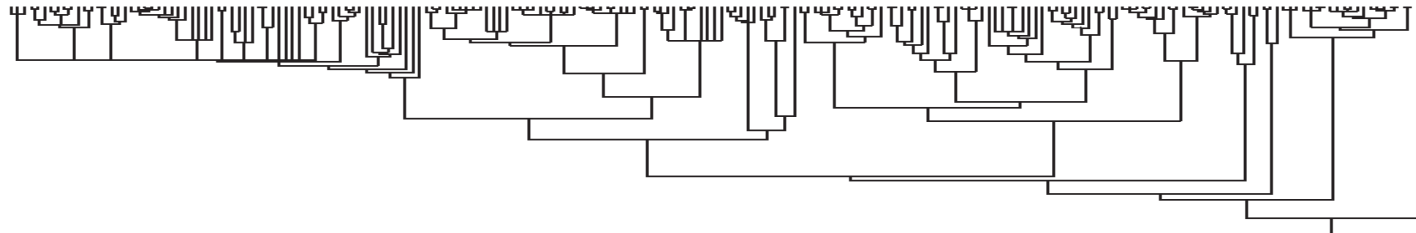
³University of Antwerp, Antwerp, Belgium.

Outline

- 1) Phylogeny: big data & missing data
- 2) Species delimitation: gene trees & species tree
- 3) Signs of reticulate evolution: historical hybridization

Case study:

Pyramidula Fitzinger, 1833
(Gastropoda: Pulmonata)



RADseq data for phylogeny: case of *Pyramidula*

Single species

Pyramidula rupestris
(Draparnaud, 1801)



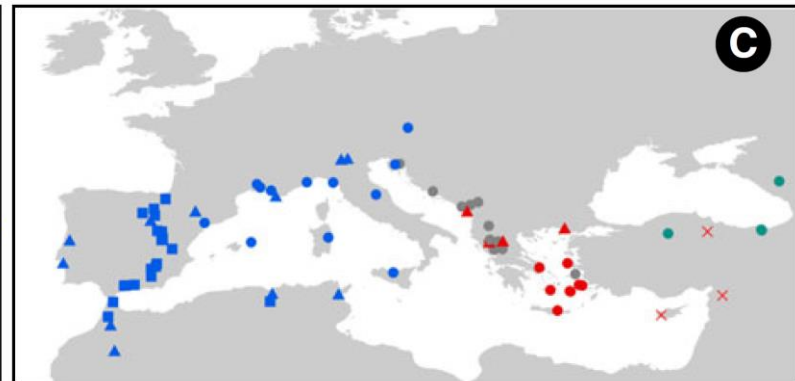
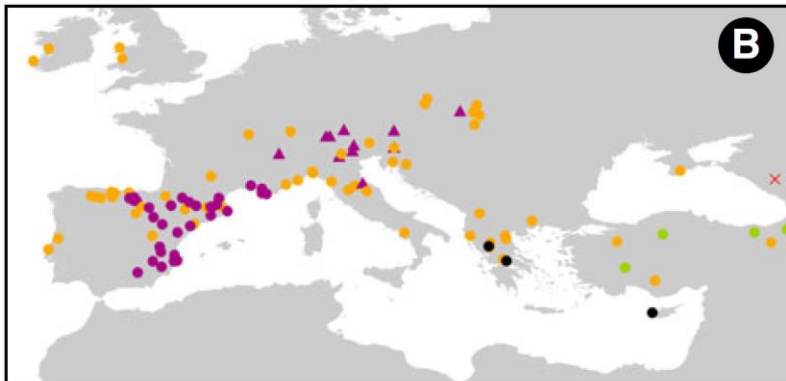
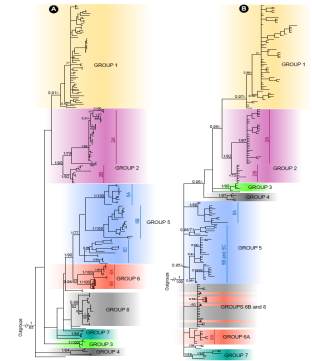
Six species

Gittenberger and Bank
(1996)



Nine putative species

Razkin et al. (in press)



Joint Experimental Molecular Unit

Africa
TERVUREN

JEMU

muséum

RADseq data for phylogeny

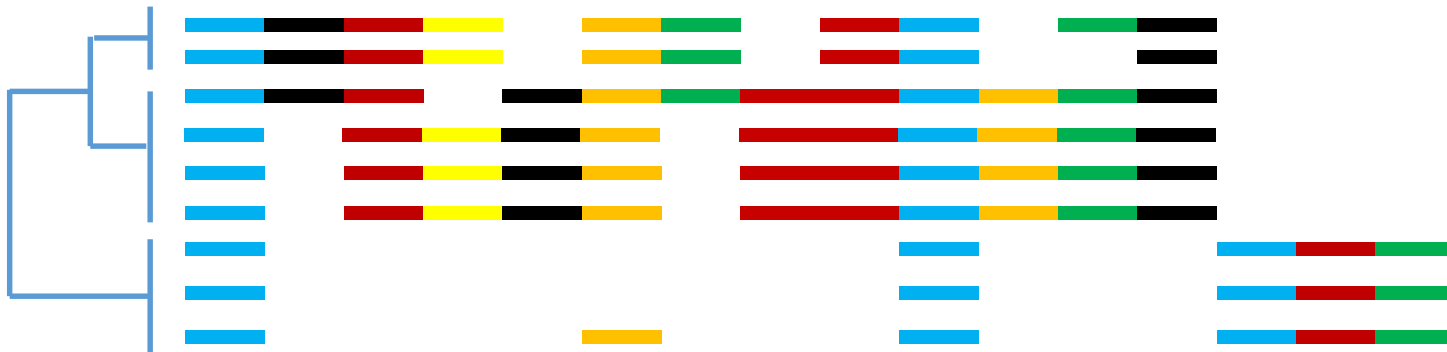
More data with RADseq...

Sanger: 211 individuals x 2 493 bp

RADseq: 25 individuals x ? hundreds of thousands of bp

But also more missing data...

Phylogeny (species complexes → family level)



1) Phylogeny:

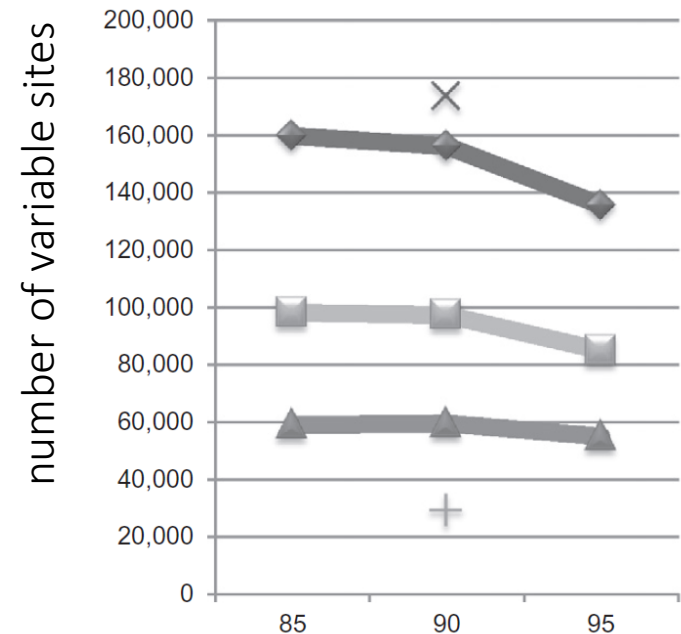
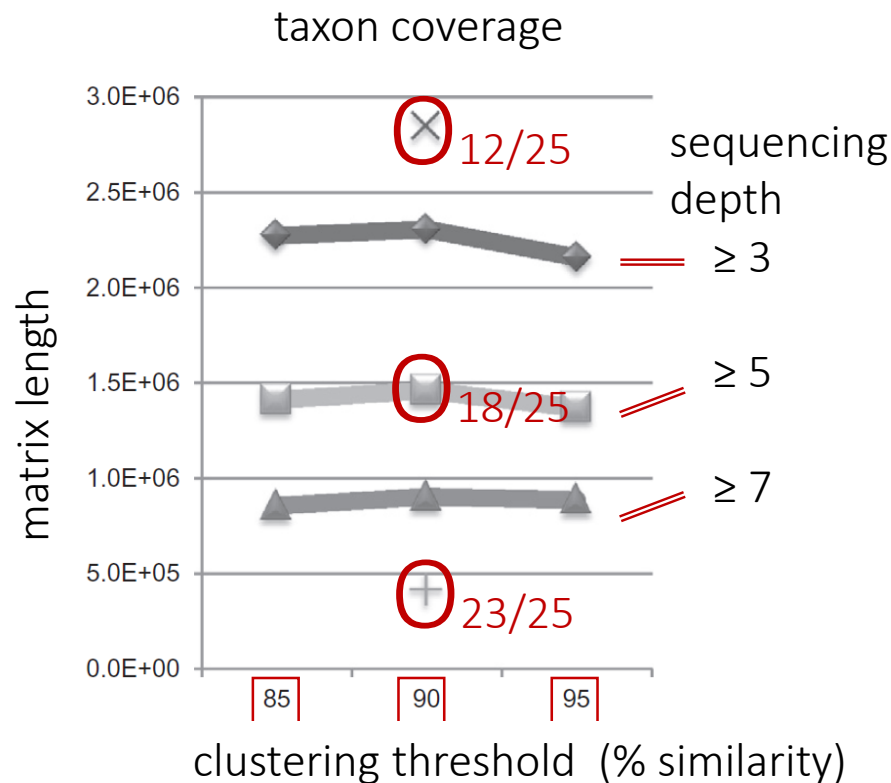
big data and missing data

RADseq data processing

25 ind., Sbfl, 2 MiSeq lanes (2x250 bp) -> 100-900 k reads per sample

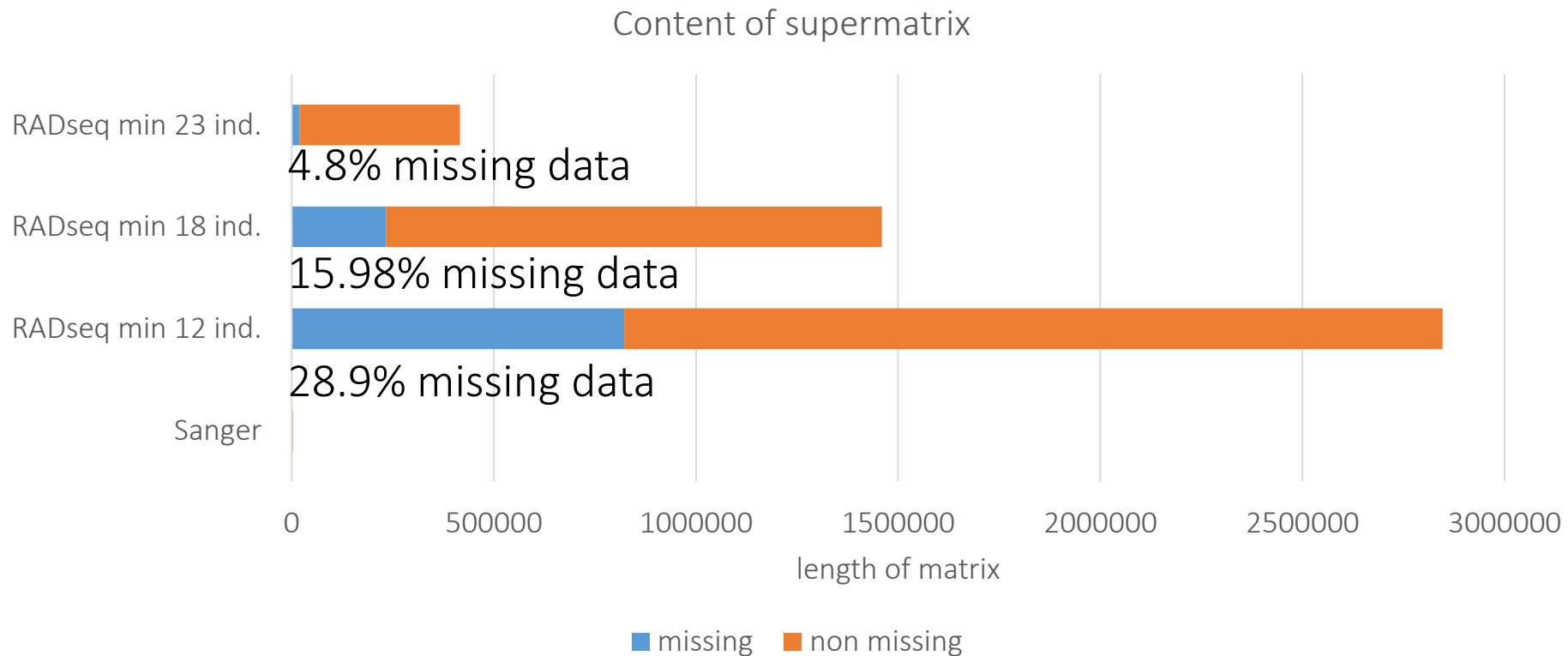
De novo pyRAD pipeline (Eaton and Ree 2013. Syst. Biol. 62: 689–706)

11 data sets obtained using different parameters:



Razkin et al. 2016. M.P. E. 101: 267–278

Big data and missing data



Phylogenomics

11 supermatrices
(concatenated RADseq data)

Lengths:

415 737 - 2 846 408

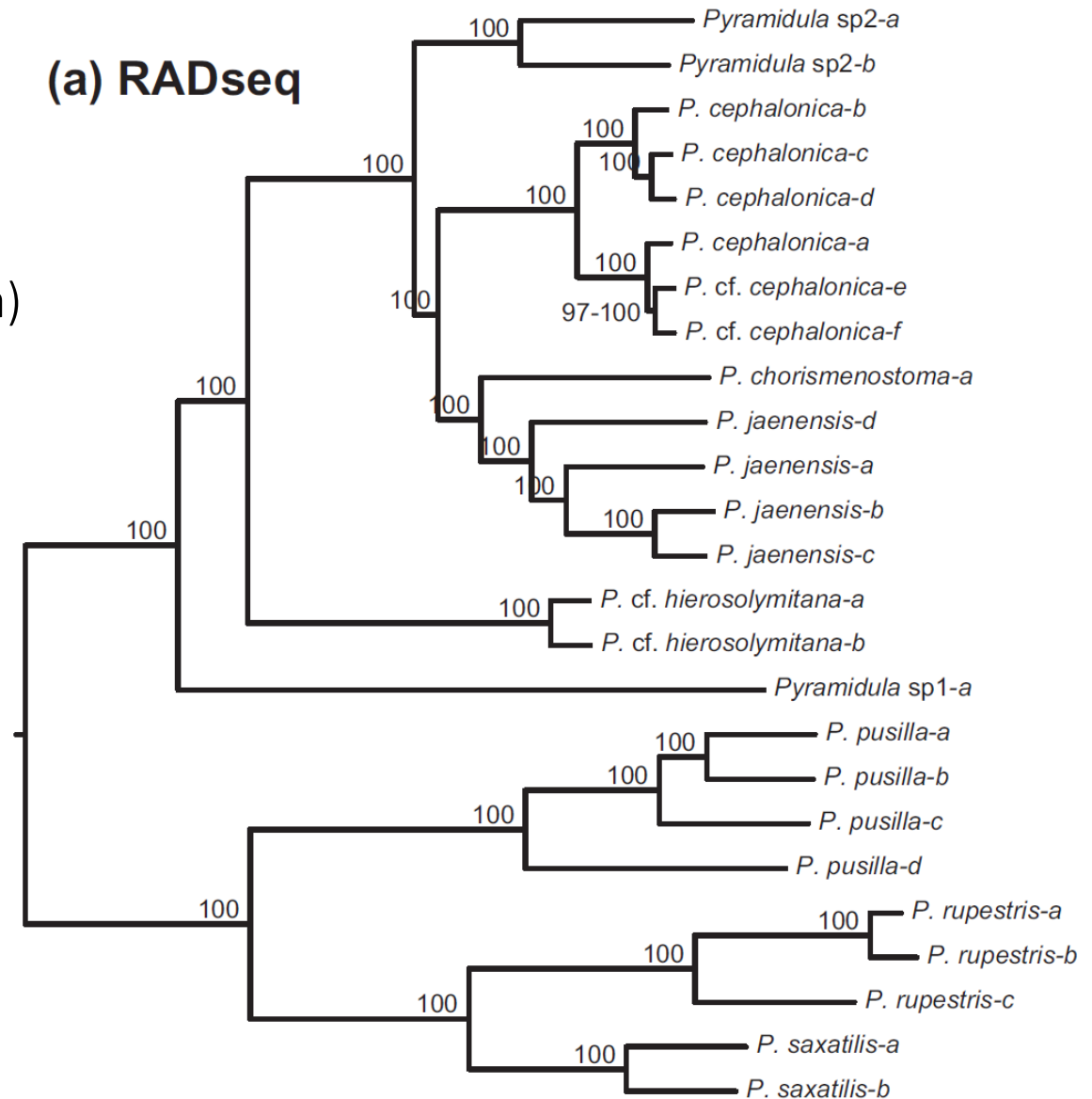
Polymorphic sites:

29 179 - 173 440

Maximum likelihood

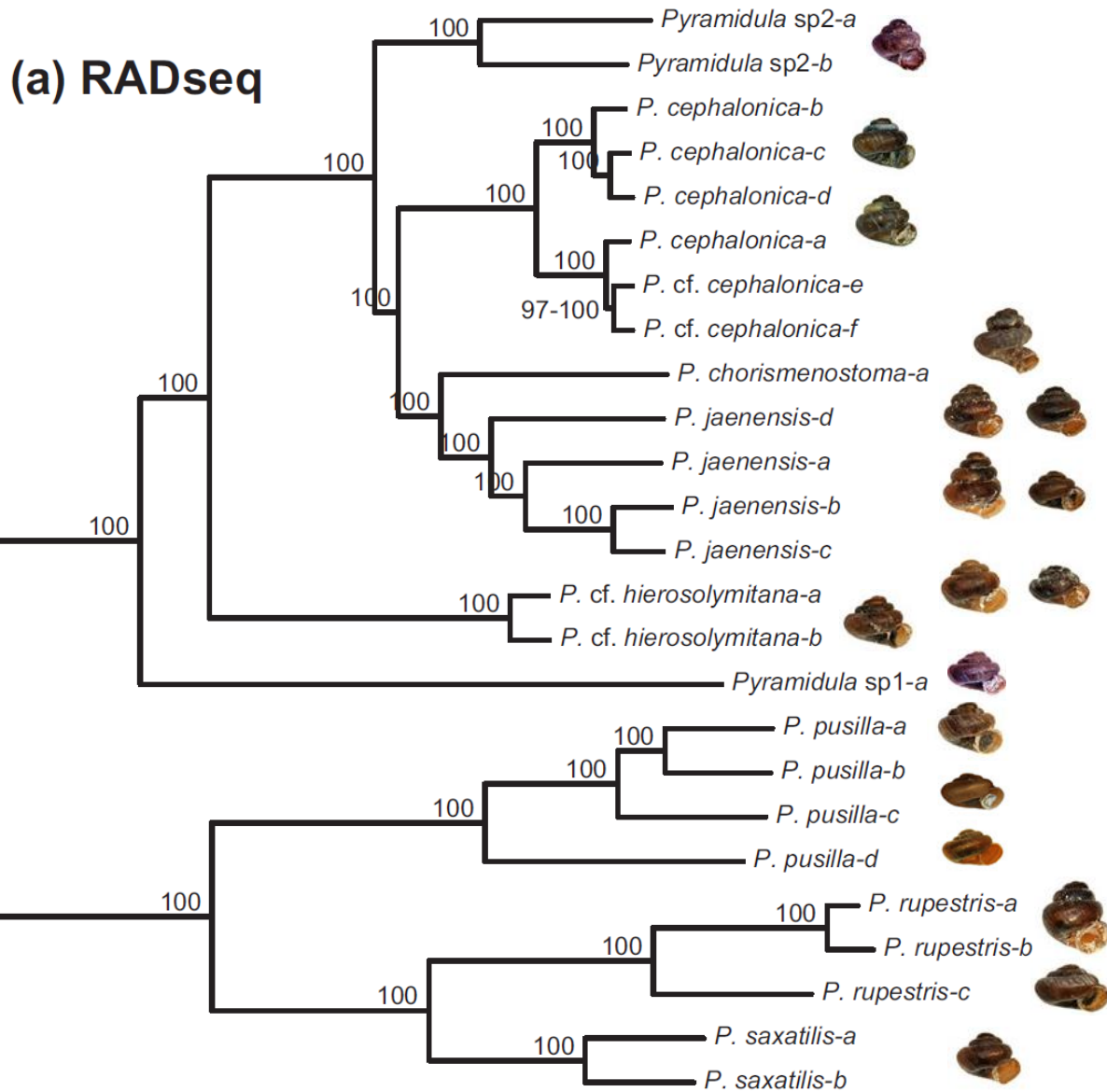
(RAxML, GTR model, 100
bootstrap replicates)

(a) RADseq



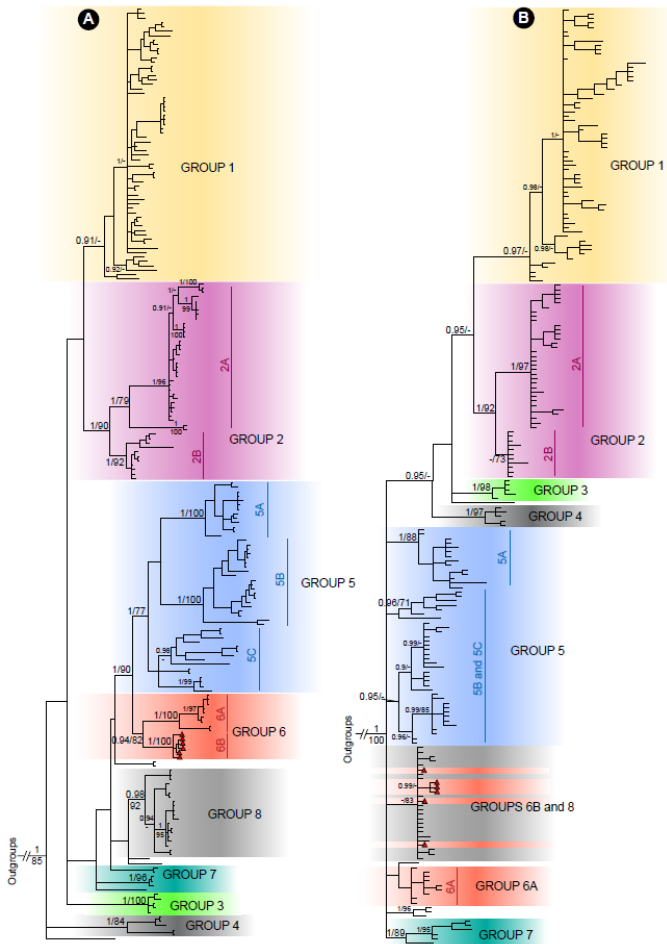
Razkin et al. 2016. M.P. E. 101: 267–278

Phylogenomics

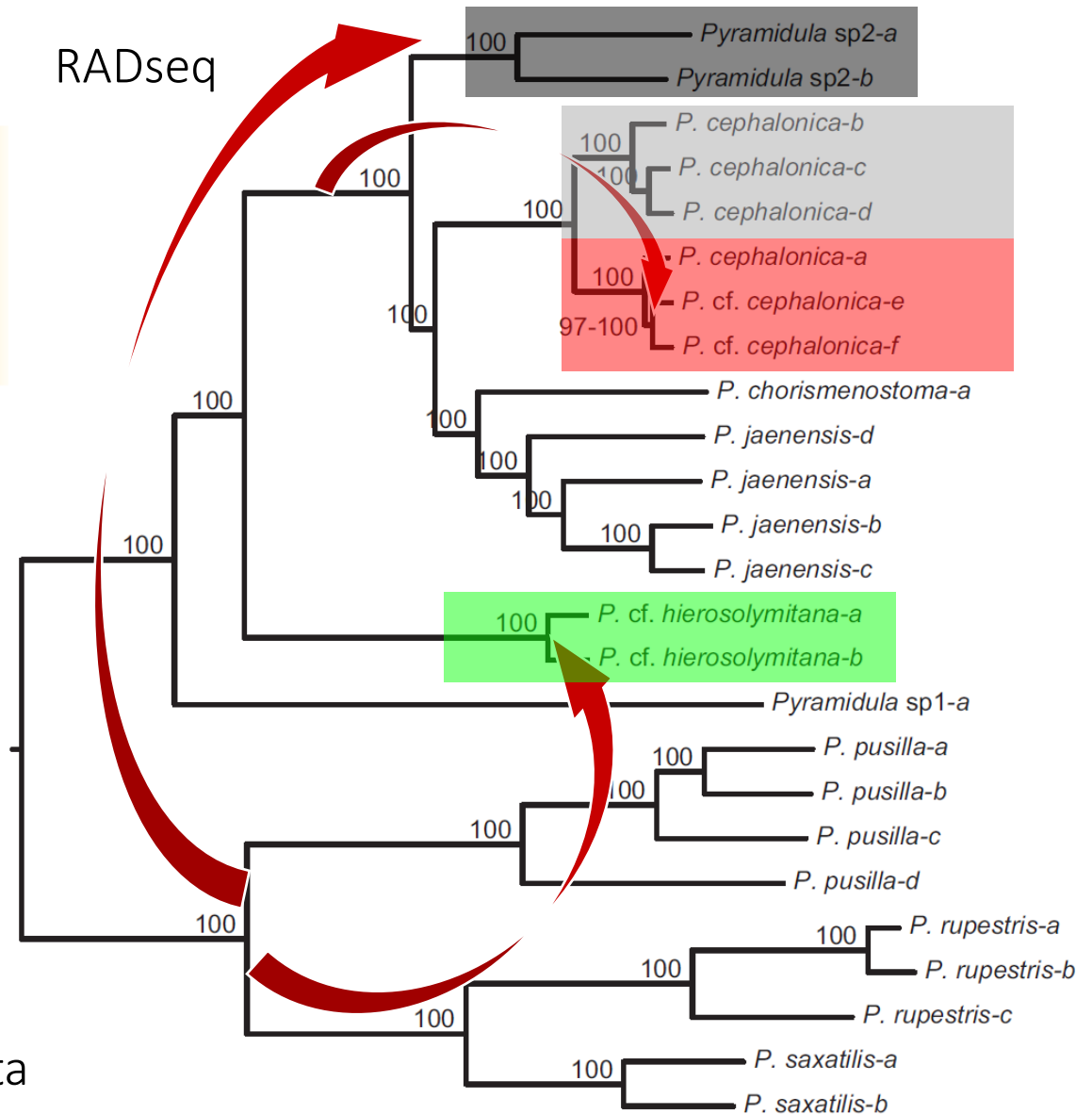


Razkin et al. 2016. M.P. E. 101: 267–278

Phylogenomics



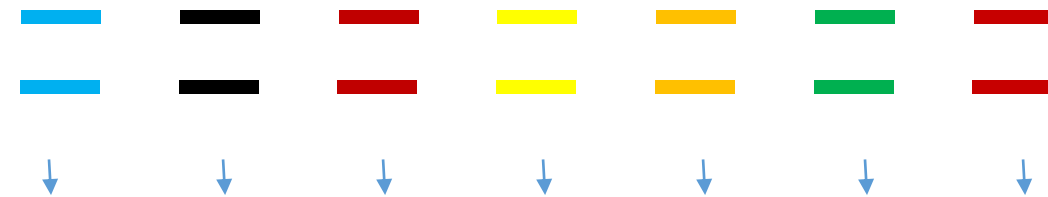
mitochondrial nuclear data



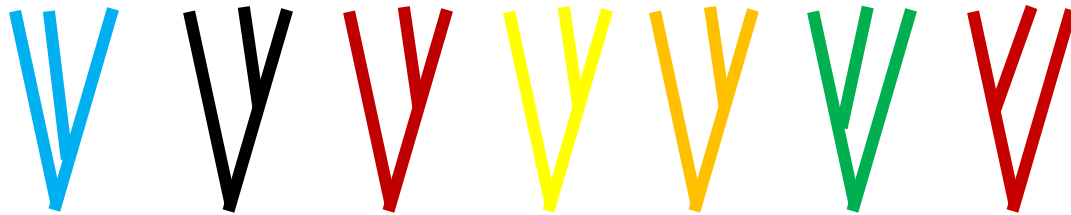
2) Species delimitation:
gene trees & species tree

Species tree

Multilocus dataset with independently segregating genes



Individual gene trees



Species tree



Species delimitation



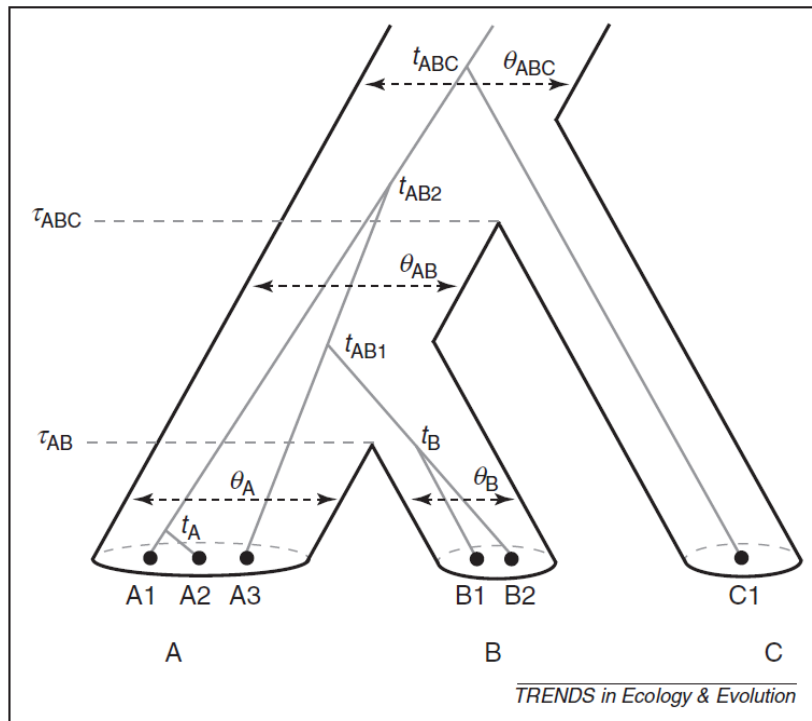
Edwards. 2009. *Evolution* 63(1):1-19
Fujita et al. 2012. *Trends Ecol. Evol.* 27:
480-488)

Species delimitation

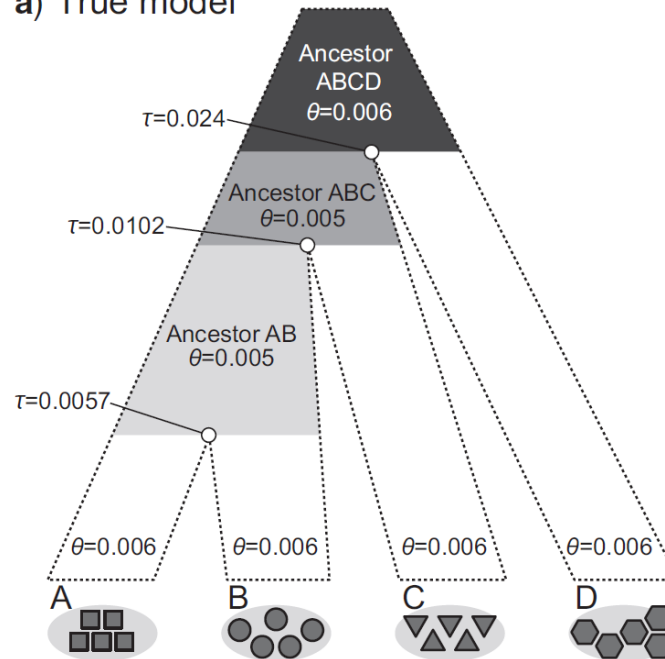
Bayes factor delimitation of species with genomic data (Leaché et al. 2014. Syst. Biol. 63(4): 534-542)

Species trees without gene trees (Bryant et al. 2012. Mol Biol Evol. 29(8): 1917-1932): SNAPP

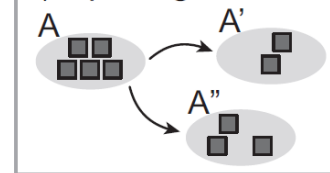
Bayes factors for species delimitation (Grummer et al. 2014. Syst. Biol. 63(2): 119-133)



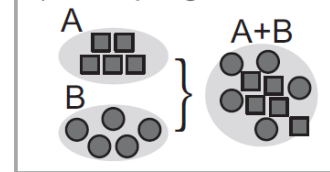
a) True model



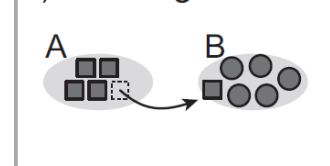
b) Splitting



c) Lumping



d) Reassignment

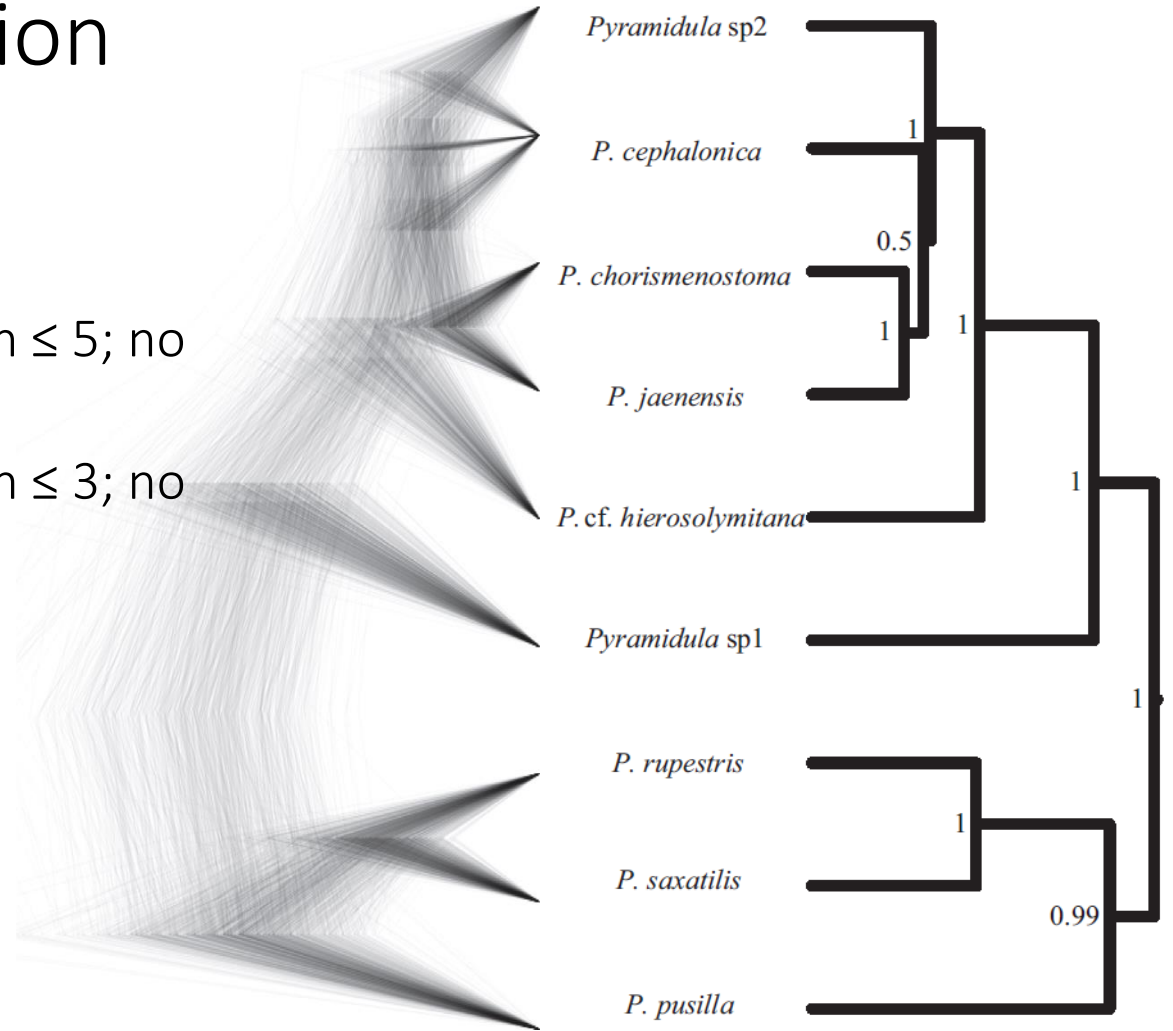


Species delimitation

SNAPP < BEAST 2

368 unlinked SNPs (seq. depth ≤ 5 ; no missing data)

875 unlinked SNPs (seq. depth ≤ 3 ; no missing data)

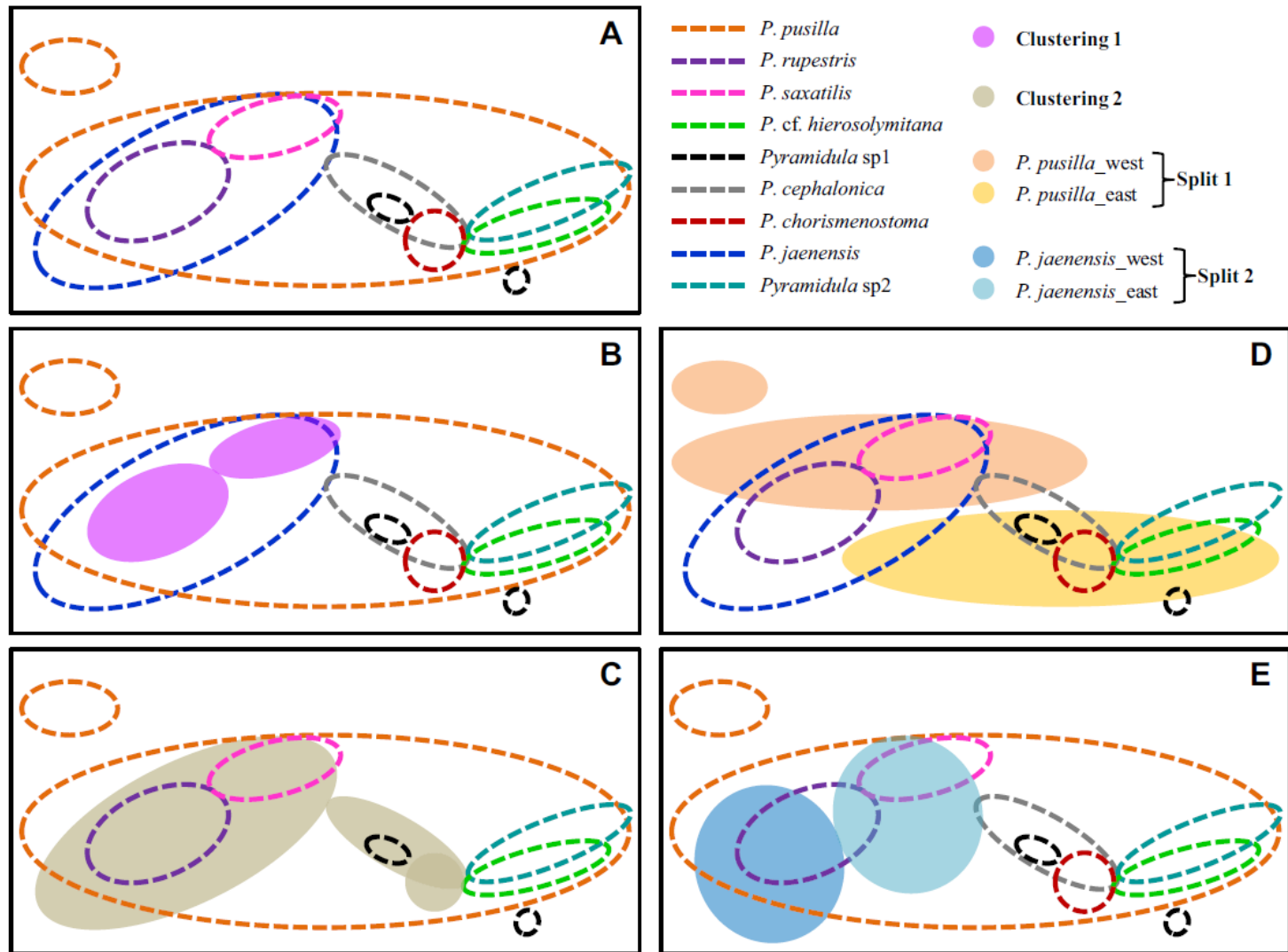


Density diagram of all
species trees

Consensus topology

Razkin et al. 2016. M.P. E. 101: 267–278

Species delimitation



Razkin et al.
2016. M.P. E. 101:
267–278

3) Signs of reticulate evolution:
historical hybridization

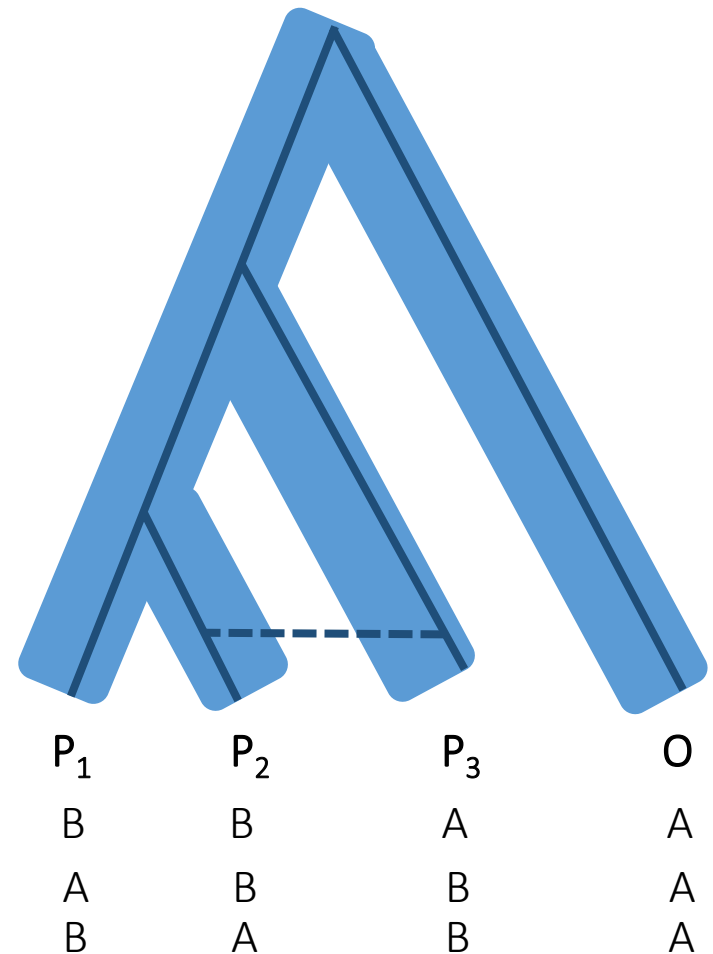
Traces of historical hybridization

Patterns of linkage disequilibrium

Spatially explicit simulations

Historical samples (ancient DNA)

Asymmetry in the occurrence of two
incongruent allele patterns (Green et al.
2010. Science 328: 710-722; Durand et al. 2011.
Mol. Biol. Evol. 28: 2239-2252)

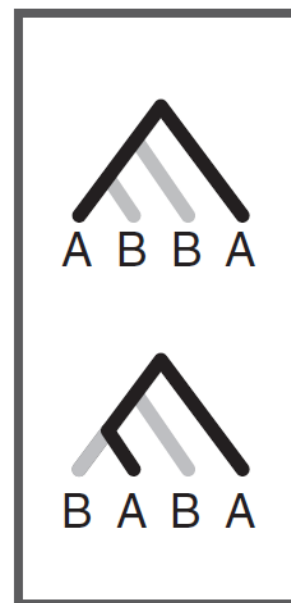
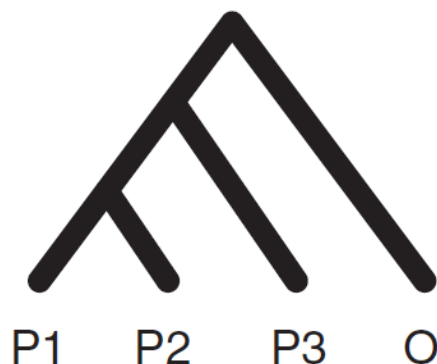


Traces of historical hybridization

H_0 : No horizontal gene transfer

Assumptions:

- One substitution $\leftarrow$ one mutation in branch
- Random mating
- Instantaneous unidirectional admixture model



measure D

Patterson's four-taxon D-statistic test (Green et al. 2010. Science 328: 710-722; Durand et al. 2011. Mol. Biol. Evol. 28: 2239-2252)

Test 1: ABBA
 8/24 ($\alpha = 0.05$)
 1/24 ($\alpha = 0.01$)

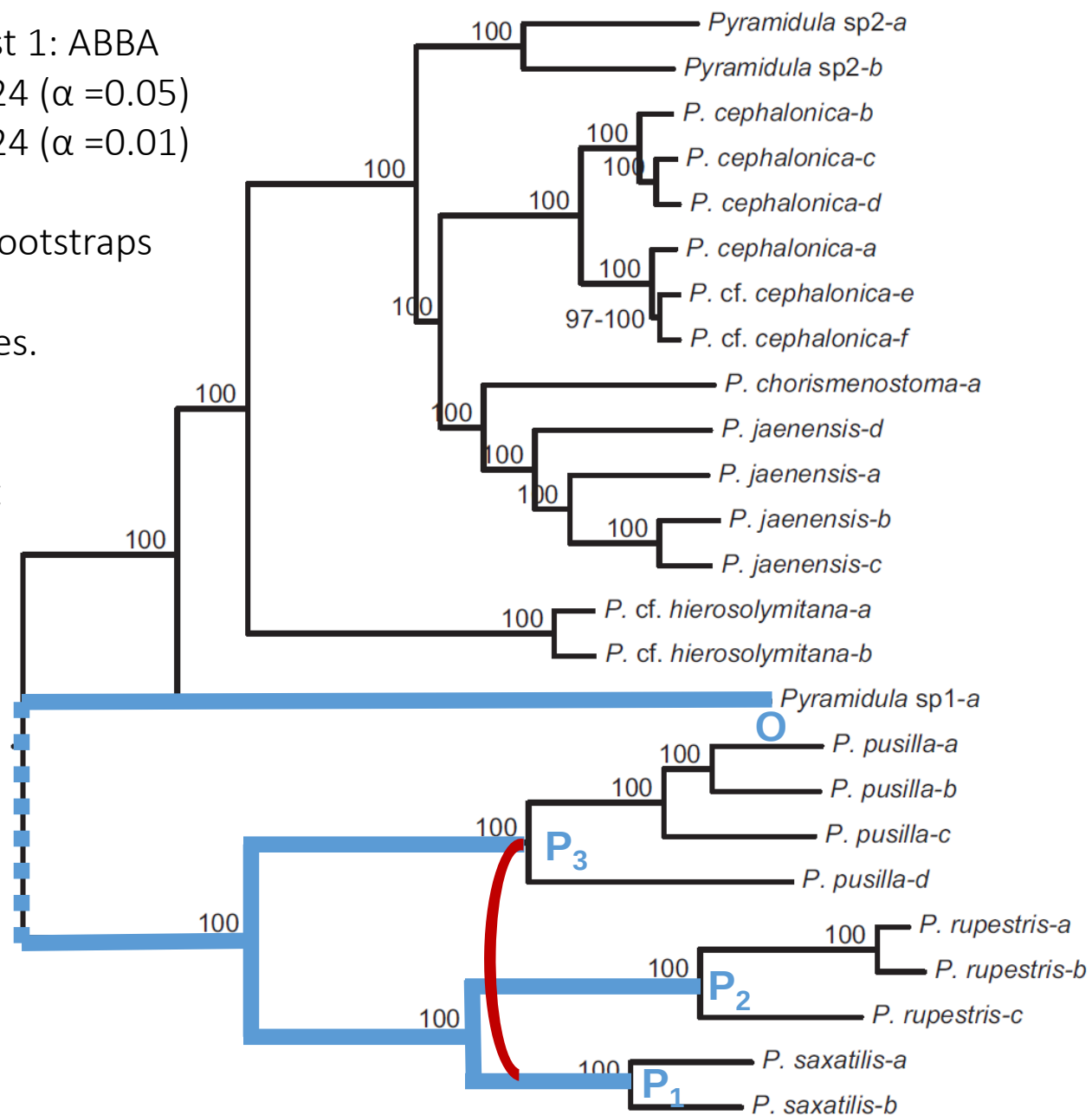
PyRAD with 1000 Bootstraps

8 tests for 8 subtrees.

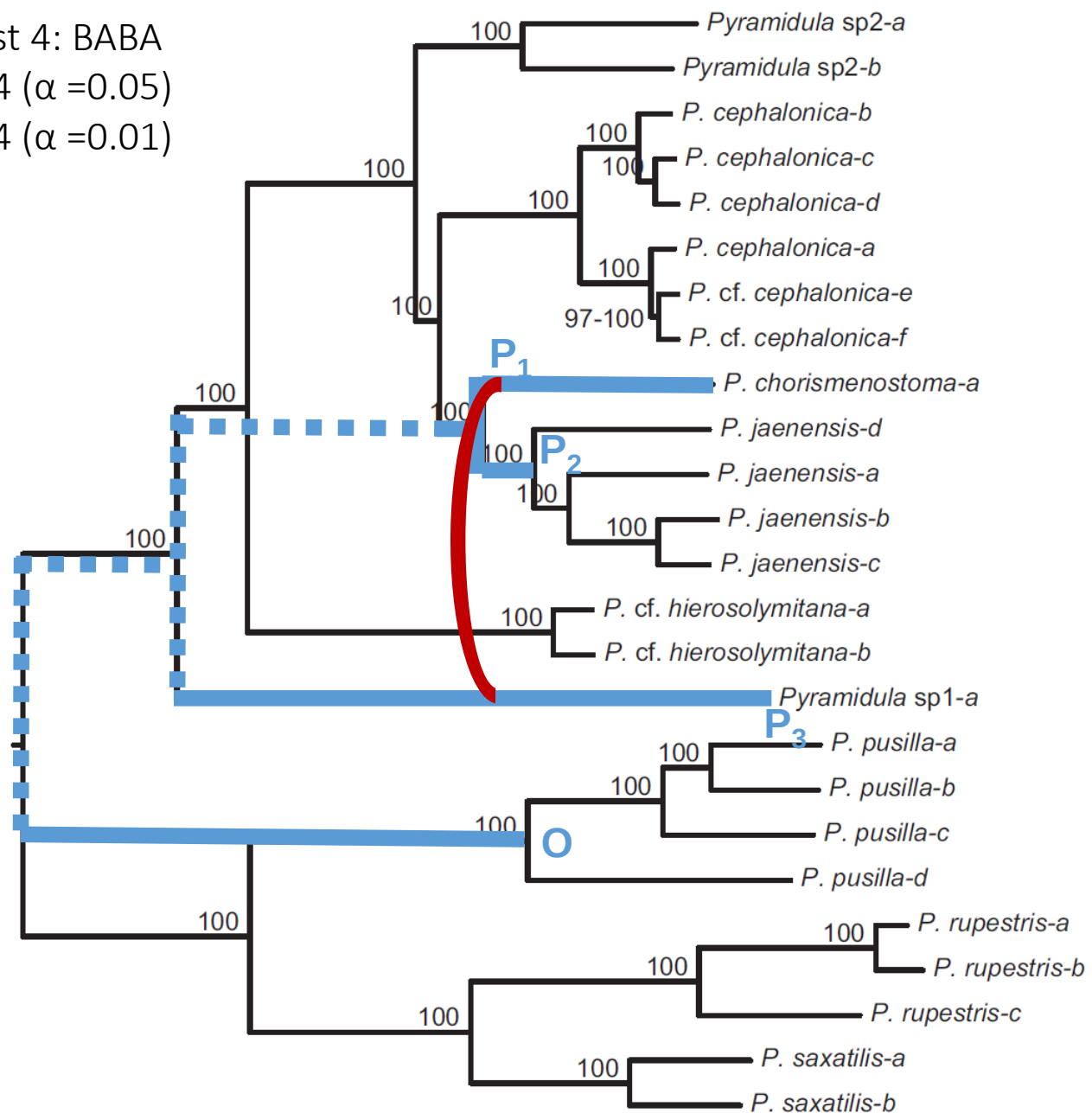
In each test:

1302-3040 SNPs

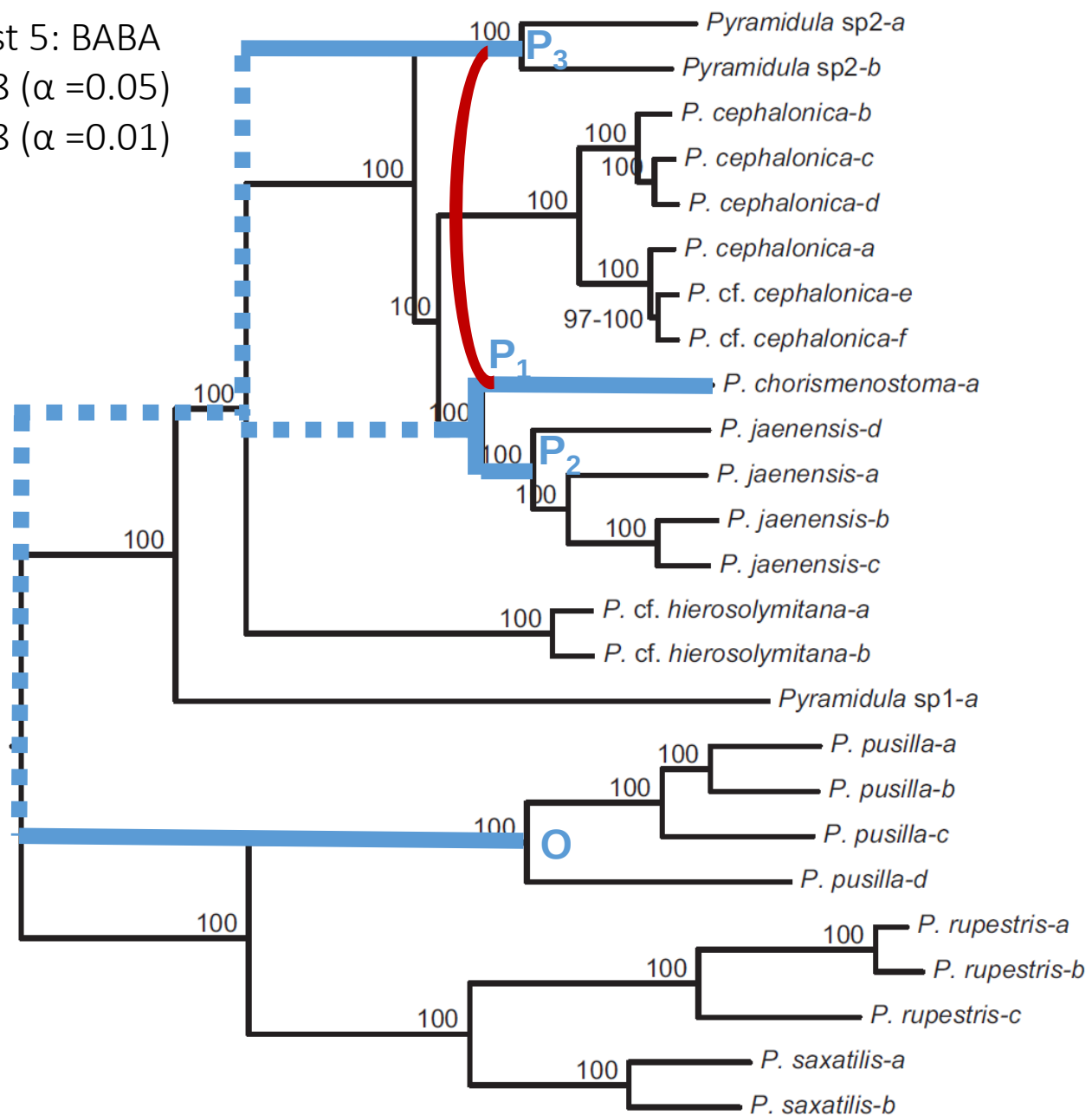
0.07-1.6 discordant



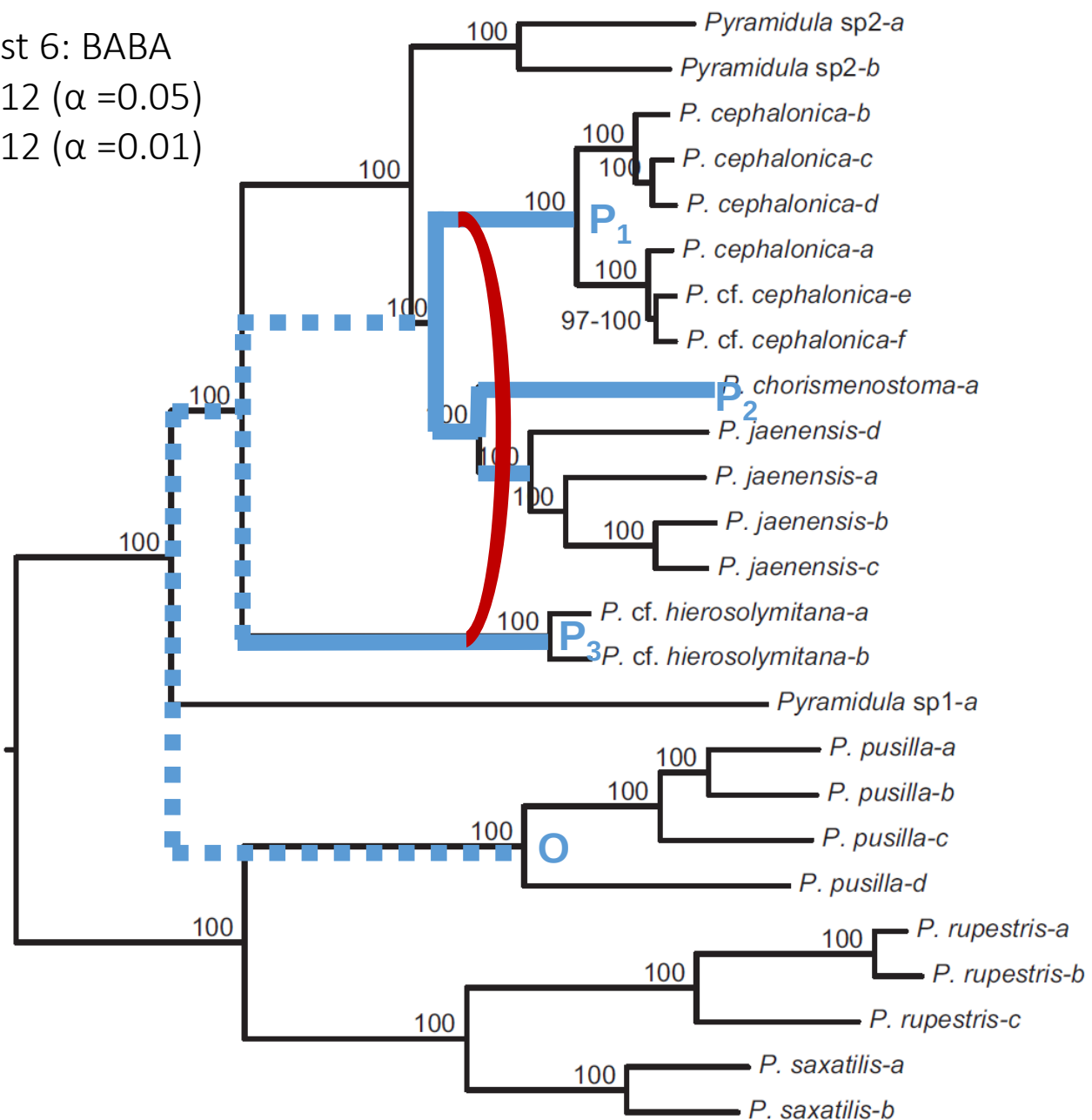
Test 4: BABA
 2/4 ($\alpha = 0.05$)
 0/4 ($\alpha = 0.01$)



Test 5: BABA
 2/8 ($\alpha = 0.05$)
 1/8 ($\alpha = 0.01$)



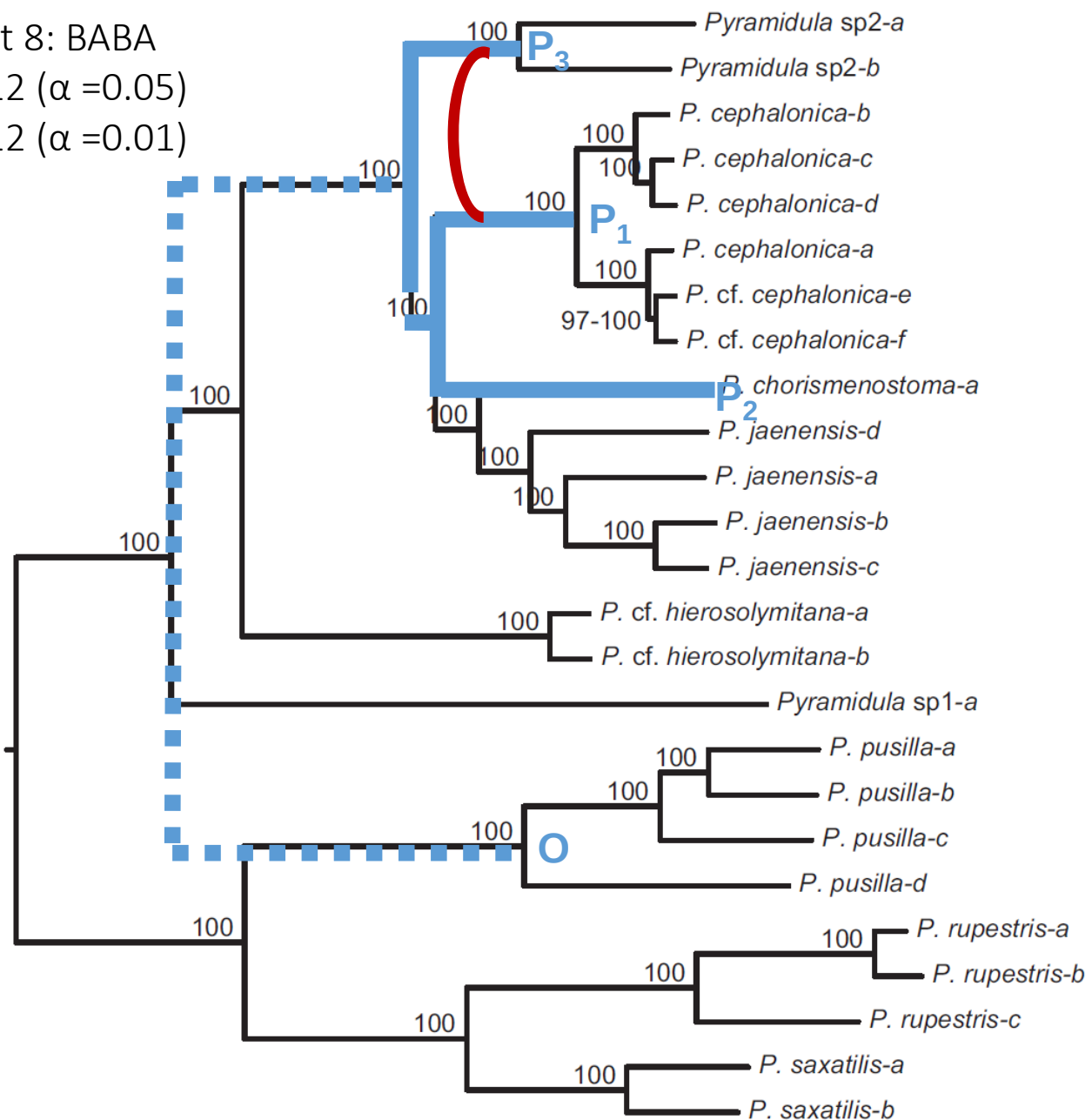
Test 6: BABA
 6/12 ($\alpha = 0.05$)
 1/12 ($\alpha = 0.01$)



Test 8: BABA

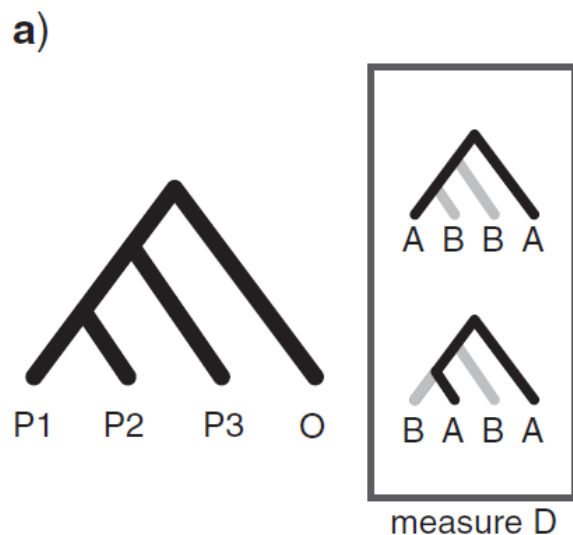
9/12 ($\alpha = 0.05$)

1/12 ($\alpha = 0.01$)

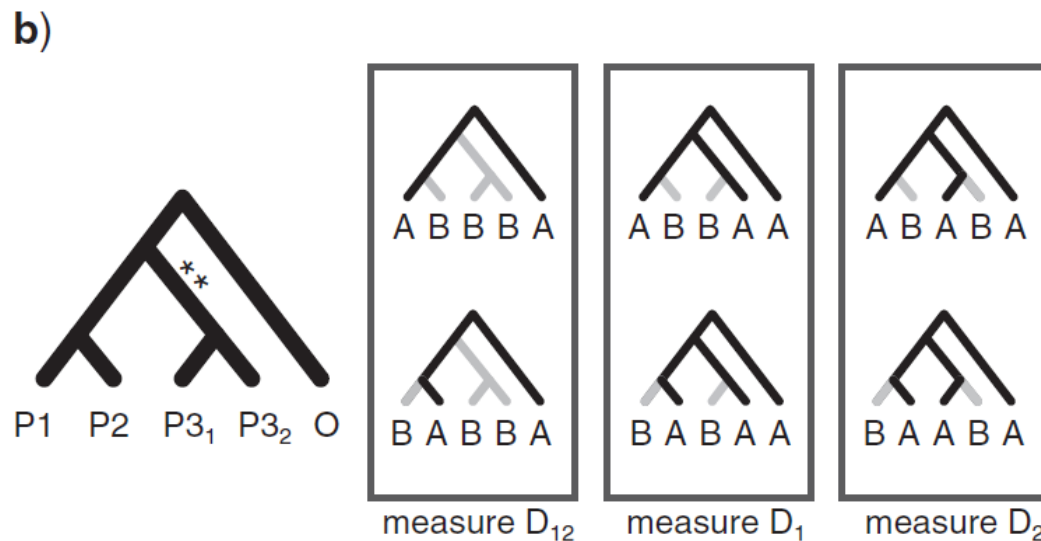


Traces of historical hybridization

a) Patterson's four-taxon D-statistic test



b) Five-taxon partitioned D-statistic test



Green et al. 2010. Science 328:
710-722; Durand et al. 2011. Mol.
Biol. Evol. 28: 2239-2252

Eaton and Ree 2013. Syst. Biol. 62: 689-706

Traces of historical hybridization

Five-taxon partitioned D-statistic test

Performed on PyRAD with 1000 Bootstraps
4 tests for 4 subtrees.

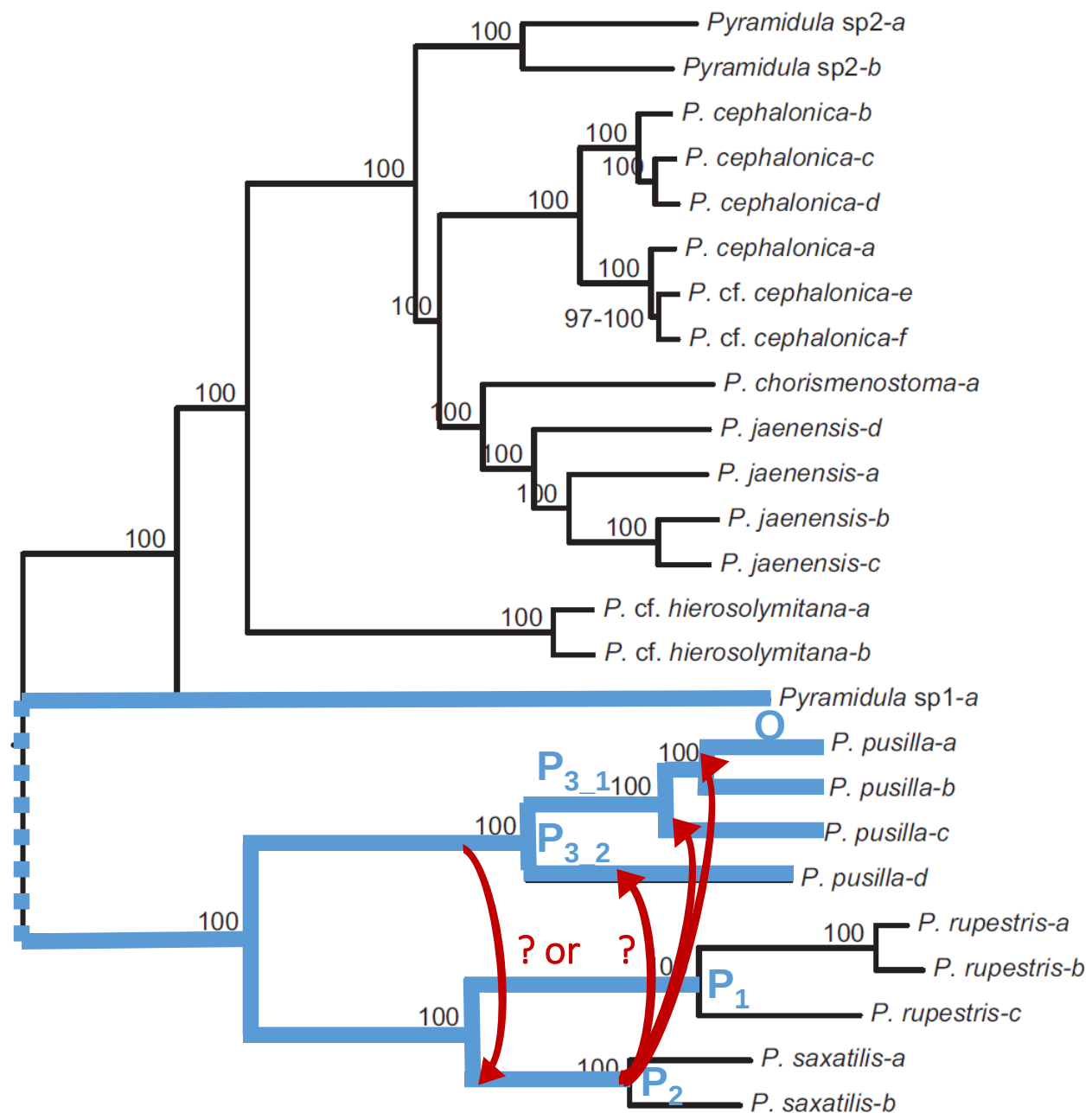
In each test:

1757-2551 SNPs

0.08-0.18% non concordant

→ >100 SNPs for D_1 but ca. 60 SNPs for D_1 or D_2

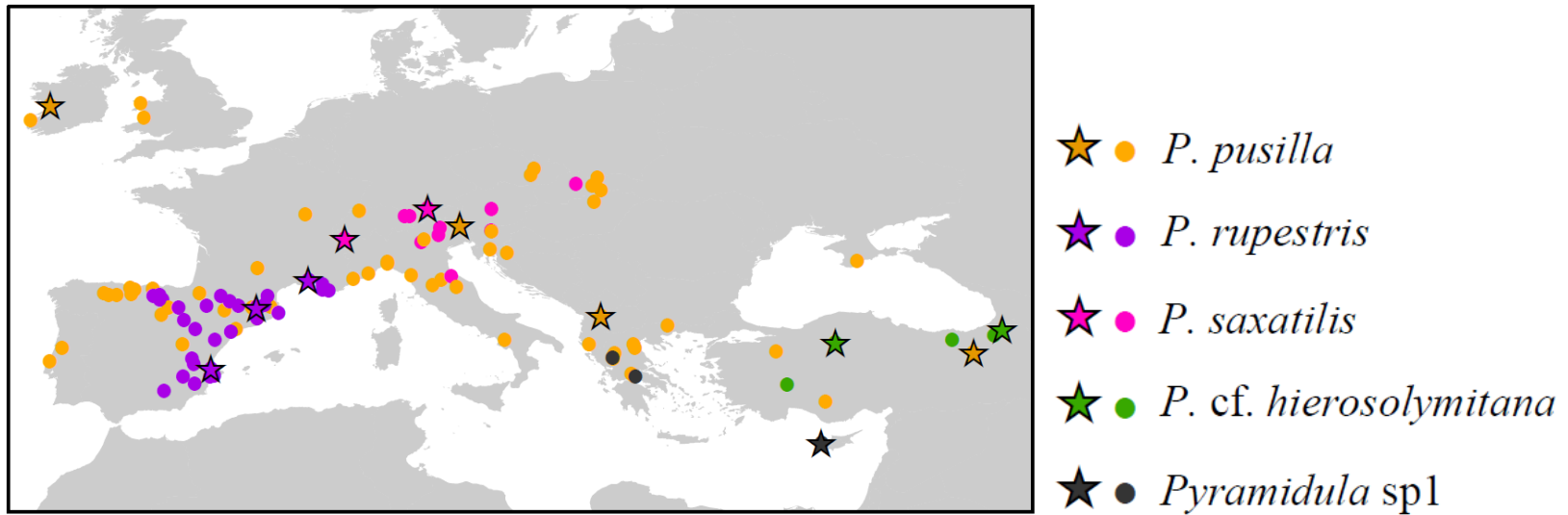
4/36 ($\alpha = 0.01$)

 $2/36 (\alpha = 0.01)$
$$2/36 (\alpha = 0.01)$$


Traces of historical hybridization

Additional observations supporting possible historical gene flow between *P. pusilla* & *P. saxatilis*:

- Overlapping distribution,
- Similar ecological niche
- Apparent longer diversification process in this clade



Razkin et al. 2016. M.P. E. 101: 267–278

Conclusion

RADseq data allowed to investigate a non-model complex of species using three complementary approaches to provide:

- a robust phylogeny,
- confirm a species delimitation hypothesis and
- test for interspecific hybridization

Thank you

Acknowledgements:

Organizers of this RADseq day



Belgian Science Policy (BELSPO)



JEMU and the Molecular DNA labs of RBINS and RMCA



Belgian Network for DNA barcoding and FWO



Department of Education of the Government of Navarra



Genomics Core, KULeuven



References (1/2)

Pyramidula taxonomy

Gittenberger & Bank. 1996. Basteria 60: 71–78

Razkin et al. 2016. Molecular Phylogenetics and Evolution 101: 267–278

Razkin et al. In press. Zoologica Scripta.

Species tree

de Oliveira Martins & Posada. 2015. doi: 10.1101/023861

Edwards. 2009. Evolution 63(1):1–19

Heled & Drummond. 2010. Mol. Biol. Evol. 27(3):570–580. 201

Species delimitation

Bryant et al. 2012. Mol Biol Evol 29(8): 1917–1932 →SNAPP MCMC FSM/ISM no branch mut.

Fujita et al. 2012. Trends Ecol Evol 27: 480–488

Grummer et al. 2014. Syst. Biol. 63(2): 119–133

Leaché et al. 2014. Syst Biol 63(4): 534–542

References (2/2)

Phylogeny from concatenated RADseq data

Cariou et al. 2013. Ecology and Evolution 3(4): 846–852

Viricel et al. 2014. Molecular Ecology Resources 14: 597–605

Wagner et al. 2013. Molecular Ecology 22: 787–798

Interspecific hybridation

Green et al. 2010. Science 328: 710–722

Durand et al. 2011. Mol. Biol. Evol. 28: 2239–2252

Eaton and Ree 2013. Syst. Biol. 62: 689–706