

DNA-based species delimitation

- Phylogenetic species concept based on tree topologies

∅ How to set species bound

∅ **Automatic species**

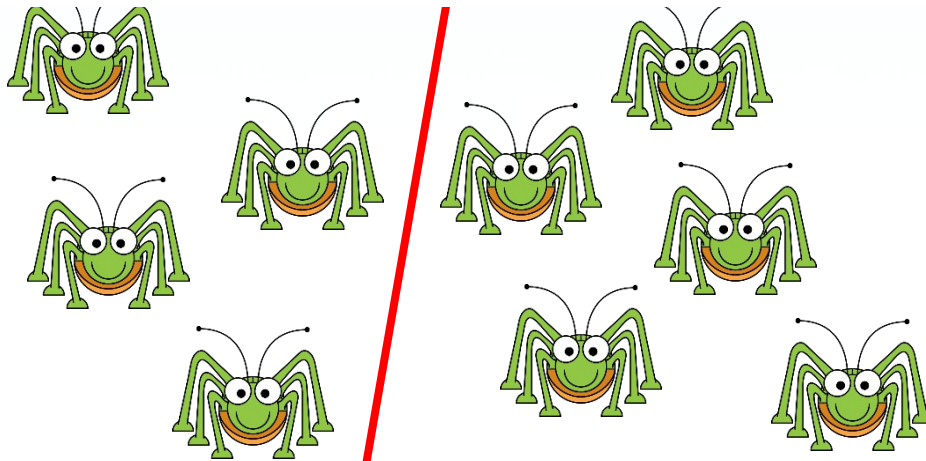
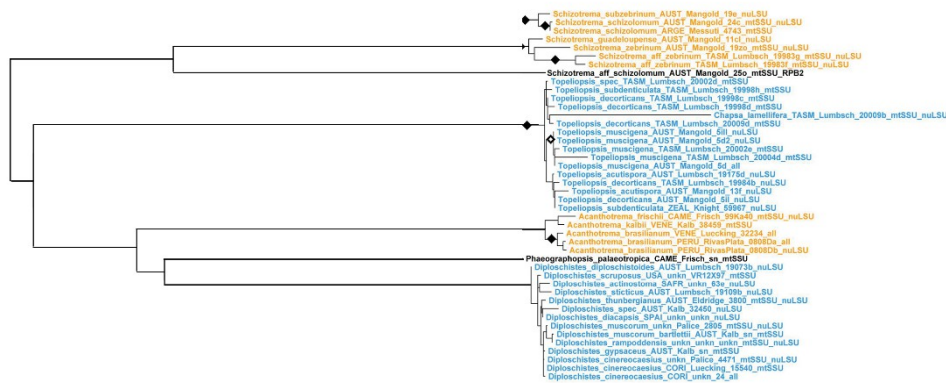
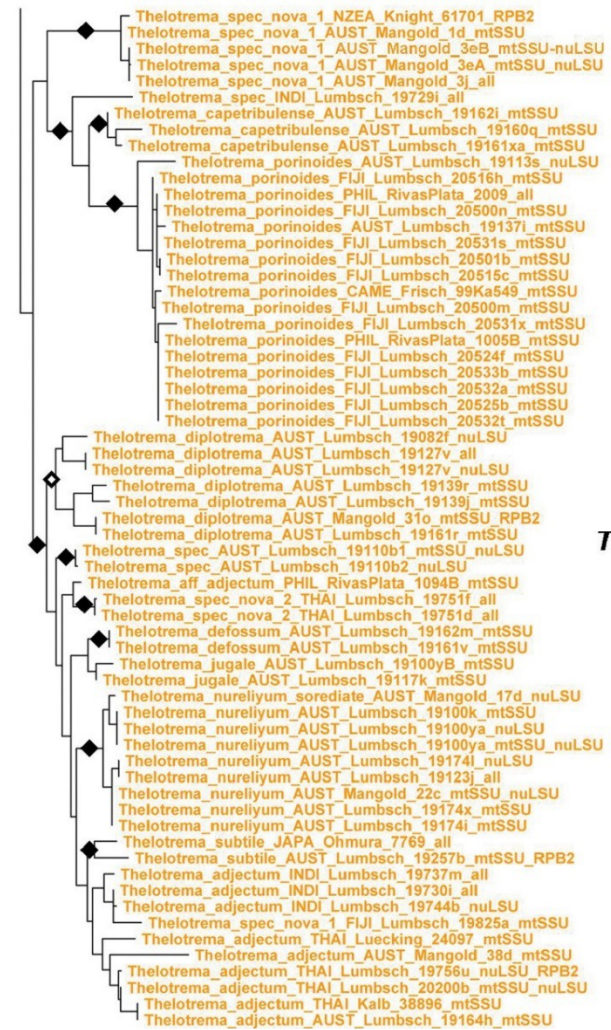


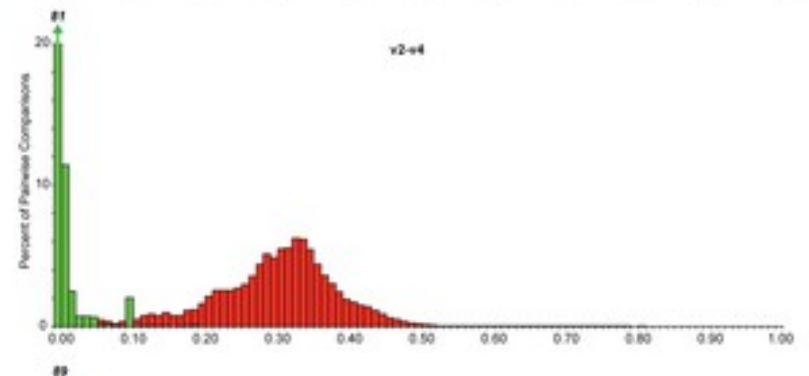
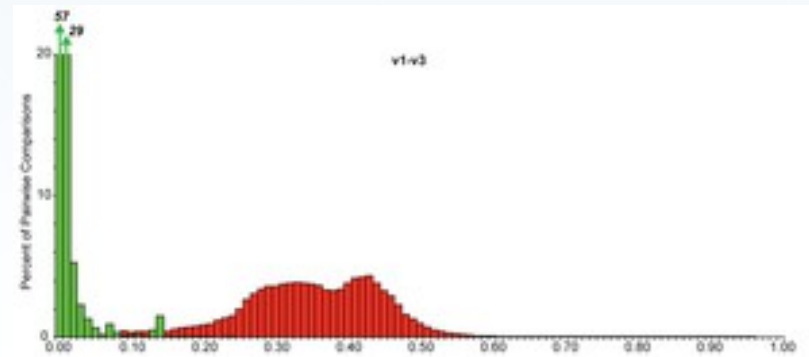
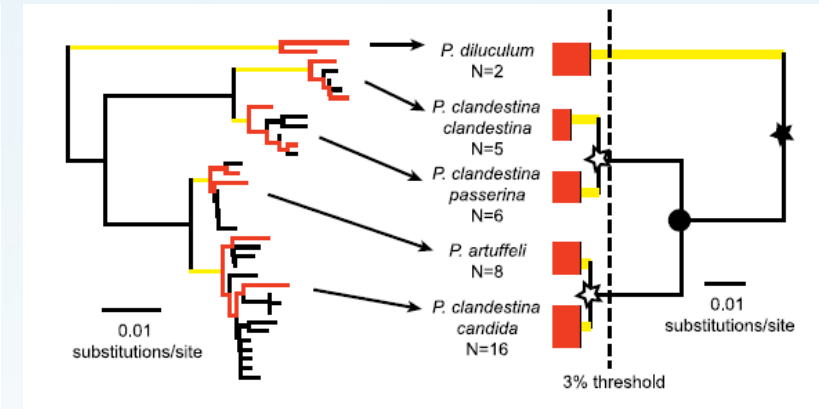
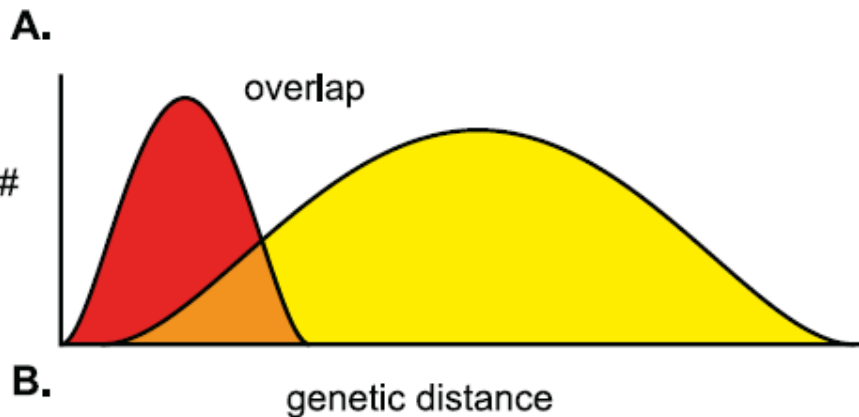
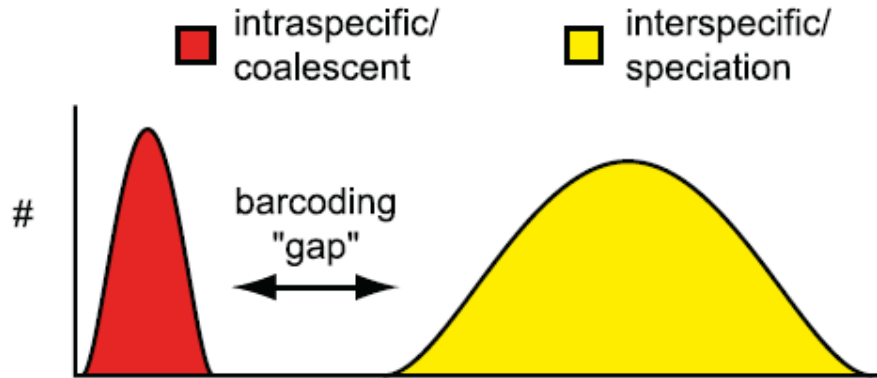
Fig. 8



Thelotrema

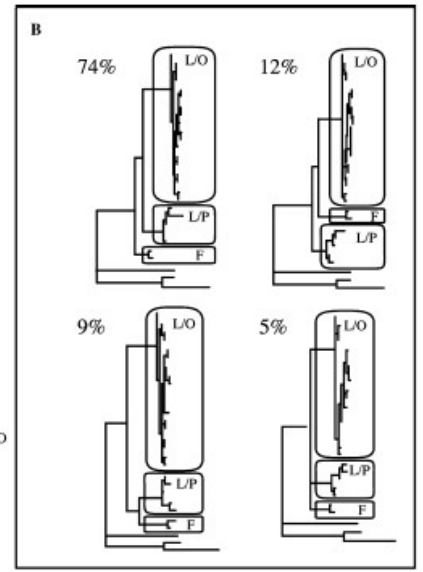
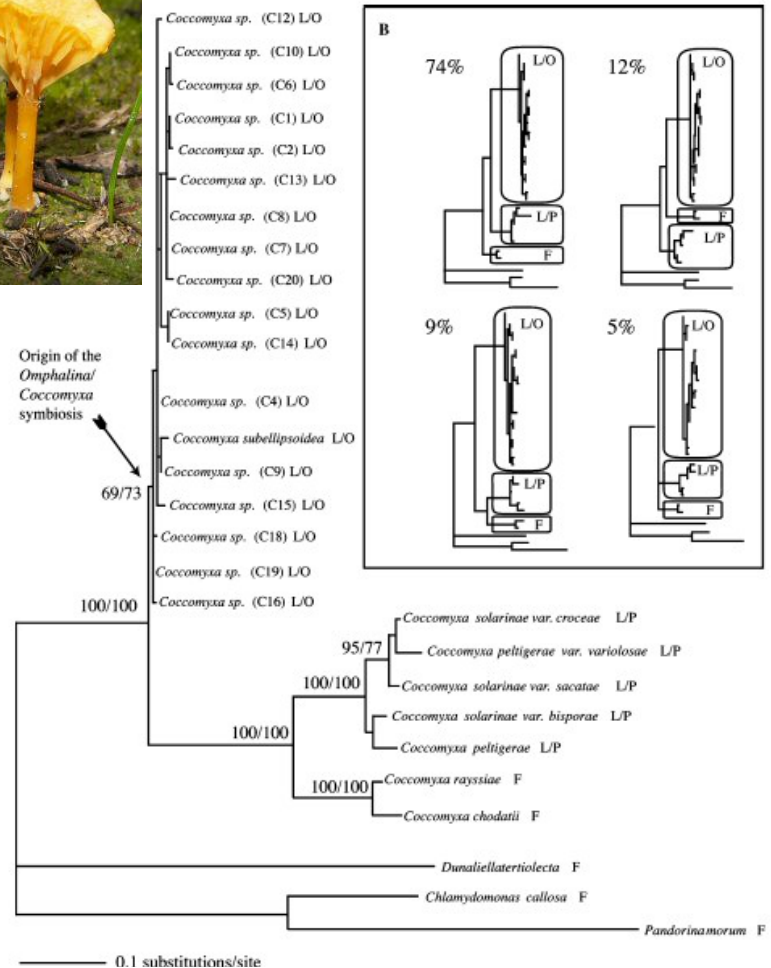
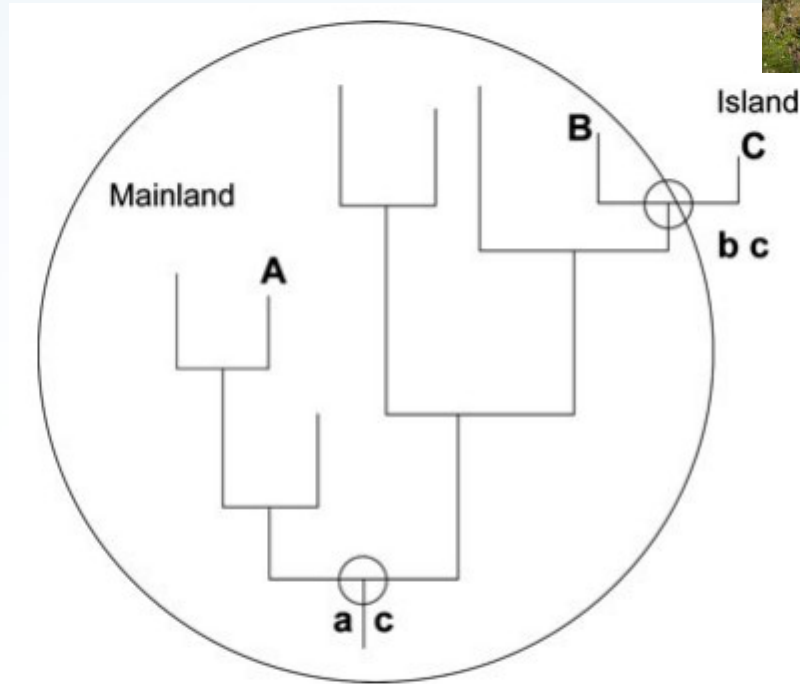
DNA barcoding

- Species boundaries recognized by identifying



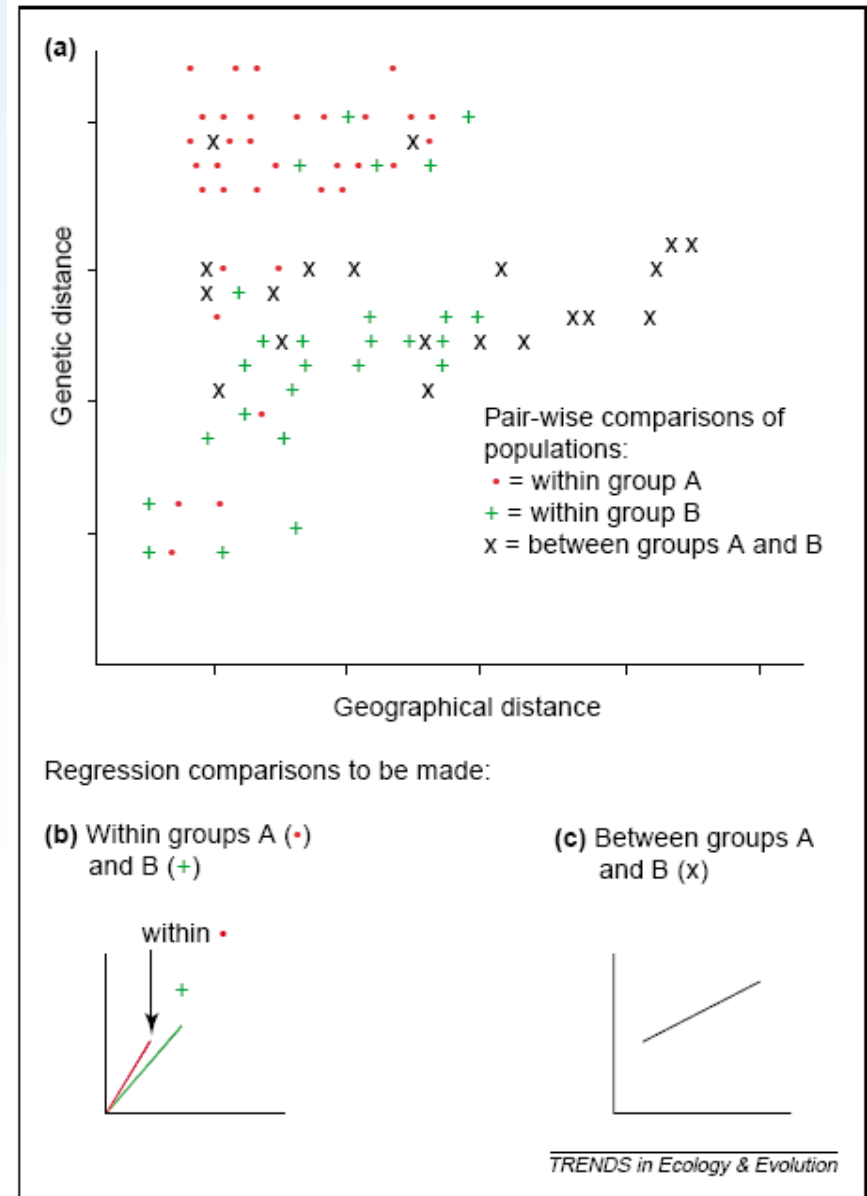
DNA barcoding

- Uneven mutation rates
 - ∅ tropics vs temperate
 - ∅ islands vs continents
 - ∅ free-living vs symbio



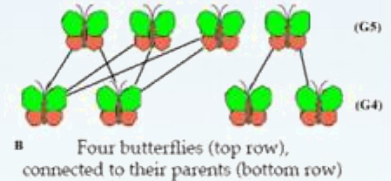
Good & Wake method

- test of *a priori* defined species
- linear regression of genetic and geographic distances should in single species go through the graph origin (gene flow with isolation-by-distance)
- Different regression indicates the presence of two distinct, genetically isolated species



Coalescence-based methods

- Linking phylogenetics and population genetics
 - ∅ Identification of independently evolving lineages

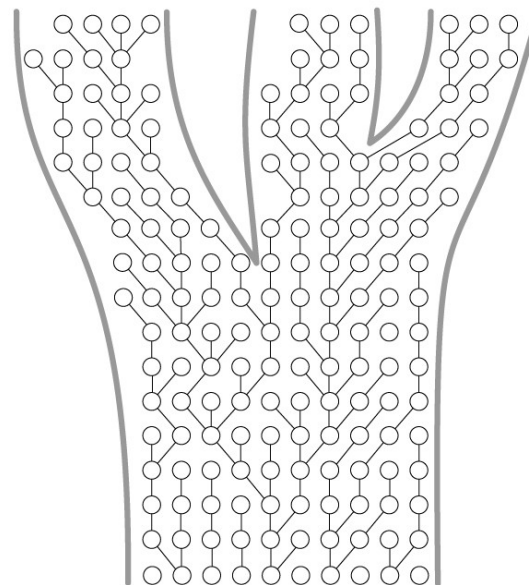


A **B** **C**

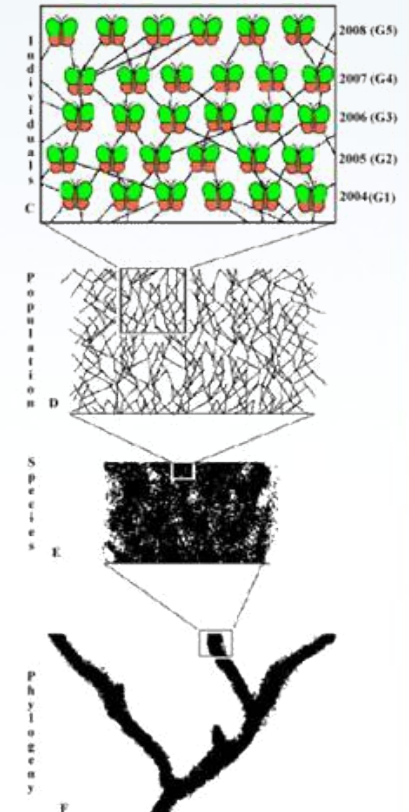


Species phylogeny

A **B** **C**

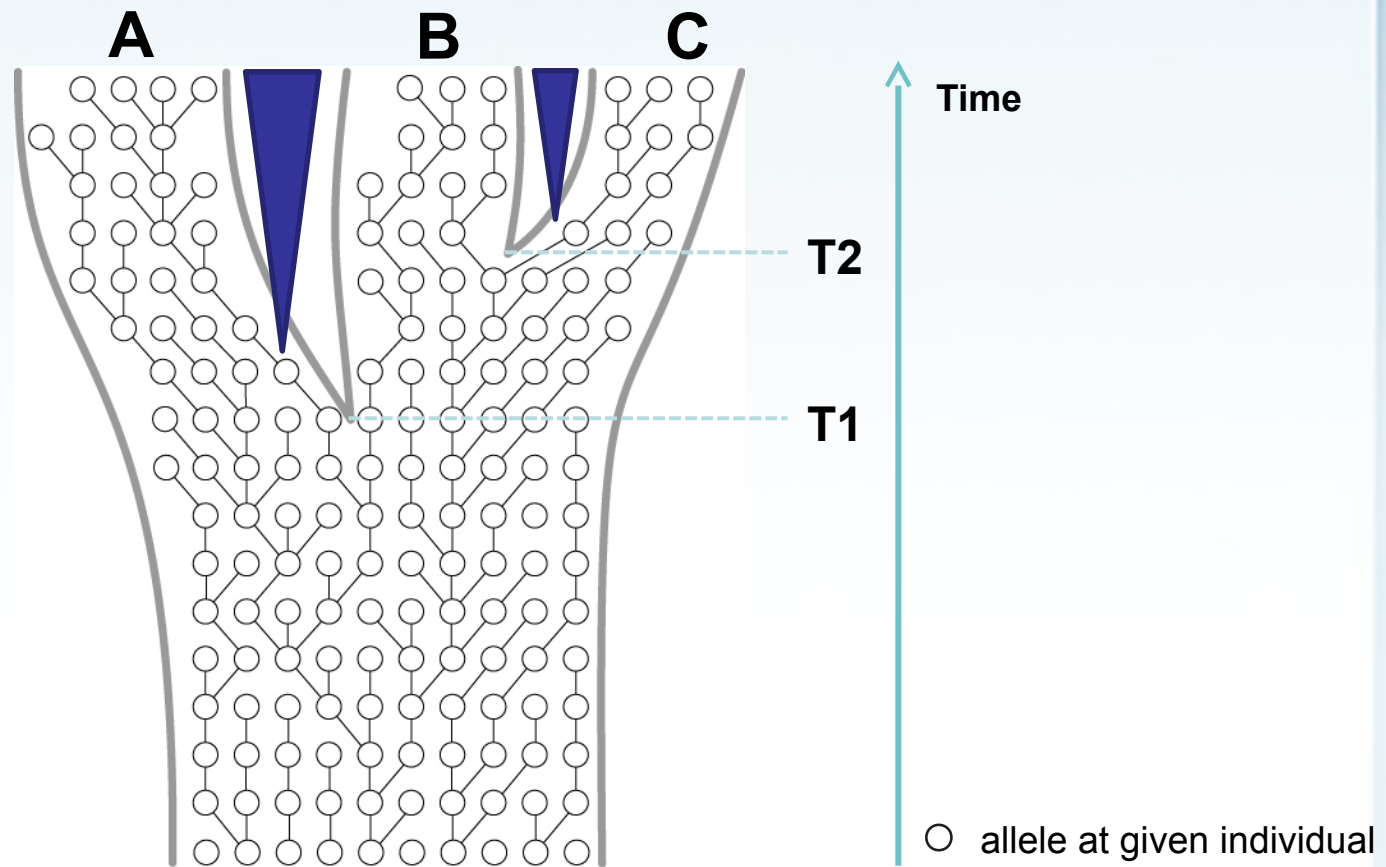


Population genetics



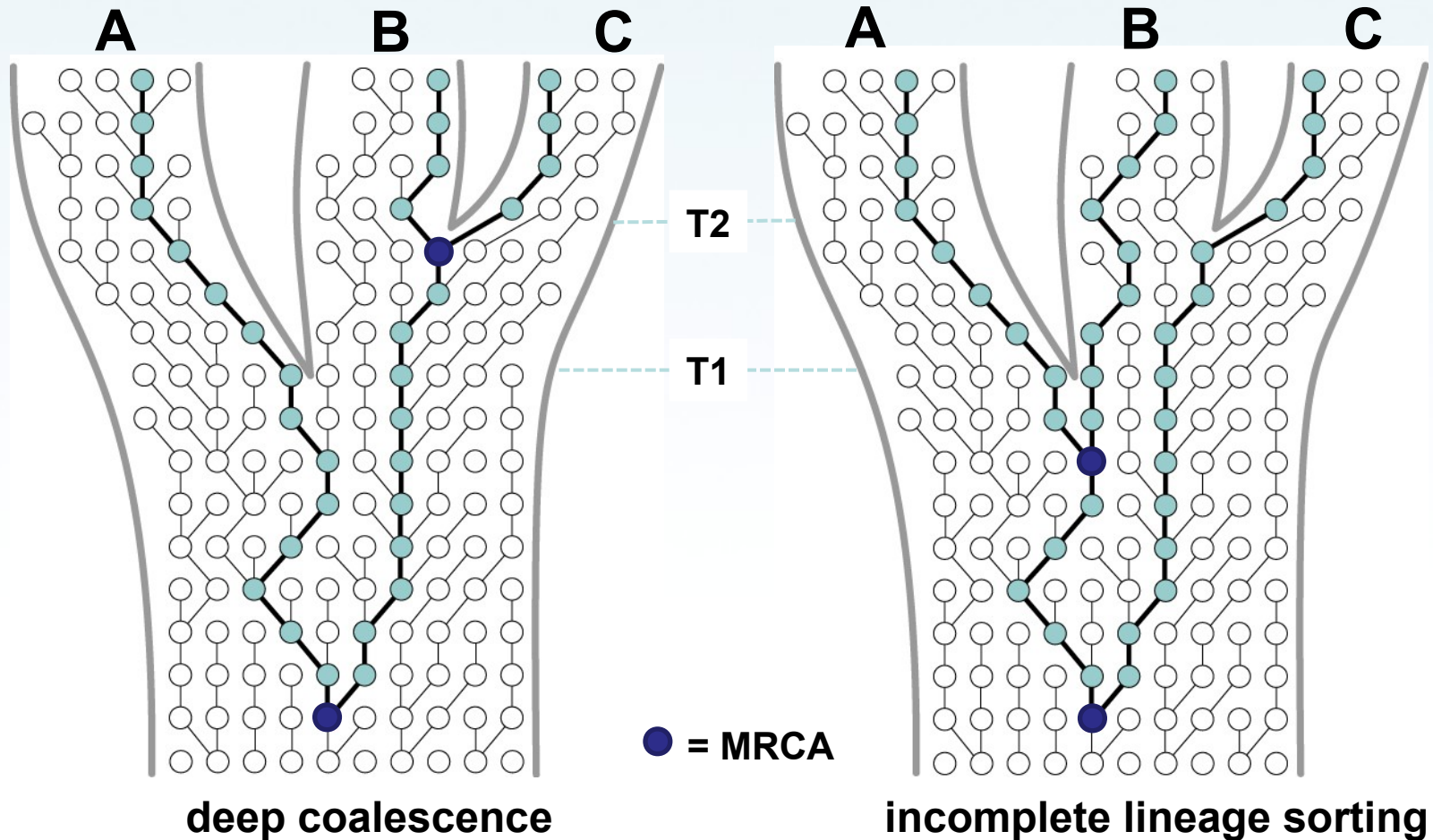
Coalescence-based methods

- Coalescence processes (Wright-Fisher)
 - ∅ Allele transfer among the generations
 - ∅ Allele frequency can change in each generation



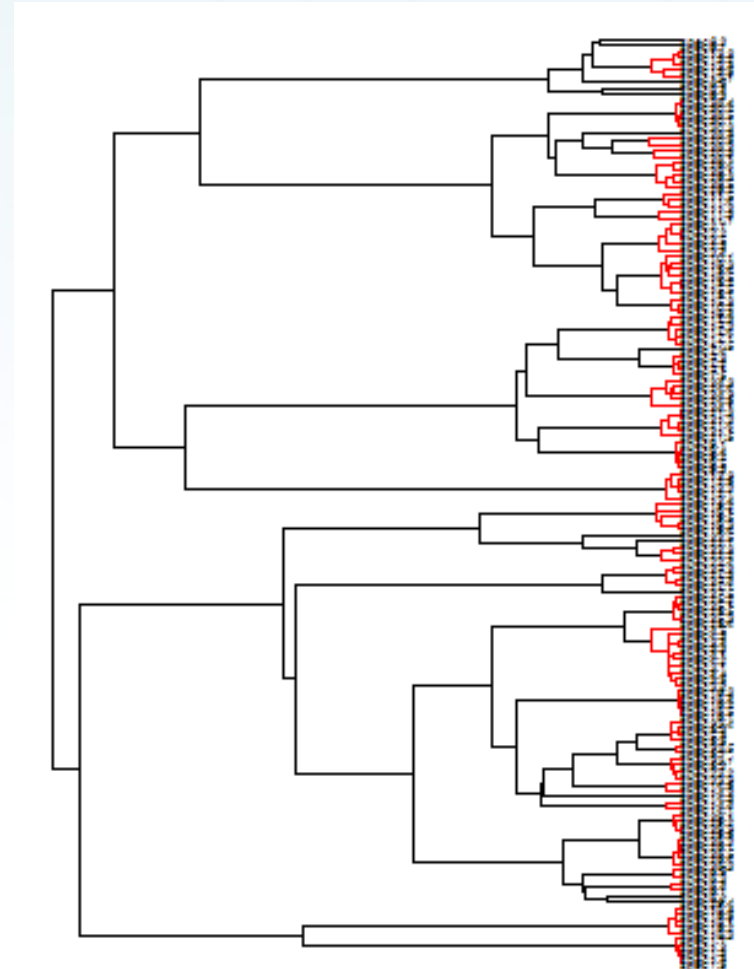
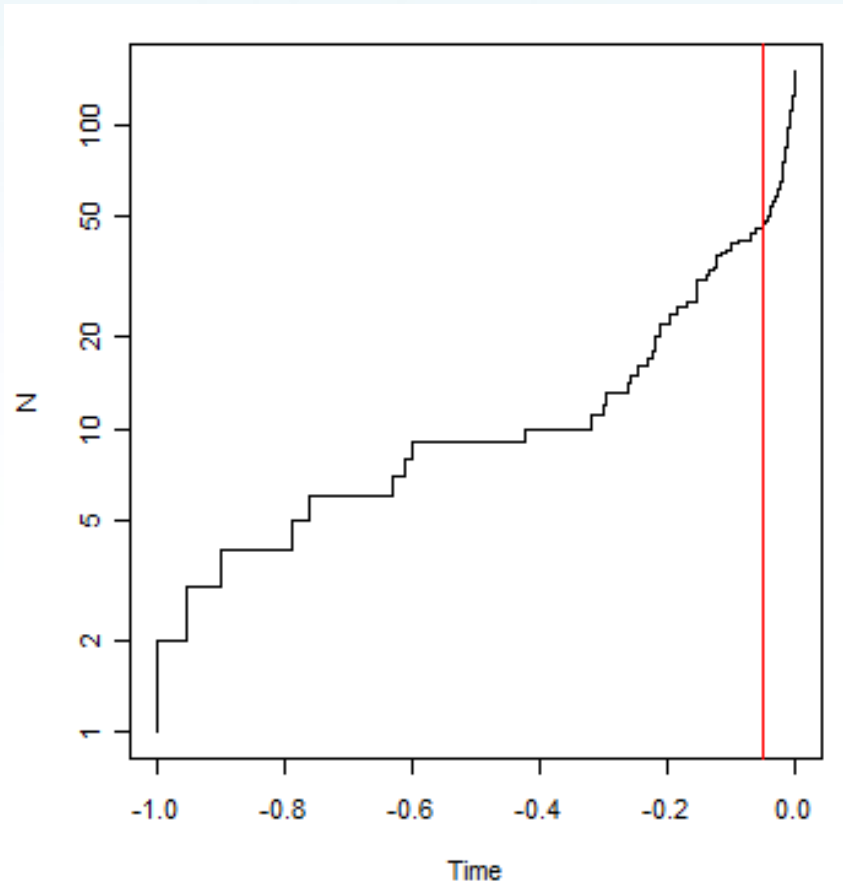
Coalescence-based methods

- Phylogeny at the level of populations and species



GMYC method

- Differences between species-level and population-level evolutionary processes (fitting Yule and coalescence models)
- Detecting the shift between



GMYC method

- Detecting the shift between interspecific and intraspecific branching

∅ Statistical tests

```
>summary(result)
```

```
Result of GMYC species delimitation
```

```
method: single
```

```
likelihood of null model: 858.9326
```

```
maximum likelihood of GMYC model: 872.3755
```

```
likelihood ratio: 26.88581
```

```
result of LR test: 6.220952e-06***
```

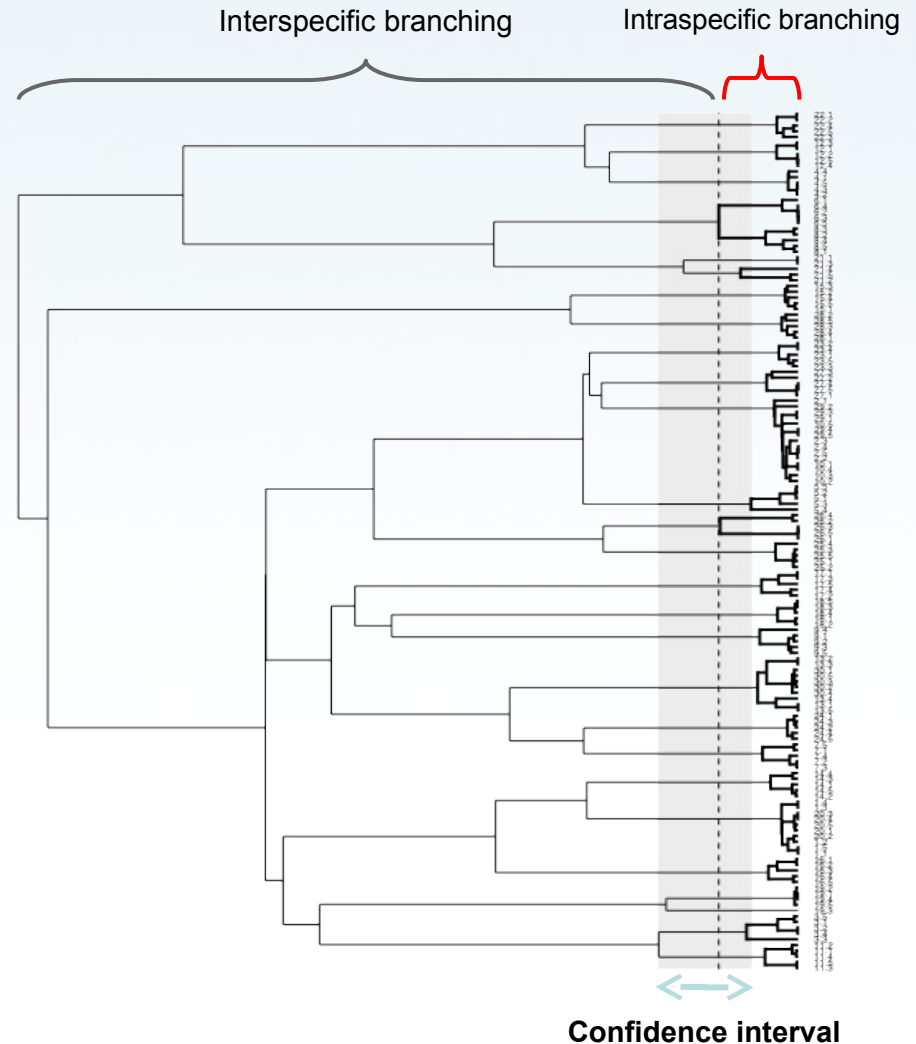
```
number of ML clusters: 33
```

```
confidence interval: 31-34
```

```
number of ML entities: 33
```

```
confidence interval: 31-34
```

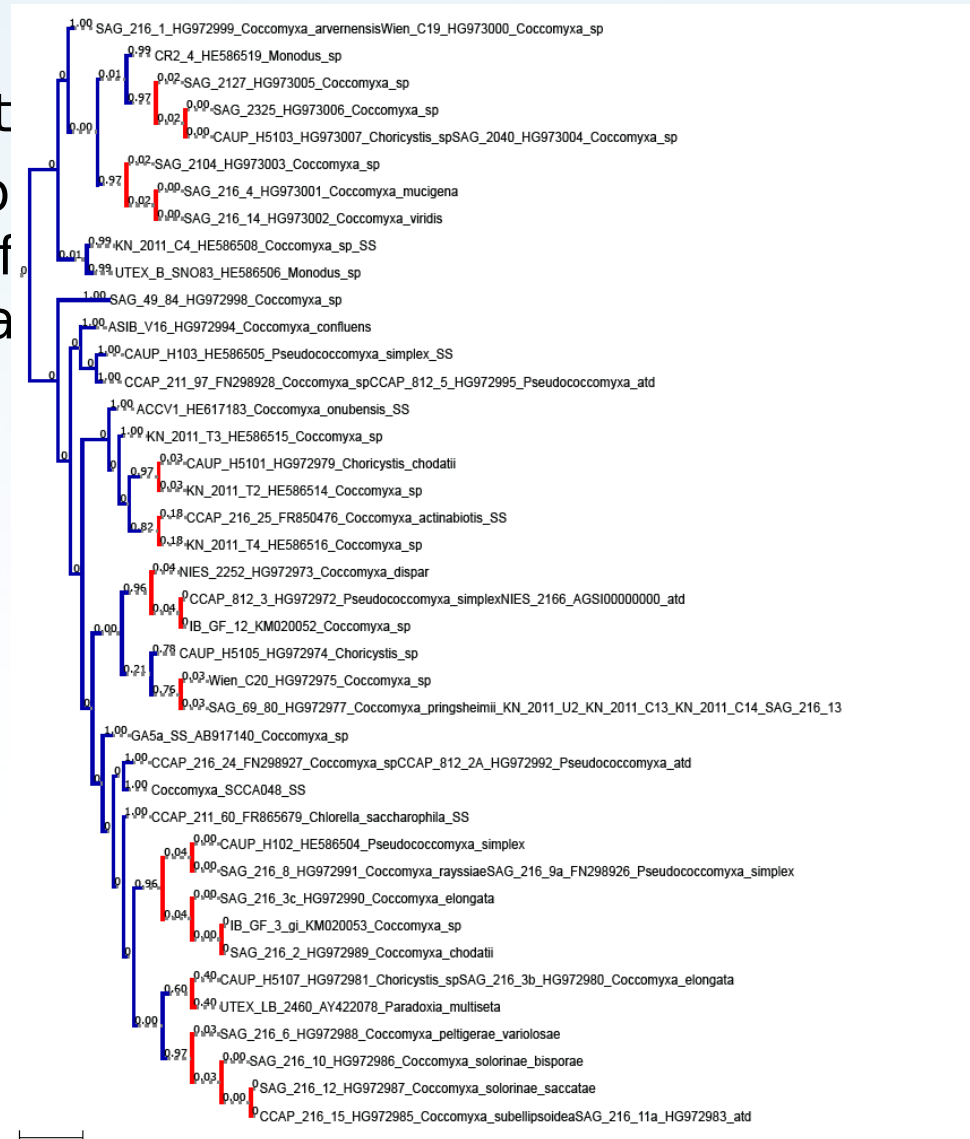
```
threshold time: -0.05174417
```



bPTP

- Bayesian Poisson tree processes method

- Similar to GMYC
- No need of ultrametric tree
- Using directly the number of substitutions (instead of time) to simulate speciation and coalescent events

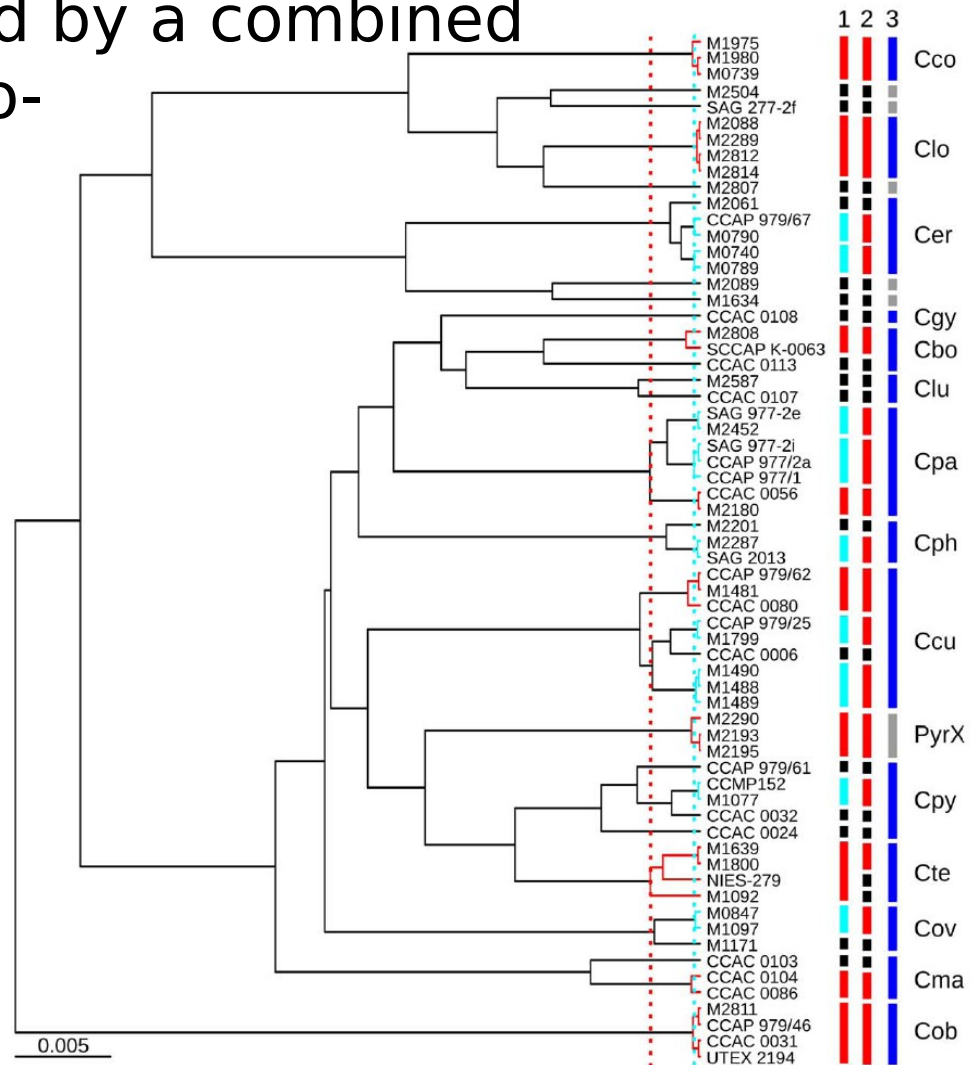


DNA-based species delimitation methods

Reference	Existing species	Species after delimitation	No. of samples	No. of loci	Discovery approaches	Validation approaches
Avila <i>et al.</i> (2006)	6	>12	293	1	Statistical parsimony (NCA)	None
Barrett & Freudenstein (2011)	3	3	162	5	Morphological cluster analysis, PCA	BPP
Burbrink <i>et al.</i> (2011)	1	1	45	3	Structurama	BPP
Camargo <i>et al.</i> (2012)	3	1	505	4	None	spedeSTEM, BPP, ABC
Carstens & Dewey (2010)	3	7	42	6	None	spedeSTEM, Bayes Factors
Carstens & Satler (2013)	1	2	82	21	Structurama, Gaussian Clustering	spedeSTEM, BPP
Duminil <i>et al.</i> (2012)	Unknown	Unstated	103	7	Morphometric clustering; structure	None
Esselstyn <i>et al.</i> (2012)	13	18–19	413	1	GMYC	None
Florio <i>et al.</i> (2012);	1	2	111	1	Canonical variates analysis	None
Flot <i>et al.</i> (2010)	1		74	3	Haplowebs	None
Hamilton <i>et al.</i> (2011)	4	3	147	1	Combo WP and barcoding gap, monophyly, GMYC	
Kelly <i>et al.</i> (2008)	39	1	114	1	WP	None
Leaché & Fujita (2010)	1	3	51	6	Structure	BPP
Leavitt <i>et al.</i> (2012)	19	2	414	6	Structure	BPP, mean genetic distance
Leliart <i>et al.</i> (2009)	19	13	175	1	GMYC, statistical parsimony (NCA)—clades that exceed 95% cut off	None
Niemiller <i>et al.</i> (2012)	1	19	135	9	O'Meara clustering	BPP
Pons <i>et al.</i> (2006)	24	54	468	1	Parsimony network, PAA, CHA, WP, GMYC	None
Puillandre <i>et al.</i> (2009)	1	4	44	2	Elliptic Fourier analysis on shape to the mollusc shell; qualitative phylogenetic evidence	None
Puillandre <i>et al.</i> (2012)	43	27	1000	2	GMYC, ABGD	None

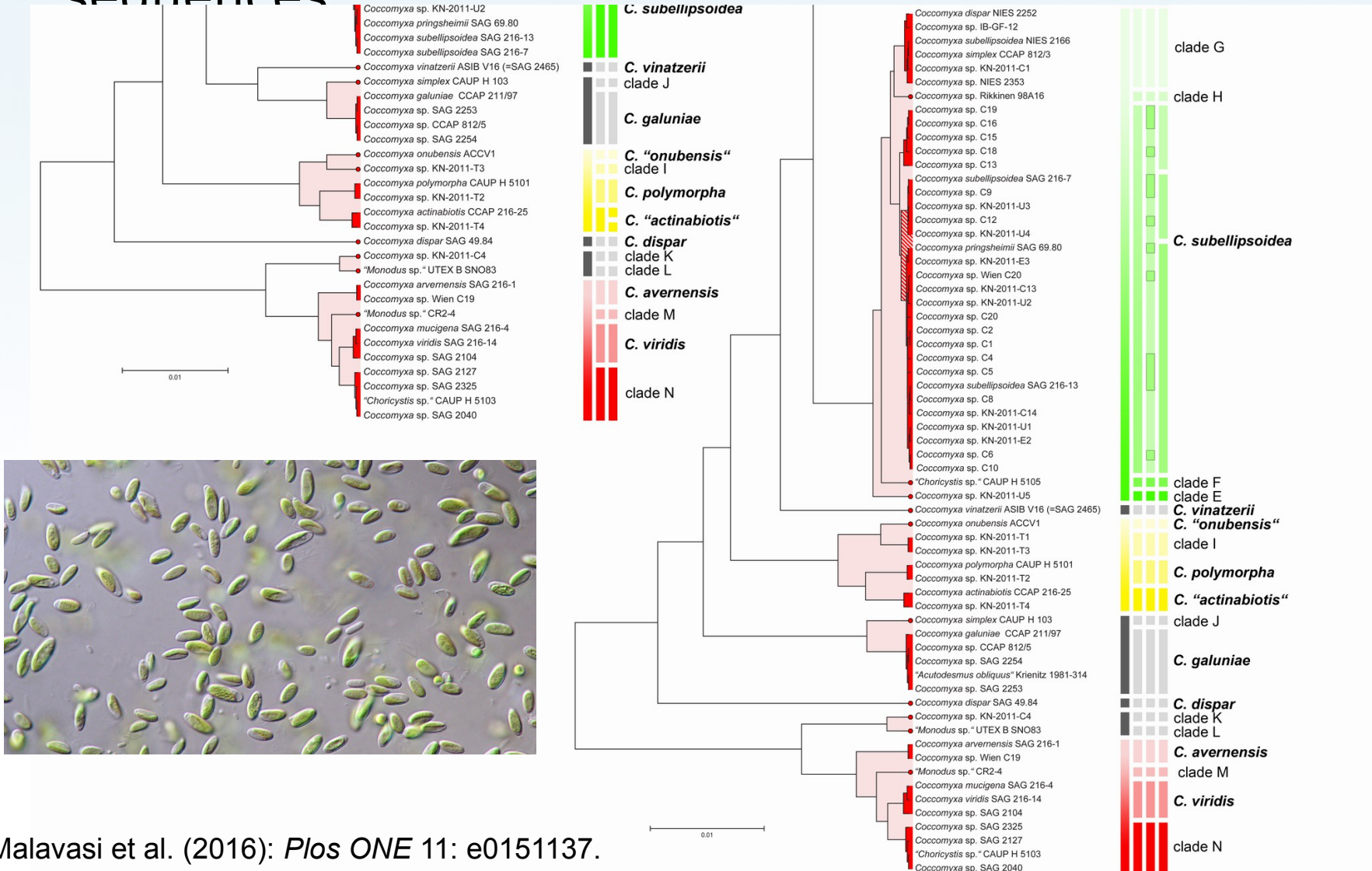
Pitfalls of DNA-based species delimitation

- Incongruence between species delimited by GMYC and by those delimited by a combined molecular and morphological approach



Pitfalls of DNA-based species delimitation

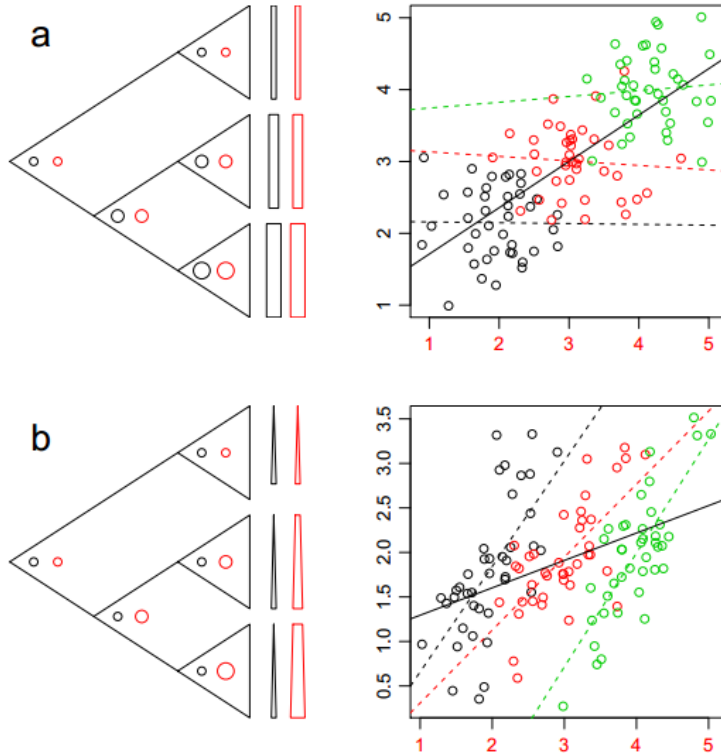
- Problems with taxon sampling, using identical sequences



Trait evolution, PGLS

Tomáš Fér

Is correlation phylogenetically correct? ...phylogenetic autocorrelation...



- strong correlation between two traits due to relatively few independent comparisons
- disappears after phylogenetic correction
- strong correlation between traits masked by phylogenetic differences among groups

- phylogenetic independent contrast (PIC) – Felsenstein (1985)
- phylogenetic generalized least square (PGLS)

PGLS (phylogenetic generalized least square)

- linear regression: $y = \beta X + \varepsilon$
 - y – response variable
 - β – estimated coefficients
 - X – explanatory variable
 - ε – residuals
- in case of related taxa y and X are not independent
- solution – include covariance matrix (V) which defines evolutionary distance between samples

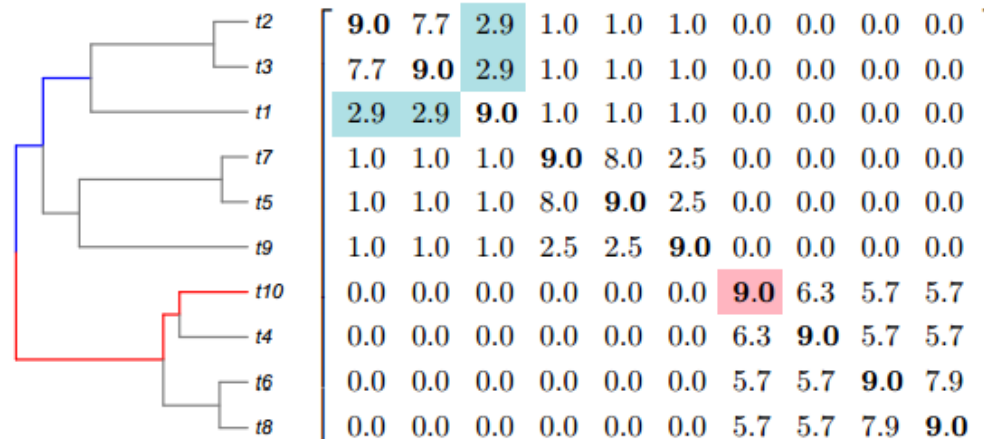


Figure 2: A phylogenetic tree of 10 taxa and the variance covariance matrix (V) of that phylogeny. The diagonal of the matrix (bold values) shows the path length from each tip to the root (example in red). Off diagonal values show the shared path length for a given pair of tips (example in blue).

- *Brownian motion* of trait evolution is expected, i.e., traits evolve proportionally to branch lengths
- not always suitable – three different types of transformation for better model fitting

PGLS – scaling parameters

- lambda (λ) – *phylogenetic association*
 - internal branch length multiplied by a constant
 - $\lambda = 0$: internal branches zero, i.e., variability independent on phylogeny
 - $\lambda = 1$: no scaling, variability copies phylogeny
 - $\lambda > 1$: more covariance than expected by *Brownian motion*
- delta (δ) – *tempo of evolution*
 - power all value to ‘ δ ’ – transform sum of shared distances individuals, i.e., scaling ‘*root to tip*’ distance
 - detection of acceleration of trait evolution
 - $\delta < 1$: early changes important (adaptive radiation), slow changes among closely related
 - $\delta = 1$: gradual evolution, i.e., proportional to branch length
 - $\delta > 1$: later changes important, i.e., acceleration of changes during evolution (species-specific adaptation)
- kappa (κ) – *mode of evolution*
 - power all branch lengths to ‘ κ ’, i.e., scaling of all branches
 - $\kappa < 1$: branch length becoming identical (longer branches shortened more), i.e. variability accumulated only if groups are different
 - $\kappa = 0$: evolution independent on branch length – punctualism
 - $\kappa = 1$: gradualism

PGLS – scaling parameters

Three scaling parameters and their interpretation when applied to trait evolution on a phylogeny

Parameter	Action	0	<1	1	>1
λ (lambda)	Assess contribution of phylogeny	star phylogeny (species independent)	phylogenetic history has minimal effect	default phylogeny	not defined
κ (kappa)	Scale branch lengths in tree	punctuational evolution	stasis in longer branches	default gradualism	longer branches more change
δ (delta)	Scale total path (root to tip) in tree	not defined	temporally early change important (adaptive radiation)	default gradualism	temporally later change (species-specific adaptation)

PGLS – scaling parameters

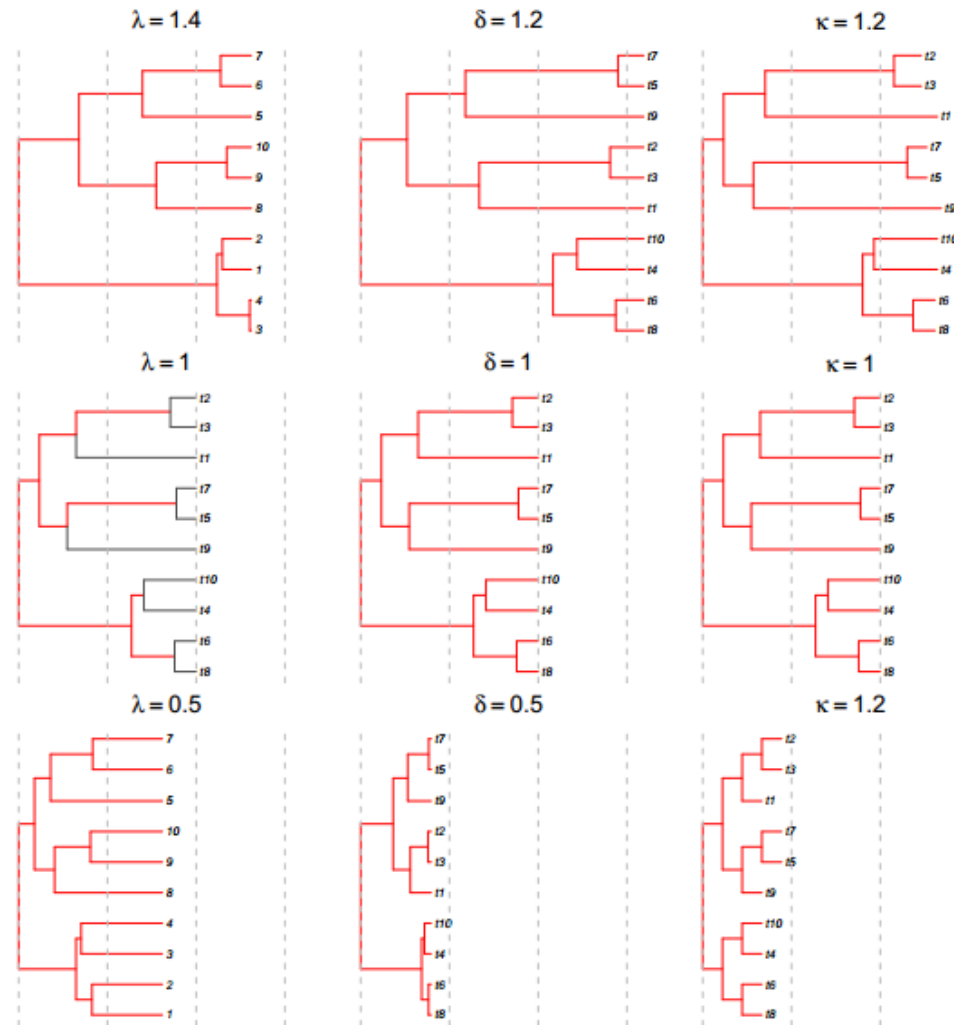
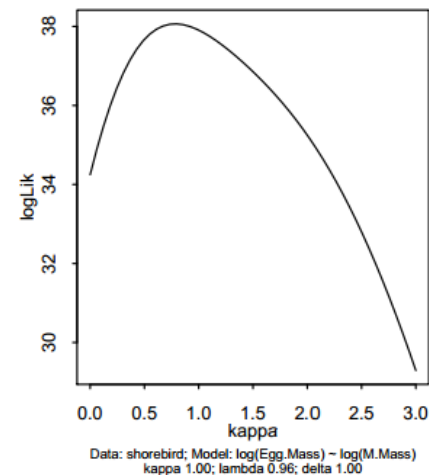
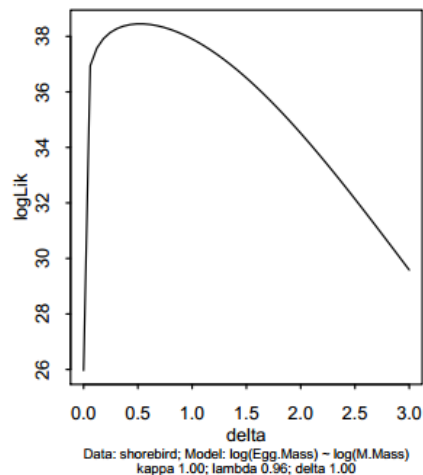
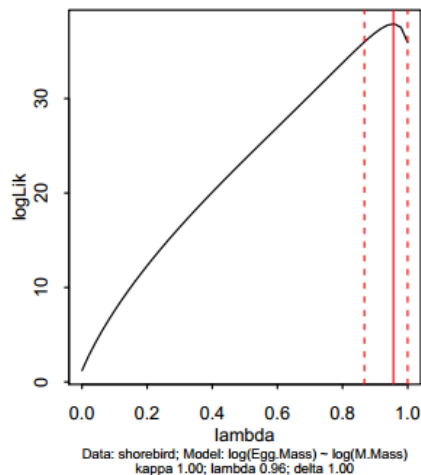


Figure 3: Examples of λ , δ and κ branch length transformations. The branches affected by a given transformation are shown in red.

PGLS

- fitting trait evolution or correlation of two traits with respect to phylogeny
- parameter estimation (λ , δ and κ) for better fit
 - ML – caper (funkce pgls) – value + confidence interval
 - Bayesian – BayesTraits
- parameters can be estimated simultaneously – interpretation of complex model?



Literature

- Pagel M. (1992): *A method for the analysis of comparative data*. Journal of Theoretical Biology, 156 (4):431-442
- Pagel M (1997): *Inferring evolutionary processes from phylogenies*. Zoologica Scripta 26, 331-348
- Freckleton R.P., Harvey P.H. & Pagel M. (2002): *Phylogenetic analysis and comparative data: A test and review of evidence*. American Naturalist, 160:712-726
- Nunn C. L. (2011): *The Comparative Approach in Evolutionary Anthropology and Biology*. The University of Chicago Press. 380 pp.