DNA-based species delimitation

Phylogenetic species concept based on tree topologies

Ø How to set species bound

Ø Automatic species





DNA barcoding

Species boundaries recognized by identifying









DNA barcoding

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







Zoller & Lutzoni (2003): Mol. Phyl. Evol. 29: 629-640

Good & Wake method

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



Sites & Marshall (2012): Trends Ecol. Evol. 18: 462-470

Coalescence-based methods

Linking phylogenetics and population genetics
 Ø Identification of independently evolving
 lineages



Species phylogeny

Population genetics

(G5)

 Four butterflies (top row), connected to their parents (bottom row)

Coalescence-based methods

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



Leliaert et al. (2014): Eur. J. Phycol

Coalescence-based methods

Phylogeny at the level of populations and species



GMYC method

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





GMYC method

Detecting the shift between interspecific and intraspecific branching

Ø Statistical tests



Interspecific branching

Confidence interval

Intraspecific branching

bPTP

- Bayesian Poisson tree processes method
 - Similar to GMYC
 - No need of ultrametric t
 - Using directly the numb substitutions (instead of, time) to simulate specia and coalescent events

```
SAG_216_1_HG972999_Coccomyxa_arvernensisWien_C19_HG973000_Coccomyxa_sp
         99 CR2 4 HE586519 Monodus sp
             <sup>02</sup>-SAG_2127_HG973005
                0,00 SAG 2325 HG973006
                <sup>D,OP</sup>CAUP H5103 HG973007_Choricystis_spSAG_2040_HG973004_Coccomyxa_sp
       0.02-SAG_2104_HG973003_Coccomyxa
<sup>2,29</sup> KN_2011_C4_HE586508_Coccomyxa_sp_3
0,99 UTEX_B_SNO83_HE586506_Monodus_sp
1.00 SAG_49_84_HG972998_Coccomyxa_
1.00 ASIB_V16_HG972994_Coccomyxa_confluent
   <sup>1,00</sup>,CAUP H103 HE586505 Pseudococcomyxa simplex SS
  1,00 CCAP_211_97_FN298928_Coccomyxa_spCCAP_812_5_HG972995_Pseudococcomyxa_atd
     <sup>1,QQ</sup> ACCV1_HE617183_Coccomyxa_onubensis_SS
      <sup>1.00</sup> KN_2011_T3_HE586515_Coccomyxa_sp
            <sup>0.03</sup>-CAUP H5101_HG972979_Choricystis_chodatii
            <sup>0.03</sup>-KN 2011_T2_HE586514_Cocc
              <sup>18</sup> CCAP_216_25_FR850476_Coccomyxa_
                                                     _actinabiotis_SS
            <sup>0,18</sup>KN 2011 T4 HE586516 Coccomvxa
           <sup>0,04</sup> NIES_2252_HG972973_Coccomyxa_dispa
                PCCAP 812 3 HG972972 Pseudococcomvxa simplexNIES 2166 AGSI00000000 atd
                PIB GF 12 KM020052 Coccomyxa sp
           P-78 CAUP H5105 HG972974 Choricvstis
                P.03-Wien_C20_HG972975_Cocco
                <sup>0,03</sup>-SAG 69 80 HG972977_Coccomyxa_pringsheimii_KN_2011_U2_KN_2011_C13_KN_2011_C14_SAG_216_13
   1.00-GA5a_SS_AB917140_Coccomyxa_sp
       <sup>1,00</sup> CCAP_216_24_FN298927_Coccomyxa_spCCAP_812_2A_HG972992_Pseudococcomyxa_atc
      1.00 Coccomyxa_SCCA048_SS
       CCAP_211_60_FR865679_Chlorella_saccharophila_SS
                     P-CAUP H102 HE586504 Pseudococcomyxa simple;
                       SAG 216 8 HG972991 Coccomvxa ravssiaeSAG 216 9a FN298926 Pseudococcomvxa simplex
                       PIB_GF_3_gi_KM020053_Coccomyxa
                      SAG_216_2_HG972989_Coccomyxa_chodat
                   <sup>40</sup>CAUP_H5107_HG972981_Choricystis_spSAG_216_3b_HG972980_Coccomyxa_elongata
                   <sup>40</sup> UTEX LB 2460 AY422078 Paradoxia multiseta
                  <sup>0,03</sup>SAG 216 6 HG972988 Coccomyxa peltigerae variolosae
                      <sup>0,00</sup>-SAG_216_10_HG972986_Coccomyxa_solorinae_bisporae
                           PSAG_216_12_HG972987_Coccomyxa_solorinae_saccata
                           CCAP_216_15_HG972985_Coccomyxa_subellipsoideaSAG_216_11a_HG972983_atd
```

DNA-based species delimitation methods

Reference	Existing species	Species after delimitation	No. of samples	No. of loci	Discovery approaches	Validation approaches
Avila et al. (2006)	6	>12	293	1	Statistical parsimony (NCA)	None
Barrett & Freudenstein (2011)	3	3	162	5	Morphological cluster analysis, PCA	BPP
Burbrink et al. (2011)	1	1	45	3	Structurama	BPP
Camargo et al. (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 et al. (2012)	Unknown	Unstated	103	7	Morphometric clustering; structure	None
Esselstyn et al. (2012)	13	18-19	413	1	GMYC	None
Florio et al. (2012);	1	2	111	1	Canonical variates analysis	None
Flot et al. (2010)	1		74	3	Haplowebs	None
Hamilton et al. (2011)	4	3	147	1	Combo WP and barcoding gap, monophyly, GMYC	
Kelly et al. (2008)	39	1	114	1	WP	None
Leaché & Fujita (2010)	1	3	51	6	Structure	BPP
Leavitt et al. (2012)	19	2	414	6	Structure	BPP, mean genetic distance
Leliart et al. (2009)	19	13	175	1	GMYC, statistical parsimony (NCA)—clades that exceed 95% cut off	None
Niemiller et al. (2012)	1	19	135	9	O'Meara clustering	BPP
Pons et al. (2006)	24	54	468	1	Parsimony network, PAA, CHA, WP, GMYC	None
Puillandre et al. (2009)	1	4	44	2	Elliptic Fourier analysis on shape to the mollusc shell; qualitative phylogenetic evidence	None
Puillandre et al. (2012)	43	27	1000	2	GMYC, ABGD Carstens et al. (2013)	: Mol. Ecol. 22: 4369

Pitfalls of DNA-based species delimitation

Incongruence between species delimited by GMYC • and by those delimited by a combined 123 M1975 M1980 molecular and morpho-M0739 M2504 SAG 277-2f M2088 M2289 logical aproach M2812 M2814 M2807 M2061





Hoef-Emden (2012): Plos ONE 7(8): e43652.

Pitfalls of DNA-based species delimitation

Problems with taxon sampling, using identical



Trait evolution, PGLS

Tomáš Fér

Is correlation phylogenetically correct? ...fylogenetic autocorrelation...



- strong correlation between two traits due to relatively few independent comparisons
- disappearmizí after phylogenetic correction
- strong correlation between traits masked by phylogenetic differences among groups
- phylogenetic independent contrast (PIC) Felsenstein (1985)
- phylogenetic generalized least square (PGLS)

http://cran.r-project.org/web/packages/caper/vignettes/caper.pdf

PGLS (phylogenetic generalized least square)

- linear regression: $y = \beta X + \varepsilon$
 - y –response variable
 - β estimated coefficients
 - X –explanatory variable
 - ε reziduals
 - 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 (\mathbf{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 lenghts
- not always suitable three different types of transformation for better model fitting

http://cran.r-project.org/web/packages/caper/vignettes/caper.pdf

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
 - δ < 1: early changes important (adaptive radiation), slow changes among closely related
 - $\delta = 1$: gradual evolution, i.e., proportional to branch length
 - δ > 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
 - κ < 1: branch length becoming identical (longer branches shortened more), i.e. variability accumulated only if groups are different
 - κ = 0: evolution independent on branch length punctualism
 - κ = 1: gradualism

http://cran.r-project.org/web/packages/caper/vignettes/caper.pdf http://www.evolution.rdg.ac.uk/Files/ContinuousManual.pdf

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					
к (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)					

http://www.evolution.rdg.ac.uk/Files/ContinuousManual.pdf

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.

http://cran.r-project.org/web/packages/caper/vignettes/caper.pdf

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?



http://cran.r-project.org/web/packages/caper/vignettes/caper.pdf

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.